

The Telecommunications and Data Acquisition Progress Report 42-100

October-December 1989

E. C. Posner
Editor

February 15, 1990



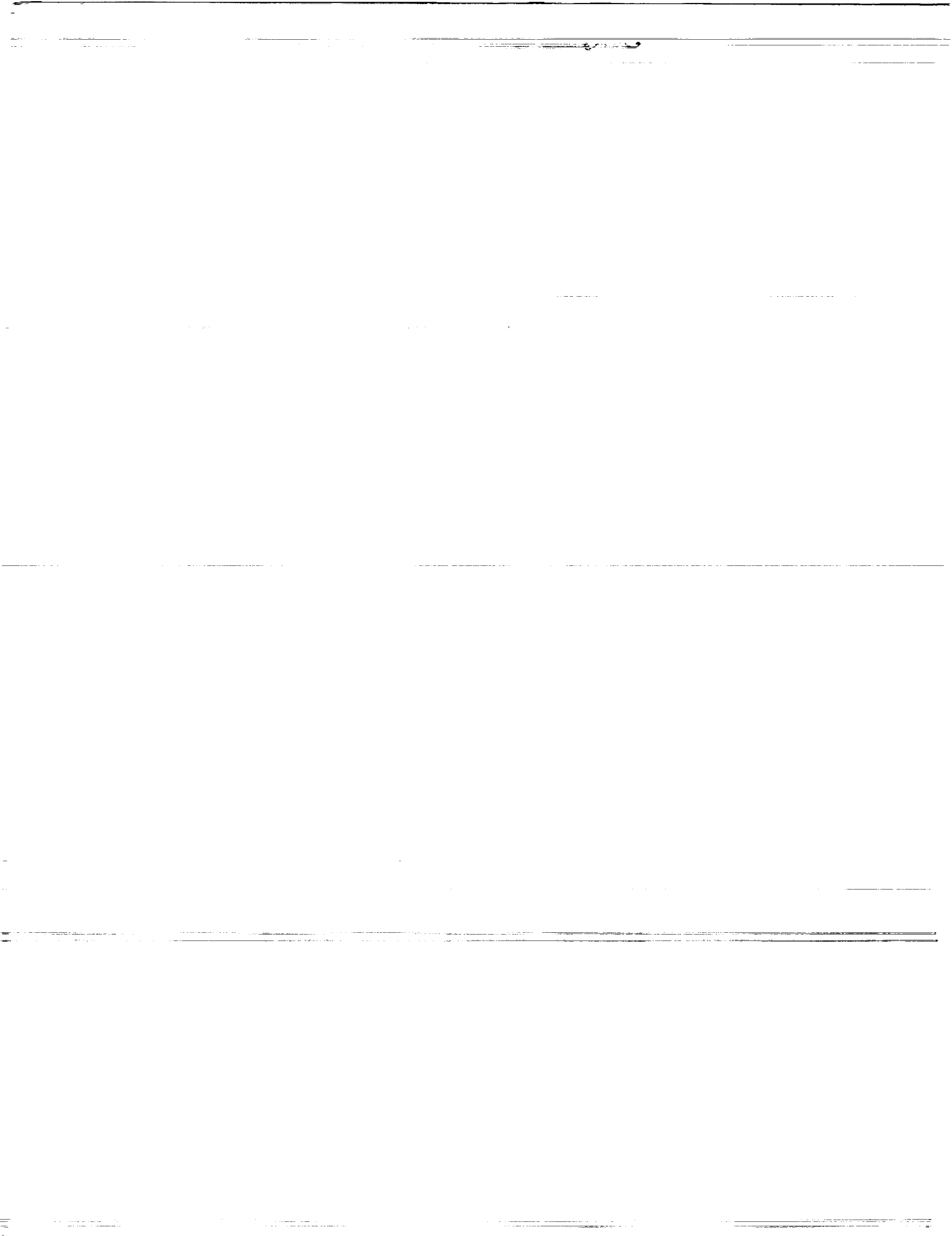
National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

(NASA-CR-186479) THE TELECOMMUNICATIONS AND
DATA ACQUISITION REPORT Progress Report,
Oct. - Dec. 1989 (JPL) 335 p CSCL 17B

N90-21880
--THRU--
N90-21905
Unclas

G3/32 0272316



The Telecommunications and Data Acquisition Progress Report 42-100

October–December 1989

E. C. Posner
Editor

February 15, 1990



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) and its associated Ground Communications Facility (GCF) in planning, in supporting research and technology, in implementation, and in operations. Also included is TDA-funded activity at JPL on data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Operations (OSO). The TDA Office also performs work funded by two other NASA program offices through and with the cooperation of the Office of Space Operations. These are the Orbital Debris Radar Program (with the Office of Space Station) and 21st Century Communication Studies (with the Office of Exploration).

In the search for extraterrestrial intelligence (SETI), the *TDA Progress Report* reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry. These three programs are performed for NASA's Office of Space Science and Applications (OSSA), with support by the Office of Space Operations for the station support time.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund which involve the TDA Office are included.

This and each succeeding issue of the *TDA Progress Report* will present material in some, but not necessarily all, of the following categories:

OSO Tasks:

- DSN Advanced Systems
 - Tracking and Ground-Based Navigation
 - Communications, Spacecraft-Ground
 - Station Control and System Technology
 - Network Data Processing and Productivity
- DSN Systems Implementation
 - Capabilities for Existing Projects
 - Capabilities for New Projects
 - New Initiatives
 - Network Upgrade and Sustaining
- DSN Operations
 - Network Operations and Operations Support
 - Mission Interface and Support
 - TDA Program Management and Analysis
 - Communications Implementation and Operations
 - Data and Information Systems
 - Flight-Ground Advanced Engineering

OSO Cooperative Tasks:

- Orbital Debris Radar Program
- 21st Century Communication Studies

OSSA Tasks:

Search for Extraterrestrial Intelligence

Goldstone Solar System Radar

Radio Astronomy

Discretionary Funded Tasks

Contents

OSO TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

Precise Estimation of Tropospheric Path Delays With GPS Techniques	15
S. M. Lichten	
NASA Code 310-10-61-84-04	
The Determination of Maximum Deep Space Station Slew Rates for a High Earth Orbiter	13
J. A. Estefan	
NASA Code 310-10-63-90-01	
Microwave Analog Fiber-Optic Link for Use in the Deep Space Network	21
R. T. Logan, Jr., G. F. Lutes, and L. Maleki	
NASA Code 310-10-62-16-00	
Galileo Earth Approach Navigation Using Connected-Element Interferometer Phase-Delay Tracking	34
S. W. Thurman	
NASA Code 310-10-63-50-00	
Using Connected-Element Interferometer Phase-Delay Data for Magellan Navigation In Venus Orbit	48
S. W. Thurman and G. Badilla	
NASA Code 310-10-63-50-00	

COMMUNICATIONS, SPACECRAFT-GROUND

Voltage Source AC-to-DC Converters for High-Power Transmitters	55
R. Cormier	
NASA Code 310-20-64-22-00	
Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations	69
V. Galindo-Israel, W. Imbriale, K. Shogen, and R. Mittra	
NASA Code 310-20-65-04-00	
A Real-Time Signal Combining System for Ka-Band Feed Arrays Using Maximum-Likelihood Weight Estimates	81
V. A. Vilnrotter and E. R. Rodemich	
NASA Code 310-30-70-89-01	
Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System	100
D. L. Johnson, S. M. Petty, J. J. Kovatch, and G. W. Glass	
NASA Code 310-20-66-53-00	
Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels	111
T. Y. Otoshi and M. M. Franco	
NASA Code 310-30-69-88-04	
A Functional Description of the Advanced Receiver	131
S. Hinedi	
NASA Code 310-30-70-84-02	
QPSK Carrier-Acquisition Performance in the Advanced Receiver II	150
S. Hinedi and B. Shah	
NASA Code 310-30-70-84-02	

Digital Doppler Extraction Demonstration With the Advanced Receiver	160	513
S. Hinedi, R. Bevan, H. Del Castillo, P. Kinman, D. Chong, and R. Labelle		
NASA Code 310-30-70-04-02		
Exact Closed-Form Expressions for the Performance of the Split-Symbol Moments Estimator of Signal-to-Noise Ratio	174	514
S. Dolinar		
NASA Code 310-30-71-83-04		
A VLSI Decomposition of the deBruijn Graph	180	515
O. Collins, S. Dolinar, R. McEliece, and F. Pollara		
NASA Code 310-30-71-83-02		
Validity of the Two-Level Model for Viterbi Decoder Gap-Cycle Performance	191	516
S. Dolinar and S. Arnold		
NASA Code 310-30-71-83-04		
Finding the Complete Path and Weight Enumerators of Convolutional Codes	203	517
I. Onyszchuk		
NASA Code 310-30-71-83-04		
An Adaptive Vector Quantization Scheme	214	518
K.-M. Cheung		
NASA Code 310-30-71-83-02		
A VLSI Design for a Systolic Viterbi Decoder	221	519
T. K. Truong, E. Satorius, M. T. Shih, and I. S. Reed		
NASA Code 310-30-71-87-02		

DSN Systems Implementation CAPABILITIES FOR EXISTING PROJECTS

Radar RFI at Goldstone DSS 12 and DSS 16	234	520
S. D. Slobin and T. K. Peng		
NASA Code BG 314-40-41-81-11		
Improving a Data-Acquisition Software System With Abstract Data Type Components	248	521
S. D. Howard		
NASA Code 314-40-41-91-93 and 314-40-22-60-12		
Structural Fatigue in the 34-Meter HA-Dec Antennas	252	522
R. A. Van Hek and B. P. Saldua		
NASA Code 314-40-44-14-46		

CAPABILITIES FOR NEW PROJECTS

Mark IV-A DSCC (Magellan-Era) Telemetry System Description	263	523
D. L. Ross		
NASA Code 314-40-41-81-13		
A Higher Density VLBI Catalog for Navigating Magellan and Galileo	274	524
J. S. Ulvestad, O. J. Sovers, and C. S. Jacobs		
NASA Code 314-30-57-34-17 and 314-30-51-34-34		
An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation	301	525
D. L. Trowbridge, J. R. Loreman, T. J. Brunzie, and R. Quinn		
NASA Code 314-30-66-11-09		
Author Index, 1989	320	526

Precise Estimation of Tropospheric Path Delays With GPS Techniques

S. M. Lichten

Tracking Systems and Applications Section

Tropospheric path delays are a major source of error in deep space tracking. However, the tropospheric-induced delay at tracking sites can be calibrated using measurements of Global Positioning System (GPS) satellites. A series of experiments has demonstrated the high sensitivity of GPS to tropospheric delays. A variety of tests and comparisons indicates that current accuracy of the GPS zenith tropospheric delay estimates is better than 1-cm root-mean-square over many hours, sampled continuously at intervals of six minutes. These results are consistent with expectations from covariance analyses. The covariance analyses also indicate that by the mid-1990s, when the GPS constellation is complete and the Deep Space Network is equipped with advanced GPS receivers, zenith tropospheric delay accuracy with GPS will improve further to 0.5 cm or better.

I. Introduction

Signal delays originating in the troposphere can seriously affect the major radiometric data types used for deep space navigation. The wet troposphere in particular is one of the major error sources for Very Long Baseline Interferometry (VLBI) [1,2,3]. Uncalibrated tropospheric delays, typically 5 cm at zenith, can also limit the accuracy for Doppler and ranging systems. For example, the fluctuating troposphere component is the limiting error for Magellan navigation [2], and the systematic component, if calibrated to only 5 cm accuracy at zenith, also becomes a limiting Magellan navigation error source at low elevations.

The Deep Space Network (DSN) has obtained and is presently installing advanced Global Positioning System

(GPS) dual-band receivers for ionospheric calibrations [4]. GPS data are also used elsewhere for a wide variety of precise positioning applications, including satellite orbit determination and ground-based geodetic studies. There are presently seven developmental and four operational GPS satellites in high-Earth orbit (20,000-km altitude) and by the mid-1990s, the 21-satellite operational constellation will be complete. In some GPS applications, uncalibrated tropospheric delays can be a serious error source and must be estimated from the GPS data along with other adjusted parameters [5]. Substantial improvement in GPS orbit and ground-station coordinate accuracy has resulted from the use of a sequential square-root filter estimation approach for tropospheric calibration [5,6,7]. It follows that stochastic GPS estimates of the tropospheric path delay at DSN sites could also be provided by the same DSN GPS receivers that are used for the ionospheric calibrations. In

this article, we report on recent results based on GPS experimental data acquired between 1985 and 1988. These data have been used previously to determine high-accuracy GPS orbits and ground-station coordinates as part of a demonstration of GPS positioning techniques [6,7]. In the present study, however, attention is focused on the accuracy of the time-varying GPS tropospheric delay estimates determined along with the precise GPS orbits. Potential benefits for calibration of deep-space radiometric data are discussed.

II. Experimental Data

The data used for the GPS troposphere study were obtained in GPS experiments in November 1985 and January 1988. Most of the results are based on the 1988 experiment. In the November 1985 experiment, data from nine GPS receivers deployed in North America were used, spanning two weeks. The nine sites included the locations of six VLBI observatories [Hatcreek, CA; Mojave, CA; Owens Valley Radio Observatory (OVRO), CA; Fort Davis, TX; Richmond, FL; and Haystack, MA] and three sites in Mexico. For a more detailed description of the 1985 experiment, see [8]. In the 1988 experiment, data from several dozen locations in the South Pacific; North, Central, and South America; and Europe were collected over a three-week period [9]. The 1988 troposphere results shown in this article are based on data in a one-week interval from North America, South America, and Europe. The 1988 North American sites included five of the six VLBI sites occupied in the earlier 1985 experiment.

In the November 1985 experiment, most stations collected data for about 8 hrs each day. Since a portion of this period had unfavorable geometry, typically 5–8 hrs of high-quality tropospheric delay estimates were determined. In the January 1988 experiment, due to better geometry, one additional satellite, and the larger tracking network, the longer view periods enabled tropospheric delay determination over periods of up to 12 hrs. In these experiments, water vapor radiometers (WVRs) were operated at some of the Mexican and South American sites, and at Haystack and Mojave on selected days. These WVR data were used for comparison purposes to assess the GPS tropospheric delay measurements.

III. Approach and Results

The GPS techniques used to determine tropospheric delays at ground tracking sites are described in [5,6,7]. A key element in the estimation of parameters such as tropospheric path delays and station coordinates using GPS is

the accuracy of the GPS orbits. To minimize orbit error, three reference sites were held fixed at their VLBI coordinates and all other ground site positions were estimated simultaneously along with other parameters. Multiday arcs were used to further improve orbit modeling. The zenith tropospheric delay was estimated at each ground site. For the November 1985 data, dry tropospheric delay calibrations (from surface pressure measurements, as described in [4]) were applied to the data, and, when available, WVR calibrations were also applied to correct for the wet path delay [5]. GPS corrections to these calibrations were estimated with a stochastic model that treats the zenith tropospheric delay as a random walk. If the GPS and WVR wet troposphere measurements were in agreement, the GPS corrections determined in this fashion should be zero.¹ With the January 1988 data, the procedure was varied slightly in order to further test the sensitivity of GPS observations to the troposphere: Here, *no troposphere calibrations* were applied, so the GPS data were used to estimate the *entire* (wet and dry) zenith path delay. These delays were then compared to the sum of the WVR measured delay and the dry delay (from surface measurements). The time-varying tropospheric zenith delays from GPS were estimated in a square-root filter simultaneously with other parameters such as station coordinates, GPS orbital epoch states, station and satellite clocks, and GPS carrier phase biases.

The WVR measures the brightness temperature of the atmosphere. WVRs utilize at least two radio frequencies. One frequency is near water resonance spectral features. The WVR antenna is moved through a variety of elevations (tipping curves) so that the elevation dependence of the water vapor content can be determined, and from that the water vapor content can be determined based on a homogeneous model of the atmosphere. A description of the method for determining wet path delays from WVRs can be found in [5], which also contains numerous other references both for the instruments themselves and for the water vapor content retrieval algorithms. The accuracy of modern three-frequency WVRs is believed under good conditions to be better than 1 cm. However, the accuracy of wet tropospheric delays determined from WVR measurements can be compromised by various hardware effects (electronic drifts, erroneous antenna temperatures); uncertainties arising in the retrieval algorithm (see [5] and references therein); and in some cases can be affected by

¹ Another approach involved applying only dry calibrations and then comparing the GPS estimate of the wet delay to the WVR estimate of the wet delay. This produces essentially the same result as the method described above (applying wet and dry calibrations, then estimating a correction with GPS), as long as the a priori covariance on the GPS-determined tropospheric delay parameters is unconstrained (large).

the amount of liquid water present in the form of rain, fog, or clouds. The retrieval algorithm [10] used in this study incorporates surface meteorological data to constrain the temperature profile of the atmosphere. The WVR tipping data were combined to produce an equivalent zenith wet path delay. This wet delay and the dry zenith delay (the latter determined from surface barometric pressure data), when used to calibrate GPS data, were mapped to the appropriate GPS line-of-sight elevations with an analytic formula [11]:

$$\rho = \rho_z R_d(\theta) + \rho_{zw} R_w(\theta) \quad (1)$$

where ρ_z is the zenith tropospheric path delay, R is the analytic mapping function, θ is the elevation angle, and w or d refers to the wet or dry components. To first order, both functions R_d and R_w behave as $1/\sin(\theta)$.

GPS data are also sensitive to tropospheric path delays since all the GPS signals pass through the troposphere prior to ground reception. The GPS receivers have the benefit of tracking multiple GPS satellites simultaneously through a variety of different elevations. The GPS carrier-phase data type has several-millimeter precision and when continuously tracked, which is the normal procedure, is sensitive to subcentimeter atmospheric fluctuations. By fitting tropospheric delay parameters to the induced signatures in the pre-fit residuals, GPS data can determine an effective zenith delay. Experiments with GPS data [5,6,7] as well as covariance analysis [12] have demonstrated that estimates of GPS orbits and of GPS-determined ground station coordinates are very sensitive to uncalibrated tropospheric delays. For example, wet tropospheric delays would be the dominant error for GPS orbits [13] and for ground baselines [12] if the data were uncalibrated and no troposphere parameters were estimated.

Figure 1 shows the difference between the GPS and WVR wet zenith troposphere delay estimates at two Mexican sites. The Mazatlan WVR was an R04 model, which uses two frequencies (20.7 and 31.4 GHz), while the Cabo San Lucas model was a J01 model, which is a newer instrument with three frequencies (20.7, 22.3, and 31.4 GHz), has lower noise, and is believed to be more reliable than the R04. The Mazatlan GPS estimates are in good agreement with the WVR, with better than 1-cm root-mean-square (rms) difference, although there appears to be a slight bias offset between the WVR and GPS estimates. Note that an error in the dry calibration would be absorbed into the GPS wet troposphere estimate, thereby leading to a potential bias between the GPS and WVR wet delays. Since the accuracy of the dry calibration is believed to be several millimeter [5], such biases would

presumably be small. The comparison at Cabo San Lucas shows excellent agreement between GPS and WVR with just a few millimeter difference. Figure 2 compares the GPS-determined *total* tropospheric zenith delay with the sum of WVR wet+dry calibrations for Mojave, CA, and for Limon, Costa Rica. In both these cases, the WVR and GPS zenith delays agree to one centimeter or better (rms). Again, there is a small (several-millimeter) bias offset between the GPS and WVR+dry estimates of the tropospheric delay. The WVRs used at Mojave and at Limon were both two-frequency models (type R07, which is an older model, and type D1, which is a newer model). Figure 3 shows comparisons of total tropospheric delay with GPS and the WVR+dry zenith delay for Haystack, MA, where a modern three-frequency J01 WVR was used. In addition to showing subcentimeter agreement, it is notable that Fig. 3 shows that *both* the GPS and WVR techniques seem to track the *same* rapid variations in zenith delay on both days. These variations of several centimeters over several hours are unusually large and were rarely seen even in much more humid sites in South and Central America where the total wet path delays were often larger than 30 cm. Nearly all the variation seen in Fig. 3 is due to changes in the wet troposphere.

Figure 4 compares daily baseline repeatability in the Gulf of California using GPS troposphere calibrations and using WVR troposphere calibrations. In a majority of the baselines, the GPS calibrations improved the rms baseline scatter over the rms scatter with WVR calibrations, particularly in the vertical component. These results suggest that calibrations based on GPS estimates of tropospheric delays, in some cases, may be more accurate than those obtained from WVRs and separate dry tropospheric calibrations. One advantage of the GPS tropospheric delay estimates is that they can absorb and correct for any dry calibration errors, while the WVR measures only the wet component and the final calibration will still include any dry calibration error. Another factor that may be relevant to Fig. 4 is that the GPS troposphere estimates are averages along the same lines of sight from which GPS observations were used to determine the baselines, while the WVR and dry calibrations are essentially averages over the whole sky. Due to spatial inhomogeneities, it might be expected that the GPS calibrations would be more correct to use with the GPS data if their intrinsic accuracy were comparable to the accuracy of the WVR calibrations. One additional possibility, which cannot be evaluated presently, is that the GPS tropospheric delay parameters are absorbing some other, unknown systematic error in the GPS measurement system. In light of the close agreement between WVR and GPS troposphere calibrations, however, the author believes that this latter possibility is unlikely.

IV. Discussion

A. Elevation Dependence

In the data analyzed here, GPS scans were made as low as 7 deg from the horizon. The WVR elevation cutoff, however, was about 25 deg. The relatively high WVR cutoff is necessary to avoid ground spillover contamination of the signal (the WVR beamwidths were about 8 deg). For DSN applications, this elevation cutoff may be a concern since intercontinental VLBI observations often require low elevation measurements (below 10 deg) for at least one site. It is desirable to make tropospheric delay calibrations in these cases using data at low elevations. The ultimate low elevation limit for GPS observations is probably determined by ground multipath. Approximately 2 percent of the GPS experimental data analyzed here were below 15 deg. Ground-station coordinates estimated with and without the low elevation data agreed to about 1.5 cm or better. When the GPS data were reprocessed and the low elevation data were *excluded*, daily baseline repeatability *worsened* slightly (by ~ 1 cm) in the vertical for four out of six baselines involving the relatively wet Mexican sites in the Gulf of California. The GPS-VLBI 2000-km baseline (between Hatcreek, CA and Fort Davis, TX) comparison was also slightly *worse* when the GPS low elevation data were *excluded*. The best results were obtained using *all* the GPS data, including low elevation data. Other baselines, such as the Mojave-Owens Valley 245-km baseline in California also showed better GPS-VLBI agreement when *all* the GPS data were used. These results seem to indicate that such low elevation data can actually enhance GPS system accuracy, probably through improved troposphere calibrations. The capability to include low elevation data is a definite advantage for using GPS measurements to calibrate tropospheric delays.

B. Flexibility

The GPS technique is flexible and can be combined with other tropospheric delay calibration methods if they are available. For instance, GPS data can be used to estimate the *entire* tropospheric delay, wet+dry.² Or, if dry

calibrations are available, GPS estimates of just the wet contribution can be made. With WVRs, not only are separate and simultaneous dry calibrations required, but any error in the dry calibration will bias the final answer. However, the GPS wet tropospheric delay can be estimated as a correction on top of the dry calibration, so it will absorb any dry calibration errors easily and correct for them also by lumping them into the effective wet path delay.

GPS observations measure the mean tropospheric path delay for the lines of sight to satellites viewed simultaneously at a given time. The measurement intervals can be relatively dense (every second) or sparse (every five minutes), depending on how often the troposphere calibrations are needed. In the GPS field experiments, data rates were typically between two and five measurements per minute, and the data were later compressed to six-minute intervals. With the currently used ground tracking networks, which tend to be rather sparse, and the relatively small number of transmitting GPS satellites, formal errors computed in the Kalman filter for the GPS tropospheric delays were between 0.5 and 2 cm with multiday arcs, with 1.25 cm a typical value. This would apply to possible *biases* in the measured troposphere delays. An upper limit on the formal uncertainty for *point-to-point variations* (GPS points are six minutes apart) in the troposphere parameters is set by the process noise model used, about 0.3 cm in this instance. This is consistent with the centimeter-level and subcentimeter agreement shown in Figs. 1 through 3, which indicate, in fact, that these formal errors, at least for the bias portion, may be conservative.

A covariance analysis was performed to determine what expected performance would be in the 1990s with a full GPS constellation (21 satellites plus three spares) and a worldwide tracking network consisting of advanced GPS receivers at the three DSN sites and seven other sites worldwide. A similar worldwide network with at least six ground sites will be operating in support of the TOPEX/POSEIDON mission [14], scheduled for launch in 1992. Furthermore, a similar geodetic worldwide GPS ground network is presently operating [15] and this network is expected to expand in the near future. The data from this network are and will continue to be distributed to interested users. The covariance analysis predicts that with a 12-hr GPS tracking arc, tropospheric zenith delays from GPS will have an accuracy of 0.2–0.5 cm over the entire data arc. Figure 5 shows an error budget for the tropospheric delay accuracy at a representative time in the middle of a 12-hr GPS pass based on a consider error analysis in which the DSN station locations, the geocenter, and gravitational constant (GM) are considered as systematic (unadjusted) error sources. The expected root-sum-square

² When the *entire* delay was estimated in this study with a single parameter, the dry mapping function was used. In principle, this is not entirely correct, since the dry mapping function would apply to ~ 90 –95 percent of the total delay and the wet mapping function should be used for the remaining ~ 5 –10 percent portion due to water vapor. A more exact approach would be to estimate a wet delay parameter and a dry delay parameter with approximate nominal values of 200 cm and 15 cm, respectively. A calculation (D. Tralli, 1989, personal communication), however, shows that the error resulting from using the dry mapping function for the entire delay was a few millimeters or less for the geometries in the GPS experiments in this study, and in fact was too small to be detected.

(RSS) error is about 0.3 cm for the zenith tropospheric path delay estimate returned every 10 min. This assumes that the GPS-determined DSN baselines are accurate to 3 cm per component and that GPS can determine the geocenter to 5 cm per component (see Table 1). Present-day accuracy for GPS baselines of several thousand kilometers is about 2 cm [8], and it is expected that several-centimeter accuracy will be achieved over intercontinental baselines in the near future with GPS. GPS measurements presently show little sensitivity to the geocenter because of the sparseness of the current constellation, but by the mid-1990s, it is expected that the geocenter will be determined to better than 5 cm from GPS observations [16].

There are other possibilities for using GPS tropospheric delay estimates in conjunction with other techniques. As suggested in [17], it may be promising to combine GPS calibrations with WVR data for determination of fluctuating dry and wet path delays. Another approach utilizes GPS data to calibrate out possible biases that can affect WVR measurements. Or, as Figs. 1 through 3 suggest, in the absence of WVRs, GPS technology alone can provide centimeter-level tropospheric delay calibrations even with only a partial GPS constellation and ground network. This represents about a factor of five improvement over available calibrations at the DSN from surface data [2]. The ultimate, yet-to-be determined limitation on the GPS calibrations for deep-space tracking will probably be due to the fact that the GPS lines of sight do not, in general, coincide with the lines of sight to the spacecraft of interest [17].

C. Operational Considerations

Two major considerations for operational tropospheric calibration at the DSN are the amounts of temporal and spatial information on the troposphere provided by various different calibration techniques.

The GPS troposphere estimates provide a thorough time history of zenith tropospheric delays since the tropospheric parameters are adjusted stochastically in a square-root Kalman filter using continuous GPS data. When a worldwide tracking network is used, however, there can be some delay expected in bringing all the data together for processing. For applications where near-real-time turnaround is needed (such as a planetary mission encounter), there are several possibilities for using GPS. Technology is presently being developed by the National Aeronautics and Space Administration (NASA) for continuously operating GPS networks [18], which return GPS orbits and related parameter estimates in less than a day. The new technology involves new hardware and software

inside the GPS receivers that can handle in real time much of the processing that is currently done after data from different sites are brought together. Although these continuously operating networks are being designed for monitoring crustal motion and advanced earthquake detection, the same technology could be used for near-real-time monitoring of tropospheric delays and satellite navigation. Since GPS orbit prediction with better than 1-m accuracy has been demonstrated [8], such predictions might make simultaneous orbit/troposphere estimation unnecessary, considerably reducing the amount of calculation needed to determine tropospheric delays from GPS observations at the DSN. Considerations such as these will be studied in future analyses.

In principle, a WVR can be operated in a co-pointing mode to better calibrate line-of-sight tropospheric path delays in the direction of the spacecraft being tracked. The accuracy of line-of-sight WVR calibrations, however, has yet to be demonstrated. In practice, however, line-of-sight WVR calibrations will not be effective below the minimum elevation angle for the WVR, which is presently about 25 deg due to the large radiometer beamwidth. This reduces considerably the scope for using co-pointing WVRs since most intercontinental VLBI observations and many one- or two-way DSN tracking observations are made at lower elevations. It is expected that low-elevation WVR performance will improve in the future with the development of WVRs with narrower beamwidths. One relevant future study would compare the accuracy of GPS tropospheric delay calibrations and WVR calibrations for VLBI and/or conventional DSN observations made at low elevation. There may be some advantage to using the GPS data due to the inclusion of low elevation data. A co-pointing WVR would have the advantage of being directed towards the general area of the sky of the target spacecraft, but would have the disadvantage of having to use data from higher elevations only. In each case, a compromise is made, and a detailed study of the trade-offs would be desirable. Perhaps the ultimate troposphere calibration system would utilize some combination of GPS and pointed WVR measurements, appropriately weighted. There are also methods for minimizing the error in the GPS calibration by mapping troposphere delays from the GPS closest to the target spacecraft [17].

V. Summary and Conclusions

A demonstration of a centimeter-accurate estimation of time-varying tropospheric path delays using GPS observations and a square-root Kalman filter has been completed. The accuracy of the GPS zenith delays is compara-

ble to that of delays determined from WVRs. The WVR- and GPS-determined zenith tropospheric delays agree at the centimeter-level or better, a five-fold improvement over present-day calibrations routinely available at the DSN. The GPS results are consistent with the formal errors from covariance analysis, and as the GPS constellation is filled out in the early 1990s with worldwide ground tracking, the GPS zenith tropospheric delay estimates should further improve to better than 0.5 cm.

The use of GPS as a tropospheric calibration system for the DSN is attractive for a number of reasons. These include: the use of GPS hardware already procured and used at the DSN for ionospheric calibrations; high-precision and complete sky coverage of GPS; flexibility of GPS-based calibrations, with a capacity for both wet and dry tropospheric calibrations; continuous tracking of GPS, and therefore continuous return of fluctuating tropospheric delay estimates as a function of time; high sensitivity of GPS data to the troposphere and the capability

to measure path delays in the presence of clouds, fog, and even rain; and the possibility for eventual near-real-time turnaround with advanced receivers currently being developed elsewhere by NASA. Although the GPS calibrations require the presence of a ground tracking network, such networks are presently operating and will be considerably expanded and improved by the time the GPS constellation is complete. The data from these networks are presently distributed to users worldwide, and in the future could be used to complement the primary GPS data collected at the DSN. The primary disadvantage of GPS troposphere calibrations is expected to be that GPS lines of sight and the line of sight to the spacecraft being tracked will not, in general, coincide.

Since GPS data are easily obtained and GPS-based tropospheric calibration appears to be so promising, it should be pursued as a potentially important technique at the DSN that can substantially reduce the effect of tropospheric delay errors on deep-space observations.

References

- [1] C. D. Edwards, "The Effect of Spatial and Temporal Wet-Troposphere Fluctuations on Connected Element Interferometry," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 47-57, May 15, 1989.
- [2] C. D. Edwards, "A Statistical Study of the Differential Very Long Baseline Interferometry Delay Rate Error Budget," paper AAS 89-403, AAS/AIAA Astrodynamics Specialist Conference, San Diego, California, pp. 1-19, Aug. 7-10, 1989.
- [3] R. N. Treuhaft and G. E. Lanyi, "The Effect of the Dynamic Wet Troposphere on Radio Interferometric Measurements," *Radio Science*, vol. 22, pp. 251-265, 1987.
- [4] G. E. Lanyi and T. Roth, "A Comparison of Mapped and Measured Total Ionospheric Electron Content Using Global Positioning System and Beacon Satellite Observations," *Radio Science*, vol. 23, pp. 483-492, July-Aug. 1988.
- [5] D. M. Tralli, T. H. Dixon, and S. A. Stephens, "Effect of Wet Tropospheric Path Delays on Estimation of Geodetic Baselines in the Gulf of California Using the Global Positioning System," *Journal Geophys. Research*, vol. 93, no. B6, pp. 6545-6557, June 10, 1988.
- [6] S. M. Lichten and J. S. Border, "Strategies For High Precision GPS Orbit Determination," *Journal of Geophysical Research*, vol. 92, pp. 12751-12762, Nov. 10, 1987.

- [7] S. M. Lichten, "Estimation and Filtering Techniques for High-Accuracy GPS Applications," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1-20, May 15, 1989.
- [8] S. M. Lichten and W. I. Bertiger, "Demonstration of Sub-Meter GPS Orbit Determination and 1.5 Parts in 10⁸ Three-Dimensional Baseline Accuracy," *Bulletin Geodesique*, vol. 63, pp. 167-189, 1989.
- [9] R. E. Neilan, T. H. Dixon, T. K. Meehan, W. G. Melbourne, J. A. Scheid, J. N. Kellogg, and J. L. Stowell, "Operational Aspects of CASA UNO'88—The First Large Scale International GPS Geodetic Network," *IEEE Transactions on Instrumentation and Measurement*, vol. 38, pp. 648-651, April 1989.
- [10] S. E. Robinson, "A New Algorithm for Microwave Delay Estimation from Water Vapor Radiometer Data," *TDA Progress Report 42-87*, vol. July-September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 149-157, November 15, 1986.
- [11] G. Lanyi, "Troposphere Calibration in Radio Interferometry," *Proceedings of the International Symposium on Space Techniques for Geodynamics*, J. Somogyi and C. Reigber (eds.), IAG/COSPAR, p. 184, July 1984.
- [12] T. P. Yunck, W. G. Melbourne, and C. L. Thornton, "GPS-Based Satellite Tracking System for Precise Positioning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. GE-23, pp. 450-457, July 1985.
- [13] W. Bertiger, S. C. Wu, J. S. Border, S. M. Lichten, B. G. Williams, and J. T. Wu, "High Precision GPS Orbit Determination Using March 1985 Demonstration Data," paper AIAA-86-0089, AIAA 24th Aerospace Sciences Meeting, New York, New York, pp. 1-6, January 6-9, 1986.
- [14] S. C. Wu, T. P. Yunck, and G. A. Hajj, "Toward Decimeter TOPEX Orbit Determination Using GPS," paper AAS89-359, AAS/AIAA Astrodynamics Conference, San Diego, California, pp. 1-13, Aug. 7-10, 1989.
- [15] M. S. Schenewerk and G. L. Mader, "Orbit Determination and Improvement at NGS," *EOS, Transactions of the AGU*, vol. 69 (presented at the Spring 1988 AGU Meeting, May 1989, Baltimore, Maryland), p. 323, April 19, 1988.
- [16] R. P. Malla and S. C. Wu, "Recovery of a Geocentric Reference Frame Using the Present-day GPS System," paper AAS89-356, AAS/AIAA Astrodynamics Conference, San Diego, California, pp. 1-10, Aug. 7-10, 1989.
- [17] R. N. Treuhaft and G. M. Resch, "Tropospheric Monitoring Technology for Gravity Wave Experiments," *Relativistic Gravitational Experiments in Space*, R. Hellings (ed.), National Aeronautics and Space Administration Conference Publication CP3046, Workshop on Relativistic Gravitational Experiments, Annapolis, Maryland, pp. 179-185, June 1989.
- [18] U. J. Lindqwister, G. Blewitt, and T. P. Yunck, "Continuously Operating GPS Networks (Co-op Nets)," *EOS, Transactions of the AGU*, vol. 70, AGU Meeting, Baltimore, Maryland, p. 301, April 11, 1989.

Table 1. Assumptions for covariance analysis

Estimated parameters	A priori uncertainty
GPS orbital states	2 m per component (position) 0.2 mm/s per component (velocity)
GPS solar pressure parameters	0.25 scale for x, z coefficients 10^{-12} km/s Y-bias parameter
Non-DSN station coordinates	10 cm per component
Clocks and carrier phase biases	1 km (clocks estimated as white noise)
Wet zenith tropospheric path delay	10 cm (bias portion) 1.2 cm/ $\sqrt{\text{hr}}$ random walk stochastic model
Consider (unadjusted) parameters	Consider sigma
DSN station coordinates	3 cm per component
Knowledge of geocenter	5 cm per component
GM	1 part in 10^8
Data Noise: 5-cm pseudorange 0.5-cm carrier phase 1 meas/10 min	

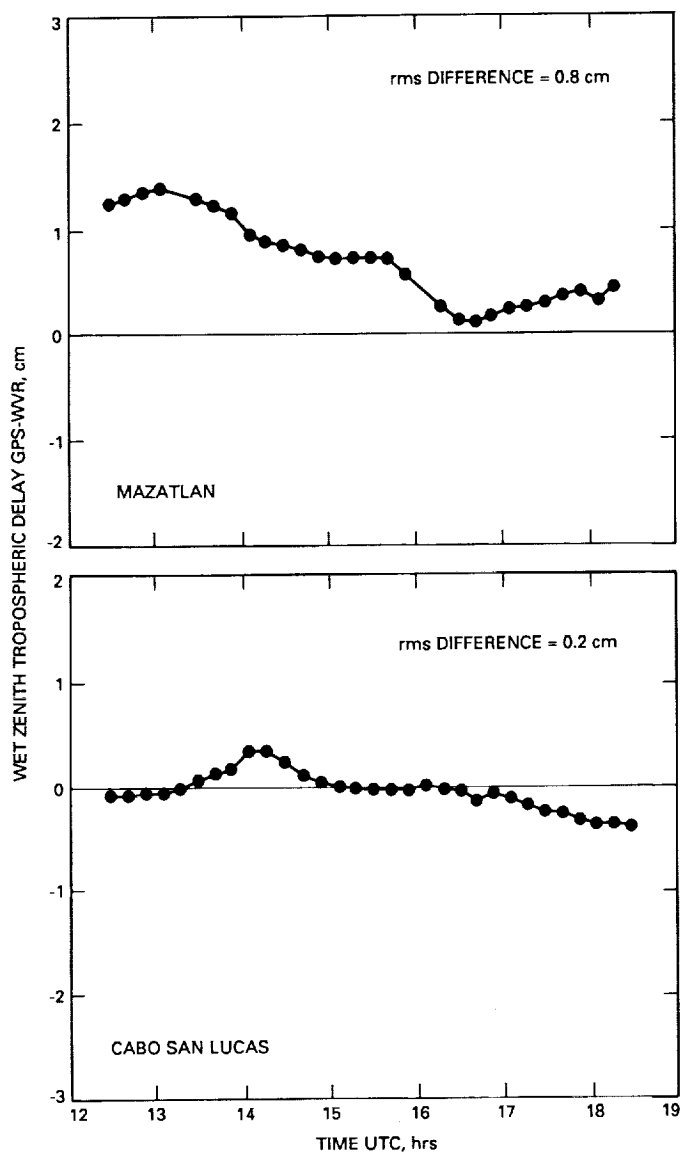


Fig. 1. Difference between GPS and WVR estimates for the zenith wet tropospheric delay estimated on November 22, 1985 at Mazatlan and on November 18, 1985 at Cabo San Lucas. Both sites are in relatively humid locations in Mexico, and during the experiment, total wet path delays of 30–40 cm were not uncommon.

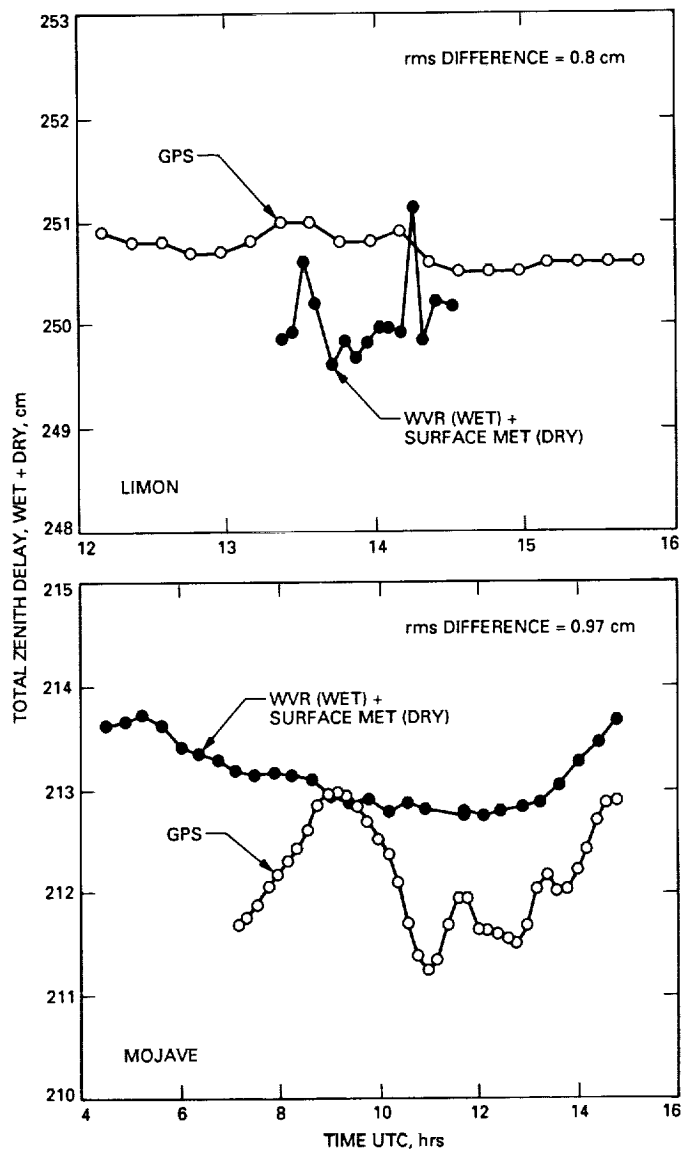


Fig. 2. Comparison of estimates of total tropospheric path delays determined using GPS and using the sum of WVR-measured wet delays plus dry delays from surface data. The comparison is shown for Limon, Costa Rica (a relatively humid site) and for Mojave, California (a relatively dry site) both on January 21, 1988.

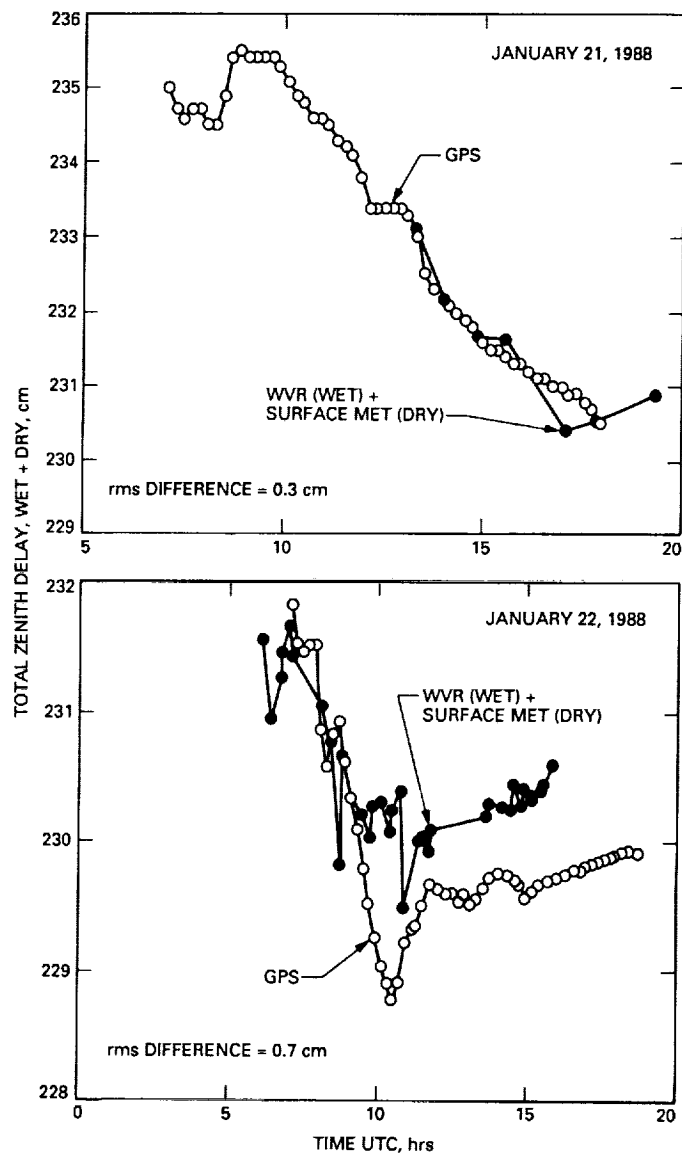


Fig. 3. Comparison of *total* tropospheric path delays from GPS and WVR+dry measurements, similar to Fig. 2, shown for the site near Haystack Observatory in Massachusetts on January 21 and 22, 1988. The subcentimeter agreement between different tropospheric delay measurement methods is sustained through periods of rapid fluctuations in the wet path delay.

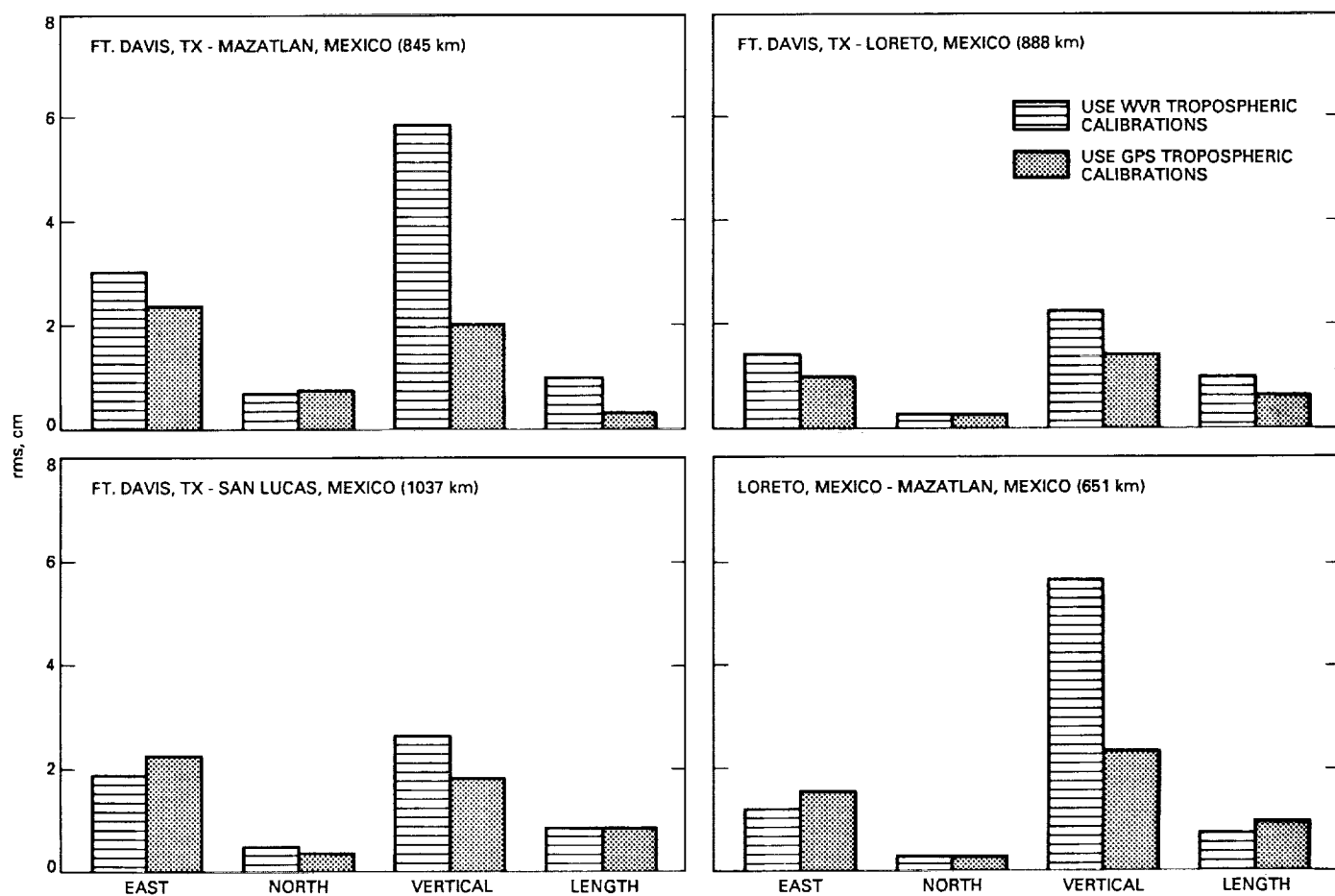


Fig. 4. Ground baseline daily repeatability from two-week GPS orbit solution arcs comparing two methods of tropospheric delay calibration: (1) using GPS, and (2) using WVRs. In most instances, the GPS calibrations resulted in lower rms scatter in the baseline estimates.

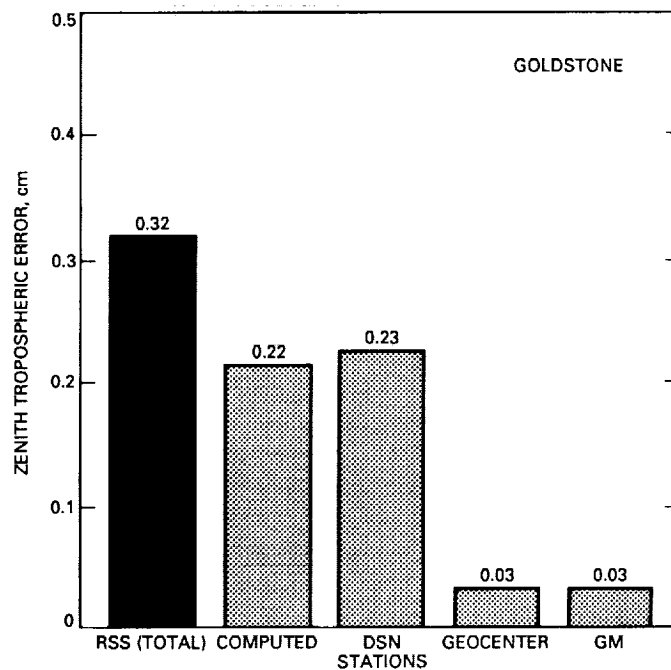


Fig. 5. Predicted error budget for zenith troposphere determination at a DSN site (Goldstone) based on the assumptions in Table 1. The "computed" error is from the square-root Kalman filter and implicitly includes effects from GPS orbits, data noise, geometry, and other estimated parameters. The other error sources are from the consider analysis and reflect quantities which are not expected to be estimated simultaneously with the tropospheric parameters: DSN station coordinates, relative location of the geocenter, and the value of GM (Earth's mass).

52-32
272318
88
N90-21882

The Determination of Maximum Deep Space Station Slew Rates for a High Earth Orbiter

J. A. Estefan
Navigation Systems Section

As developing national and international space ventures, which seek to employ NASA's Deep Space Network (DSN) for tracking and data acquisition, evolve, it is essential for navigation and tracking system analysts to evaluate the operational capability of Deep Space Station antennas. To commission the DSN for use in tracking a highly eccentric Earth orbiter could quite possibly yield the greatest challenges in terms of slewing capability; certainly more so than with a deep-space probe. This article focuses on the determination of the maximum slew rates needed to track a specific high Earth orbiter, namely, the Japanese MUSES-B spacecraft of the Very Long Baseline Interferometry Space Observatory Program. The results suggest that DSN 34-m antennas are capable of meeting the slew rate requirements for the nominal MUSES-B orbital geometries currently being considered.

I. Introduction

The Institute for Space and Astronautical Science (ISAS) has proposed the implementation of an orbiting radio telescope observatory mission known as the Very Long Baseline Interferometry (VLBI) Space Observatory Program (VSOP) for launch in the mid-1990s timeframe. The program is structured to be an international collaboration of the VLBI radio astronomy communities of the United States, Europe, and Australia, as well as other interested nations, with the intention of providing a worldwide observatory network.

ISAS and NASA will undertake a cooperative effort for precise tracking of VSOP's MUSES-B spacecraft by commissioning NASA's Deep Space Network (DSN) together with various Japanese ground stations.

Due to the highly eccentric orbit proposed for MUSES-B, concern arose as to whether or not DSN 34-m antennas would be capable of slewing fast enough to follow the spacecraft. An effort was undertaken to formulate an analytic model for the determination of antenna slew angles (i.e., azimuth and elevation angles), as well as antenna slew rates, and produce a computer program for generating their corresponding time histories. Analysis of different scenarios together with knowledge of the operational slewing capability of the Deep Space Station (DSS) antennas would thereby yield tracking feasibility results.

This article details the modeling and analysis process of this effort as well as results and implications of a few selected mission scenarios. Results indicate that DSN 34-m antennas are indeed capable of meeting slewing

requirements for MUSES-B's nominal mission orbit and, more importantly, further preliminary studies can now be readily performed for future planned missions.

II. Modeling Assumptions and Analysis

In order to arrive at an analytical formulation of the problem, the spacecraft orbit was modeled with two-body dynamics. The Earth's precession of equinoxes was *not* modeled. Classical orbital elements were assumed time-invariant as were the cylindrical coordinates of the station position.¹ The derivation of the azimuth and elevation angles in a (north, east, down) topocentric-horizon reference frame yielded inverse trigonometric functions of the station-to-spacecraft slant range vector. The actual equations are omitted here for brevity. Once the analytical expressions for the slant range vector were derived, analytical expressions for the time-rates-of-change of azimuth and elevation could be determined. It should be noted that expressions for the slant range vector components were derived as functions of the time-varying parameters of station right-ascension, $\theta(t)$, and spacecraft eccentric anomaly, $E(t)$; together with the time-invariant parameters of the classical elements, station position components, and associated direction-cosine terms.

Despite the fact that a simple model was used for this study, the analytical expressions for azimuth and elevation angle determination and their corresponding time-rates-of-change were too lengthy to obtain a time history of results by hand. The expressions were mechanized in a computer program to facilitate rapid evaluation of different cases. Input/Output specifications are listed below.

A. Inputs

The inputs are r_s , λ_{sta} , z_0 , θ_{g0} , ω_\oplus , a , e , Ω , i , ω , μ_\oplus , E_0 , t_0 , dt , t_{STOP} where

- r_s = station distance from Earth's spin axis (km)
- λ_{sta} = station longitude (east of prime meridian) (deg)
- z_0 = station distance above equatorial plane (km)
- θ_{g0} = Greenwich mean sidereal time (sec)
- ω_\oplus = Earth's rotation rate (rad/sec)
- a = semi-major axis (km)
- e = eccentricity
- Ω = longitude of the ascending node (deg)

i = inclination (deg)

ω = argument of perigee (deg)

μ_\oplus = Earth's gravitational parameter (km³/sec²)

E_0 = initial spacecraft eccentric anomaly (sec)

t_0 = specified epoch time (sec)

dt = time increments (sec)

t_{STOP} = program stop time (sec)

B. Outputs

The outputs are t , $\theta = \theta(t)$, $E = E(t)$, $Az = Az(\theta, E, t; p)$, $El = El(\theta, E, t; p)$, $\dot{t} = \dot{\theta}(t)$, $\dot{E} = \dot{E}(t)$, $\dot{Az} = \dot{Az}(\theta, E, t; p)$, $\dot{El} = \dot{El}(\theta, E, t; p)$ (\dot{Az})_{max}, (\dot{El})_{max} where

t = time (hrs)

θ = station right-ascension (deg)

E = spacecraft eccentric anomaly (deg)

Az = azimuth angle (deg)

El = elevation angle (deg)

\dot{Az} = azimuth rate (deg/sec)

\dot{El} = elevation rate (deg/sec)

p = input parameter set

III. Results and Implications

A case-by-case synopsis is provided for three relevant mission scenarios. Tables of input parameters are included as are plotted slew angle time histories. The referenced epoch was taken to be January 1, 1995 at 0^{hr}0^{min}0^{sec} and duration of each run was seventy-two hours. The orbit period in all cases was approximately 6.06 hours.

A. Case 1: (MUSES-B Nominal Orbit; DSS 10)²

The Deep Space Station represented in this case is located at Goldstone, California. Input parameters are listed in Table 1 while plotted results of slew angles are illustrated in Figs. 1 and 2.

Maximum slew rates for this case were found to be approximately

0.14 deg/s in azimuth

0.10 deg/s in elevation

B. Case 2: (MUSES-B Nominal Orbit; DSS 40)

This scenario was identical to Case 1 with the exception that the Deep Space Station represented is located at

¹ T. Moyer, "Station Location Sets Referred to the Radio Frame," JPL IOM 314.5-1334 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 24, 1989.

² A. Konopliv, "Preliminary Orbit Determination Analysis for the VSOP Mission," JPL IOM 314.5-648 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 9, 1989.

Canberra, Australia. The input parameters are listed in Table 2. Plotted results of slew angles are illustrated in Figs. 3 and 4.

0.14 deg/s in azimuth
0.08 deg/s in elevation

Maximum slew rates for this case were found to be approximately

0.14 deg/s in azimuth
0.09 deg/s in elevation

C. Case 3: (MUSES-B Nominal Orbit; DSS 60)

This scenario was also identical to Case 1 with the exception that the Deep Space Station represented is located at Madrid, Spain. The input parameters are listed in Table 3. Plotted results of slew angles are illustrated in Figs. 5 and 6.

Maximum slew rates for this case were found to be approximately

IV. Conclusions

Maximum slew rates for a DSN 34-m antenna are roughly 0.40 deg/s in both azimuth and elevation *and*, hence, the 34-m antenna would certainly be capable of meeting slewing requirements for any of the case scenarios described above.

Station view periods can also be determined by examining elevation angle time histories that are generated by the computer program. Fig. 7 provides an example of this capability, which represents Cases 1 through 3 employing an elevation cutoff angle of ten degrees above the local horizon.

Acknowledgment

The author would like to express a note of appreciation for those who assisted in this work effort. In particular, a warm thanks to Alex Konopliv for providing initial guidance to problem modeling. A very special thanks to Sam Thurman for providing invaluable guidance and assistance in all aspects of the study.

Table 1. Input parameter set for Case 1

$\frac{r_s(\text{km})}{5204.00}$	$\frac{\lambda_{sta}(\text{deg})}{243.1105}$	$\frac{z_0(\text{km})}{3677.05}$
$\frac{\theta_{g0}(\text{deg})}{100.175}$	$\frac{\omega_{\oplus}(\text{rad/sec})}{7.292116 \times 10^{-5}}$	
$\frac{a(\text{km})}{16878.00}$	$\frac{e}{0.5629}$	$\frac{\Omega(\text{deg})}{0.00}$
$\frac{i(\text{deg})}{46.00}$	$\frac{\omega(\text{deg})}{0.00}$	
$\frac{\mu_{\oplus}(\text{km}^3/\text{sec}^2)}{3.9860045 \times 10^5}$	$\frac{E_0(\text{deg})}{0.00}$	
$\frac{t_0(\text{sec})}{0.00}$	$\frac{dt(\text{sec})}{1200.00}$	$\frac{t_{STOP}(\text{sec})}{2.592 \times 10^5}$

Table 2. Input parameter set for Case 2

$\frac{r_s(\text{km})}{5205.25}$	$\frac{\lambda_{sta}(\text{deg})}{148.9813}$	$\frac{z_0(\text{km})}{-3674.75}$
$\frac{\theta_{g0}(\text{deg})}{100.175}$	$\frac{\omega_{\oplus}(\text{rad/sec})}{7.292116 \times 10^{-5}}$	
$\frac{a(\text{km})}{16878.00}$	$\frac{e}{0.5629}$	$\frac{\Omega(\text{deg})}{0.00}$
$\frac{i(\text{deg})}{46.00}$	$\frac{\omega(\text{deg})}{0.00}$	
$\frac{\mu_{\oplus}(\text{km}^3/\text{sec}^2)}{3.9860045 \times 10^5}$	$\frac{E_0(\text{deg})}{0.00}$	
$\frac{t_0(\text{sec})}{0.00}$	$\frac{dt(\text{sec})}{1200.00}$	$\frac{t_{STOP}(\text{sec})}{2.592 \times 10^5}$

Table 3. Input parameter set for Case 3

$\frac{r_s(\text{km})}{4862.45}$	$\frac{\lambda_{sta}(\text{deg})}{355.752}$	$\frac{z_0(\text{km})}{4115.11}$
$\frac{\theta_{g0}(\text{deg})}{100.175}$	$\frac{\omega_{\oplus}(\text{rad/sec})}{7.292116 \times 10^{-5}}$	
$\frac{a(\text{km})}{16878.00}$	$\frac{e}{0.5629}$	$\frac{\Omega(\text{deg})}{0.00}$
$\frac{i(\text{deg})}{46.00}$	$\frac{\omega(\text{deg})}{0.00}$	
$\frac{\mu_{\oplus}(\text{km}^3/\text{sec}^2)}{3.9860045 \times 10^5}$	$\frac{E_0(\text{deg})}{0.00}$	
$\frac{t_0(\text{sec})}{0.00}$	$\frac{dt(\text{sec})}{1200.00}$	$\frac{t_{STOP}(\text{sec})}{2.592 \times 10^5}$

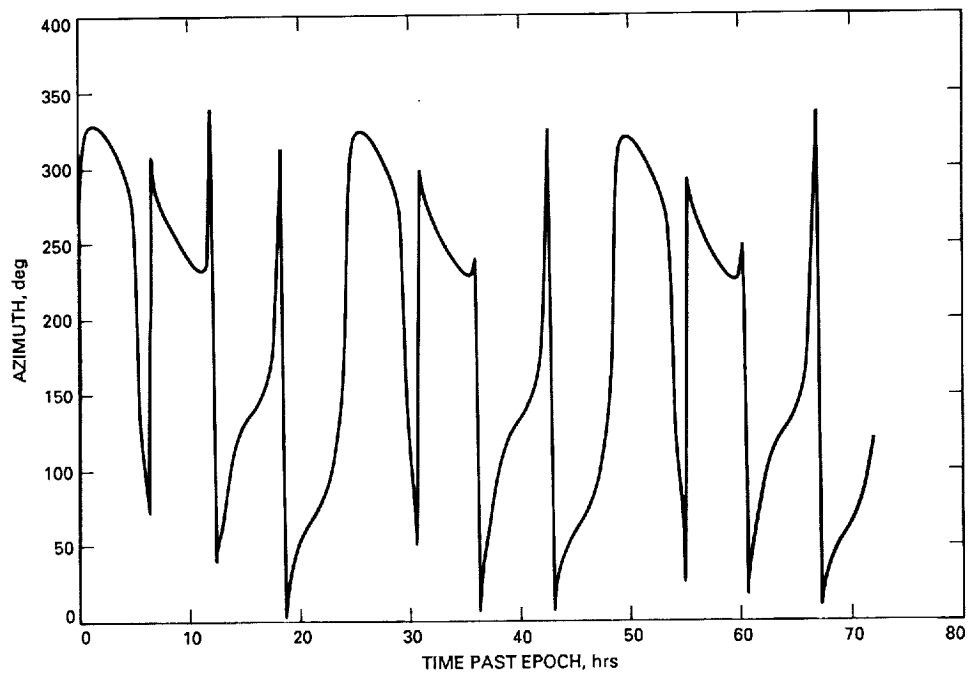


Fig. 1. Azimuth versus time (Case 1).

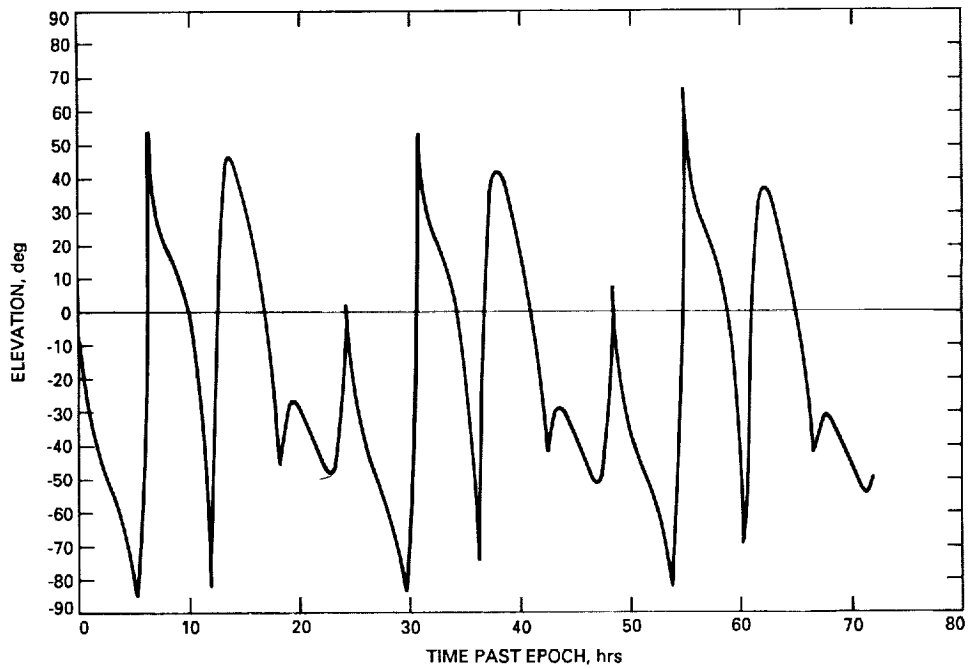


Fig. 2. Elevation versus time (Case 1).

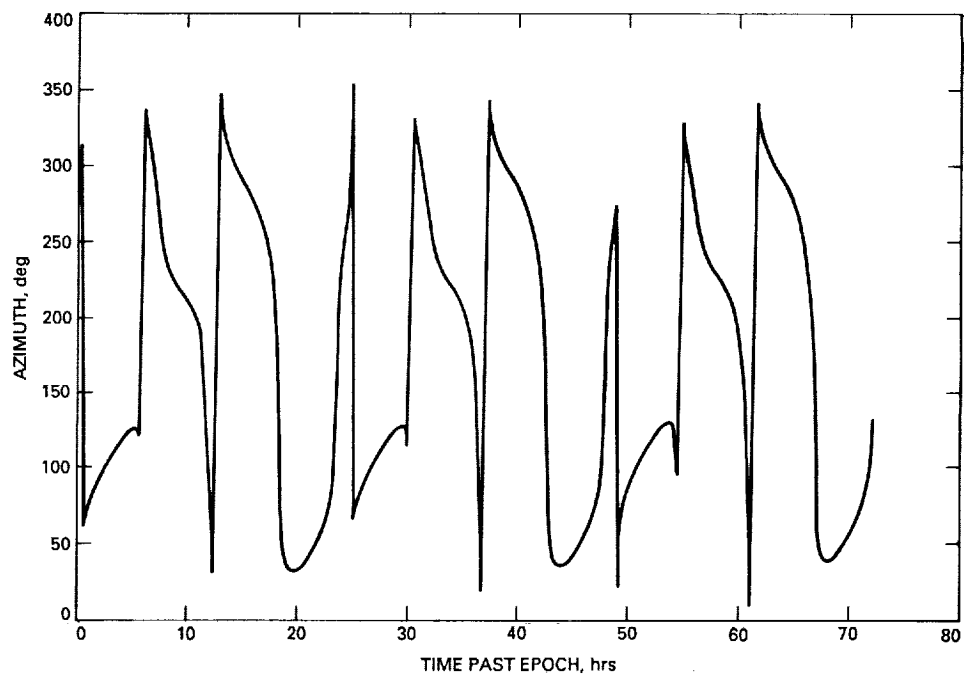


Fig. 3. Azimuth versus time (Case 2).

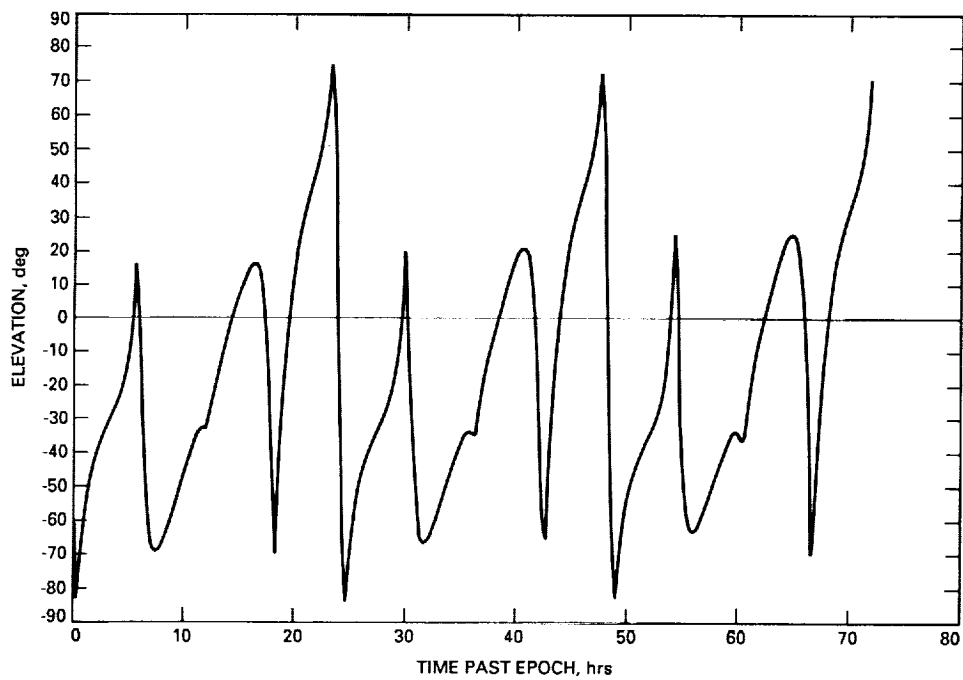


Fig. 4. Elevation versus time (Case 2).

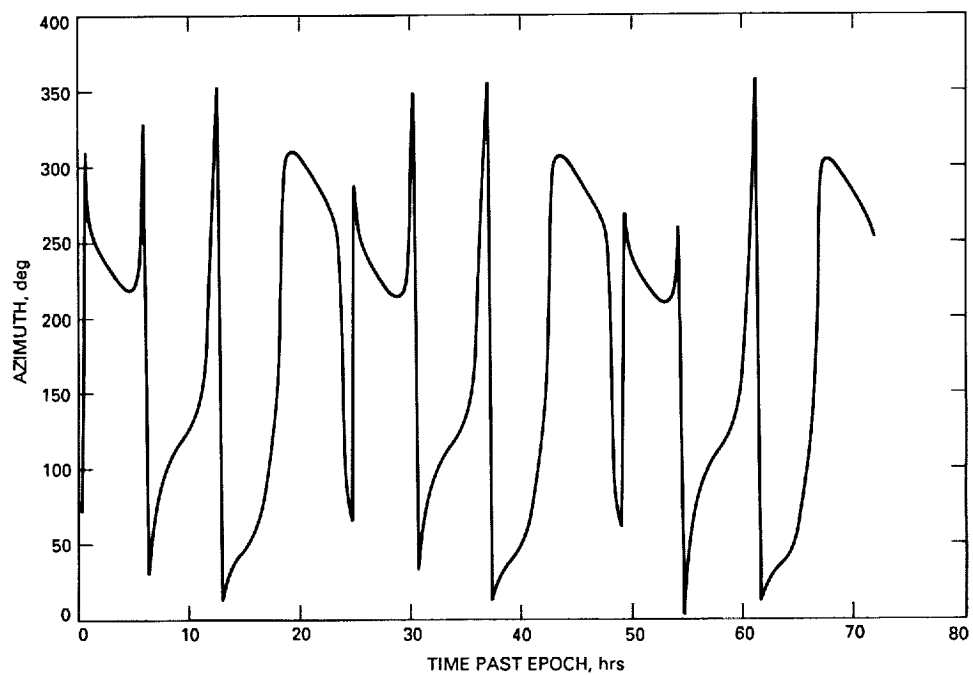


Fig. 5. Azimuth versus time (Case 3).

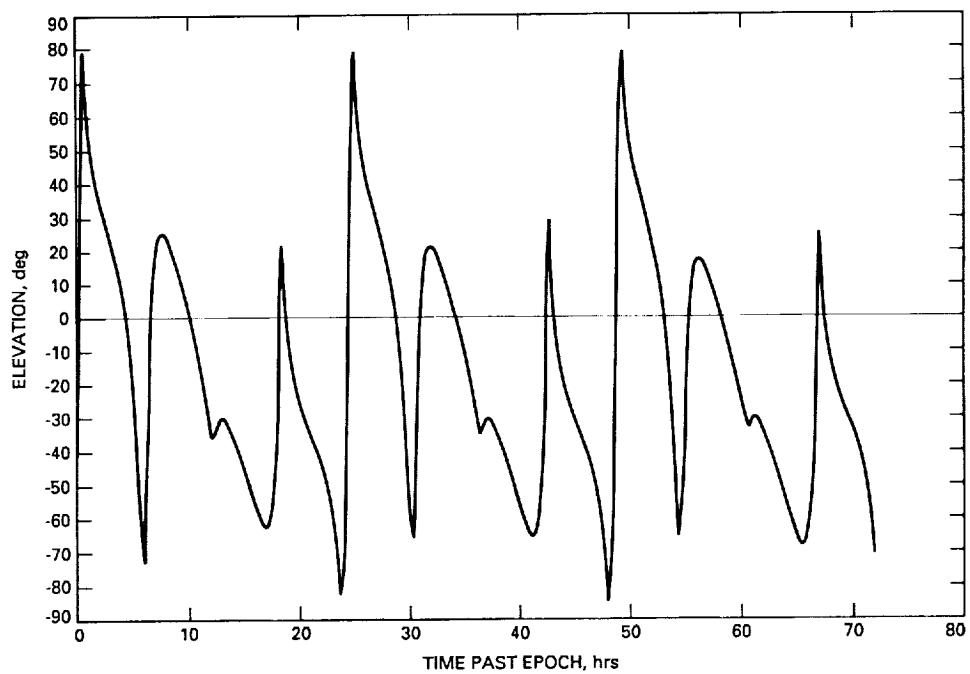


Fig. 6. Elevation versus time (Case 3).

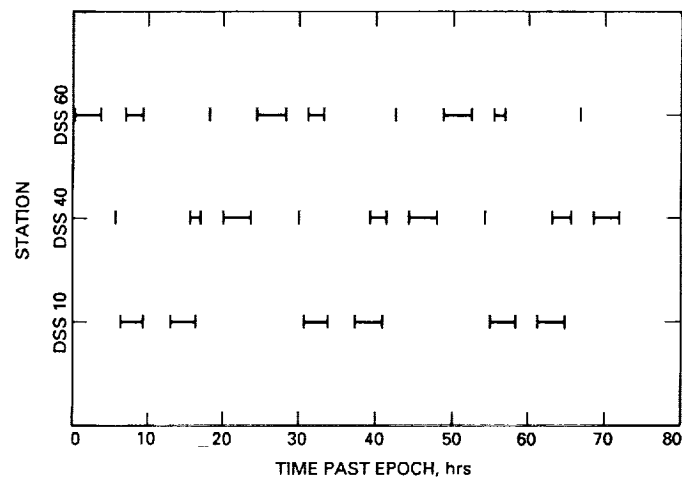


Fig. 7. DSS view periods with 10-deg lower elevation cutoff, epoch 1995-01-01 00:00:00.0000.

Microwave Analog Fiber-Optic Link for Use in the Deep Space Network

R. T. Logan, Jr., G. F. Lutes, and L. Maleki
Communications Systems Research Section

A novel fiber-optic system with dynamic range of up to 150 dB-Hz for transmission of microwave analog signals is described. The design, analysis, and laboratory evaluations of this system are reported, and potential applications in the NASA/JPL Deep Space Network are discussed.

I. Background and Introduction

Commercially available fiber-optic links for microwave analog signal transmission do not meet the most stringent dynamic-range requirements for use in the NASA/JPL Deep Space Network (DSN). At present, the performance of the fiber-optic transmitters limits the capacity of these links. Although semiconductor-laser-based fiber-optic systems are commercially available with frequency response to 10 GHz [1], the output power, intensity noise, and modulation linearity of the semiconductor laser sources limit system dynamic range to about 125 dB-Hz [2]. While this may be adequate for many applications, it falls short of the most stringent requirements of the DSN (which will be discussed in detail below), where dynamic range on the order of 150 dB-Hz is required.

The dynamic range of a fiber-optic system is determined by the optical power, the modulation depth, and the noise floor. Commercially available microwave fiber-optic transmitters employ semiconductor lasers with optical power typically 1-3 mW. Intensity modulation is achieved by direct modulation of the bias current with the radio frequency (RF) or microwave signal. The lin-

earity of the laser typically limits modulation depth to less than 50 percent, before harmonic distortion becomes too great. Semiconductor lasers exhibit a pronounced amplitude noise increase at microwave frequencies. Therefore, the dynamic range possible for microwave fiber-optic systems is limited by laser power, modulation depth, and laser intensity noise to approximately 125 dB-Hz.

To increase dynamic range, one or all of the following must be done:

- (1) Increase the optical power
- (2) Increase the modulation depth
- (3) Reduce the system noise floor

Commercial manufacturers of microwave fiber-optic transmitters have attempted to maximize all three of these parameters, but still are 25 to 30 dB away from the maximum dynamic-range requirements of the DSN.

Another way to achieve extreme levels of dynamic range (> 150 dB) in a fiber-optic system is by external intensity modulation of a high-power, low-noise laser source.

Until recently, reliable and field-worthy solid-state lasers did not have adequate power and low-enough-intensity noise to make such a system feasible. External intensity modulators were available, but had high excess optical loss. This made construction of a field-worthy system infeasible, since impractically high power levels would be required to offset the optical losses.

Recent advances in all-solid-state high-power, ultralow-noise, narrow-linewidth diode-pumped neodymium : yttrium-aluminum-garnet (Nd:YAG) lasers and microwave optical modulators with low insertion loss were recognized by the authors as possibly providing the building blocks for a system which approaches the dynamic range and bandwidth performance levels required in the DSN. Recent tests of the key components and the following analysis confirm that such a system is feasible at this time.

In order to extend the performance of the frequency-reference distribution systems already in place, and to address new applications for fiber optics in the DSN, a novel fiber-optic link for RF and microwave analog signal transmission was designed, analyzed, constructed, and tested. This system offers extremely high dynamic range of 150 dB-Hz (150 dB in a one-hertz bandwidth) with frequency response to 18 GHz from commercially available components. The system employs intensity modulation of a high-power, low-noise solid-state Nd:YAG laser with a lithium niobate (LiNbO_3) electro-optic modulator, an improved type of single-mode optical fiber with low thermal coefficient of delay (TCD), and a commercially available optical receiver.

This article describes the design of the link, and the calculated and measured performance of the two key components: the intensity modulator and the Nd:YAG laser. The expected dynamic range of a fiber-optic system employing these components is then predicted with a system model. Two Nd:YAG lasers and a modulator were loaned to JPL for evaluation by the manufacturers. A laboratory-bench fiber-optic link was constructed from these components. Measurements performed on this link match the predictions of the model. Next, the reliability of the various components is discussed, and various applications of these links in the DSN are examined.

In particular, a study of the suitability of such a fiber-optic link for the transmission of the S-band (2.4 GHz) and X-band (8.4 GHz) outputs of the low-noise amplifiers of DSN antennas to a remote signal processing center (SPC) is summarized. The delay stability possible with such a link using a new type of temperature-compensated low-

TCD fiber is illustrated by a calculation. Finally, remaining questions and ongoing research efforts are discussed.

II. Fiber-Optic System Design

In the course of development of ultrastable fiber-optic systems for precise frequency-reference signal distribution, the authors have recognized that recent developments in opto-electronic component technology now make feasible the construction of a microwave analog fiber-optic system with the following desirable characteristics:

- (1) High dynamic range, up to 150 dB-Hz
- (2) Frequency response to 18 GHz now, and higher in the future
- (3) Low intermodulation distortion
- (4) Low group-delay distortion
- (5) High group-delay stability, < 1 picosecond over intracomplex distances
- (6) High reliability

Such a system is now feasible due to recent technological advances of the following key components:

- (1) Optical sources: semiconductor, laser-diode-pumped, solid-state lasers with high output power, narrow-linewidth single-frequency operation, and extremely low noise
- (2) Optical intensity modulators: wide frequency response (to 18 GHz), low optical insertion loss
- (3) Low thermal coefficient of delay (TCD) optical fiber

A block diagram of the proposed fiber-optic system is shown in Fig. 1. A semiconductor-laser-diode-pumped solid-state Nd:YAG laser generates the optical carrier at 1318.5 nm. An optical isolator prevents reflected optical power from entering the Nd:YAG laser, since reflected optical power destabilizes the optical output. After the isolator, an electro-optic modulator intensity-modulates the optical carrier with the electrical microwave signal. A single-mode low-TCD optical fiber carries this modulated optical signal to an optical receiver which converts the optical signal back to the original electrical microwave signal.

Questions remain about the phase and amplitude linearity of the electro-optic modulators in the passbands of interest (S- and X-band) for the DSN. There were not sufficient time and equipment to fully characterize these parameters with the borrowed components. However, it is

felt that adequate performance can be achieved over the narrow frequency bands used by the DSN with proper microwave matching of the electro-optic modulator to the signal source. This option is currently being investigated.

III. Diode-Pumped-Nd:YAG Lasers

Two samples of the semiconductor-laser-diode-pumped Nd:YAG lasers were loaned to JPL by two manufacturers: Amoco Laser (Naperville, Illinois) and Lightwave Electronics (Mountain View, California). The Amoco laser is an end-pumped linear-cavity type which lases in two longitudinal modes. The Lightwave laser employs a novel nonplanar ring cavity which exhibits true single-mode operation. The output power of both lasers was advertised to be 25 mW at 1318.5 nm; each laser slightly exceeded the output power specification.

The noise performance of these lasers was tested and found to be shot-noise limited. For a shot-noise-limited system, the noise power spectrum is white (flat with frequency), and depends on the optical power level [2]:

$$P_{\text{shot}} = 2eI_o R_L \quad (\text{W/Hz}) \quad (1)$$

where

$$I_o = rP_o 10^{-(P_L/10)} \quad (\text{A}) \quad (2)$$

is the average photocurrent induced by the incident laser power, r is the responsivity of the photodiode in amps per watt, P_o is the optical output power of the laser in watts, P_L is the optical loss between the laser and photodiode in optical dB, e is the electronic charge, and R_L is the load resistance.

The noise was flat in frequency with the exception of a slight hump at about 300 kHz due to the relaxation resonance of the laser cavity. This noise was 3 dB higher than the shot-noise floor for the Amoco laser, and was found to vary substantially in amplitude as a function of back reflections for the Lightwave laser. The use of a 35-dB optical isolator, a routine engineering solution, was found to eliminate the effects of optical reflections on the laser output.

In an actual system, shot noise will predominate until the optical losses become high and the thermal noise of the receiver becomes significant. The total noise power as a function of optical loss is

$$P_{\text{noise}}(P_L) = P_{\text{shot}}(P_L) + P_{\text{th}} \quad (\text{W/Hz}) \quad (3)$$

where P_{th} is the thermal-noise power. For a matched load, this is

$$P_{\text{th}} = kTF \quad (\text{W/Hz}) \quad (4)$$

where k is Boltzmann's constant, T is the temperature in kelvins, and F is the noise figure of the amplifier following the photodiode.

IV. Electro-Optic Intensity Modulators

Electro-optic (E-O) modulators suitable for use in high-dynamic-range fiber-optic links have recently become available. These modulators have good linearity, with less than 5 percent total harmonic distortion at 80 percent modulation index. At present, the modulation bandwidth is as high as 18 GHz.

Waveguide versions of the Mach-Zehnder interferometer E-O modulator provide the highest modulation frequency response for use in fiber-optic systems [3]. The interferometer is fabricated by diffusing titanium into high-purity lithium niobate substrates to form optical waveguides.

The E-O modulator consists of an input optical waveguide which splits into two parallel waveguides located under a metallic microwave-transmission line. The dielectric for the microwave transmission line is the lithium niobate substrate, which contains the optical waveguides. After passing under the microwave transmission line, the two optical waveguides converge to form a single optical waveguide which exits the device.

Modulation is accomplished by generating a difference in delay between the optical signals traveling in the two parallel optical waveguides. The electric field of the microwave signal changes the dielectric constant of the lithium niobate via the linear electro-optic (Pockels) effect. A change in the dielectric constant of the optical waveguide results in a change in propagation velocity and therefore a change in delay of the optical signal. When the two optical signals are added back together in the output optical waveguide, interference between the two signals results in an amplitude-modulated optical output signal. If the two optical signals which are combined in the output waveguide are in phase, they add in power. If they are out of phase, they cancel and no power exits the device.

Since the delay change in the waveguide is linear with voltage, there is a linear change in the phase difference between the two optical signals, which are then added

vectorially. This results in a well-defined cosine-squared intensity-modulation response.

Figures 2 and 3 illustrate the calculated modulation response for single-tone modulation of various indices, and the calculated FFT (fast Fourier transform) of the response. Since the modulator response is symmetric, the even harmonics are theoretically zero. From Fig. 3, at 81 percent modulation, the third and higher harmonics are down more than 28 dB below the fundamental, corresponding to less than 5 percent total harmonic distortion (THD) for the highest level of input signal.

The linearity of a modulator manufactured by GEC-Marconi, and loaned to JPL by Hoechst Celanese was tested by measuring the harmonic content as a function of modulation index. The measured results match the theoretical predictions extremely well: modulation depth of 80 percent was achieved with third harmonics more than 28 dB below the fundamental. Second harmonics were down better than 50 dB, but were strongly influenced by the biasing of the modulator, as expected.

V. Model for Dynamic-Range Predictions

The Institute of Electrical and Electronics Engineers (IEEE) dictionary defines the dynamic range of a communications system as the ratio of the saturation signal level to the minimum acceptable signal level. This is the definition we shall use. The saturation signal level is set by the maximum acceptable level of harmonic content. We shall take the minimum acceptable signal level to be equal to the equivalent input noise (EIN) of the fiber-optic link.

The dynamic range is thus the ratio of two signal levels at the link output: (1) the output signal level produced by an input signal equal to the saturation level of the optical modulator, and (2) the noise floor at the output. The output signal level for a given input level is calculated using standard formulae for amplitude-modulated fiber-optic systems [2]. The radio frequency (RF) signal power out of the fiber-optic system is a function of the modulation index and the optical losses:

$$P_{rf}(P_L) = 1/2[mr]^2[P_o \cdot 10^{-(P_L/10)}]^2 R_L \quad (5)$$

where

m = the optical modulation index

r = detection efficiency of the photodiode (responsivity)

P_o = the average optical power (at pigtailed laser output)

R_L = the load resistance across the photodiode

The optical loss is

$$P_L \text{ (dB)} = P_{\text{fixed}} \text{ (dB)} + P_{\text{fiber}} \text{ (dB/km)} \cdot L \text{ (km)} \quad (6)$$

where L is the fiber length in kilometers. The loss consists of fixed insertion losses of the various components in the optical path (P_{fixed}), such as the modulator insertion loss, connectors and splice losses, and P_{fiber} , the variable loss due to the length of the fiber.

For a given input optical power level and maximum modulation index, the dynamic range of the fiber-optic link is a function of the optical losses. The losses increase with fiber length, so the dynamic range decreases:

$$\begin{aligned} D(L) &= 10 \log[P_{rf,\max}(L)/P_{\text{noise}}(L)] \text{ (dB-Hz)} \\ &= 10 \log[P_{rf,\max}(L)/(P_{\text{shot}}(L) + P_{\text{thermal}})] \text{ (dB-Hz)} \end{aligned} \quad (7)$$

In a high-performance, low-loss system, shot noise would be more significant than the thermal noise, so the system is said to be "shot-noise-limited." Considering the dynamic range when the system is shot-noise-limited, the last equation simplifies to:

$$\begin{aligned} D \text{ (dB-Hz)} &= 10 \log[(m_{\max}^2 r)/4e] \text{ (dB-Hz)} \\ &\quad + P_o \text{ (dB)} - P_L \text{ (dB)} \end{aligned} \quad (8)$$

where m_{\max} is the maximum acceptable modulation index (limited by the acceptable harmonic distortion level), and as before, r is the photodiode responsivity, P_o is the optical power level, and P_L the optical loss. From this expression, it is clear that if the maximum acceptable modulation index is fixed, the dynamic range of the fiber-optic link can only be increased by employing a more powerful optical source or by reducing the optical losses. The system assembled in the laboratory had the following parameters:

m_{\max} = 80 percent modulation, corresponding to less than 5 percent THD at the saturation level

P_o = 25 mW (14 dBm) optical power at the laser output

r = photodiode responsivity: 0.4 A/W

P_L = optical loss of the system: 20 dB

This corresponds to a predicted value for the dynamic range of:

$$D = 140.1 \text{ dB-Hz} \quad (9)$$

which agreed extremely well with the observed value. The optical losses in the laboratory system were 10 dB higher than the design levels, due to problems in coupling the optical power into the modulator for lack of the proper optical components. This is not a serious problem, and can be solved with proper optics. If the losses had been only 10 dB, the dynamic range would have been 150 dB-Hz.

This result is significant, however, because it validates the model for dynamic-range predictions. From this simple model, it is apparent that improved performance can be obtained by using a higher power laser, a higher responsivity photodiode, and decreasing the optical losses.

To get an idea of how an actual system would perform, a prediction can be made using the model with values for commercially available components which were not on hand in the lab. Typical single-mode fiber has very low loss: $P_{\text{fiber}} = 0.4 \text{ dB/km}$. Typical values for the fixed-system optical losses, as measured by JPL, for the various optical components are:

modulator	8	dB at the bias point
2 connectors	2	dB
4 fusion splices	0.4	dB
$\Rightarrow P_{\text{fixed}} =$	10.4	dB

Lasers with output power of up to 175 mW and photodiodes with responsivities of 0.7 A/W are commercially available. Figure 4 is a plot of the calculated fiber-optic system dynamic range as a function of fiber length for a hypothetical practical system with the following conservative parameters:

- (1) A conservative value of $P_{\text{fixed}} = 12 \text{ dB}$ for a typical system fixed optical loss
- (2) $P_o = 30 \text{ mW}$ (14.8 dBm) coupled into the modulator
- (3) $r = 0.7 \text{ A/W}$, commercially available photodiode responsivity
- (4) $T = 300 \text{ K}$
- (5) $F = 3 \text{ dB}$, noise figure of the first amplifier

From Fig. 4, it is apparent that system dynamic range of 148 dB-Hz is maintained for distances up to 3.5 kilometers.

The practical upper bound on the optical power level is determined by the saturation-current level of the photodiode in the optical receiver. Commercially available high-speed photodiodes are adequately linear up to about 2 mA of average photocurrent. This corresponds to an incident average optical power level of $1.43 \text{ mW} = 1.6 \text{ dBm}$ for photodiode saturation. For the system modeled in the above calculation with 14.8 dBm optical power at the input to the modulator and a 3.5 km fiber length, the optical losses would be 13.4 dB, so the optical incident power is $1.4 \text{ dBm} = 1.3 \text{ mW}$, which is within the linear operating range of the photodiode.

Further advances in photodiode saturation levels may enable more powerful optical sources to be used, which will increase dynamic range accordingly. Also, increasing the load impedance, R_L , of the receiver would relax the requirement on optical power, or equivalently, allow shot-noise-limited performance with longer lengths of fiber. This may become possible as lower capacitance, higher speed photodiodes become available, and with improved microwave receiver designs.

VI. Components and Reliability

Most of the devices which would be used in the proposed fiber-optic system have proven reliability performance. PIN photodiodes are known to have mean-time-before-failure (MTBF) of 10^5 to 10^6 hours. There is no known mechanism which would result in early failure of the electro-optic modulator if the modulator were hermetically packaged. Researchers have found no degradation in these devices over time, but extensive reliability testing has not been performed. The lifetime of optical fiber has been estimated to be greater than 100 years if properly protected by the surrounding cable.

The semiconductor-laser-diode-pumped Nd:YAG laser is the least reliable assembly in the proposed system. The lifetime of current lasers of this type is in the range of 5 to 20 thousand hours. However, the manufacturers have not yet addressed the reliability aspects of their designs. There are several approaches that have been suggested to increase the lifetime of these lasers. They are:

- (1) Operate the pump laser at a lower temperature
- (2) Underrate the pump laser more
- (3) Identify a more reliable pump laser and use it
- (4) Use multiple pump lasers which are switched in one at a time when the previous laser deteriorates to a predetermined level

A reasonable MTBF (>50,000 hours) should be readily achievable if engineering effort is devoted to this problem.

VII. Fiber-Optic Link Requirements for Low-Noise Amplifier Output Transmission

This section will give the requirements for transmission of the microwave output of the maser and high-electron mobility transistor (HEMT) low-noise amplifiers (LNAs) used in the NASA/JPL DSN. This will be followed by an analysis showing how the proposed fiber-optic links can meet the requirements. Finally, the improvement to system stability by removal of the RF to intermediate frequency (IF) equipment from the cone to a vibration-free, temperature-stable environment will be discussed.

The construction of such a system requires a thorough understanding of the characteristics of the front end of a DSN receiver. In the analysis that follows, the noise temperature and saturation levels of both the maser and HEMT LNAs are used to determine the specifications for a transmission system which can pass these signals without degrading the system-noise temperature.

The front-end of every DSN antenna is composed of either a maser or HEMT low-noise amplifier. The signal levels for each amplifier type are as follows:

Amplifier	Noise Temp./ N_o	Saturation Level	Gain
maser	15 K/-186 dBm/Hz	-86 dBm	45 dB
HEMT	30 K/-183 dBm/Hz	-66 dBm	45 dB

The fiber-optic system should ideally be able to handle the worst case of signal extremes for either amplifier. This worst case corresponds to the lowest noise level and the highest saturation level. As indicated above, the maser has the lowest noise floor, while the HEMT has the highest saturation level. Each of the LNAs has 45 dB of gain, so that the noise level at the output of the LNA is -186 dBm + 45 dB = -141 dBm/Hz. The saturation level at the output is -66 dBm + 45 dB = -21 dBm.

The required dynamic range of the fiber-optic system is determined using the data for the LNAs and the IEEE definition for dynamic range discussed previously. The ratio of the saturation level and the minimum acceptable signal level determines the available dynamic range of the front end on a per-hertz basis:

$$P_{\text{sat}}/N_o = -66 \text{ dBm} - (-186 \text{ dBm/Hz}) = 120 \text{ dB-Hz} \quad (10)$$

The equivalent input noise of the fiber-optic system must be sufficiently below the output noise level of the front-end amplifier chain so that the overall system-noise temperature is not excessively degraded. A conservative estimate is that the fiber-optic system input noise must be ≈ 28 dB below the noise of the amplifier output so that the overall system-noise temperature is degraded by about 0.15 percent. This is much less than the 1 percent maximum acceptable degradation for the whole system.

Also, the fiber-optic system must handle signals up to the saturation level of the LNA. Any signal at this level will already be saturating the LNA, so excessive dynamic range above this level is not required from the fiber-optic system. The maximum drive level of the fiber-optic system is determined by the maximum allowable harmonic content in the signal. For single-tone modulation, system performance with less than 5 percent total harmonic distortion requires that the largest harmonic must be 26 dB below the fundamental tone, with all other harmonics substantially below this level (at least 10 dB). When the LNA reaches its 1 dB compression point, the harmonic content added by the fiber-optic system must be at least 26 dB below the signal level.

These requirements for the fiber-optic system define the dynamic range needed to pass the LNA output signal while degrading the system-noise temperature by less than 0.15 percent. The dynamic-range requirement for the fiber-optic system is:

$$\begin{aligned} D_{\text{FO}} &= P_{\text{sat}} - (N_o - 28 \text{ dB}) \\ &= -66 \text{ dBm} - (-186 \text{ dBm/Hz} - 28 \text{ dB}) \\ &= 148 \text{ dB-Hz} \end{aligned} \quad (11)$$

From Fig. 4, conservative calculations indicate that this requirement can be met with commercially available components for link lengths up to 3.5 kilometers.

The number calculated for D_{FO} applies throughout the 500-MHz bandpass at both S- and X-band. The number is given as a dB-Hz specification to preserve full generality, since individual applications have different bandwidths. To derive the specification for a given application, one only need divide D_{FO} by the bandwidth of the application.

For example, a wideband application employing the full 500-MHz bandwidth of the LNA, such as very long-

baseline interferometry (VLBI), would require dynamic range in dB of

$$\begin{aligned} D_{\text{FO-VLBI}} &= D_{\text{FO}} - 10 \log(500 \text{ MHz}) \\ &= D_{\text{FO}} - 87 \text{ dB-Hz} \\ &= 62 \text{ dB} \end{aligned} \quad (12)$$

for the 500-MHz bandpass. Conversely, a telemetry application may have a 10-MHz bandpass, so the dynamic range required is

$$\begin{aligned} D_{\text{FO-TLM}} &= D_{\text{FO}} - 10 \log(10 \text{ MHz}) \\ &= 79 \text{ dB} \end{aligned} \quad (13)$$

Both applications require the same per-hertz dynamic range, which is fundamentally limited by the LNA. Therefore, if the fiber-optic system can handle the full per-hertz dynamic-range requirement set by the LNA, it will handle any application which currently uses that LNA.

For a two-kilometer link, the system signal and noise levels are as indicated in Fig. 5. The additional gain after the LNA is required to raise the noise level of the LNA above the input noise of the fiber-optic system, but since the LNA has 45 dB of gain, a 3-dB noise figure for this amplifier will degrade the system temperature by 0.1 percent. The level at the output of the fiber-optic system is the same as the level at the output of the LNA when the additional gain equals the insertion loss of the fiber-optic link. This makes the interface to the first mixer transparent, and no additional gain is required before the first mixer.

VIII. Low Thermal Coefficient of Delay Fiber-Optic Cable

Low thermal coefficient of delay (TCD) single-mode optical-fiber cable will be used with a simple form of thermal control to achieve near-zero change in group delay with changes in temperature. This special fiber has a TCD near zero (<0.01 ppm/deg C) between 10 deg C and 20 deg C which increases to 0.4 ppm/deg C at about 35 deg C. The near-zero TCD of the cable over a wide temperature range simplifies thermal control of the cable. If the temperature of the cable is held near 15 deg C, small changes (± 5 deg C) in temperature will have little effect on the group delay.

All of the fibers requiring high thermal stability can be contained in a single cable less than 1 inch in diameter. The temperature of this cable could be maintained near

15 deg C on the antenna by installing it with a chilled liquid line enclosed in insulation. The cable could be buried underground between the antenna and the signal processing center (SPC). The temperature at 1.5 meters depth at Goldstone has been found to be about 10 deg C and stable to millidegrees C.

The change in group delay of a cable is

$$dD_g = [K_{\text{TCD}}(L)dT]/[10^6 V_g] \quad (14)$$

where

K_{TCD} = the thermal coefficient of delay for the cable at the operating temperature in ppm/deg C

L = the length of cable subjected to the temperature change

dT = the temperature change on the cable

V_g = the group velocity in the cable

For the cable run between SPC-10 and the Deep Space Station (DSS)-15 antenna, the total group-delay change in the cable as a result of temperature change can be calculated based on the following conservative assumptions:

- (1) total cable length of 500 meters
- (2) fifty meters of cable in the antenna cooled to 15 deg C (TCD = 0.01 ppm/deg C) and subjected to ± 3 deg C change in temperature
- (3) ten meters of cable in the wrap-up at 25 deg C (TCD = 0.12 ppm/deg C) and subjected to ± 5 deg C change in temperature, and
- (4) 440 meters of underground cable at 10 deg C (TCD = 0.05 ppm/deg C) and subjected to ± 0.1 deg C change in temperature

The delay instability is evaluated from Eq. (14) for dD_g , using the assumed values given for each section of cable. For the antenna section of cable,

$$dD_g = [0.01(50)6]/[10^6(2.1 \times 10^8)] = 0.014 \text{ (psec)} \quad (15)$$

for the wrap-up section of the cable,

$$dD_g = [0.12(10)10]/[10^6(2.1 \times 10^8)] = 0.06 \text{ (psec)} \quad (16)$$

and for the buried portion of the cable,

$$dD_g = [0.05(440)0.2]/[10^6(2.1 \times 10^8)] = 0.02 \text{ (psec)} \quad (17)$$

The total delay change is the sum of the changes in the three sections,

$$dD_g(\text{total}) = 0.014 + 0.06 + 0.02 = 0.094 \text{ (psec)} \quad (18)$$

This represents a diurnal phase change at 8.4 GHz of 0.34 degrees or 0.006 radians that is contributed by the cable. This corresponds to a fractional frequency deviation of

$$\begin{aligned} df/f &= dt/t = 0.094 \text{ (psec)}/86,400 \text{ (sec/day)} \\ &= 1.1 \times 10^{-18} \end{aligned} \quad (19)$$

over a 24-hour averaging time. Compare this to the stability of a hydrogen maser

$$df/f = 8 \times 10^{-16} \quad (20)$$

and it is apparent that the instability added by the fiber-optic cable due to temperature fluctuations is insignificant.

IX. Benefits to the DSN

Advanced fiber-optic systems based on these new developments could enable the final step toward a fully centralized SPC. They would enable transmission of microwave signals between the antenna cone and a nearby SPC. In the case of the receiver channel, this would permit removal of most of the equipment from the antenna cone area, resulting in a front-end area (FEA) configured as shown in Fig. 6. In this scheme, the RF-IF downconversion would be performed in a temperature-controlled and vibration-free environment at the SPC. Many of the advantages of the beam-waveguide antenna concept would be realized with this system, with the added advantage of a fully centralized SPC.

Also, the transmitter exciters could be centrally located in the fully centralized SPC for an entire Deep Space Communications Complex (DSCC). Since the exciter signal has a low dynamic range compared to the receiver channel, and operates on the same frequencies, a similar fiber-optic system could be used for both downlink and uplink. This would remove the exciter from the antenna cone, and preclude the need for stable frequency-determining elements in the antenna cone. Even the comb-generator calibration signals could be sent to the antenna cone from the SPC over these highly stable fiber links. Frequency stability and station flexibility could be enhanced by collocation of sensitive frequency determining elements in the SPC.

DSN engineers who design equipment for service in the antenna cone are well aware that the harsh operating environment limits system performance. A few of the ways in which the proposed fiber-optic links would impact the performance of the downlink are:

- (1) Temperature control—Thermal control of the RF-IF chain could be substantially improved by relocating these components to the SPC, thereby reducing drift and noise.
- (2) Vibration—The antenna cone is subjected to vibrations due to wind and other sources, which seriously impact frequency-determining elements such as quartz-crystal-oscillator-controlled phase-locked-loops by introducing close-to-carrier phase noise.
- (3) Magnetic field shielding—The deleterious effects of magnetic fields from power supplies on the maser LNA and other sensitive components would be eliminated.
- (4) Ease of servicing equipment—With the entire RF-IF chain located on the ground in the SPC, maintenance and servicing would no longer require technicians to climb the antenna; the only receiver components left in the antenna cone would be the LNA and the electro-optic modulator.

Bringing together the microwave signals from all of the antennas at the SPC opens the possibility of coherent uplink and downlink antenna arraying using any number of antennas in any configuration in real time. Novel radio science experiments utilizing the raw microwave signals from widely separated antennas will become possible. Upgrades to the receiver channels will be simplified greatly, since the raw microwave output of the LNAs will be available at the SPC. As newer and better receivers become available, they may be incorporated into the SPC with ease; no modifications to the front-end area would be required.

The central frequency and timing facility, the central monitor and control facility, and the central signal processing facility located at the fully centralized SPC could be extended to cover all stations at a complex regardless of their location. The central communications facility might also be located at the SPC, where all of the monitor and control signals are accessible. This would provide a convenient interface through which the complex could be controlled from a remote location such as JPL, via the existing international digital fiber-optic networks.

X. Summary of Concerns and Future Investigations

As may be expected, during the course of this preliminary study many questions were raised. These concerns are summarized here:

- (1) Due to lack of time and equipment, the phase and amplitude linearity of the intensity modulators was not characterized. It may be the case that a microwave impedance-matching network for the optical modulator will be required to achieve sufficient phase and amplitude linearity in the microwave passband to meet DSN requirements. However, since the percentage bandwidth required is small compared to the carrier frequency ($500 \text{ MHz}/2.4 \text{ GHz} = 20 \text{ percent}$ at S-band, 6 percent at X-band), it is felt that good microwave design can provide a modulator with sufficient phase and amplitude linearity for DSN applications.
- (2) The mean time to failure of the semiconductor-laser-pump diodes must be increased to an adequate level, or a multiple-pump scheme implemented to enable a soft failure-mode system. The laser manufacturers are already working on this problem for space-qualified systems.
- (3) The improvement in the downlink stability budget when the RF-IF conversion is moved to the SPC must be quantified.
- (4) The effect of the Nd:YAG laser linewidth on the microwave signal transmission on optical fiber for long distances (10s of kilometers) is not yet fully modeled. Preliminary calculations indicate that the bandwidth of the present optical fiber is adequate. If it is not, optical dispersion compensation schemes can be employed to increase bandwidth.
- (5) The impact of laser wavelength stability on the stability performance of the fiber-optic system must be determined.

XI. Conclusion

This article presented a system design, analytical model, and experimental test data for a novel fiber-optic system for microwave analog signal transmission. Laboratory experiments matched the predicted values extremely well, indicating that such a system is capable of dynamic range approaching 150 dB-Hz, with frequency response to 18 GHz. The dynamic range of the maser and HEMT LNA outputs of a DSN antenna were calculated from the DSN specifications, and it was shown that the fiber-optic system described has adequate dynamic range and phase stability to transmit the S-band and X-band outputs of the maser or HEMT LNA to a remote SPC several kilometers away without appreciable degradation of signal quality. This system would also be capable of transmitting the exciter signals from a fully centralized SPC to the various antennas. The phase and amplitude linearity of the fiber-optic system in the S- and X-band passbands may be inadequate at present, and should be studied in the course of continuing research. From this preliminary study, however, it is clear that the basic building blocks are in place to realize a fully centralized SPC, and that fiber optics is the enabling technology to achieve this goal.

These advanced fiber-optic systems could substantially simplify the uplink and downlink channels of the DSN, and drastically increase the flexibility and capabilities of the DSN. This idea takes the concept of the beam-waveguide antenna one step further, by not only locating most of the signal processing equipment in a controlled environment, but more importantly, by bringing together the microwave outputs from the front ends of widely separated antennas to a fully centralized SPC. This capability would enable radical changes in DSN configurations, including the possibility of coherent uplink and downlink arraying and novel radio-science experiments utilizing the raw microwave signals from widely separated antennas in real time. The fully centralized SPC could itself be remotely controlled via the international digital fiber-optic networks already in place, enabling command and control of the DSN from JPL, or any other place.

References

- [1] *3515A 10 GHz Transmitter Specification*, Ortel Corporation, Alhambra, California, 1988.
- [2] K. Y. Lau, "Signal-to-Noise Calculation for Fiber Optics Links," *TDA Progress Report 42-58*, vol. May and June 1980, Jet Propulsion Laboratory, Pasadena, California, pp. 41-48, August 15, 1980.
- [3] R. A. Becker, "Travelling-wave electro-optic modulator with maximum bandwidth-length product," *Applied Physics Letters*, vol. 45, pp. 1168-1170, 1984.

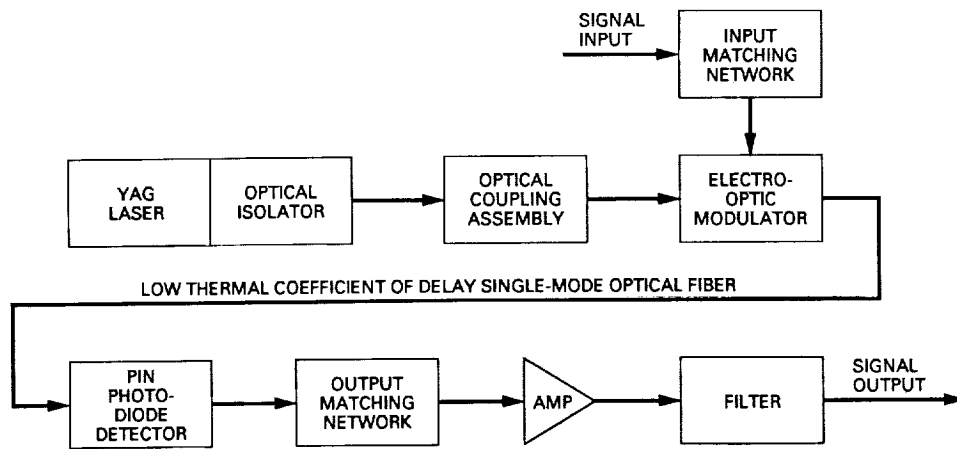


Fig. 1. Block diagram of proposed fiber-optic system.

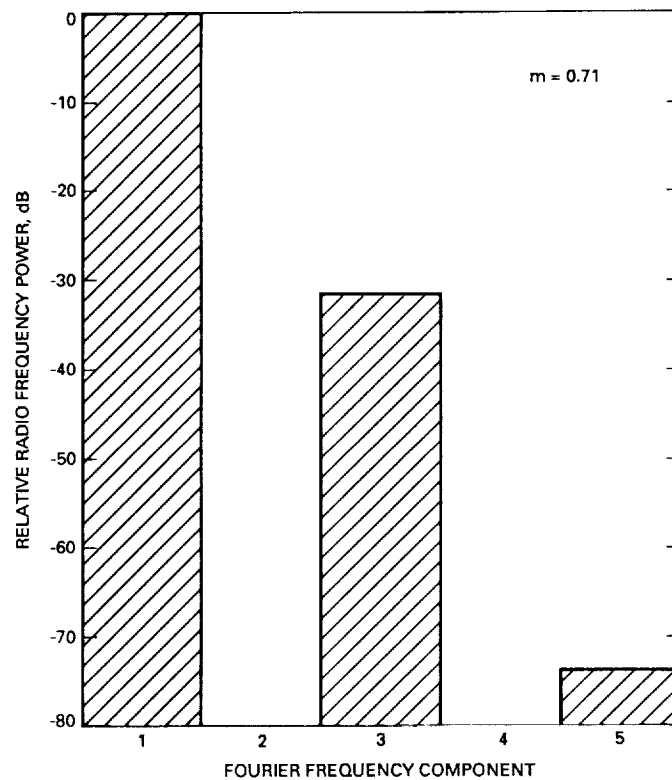


Fig. 2. Calculated fast Fourier transform of electro-optic modulator response: 71 percent modulation Index.

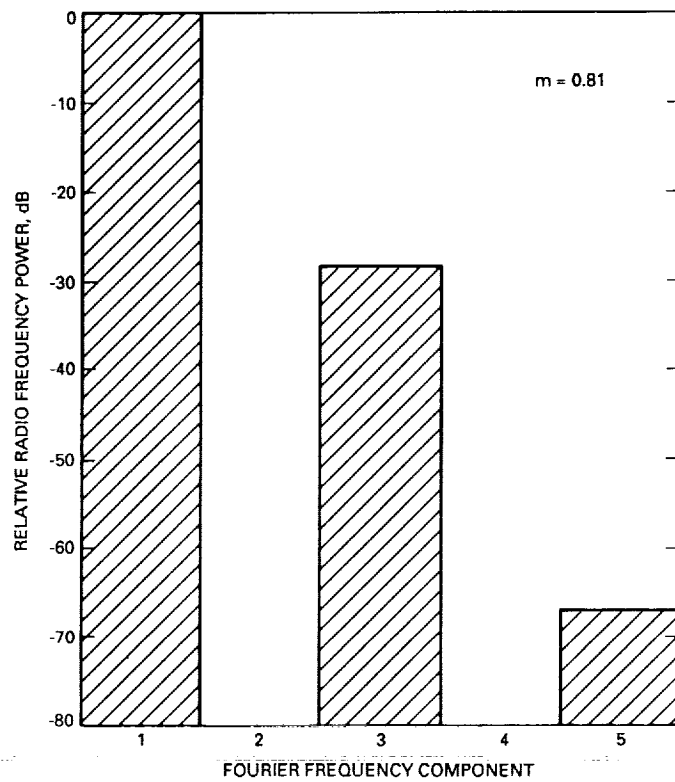


Fig. 3. Calculated fast Fourier transform of electro-optic modulator response: 81 percent modulation index.

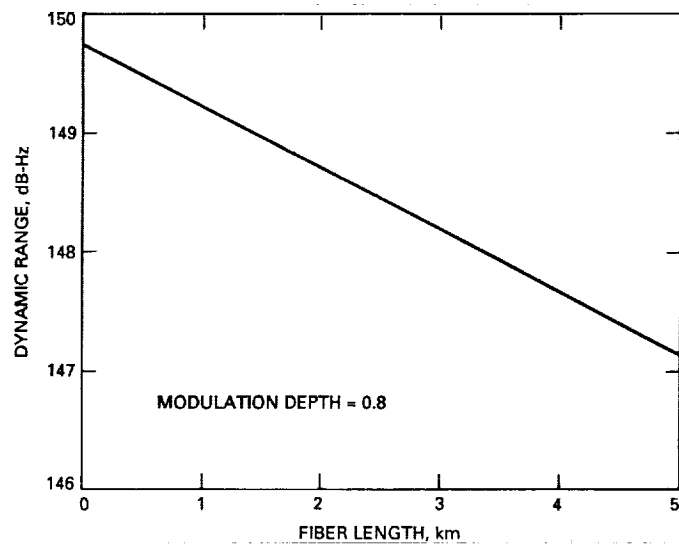


Fig. 4. Calculated fiber-optic system dynamic range versus fiber length.

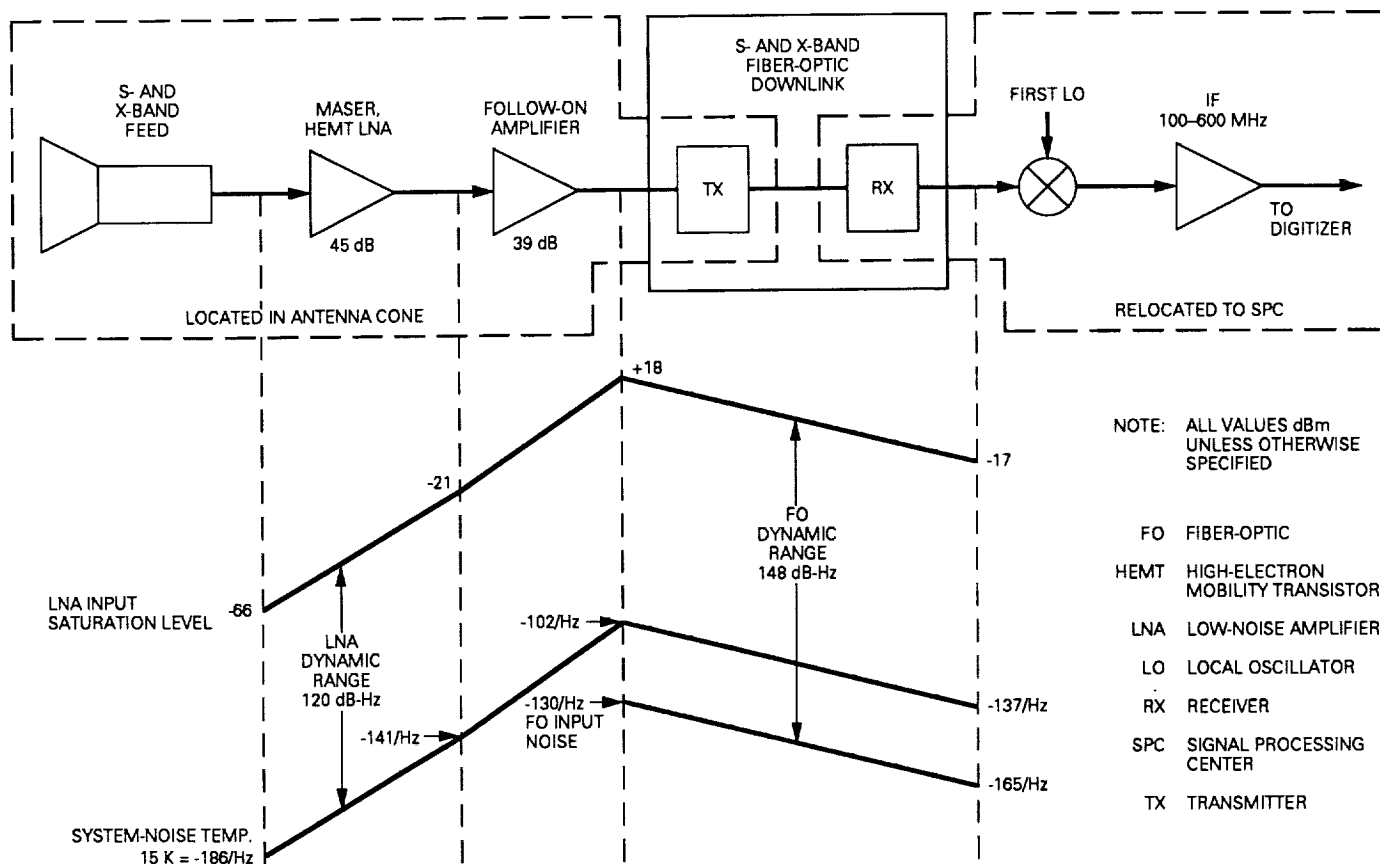


Fig. 5. Signal and noise levels of fiber-optic link for transmission of low-noise amplifier output.

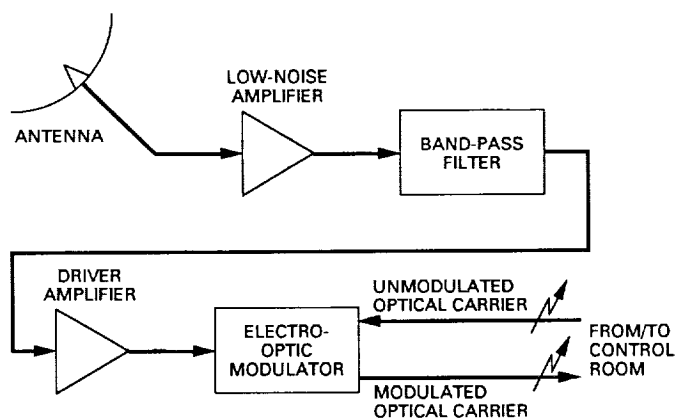


Fig. 6. New Deep Space Network front-end area configuration possible using advanced fiber optic links.

54-17
272320

148
TDA Progress Report 42-100

N90-21884

February 15, 1990

Galileo Earth Approach Navigation Using Connected-Element Interferometer Phase-Delay Tracking

S. W. Thurman
Navigation Systems Section

This article investigates the application of a Connected-Element Interferometer (CEI) to the navigation of the Galileo spacecraft during its encounter with Earth in December 1990. A CEI tracking demonstration is planned for the week of November 11 through 18, 1990, from 27 days to 20 days prior to Earth encounter on December 8. During this period, the spacecraft will be tracked daily with Deep Space Network Stations 13 and 15 at Goldstone. The purpose of this work is twofold: first, to establish and define the navigation performance expected during the tracking demonstration and, second, to study, in a more general sense, the sensitivity of orbit determination results obtained with CEI to the data density within CEI tracking passes and to important system parameters, such as baseline orientation errors and the phase-delay measurement accuracy. Computer simulation results indicate that the use of CEI data, coupled with conventional range and Doppler data, may reduce the uncertainty in the declination of the spacecraft's incoming trajectory by 15 to 66 percent compared with the operational solution using range and Doppler data only. The level of improvement depends upon the quantity and quality of the CEI data.

I. Introduction

Connected-Element Interferometry (CEI) is being developed for use as a medium-accuracy (100-400 nrad for individual measurements) angular measurement system [1]. The use of a common frequency reference distributed through a fiber-optic link to two nearby antennas would be used to make very precise measurements of the phase delay of signals from a radio source. Since the measurements are made at a single frequency, there is no need for a transponder on board the spacecraft for beat frequency generation, as required for wideband Δ Very Long Base-

line Interferometer (Δ VLBI) observations. The incoming data from each station are routed through the fiber-optic link to an on-site correlator, whose output data will then be sent immediately to JPL to be processed, along with other radio metric data, for navigation use.

The principal advantages of CEI over the present intercontinental Δ VLBI system are the simplicity of the measurement system and the speed with which the data become available for use. The long observation periods which can be obtained with two nearby antennas and the

near-real-time correlation of the incoming signals make CEI very attractive for use as a navigation tool. With CEI, it is possible to monitor system performance as the measurements are being made and to take corrective action if problems occur. The combination of many CEI measurements made over a single tracking pass make it possible to achieve angular measurement accuracies on the order of 50 nrad, which should be very useful in deep-space navigation.

This article is an expansion and continuation of earlier work done by D.W. Murrow.¹ It begins with a short description of Galileo's approach to its first of two Earth encounters, and continues with a summary of the trajectory, tracking schedule, filter model, and assumed a priori parameter uncertainties used in the orbit-determination simulations which follow. All computer simulations were performed using the Orbit Determination Program (ODP) system software. The results of Murrow's earlier study are summarized, and are used as a basis for comparison and contrast with the new simulations which follow.

II. Galileo Earth Gravity Assist One (EGA1)

This study was performed using a prelaunch reference trajectory, in which it was assumed that the spacecraft was injected into its interplanetary trajectory on October 10, 1989. The Galileo EGA1 trajectory segment begins shortly after the spacecraft's encounter with Venus on February 9, 1990. Navigation accuracy during this time frame with two-way range and Doppler data is degraded due to the near-zero declination of the incoming trajectory, shown in Fig. 1. In addition, there are several targeting maneuvers during the approach which must be determined very accurately to meet mission requirements. The low declination of the flight path leads to a large uncertainty in the declination component of the targeting conditions, expressed in B-plane coordinates.² In the CEI tracking demonstration, data will be taken for one week, using the baseline formed by Deep Space Stations (DSSs) 13 and 15 at Goldstone, from encounter -27 days (E-27) to E-20 days (November 11-18, 1990), spanning one of the targeting maneuvers scheduled for E-25 days. CEI passes will be made daily during this time, using quasar P0528+134 to create an observable consisting of the phase delay from the spacecraft differenced with the phase delay from the quasar.

¹ D.W. Murrow, "Galileo Earth Approach Analysis Using Connected Element Interferometry Data," JPL IOM GLL-NAV-89-28 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 21, 1989.

² "Galileo Navigation Plan," Galileo Project Document 625-566, Revision A (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1989.

The quasar location relative to Galileo's flight path is also shown in Fig. 1. The DSS 13 to DSS 15 baseline is about 20 km long, with a nearly north-south orientation, and should provide strong information in declination, improving the uncertainty in this component over that obtained with range and Doppler tracking.

III. Trajectory

A. Epoch

The epoch is at 1990 February 13 08:44:09.556 (Ephemeris Time). The spacecraft state in Earth-centered Earth mean equator and equinox of B1950.0 Cartesian coordinates is:

$$X = 1.968896282785908 \times 10^7 \text{ km}$$

$$Y = -4.981570235638191 \times 10^7 \text{ km}$$

$$Z = -1.483032488173643 \times 10^7 \text{ km}$$

$$DX = 7.905736580213764 \text{ km/sec}$$

$$DY = -13.48008080659643 \text{ km/sec}$$

$$DZ = -5.834630696184871 \text{ km/sec}$$

The spacecraft's distance and speed with respect to Venus are:

$$R = 2.500 \times 10^6 \text{ km}$$

$$V = 6.363 \text{ km/sec}$$

The spacecraft's distance and speed with respect to Earth are:

$$R = 5.558 \times 10^7 \text{ km}$$

$$V = 16.68 \text{ km/sec}$$

B. Trajectory Correction Maneuver (TCM) Schedule

The schedule for the trajectory correction maneuvers follows.

Maneuver	Time, ET	Nominal Magnitude, m/sec
TCM1	1990 May 11 12:00:00	12.093
TCM2	1990 May 31 12:00:00	0
TCM3	1990 October 09 11:32:10	0
TCM4	1990 November 13 11:32:10	0
TCM5	1990 November 26 11:32:10	0

C. Closest Approach to Earth

The closest approach to Earth occurs on 1990 December 8 11:32:10.586 (ET).

IV. Baseline Tracking Schedule

In Table 1 all data (range, Doppler, and CEI) are assumed to be at S-band (2.29 GHz) frequency. The simulations use all data up to November 22, 1990, the data cut-off point for the EGA1 navigation delivery. In this study, simulated CEI data are taken with the DSS 13 to DSS 14 baseline, although the actual CEI demonstration will be performed using the DSS 13 to DSS 15 baseline. DSS 14 and DSS 15 are very near each other (200 m apart) at the Goldstone complex, so the results presented in Table 1 are equally valid for both stations.

V. Filter Model and Assumed A Priori Parameter Uncertainties

The parameter list described here is based primarily on the orbit-determination error model used by the Galileo Project.³ Range and Doppler data taken by DSS 14 (Table 1) were removed from the filter for the CEI simulations to avoid creating over-optimistic results. The baseline orientation uncertainty for the DSS 13 to DSS 14 baseline was not modeled explicitly in most of the simulations, but was accounted for in some sense by quasar-direction uncertainty. There were some cases, described in Section VI, in which station-location errors were assumed for DSSs 13 and 14. In these cases, station coordinates for DSS 13 and DSS 14 were included as "consider" parameters—parameters which influence the uncertainty in the knowledge of the estimated parameters, but are not estimated themselves.

A. Filter Type

The filter is a batch-sequential epoch state filter with a five-day batch size. The batch size is shorter for certain time periods near maneuvers.

B. Estimated Parameters

The estimated parameters are:

- (1) Spacecraft epoch state vector
- (2) Three components for each TCM occurring during the data arc—includes TCMs 1, 2, 3, and 4

- (3) Radial (with respect to the Sun) solar radiation pressure coefficient (GR)
- (4) Radial (with respect to the Earth) bias acceleration
- (5) Radial (with respect to the Earth) stochastic acceleration with an assumed time constant of five days ($ATAR$)

C. Considered Parameters

The considered parameters are:

- (1) Mass of the Earth (GM_3)
- (2) Earth's ephemeris
- (3) Transverse (with respect to the Sun) solar radiation pressure coefficients (G_X and G_Y)
- (4) Station coordinates for DSS 43 and DSS 63
- (5) Wet troposphere component at DSSs 43 and 63
- (6) Dry troposphere component at DSSs 43 and 63
- (7) Daytime ionosphere at DSS 43 and DSS 63
- (8) Nighttime ionosphere at DSS 43 and DSS 63
- (9) Quasar direction uncertainty (right ascension and declination)

D. A Priori 1σ Uncertainties for Estimated Parameters

- (1) Spacecraft epoch state:
position components: 1.0×10^8 km
velocity components: 1.0×10^8 km/sec
- (2) Trajectory correction maneuvers per component:
TCM1: 12.0 cm/sec
TCM2: 10.0 cm/sec
TCM3: 10.0 cm/sec
TCM4: 10.0 cm/sec
- (3) Radial solar radiation pressure coefficient: $\sigma_{GR} = 0.17 = 10\%$ of nominal value of GR
- (4) Radial bias acceleration: 1.0×10^{-12} km²/sec⁴
- (5) Radial stochastic acceleration: $\sigma_{ATAR} = 1.0 \times 10^{-12}$ km²/sec⁴, time constant $\tau = 5$ days

E. A Priori 1σ Uncertainties for Considered Parameters

- (1) Mass of the Earth: $\sigma_{GM3} = 0.14142$ km³/sec².
- (2) Earth's Ephemeris: statistics used are from Planetary Ephemeris DE125. The Earth's ephemeris

³ D. W. Murrow, "Galileo Orbit Determination Error Model Assumptions," JPL IOM GLL-NAV-89-16 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 3, 1989.

uncertainty relative to the Sun, at the trajectory epoch (February 13, 1990) is:

Position:	radial	0.02 km
	along track	20.58 km
	cross track	8.21 km
Velocity:	radial	0.002 mm/sec
	along track	0.058 mm/sec
	cross track	0.870 mm/sec

- (3) Transverse solar radiation pressure coefficients: $\sigma_{GX} = \sigma_{GY} = 0.0342$ (2% of nominal value of GR).
- (4) Station coordinates consist of a correlated, station-location, error covariance generated by Murrow and Nicholson.⁴ Only the sigmas of the diagonal elements are stated.
DSS 43: radial 0.48 m DSS 63: radial 0.47 m
polar 5.77 m polar 5.77 m
long. 0.59 m long. 0.58 m
- (5) Wet troposphere: 4-cm uncertainty at DSS 43 and DSS 63
- (6) Dry troposphere: 1 cm
- (7) Daytime ionosphere: 75-cm uncertainty at DSS 43 and DSS 63
- (8) Nighttime ionosphere: 15 cm, ionosphere uncertainty values quoted are for S-band (2.29 GHz)
- (9) Quasar direction uncertainty (right ascension and declination): 100 nrad for each component

F. Measurement Accuracy

All data points used in the simulations were assigned 1σ uncertainties by data type as follows:

- (1) Range: 1000 m
- (2) Doppler: 0.50 mm/sec (for a 60-sec count time)
- (3) CEI: 8.0 mm (400 nrad), 4.0 mm (200 nrad), or 1.0 mm (50 nrad)

VI. Results

The measure of performance used to express the *quality* of each simulated solution consists of three components: two representing the uncertainty in the direc-

tion and magnitude of the *impact parameter*, or B-vector, within the B-plane, and the third representing the uncertainty of the third B-plane component, the linearized time of flight (LTOF). These three parameters completely characterize the uncertainty in the targeting of the spacecraft's incoming flight path and its time of closest approach to the Earth.

Figure 2 shows the error ellipses (representing uncertainty in the magnitude and direction of the B-vector) in the B-plane for the four cases generated by Murrow.⁵ In this figure, the T-axis lies in the ecliptic plane and is perpendicular to the spacecraft's incoming velocity asymptote (which points into the paper), while the R-axis is perpendicular to both the T-axis and the velocity asymptote, completing an orthogonal coordinate system. In each of the three cases which had CEI data, the CEI points were assumed to be at one-hour intervals, and any points which had a minimum elevation of 15 degrees or less were eliminated. Figures 3, 4, and 5 show breakdowns of the computed and consider-state contributions to the total error uncertainty in each of the three performance measures; the B-vector R-component, the B-vector T-component, and the LTOF, respectively. Table 2 contains a numerical summary of these results.

During its approach, the Galileo spacecraft is flying almost directly at the Earth. An interferometric data type such as CEI will be sensitive to plane-of-sky trajectory deflections (the plane of the sky is the plane perpendicular to the Earth-spacecraft line of sight), and insensitive to along-track perturbations. As shown in Figs. 2, 3, and 4, CEI does indeed have a great deal of impact on plane-of-sky position uncertainty, reflected in the B-plane. Figure 5 shows that the CEI data had no effect on the LTOF uncertainty, as expected. The reduction in the station-location consider error seen in Fig. 3 is likely due to the reduced dependence of the solution on the Doppler data, which are very sensitive to station-location errors. In Fig. 4, CEI data apparently have some negative impact in that they exaggerate the computed error and the media effects on the range and Doppler data, resulting in some degradation in solution accuracy.

An assumed CEI data rate of 1 point/hr during a pass is probably a very conservative estimate of the data rate which can be achieved in practice. In studying the effects of troposphere fluctuations, the dominant error source in CEI

⁴ D. W. Murrow and F. T. Nicholson, "Station Location Covariance," JPL IOM GLL-NAV-88-50 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 2, 1988.

⁵ D. W. Murrow, "Galileo Earth Approach Analysis Using Connected-Element Interferometry Data," JPL IOM GLL-NAV-89-28 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 21, 1989.

phase-delay measurements, Edwards [2] has shown that successive observations separated by as little as 200–300 seconds should have a correlation coefficient of less than 0.1. The implication is that essentially uncorrelated CEI measurements may be generated as fast as is physically possible, given the capabilities of the current equipment (predicted to be about 300 seconds to perform one observation). It must be pointed out here that much less is known about random fluctuations in the ionosphere, which may have some impact on the degree of correlation between measurements taken at S-band (2.29 GHz).

In the next set of cases, simulations were generated using CEI data rates of 4, 8, and 12 points/hr for each of the measurement accuracies (8 mm, 4 mm, and 1 mm) used previously. Table 3 contains a summary of the results obtained. Figure 6 shows the B-plane error ellipses obtained with data rates of 1, 4, 8, and 12 points/hr for 4-mm (200-nrad) CEI data. Figures 7 and 8 portray the behavior of the B-vector component uncertainties as a function of the CEI data rate for each of the data weights used (1 mm, 4 mm, and 8 mm). Finally, Figs. 9 and 10 give error breakdowns of the B-vector components for the four cases run using 4-mm CEI. The principal effects of adding in more CEI data are reductions in the magnitude of the computed error and the station-location consider error in B-R, as shown in Figs. 9 and 10. The behavior of the error source contributions seen in these two figures is representative of the results obtained for 8-mm and 1-mm CEI as well.

The final set of simulations carried out included station-location errors for DSS 13 and DSS 14 as consider parameters. A correlated station-location covariance was generated for DSS 13 and DSS 14 in the same manner employed by Murrow and Nicholson,⁶ such that the resulting matrix contained a 2-cm uncertainty in each baseline component. The absolute coordinate uncertainties were as follows:

DSS 13:	radial	0.43 m	DSS 14:	radial	0.43 m
	polar	5.77 m		polar	5.77 m
	long.	0.58 m		long.	0.58 m

The proximity of the two stations results in identical values of uncertainty for the absolute position coordinates of each station.

A 2-cm baseline component uncertainty corresponds to about 1- μ rad orientation uncertainty, which should be achievable once some quasar observations from Goldstone

baselines become available. The quasar direction uncertainty of 100 nrad was retained here, even though the actual value is about 20 nrad. This was done to facilitate comparison of the results with those in which baseline component errors were not explicitly modeled, even though some "double-accounting" is admittedly taking place.

Six cases were run using different values of data rate and measurement accuracy, with results given in Table 4. Figures 11 and 12 show a comparison of the results for cases with the same tracking schedule and measurement accuracy, both with and without station-location errors for DSS 13 and DSS 14. Performance degradation was largest for cases with the highest data rate and measurement accuracy. Remarkably, though, the degradation observed was insignificant in virtually all of these cases. The "double-differencing" of the spacecraft phase delay with that of the quasar, coupled with the proximity of the two stations, cancels out station-location errors quite well, making the CEI observations nearly insensitive to baseline errors, at least at the 2-cm level.

The effects of baseline length and orientation errors on the CEI phase-delay measurement can be approximated with the following simple formula, which ignores correlations between the two:

$$\sigma_r^2 = (\sigma_B \cdot a)^2 + (B \cdot a \cdot \sigma_\phi)^2$$

where

σ_r = phase-delay measurement uncertainty (cm)

σ_B = baseline length uncertainty (cm)

σ_ϕ = baseline orientation uncertainty (rad)

a = spacecraft-quasar angle (rad)

Using this formula, a 1 σ baseline length uncertainty of 2 centimeters, coupled with a 1 σ orientation uncertainty of 1 μ rad and a spacecraft-quasar angle of 3 degrees (a typical value for Galileo and P0528+134), produces a 1 σ measurement uncertainty of about 1.5 mm for the DSS 13 to DSS 14 baseline (which is 21.5 km in length). This amount of error only starts becoming significant for the cases with an assumed CEI measurement uncertainty of 1 mm, as seen in Fig. 12. For the range of CEI measurement accuracy investigated here, baseline component errors would probably have to exceed the 5-cm level to have much impact on navigation performance.

VII. Summary

For an assumed measurement accuracy of 200 nrad, the addition of CEI data at one point/hr brings about a

⁶ D. W. Murrow and F. T. Nicholas, "Station Location Covariance," JPL IOM GLL-NAV-88-50 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 2, 1988.

40 percent improvement in the encounter targeting uncertainty in declination ($B \cdot T$ direction). If the data rate is increased to 10 to 12 points/hr, the improvement in declination uncertainty may be up to 66 percent, again using 200-nrad CEI. With an assumed measurement accuracy of 400 nrad, the level of improvement in declination uncertainty was 15 to 55 percent, indicating that the information content of the CEI data is significant even at a modest level of accuracy. Some degradation occurred in the in-plane ($B \cdot R$ direction) targeting uncertainty, which was 5 to 25 percent greater than that obtained with range and Doppler data only, depending on the CEI data rate and measurement accuracy used. This degradation is due to the influence of parameters which are not explicitly modeled, but are known to affect the radio metric data used in determining the flight path (consider parameters).

Several simulations were performed to study the effects of baseline length and orientation errors on the navigation performance obtained using CEI. With an assumed 2-cm uncertainty in the three components of the DSS 13 to DSS 14 baseline, the degradation in performance versus the earlier cases with no baseline errors was typically 6 percent or less; the largest degradation observed was 9.6 percent.

The absolute station-location errors used were nearly the same as those presently used by the Galileo Project. These results indicate that CEI data are remarkably insensitive to both absolute station-location errors and baseline component errors.

VIII. Conclusions

The CEI phase-delay tracking demonstration with the Galileo spacecraft during its approach to Earth may produce a greatly improved trajectory solution over that which will be used operationally. In the cases studied, the addition of more CEI data by increasing the data rate continued to improve the navigation performance and did not further exaggerate the effects of the consider parameters, which were small for the most part. Simulations which included baseline component uncertainties showed that, at the 2-cm level, these error sources cause no significant degradation in performance. The results as a whole indicate that CEI tracking can improve navigation performance for this trajectory with modest assumptions in measurement accuracy, station-location errors, and quasar direction errors.

References

- [1] C. D. Edwards, "Angular Navigation on Short Baselines Using Phase Delay Interferometry," *IEEE Transactions on Instrumentation and Measurement*, vol. 38, no. 2, pp. 665-667, April 1989.
- [2] C. D. Edwards, "The Effect of Spatial and Temporal Wet-Troposphere Fluctuations on Connected Element Interferometry," *TDA Progress Report 42-97*, vol. January-March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 47-57, May 15, 1989.

Table 1. S-band data

Time Period, days	Data Types	Data Rate	DSS
E-298 to E-268	Range Doppler	1 pt/day 1 pt/hr (daily passes)	14, 43, 63 14, 43, 63
E-268 to E-219	Range Doppler	2 pts/wk 2 passes/wk at 1 pt/hr	43 43
E-219 to E-207	Range Doppler	1 pt/day 1 pt/hr (daily passes)	14, 43, 63 14, 43, 63
E-207 to E-70	Range Doppler	2 pts/wk 2 passes/wk at 1 pt/hr	43 43
E-70 to E-58	Range Doppler	1 pt/day 1 pt/hr (daily passes)	14, 43, 63 14, 43, 63
E-58 to E-35	Range Doppler	2 pts/wk 2 passes/wk at 1 pt/hr	43 43
E-35 to E+1	Range Doppler	1 pt/day 1 pt/hr (daily passes)	14, 43, 63 14, 43, 63
E-27 to E-20	CEI	1 pass/day (data rates from Quasar P0528 + 134 1-12 pts/hr)	DSS 13-DSS 14 baseline
Elevation cutoff for all data = 15° Quasar P0528 + 134: right ascension = 82.03° declination = 13.49°			

Table 2. Results summary for CEI data rate of 1 point/hr

CEI Data Weight, mm	$\sigma B \cdot R$, km	$\sigma B \cdot T$, km	$\sigma LTOF$, sec
(Base Range, Doppler)	14.55	6.20	0.184
8 (400 nrad)	12.54	7.97	0.192
4 (200)	8.7	8.08	0.192
1 (50)	4.39	5.34	0.190

Table 3. Sensitivity of results of CEI data rate

CEI Data Weight, mm	Data Rate, pt/hr	$\sigma B \cdot R$, km	$\sigma B \cdot T$, km	$\sigma LTOF$, sec
8 (400 nrad)	1	12.54	7.97	0.192
8	4	8.98	8.14	0.192
8	8	7.35	8.22	0.191
8	12	6.57	8.14	0.191
4 (200 nrad)	1	8.78	8.08	0.192
4	4	6.07	7.96	0.191
4	8	5.21	7.24	0.191
4	12	4.77	6.55	0.190
1 (50 nrad)	1	4.39	5.84	0.190
1	4	3.41	3.71	0.187
1	8	3.07	3.12	0.185
1	12	2.93	2.89	0.184

Table 4. Navigation performance with CEI baseline errors included

CEI Data Weight, mm	Data Rate, pts/hr	$\sigma B \cdot R$, km	$\sigma B \cdot T$, km	$\sigma LTOF$, sec
8	4	9.06	8.17	0.192
8	12	6.60	8.18	0.191
4	4	6.15	8.03	0.191
4	12	4.82	6.55	0.190
1	4	3.52	3.86	0.187
1	12	3.04	3.16	0.184

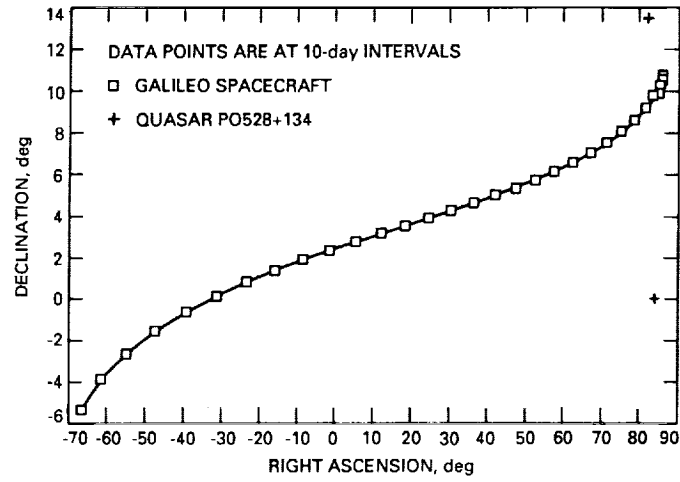


Fig. 1. Galileo EGA1 approach trajectory.

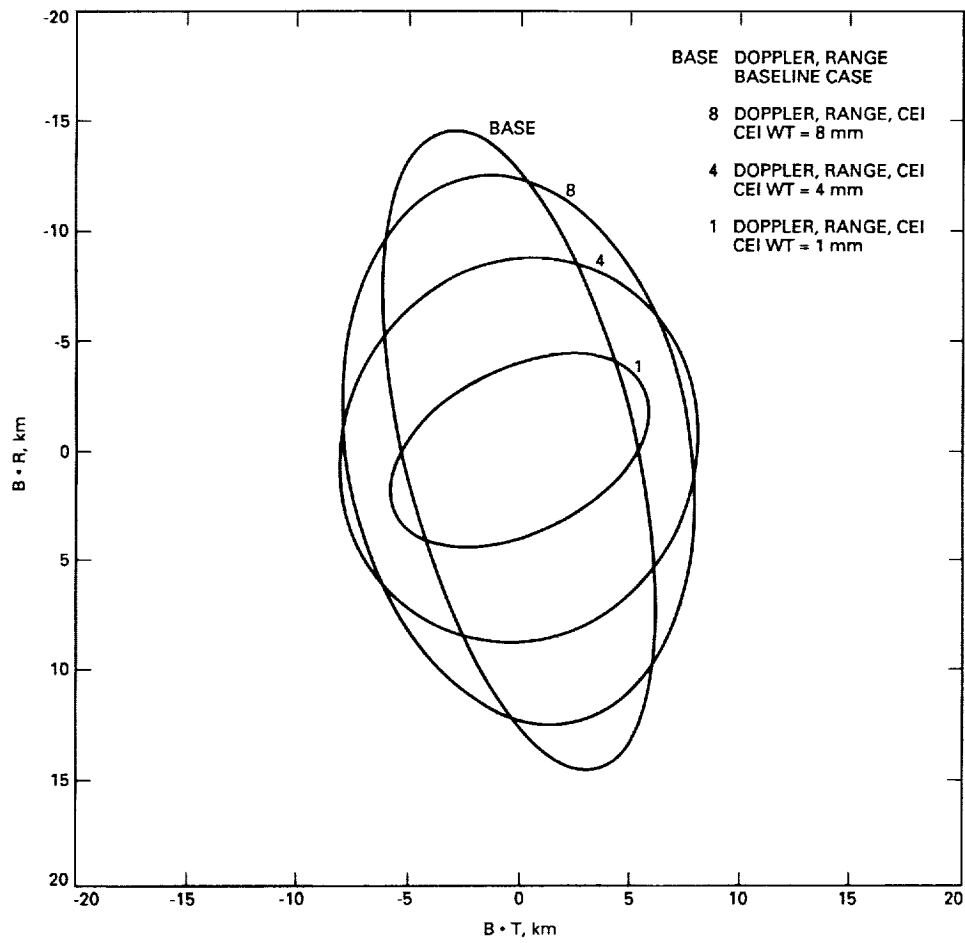


Fig. 2. The B-plane error ellipses for Galileo EGA1 (footnote 5 in text).

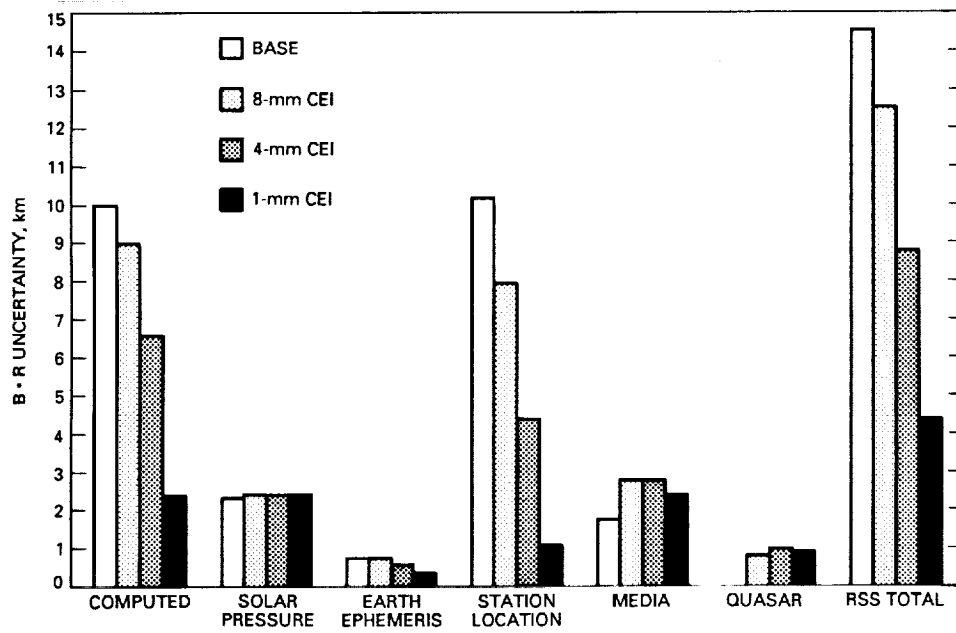


Fig. 3. Error breakdown for B-R component.

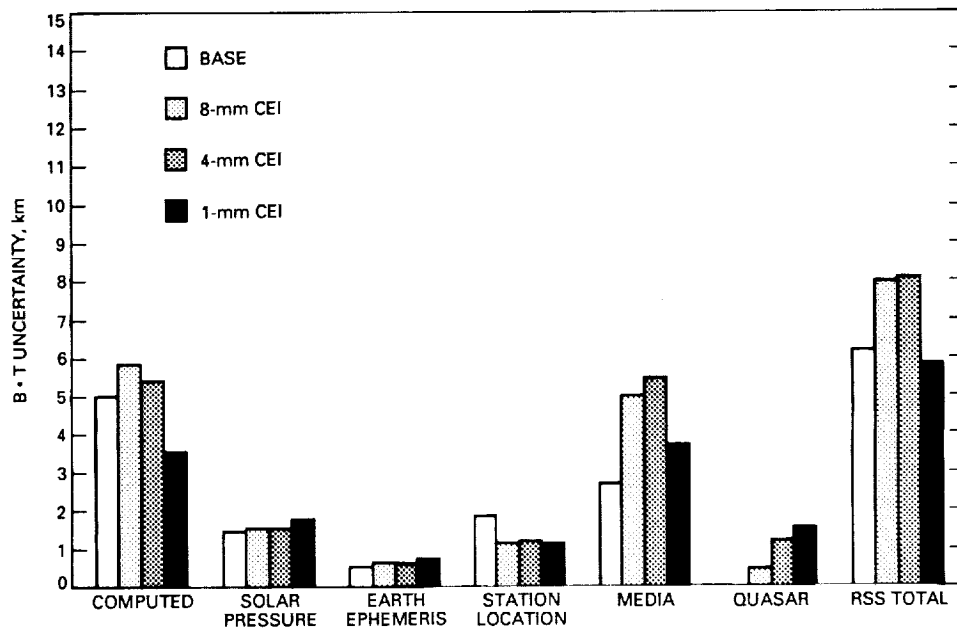


Fig. 4. Error breakdown for B-T component.

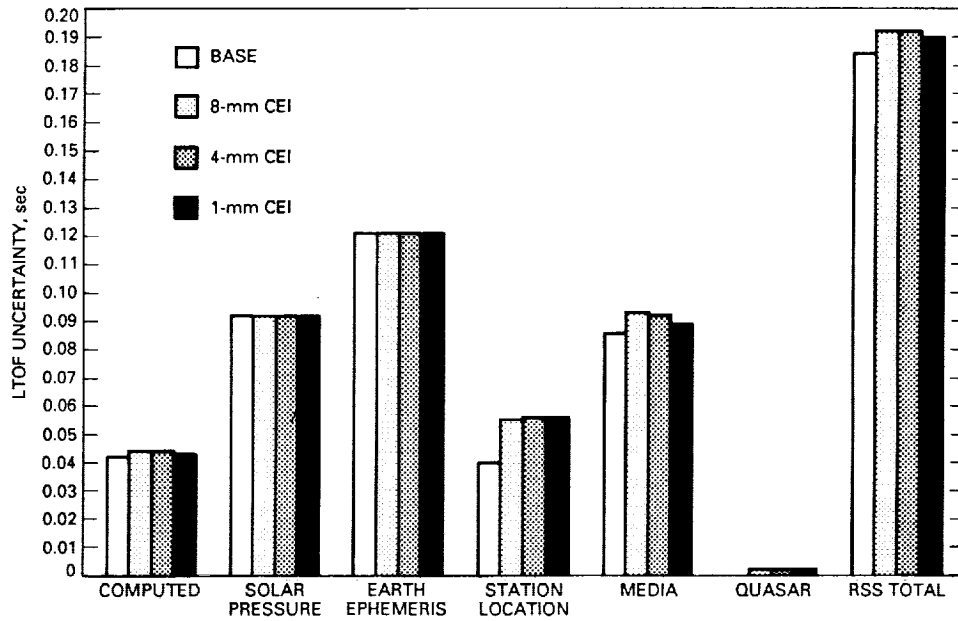


Fig. 5. Error breakdown for LTOF.

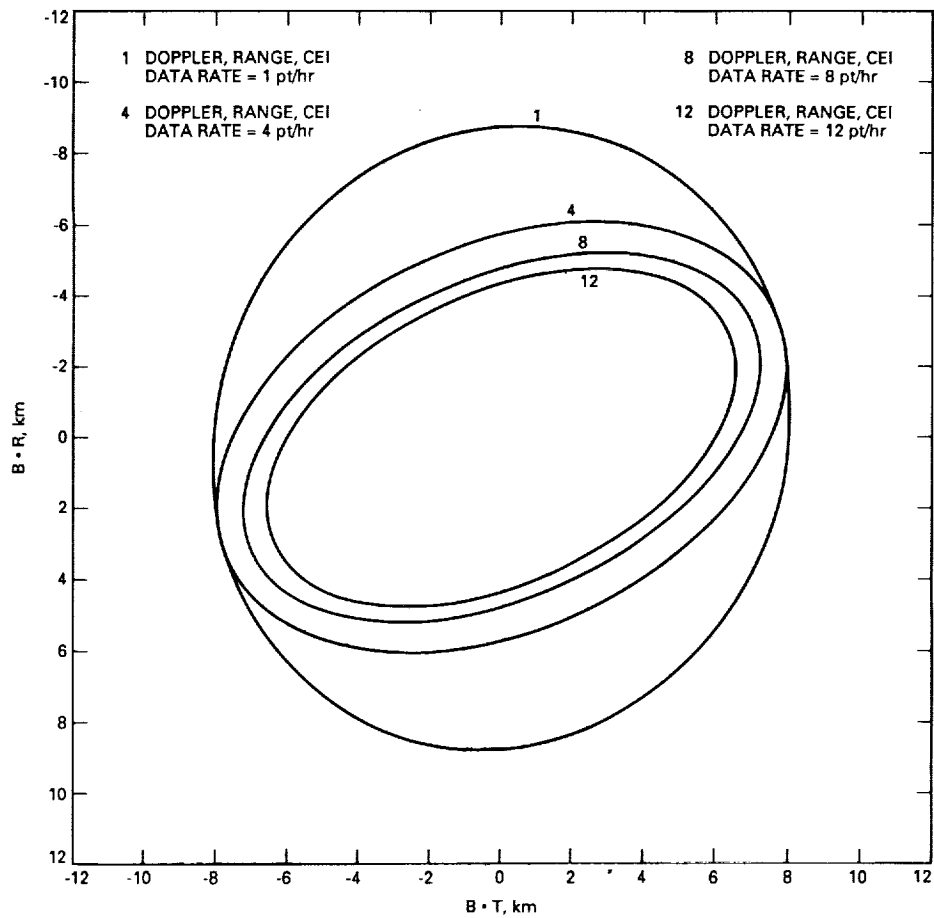


Fig. 6. The B-plane results for different data rates (4-mm CEI).

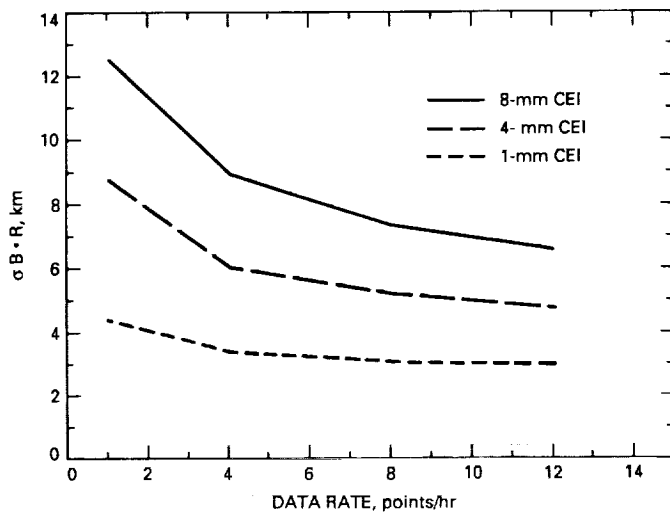


Fig. 7. The B-R uncertainty versus data rate.

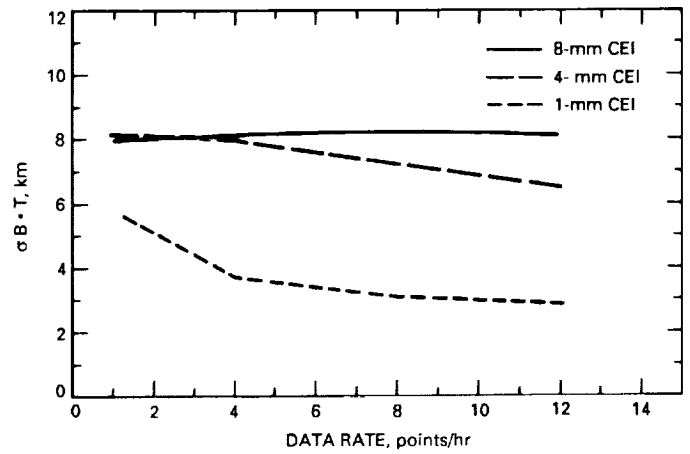


Fig. 8. The B-T uncertainty versus data rate.

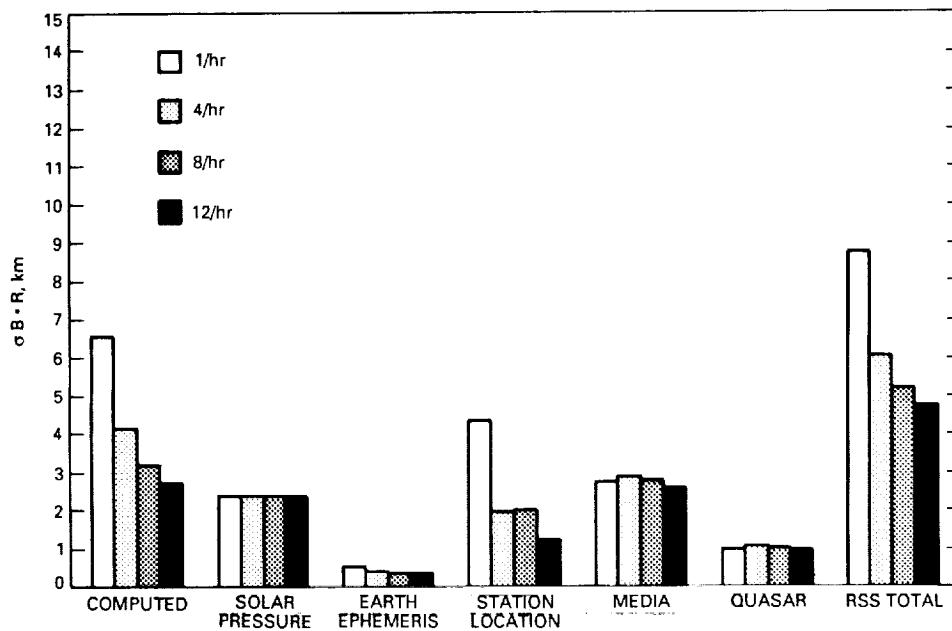


Fig. 9. Error breakdown of B-R for different CEI data rates (using 4-mm CEI).

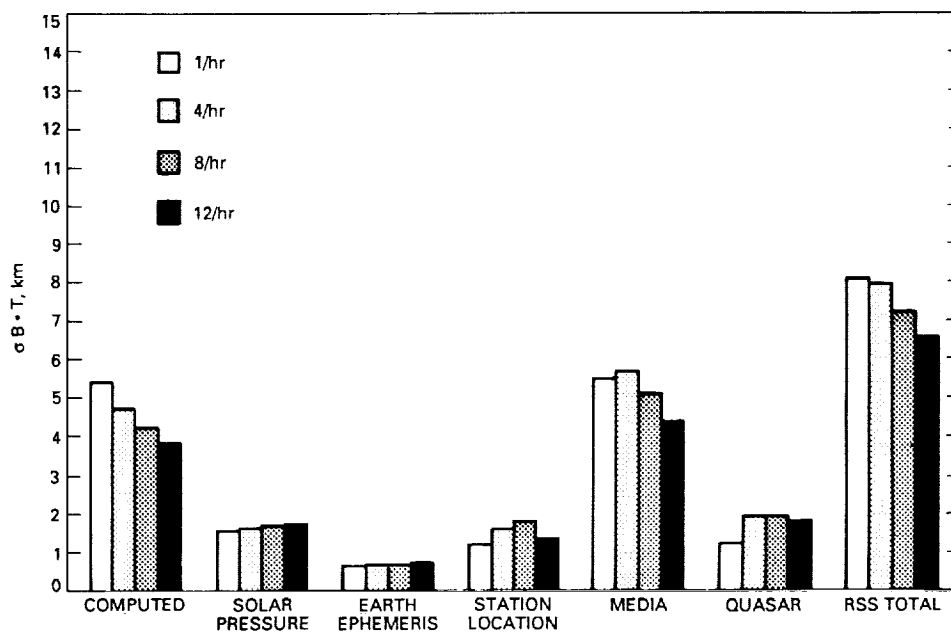


Fig. 10. Error breakdown of B-T for different CEI data rates (using 4-mm CEI).

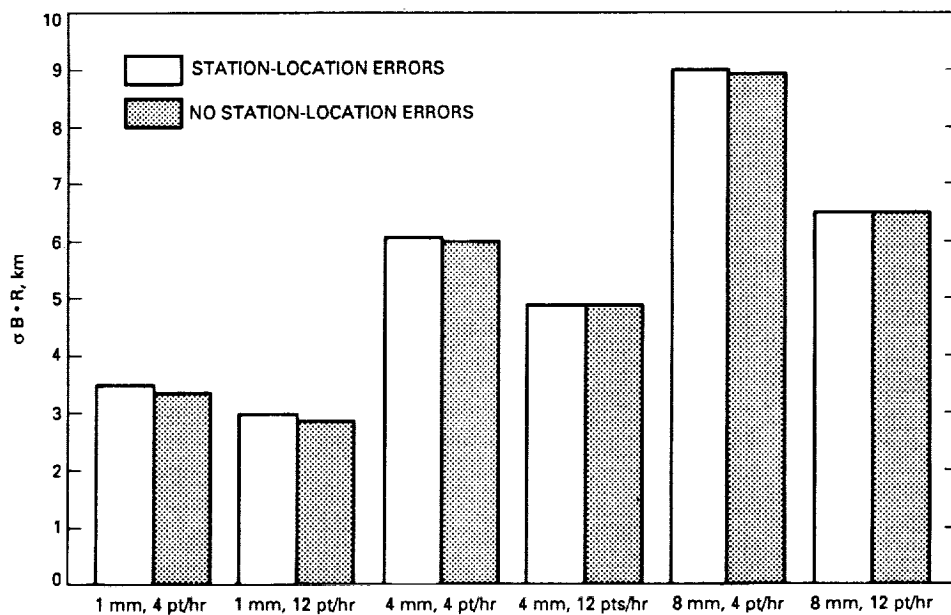


Fig. 11. Performance comparison with station location errors Included (B-R component).

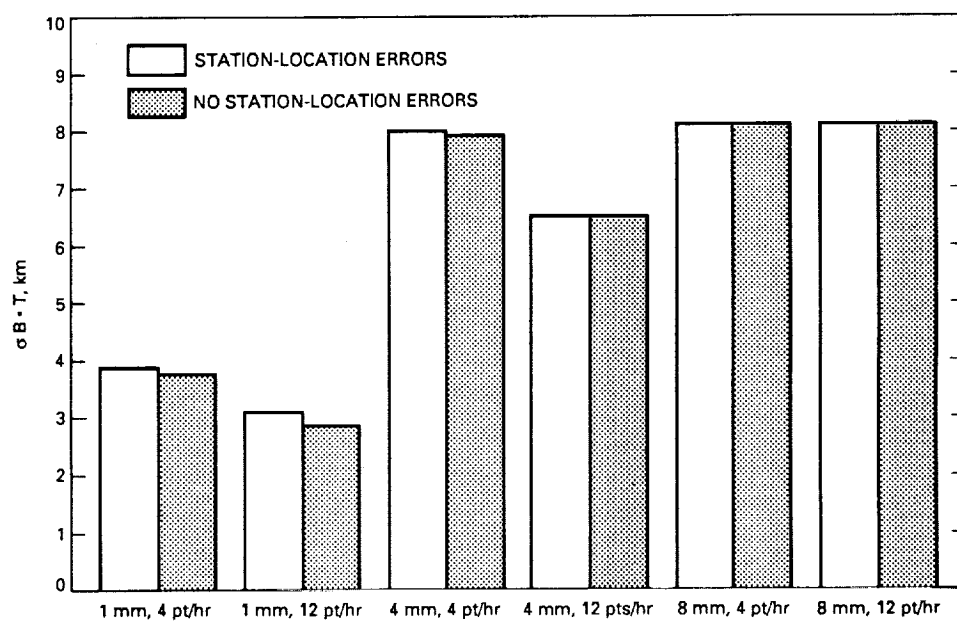


Fig. 12. Performance comparison with station location errors included (B-T component).

55-17
27232/
7B.

Using Connected-Element Interferometer Phase-Delay Data for Magellan Navigation in Venus Orbit

S. W. Thurman and G. Badilla
Navigation Systems Section

A 2110

The pointing accuracy needed to support Magellan's Synthetic Aperture Radar mapping of Venus places stringent requirements on navigation accuracy. This need is met with a combination of two-way Doppler and narrowband Δ Very Long Baseline Interferometer (Δ VLBI) data, which are capable of determining the spacecraft's orbit to the required level, typically about one-kilometer position uncertainty. Differenced Doppler (two-way Doppler minus three-way Doppler) is also capable of meeting mission navigation requirements, and serves as a backup to narrowband Δ VLBI. The Magellan Project specifies that the turn-around time for processing narrowband Δ VLBI data must be 12 hours or less, a very difficult requirement to meet operationally. In this article, the use of phase-delay data, taken from a Connected-Element Interferometer (CEI) with a 21-km baseline, for Magellan orbit determination was investigated to determine if navigation performance comparable with narrowband Δ VLBI and differenced Doppler could be achieved. CEI possesses an operational advantage over Δ VLBI data in that the observables are constructed in near-real time, thus greatly reducing the turn-around time needed to process the data, relative to the off-line system used to generate Δ VLBI observables. Unfortunately, the results indicate that CEI data are much less powerful than narrowband Δ VLBI and differenced Doppler for orbiter navigation, although there was some marginal improvement over the navigation performance obtained when only two-way Doppler data were used.

I. Introduction

Magellan will arrive at Venus in August 1990, and after a short period of time devoted to orbit trim and systems checkout, the spacecraft will begin its 243-day primary mission phase, the mapping of Venus. Navigation of Magellan will be very challenging due to the high pointing accuracy required by the Synthetic Aperture Radar (SAR) system. The SAR pointing accuracy requirements correspond to a navigation accuracy of about a one-kilometer

(1σ) position uncertainty at periapsis. Furthermore, this level of performance must be maintained over most of the primary mission, except for an 18-day period around superior conjunction, during which mapping operations are suspended.¹

¹ S. Mohan, et al., "Magellan Navigation Plan," JPL D-1480 (internal document), Revision B, Jet Propulsion Laboratory, Pasadena, California, March 23, 1988.

Navigation studies carried out early in Magellan's history² discovered that Doppler-based orbit determination, which had been sufficient for previous missions such as Mariner-Mars '71 and Viking, would not meet navigation requirements and that new radio metrics, such as narrowband Δ Very Long Baseline Interferometry (Δ VLBI) and differenced Doppler, would be needed [1]. It is now well known that the orbit determination accuracy obtained with two-way Doppler is dependent on the orbit orientation with respect to the plane of the sky, which is the plane perpendicular to the Earth-spacecraft line of sight [2,3]. Recent studies done in support of Magellan navigation planning¹ [4] show that the combination of either narrowband Δ VLBI or differenced Doppler with two-way Doppler produces improvements of a factor of 2 to 10 in orbit determination uncertainty over Doppler-only navigation.

The major drawback involved with the use of narrowband Δ VLBI data is the time required to process the raw measurement data into observables. The Magellan Project requires a turn-around time of 12 hours for each of the two narrowband Δ VLBI scans scheduled daily during the primary mission. Connected Element Interferometry (CEI), a new technique currently under development, can be used to generate medium accuracy (100–400 nrad for individual measurements) angular observations in near-real time, using relatively close (10–100 km apart) stations [5]. This study investigated the possibility that CEI angular measurements could be used to infer angular rate information, which is directly observed with narrowband Δ VLBI. It was hoped that this approach might be capable of producing orbit determination accuracy comparable with that of Δ VLBI, thus offering an alternative data type for Magellan navigation that would reduce the operational strain imposed on the Deep Space Network (DSN).

II. The Magellan Venus Mapping Orbit

Magellan orbital elements relative to Venus' equator are given in Table 1. Since Venus has a nearly spherical gravity field, the spacecraft's orbit remains nearly fixed in orientation with respect to Venus throughout the primary mission. The mapping orbit's orientation with respect to the plane of the sky does, however, change during the primary mission due to the relative motion of Venus and the Earth. The orbit geometry chosen for analysis in this study represents Magellan's orbit as seen from the Earth on December 4, 1990, 120 days after Venus Orbit

Insertion (VOI+120). Table 2 gives the orbital elements in the plane-of-the-sky coordinate system for this date.

III. Tracking Schedule

Nominally, trajectory solutions for Magellan are generated daily, based on 24 hours of continuous two-way Doppler data from Deep Space Stations (DSSs) 14, 42, and 61 and on two 10-minute Δ VLBI scans using the baselines defined by DSS 14 to 42 and DSS 14 to 61, taking place during the same 24-hour period. Doppler data are not taken for a 1-hour period centered about each periapsis, since the spacecraft's antenna is pointed at Venus during that time. Doppler data are also not taken from 10 minutes before to 5 minutes after each apoapsis while a star sighting is being done to update the spacecraft's inertial attitude reference system.³ Since Magellan's orbit period is 3.15 hours, the data going into each trajectory solution are spread over just under eight revolutions. There are certain time periods within the primary mission in which the declination of the spacecraft is too low for it to be viewed from the DSS 14 to DSS 61 baseline, thus leaving out one of the two daily Δ VLBI scans. The epoch for this investigation, December 4, 1990, was intentionally chosen during one of these periods in which only one Δ VLBI baseline is available, since it is these periods in which the mission navigation requirements can only marginally be met.

Simulated CEI data were generated using the DSS 13 to DSS 15 baseline within the Goldstone DSN complex. This baseline is 21 kilometers in length and is oriented in a near north-south direction. During the 24-hour period used in the study, the spacecraft was in view of DSS 13 to DSS 15 for just under three revolutions, with a total view time of six hours available, subject to the restrictions described above regarding time periods during which the spacecraft could not be tracked. Two different CEI data rates were used in the CEI navigation simulations, either 5 points/hour or 10 points/hour.

One case was studied which included differenced Doppler data (two-way Doppler minus three-way Doppler) in place of narrowband Δ VLBI data. The data schedule for this case contained the Doppler data from the baseline case plus one pass (approximately two hours) of differenced Doppler data using the DSS 14 to DSS 42 baseline. Differenced Doppler data are constructed by differencing the Doppler signal received at a Deep Space Station operating in a coherent (two-way) mode with the Doppler signal received at another, distant station (known as "three-way"

² J. Ellis and R. Russell, "Earth-Based Determination of a Near-Circular Orbit About a Distant Planet," JPL Technical Memorandum 391-406 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 9, 1973.

³ S. Mohan, et al., op. cit.

Doppler). In theory, the information content of differenced Doppler data is equivalent to that of narrowband Δ VLBI data. In practice, the narrowband Δ VLBI data are more accurate due to the cancellation of the effects of several error sources common to both receiving stations through the use of a quasar (extragalactic radio source, or EGRS). Because of this, it is necessary to obtain more differenced Doppler data than narrowband Δ VLBI data to get comparable navigation performance.

IV. Filter Model

A. Type

The filter used was a batch-weighted, least-squares, with the exception that narrowband Δ VLBI and differenced Doppler data were processed sequentially due to the presence of stochastic parameters, as noted below.

B. Estimated Parameters

- (1) Spacecraft epoch state vector.
- (2) Venus harmonic gravity coefficients up to degree 3, order 3.
- (3) Spacecraft drag coefficient.
- (4) Stochastic station oscillator frequency error—this parameter is only included for cases with narrowband Δ VLBI data.
- (5) Stochastic troposphere error—this parameter is also included only for cases with narrowband Δ VLBI data.
- (6) Stochastic Doppler frequency error—this parameter is only included for the cases with differenced Doppler data.

C. Considered Parameters

- (1) Venus harmonic gravity coefficients of degree and order 4–5.
- (2) Mass of Venus.
- (3) Seven mass concentrations (*mascons*), see Table 3 for a list detailing the name, location, size, and a priori uncertainty of each.
- (4) Station coordinates of DSSs 14 (Goldstone), 42 (Canberra), and 61 (Madrid).
- (5) Impulsive maneuver representing daily momentum wheel desaturation.
- (6) Quasar location uncertainty (right ascension and declination).

D. Data Accuracy

- (1) Two-way Doppler: 10.0 mm/sec (60-sec count time).
- (2) Narrowband VLBI: 0.13 mm/sec (equivalent to 1 mHz).
- (3) Differenced Doppler: 0.39 mm/sec (equivalent to 3 mHz).
- (4) CEI phase-delay: 1 mm (50 nrad, 4 mm (200 nrad), or 8 mm (400 nrad).

V. A Priori 1σ Uncertainties

A. Estimated Parameters

- (1) Spacecraft epoch state: 10^6 km in position and 1 km/sec in velocity.
- (2) Venus 3×3 gravity harmonics.⁴
- (3) Spacecraft drag coefficient: 0.5 (50-percent uncertainty).
- (4) Stochastic station oscillator frequency error: the σ of this parameter is zero except during the 10-min Δ VLBI scan, when the $\sigma = 0.3538$ mHz, with a time constant $\tau = 27.6$ sec.
- (5) Stochastic troposphere error: this parameter also has a σ of zero except during the Δ VLBI scan, during which the $\sigma = 1.223$ mHz, with a time constant $\tau = 60.7$ sec.
- (6) Stochastic Doppler frequency error: the σ is zero except during the differenced Doppler pass, during which it has a $\sigma = 0.55$ mHz, with a time constant $\tau = 3604$ sec.

B. Considered Parameters

- (1) Remaining Venus gravity harmonics up to 5×5 .⁴
 - (2) GM of Venus: $1.0 \text{ km}^3/\text{sec}^2$.
 - (3) Mascons: see Table 3.
 - (4) Station coordinates for DSSs 14, 42, and 61:
- | | | | |
|--------------|----------------------|----------------------|----------------------|
| Coordinate | 14 | 39 | 54 |
| Spin rad., m | 0.74 | 0.81 | 0.79 |
| Z-height, m | 8.70 | 8.70 | 8.70 |
| Long., deg | 1.8×10^{-5} | 2.1×10^{-5} | 2.2×10^{-5} |
- (5) Momentum wheel desaturation burn: 8.3×10^{-7} km/sec.

⁴ Ibid., Appendix E.

- (6) Quasar location uncertainty: 167 nrad in both right ascension and declination.

VI. Results

A total of eight simulation cases was run, all using the Magellan delivery of the Orbit Determination Program on a Sun 3/260 workstation. Table 4 contains a summary of the results, with navigation performance being defined in terms of the 1σ uncertainty in the longitude of the ascending node in the plane of the sky, the dominant orbit-determination error. Figure 1 shows a graphic comparison of the eight cases. This figure shows that, of the five cases run using simulated CEI data, only one yielded results that were appreciably better than the Doppler-only case. Recalling that only one scan of narrowband Δ VLBI data was included in the Δ VLBI case (taken from the DSS 14 to DSS 42 baseline; the spacecraft cannot be viewed from the DSS 14 to DSS 61 baseline at this time), Fig. 1 shows that the performance obtained using two hours of differenced Doppler data was actually better than that obtained with Δ VLBI data. This is not always the case however, when narrowband Δ VLBI scans can be taken using both the DSS 14 to DSS 42 and the DSS 14 to DSS 61 baselines.

A breakdown of the root sum square (RSS) contribution of different error source groups to the total RSS error

in longitude of the ascending node is given in Table 5. Figure 2 shows a graphic version of the data in Table 5, the four cases in the figure corresponding to the four cases in the table. Figure 2 shows clearly that the CEI data are much more sensitive to gravity field uncertainty and mascons than the Δ VLBI data. Even excluding all consider error sources, the comparison of the computed (state-only) solution uncertainty shows that narrowband Δ VLBI data are much more powerful than CEI data in this situation. Further, Fig. 1 shows that, even if the CEI data were accurate to 1 millimeter (which is very optimistic), the sensitivity of this data type to the consider errors becomes extreme.

VII. Conclusions

An error analysis was conducted to determine the usefulness of CEI phase-delay data in the navigation of the Magellan mission. The results indicate that CEI data are much less capable than narrowband Δ VLBI and differenced Doppler data in this situation, both in information content, which is reflected in the comparison of computed solution uncertainties, and in sensitivity to consider error sources such as gravity field uncertainty. Based on the results presented in this article, this data type would likely not be capable of meeting Magellan navigation requirements during any portion of the primary Venus mapping mission.

References

- [1] S. Poole, M. Ananda, and C. Hilderbrand, "Radio Interferometric Measurements for Accurate Planetary Orbiter Navigation," Paper AAS-79-121, AAS/AIAA Astrodynamics Conference, Provincetown, MA, pp. 25-27, June 1979.
- [2] L. Wood, "Orbit Determination Singularities in the Doppler Tracking of a Planetary Orbiter," *Journal of Guidance, Control and Dynamics*, vol. 9, no. 4, pp. 485-494, July-August 1986.
- [3] R. Russell and S. Thurman, "An Analytic Development of Orbit Determination for a Distant Planetary Orbiter," Paper AAS-89-402, AAS/AIAA Astrodynamics Conference, Stowe, VT, pp. 7-10, August 1989.
- [4] P. Esposito and S. Demcak, "Orbit Determination Study Results for the Venus Radar Mapper Orbiter," Paper AAS-85-414, AAS/AIAA Astrodynamics Conference, Vail, CO, pp. 12-15, August 1985.
- [5] C. Edwards, "Angular Navigation on Short Baselines Using Phase Delay Interferometry," *IEEE Transactions on Instrumentation and Measurement*, vol. 38, no. 2, pp. 665-667, April 1989.

Table 1. Magellan orbital elements relative to Venus equator

Epoch: December 4, 1990 18:20:42 UTC	
Semi-major axis:	a 10189.7
Eccentricity:	e 0.38168
Inclination:	i 85.9743°
Argument of periapsis:	ω 170.016°
Longitude of ascending node:	Ω 354.205°
Altitude at periapsis:	249.50 km
Altitude at apoapsis:	8027.0 km
Sun-Earth-probe angle:	9.0°

Table 2. Magellan orbital elements relative to the plane of the sky

Epoch: December 5, 1990 22:42:30 UTC	
Semi-major axis:	a 10189.7 km
Eccentricity:	e 0.38168
Inclination:	i 55.7623°
Argument of periapsis:	ω 84.2673°
Longitude of ascending node:	Ω 10.0999°

Table 3. Mascons considered in Magellan orbit determination^a

Name	Lat, deg	Long, deg	Radius, km	σ km ³ /sec ²
Ovda Regio	-5.4	95.6	900	0.040
Leda	42.5	55.2	600	0.012
Hestia	6.4	64.6	350	0.010
Tellus	34.0	83.2	900	0.020
Bell Regio	28.9	49.0	500	0.010
FN27	7.1	78.0	350	0.010
Maxwell	65.0	10.0	900	0.040

^aAll seven mascons have a radial position of 6000 km; for comparison, Venus has a mean radius of 6052 km.

Table 4. Summary of results for all cases

Case	Quasar-S/C Angle, deg	Node Uncertainty, 1σ
CEI 50 nrad, 10 pt/hr	26	0.25650°
CEI 200 nrad, 10 pt/hr	26	0.10530°
CEI 400 nrad, 10 pt/hr	26	0.10140°
CEI 200 nrad, 5 pt/hr	26	0.09959°
CEI 200 nrad, 10 pt/hr	3	0.08809°
Doppler only	–	0.10450°
Doppler + Δ VLBI	26	0.03679°
Doppler + differenced Doppler	–	0.02963°

Table 5. Root-sum-square error breakdown of node error for selected cases

Error Source	RSS Contribution, deg ^a			
	Case 1	Case 2	Case 3	Case 4
Computed	0.0242	0.0479	0.0500	0.0532
Gravity field	0.0195	0.0650	0.0700	0.0801
Mascons	0.0071	0.0229	0.0246	0.0281
Quasar location	0.0183	0.0265	0.0555	0.0158
Station locations	0.0003	0.0001	0.0001	0.0001
Momentum wheel desat	0.0002	0.0008	0.0009	0.0010
RSS total, deg	0.0368	0.0880	0.1053	0.1014

^aCase 1: Baseline operational case with Doppler, narrowband Δ VLBI
Case 2: Doppler, 4-mm CEI, data rate 10 pt/hr, quasar-S/C angle 3°
Case 3: Doppler, 4-mm CEI, data rate 10 pt/hr, quasar-S/C angle 26°
Case 4: Doppler, 8-mm CEI, data rate 10 pt/hr, quasar-S/C angle 26°

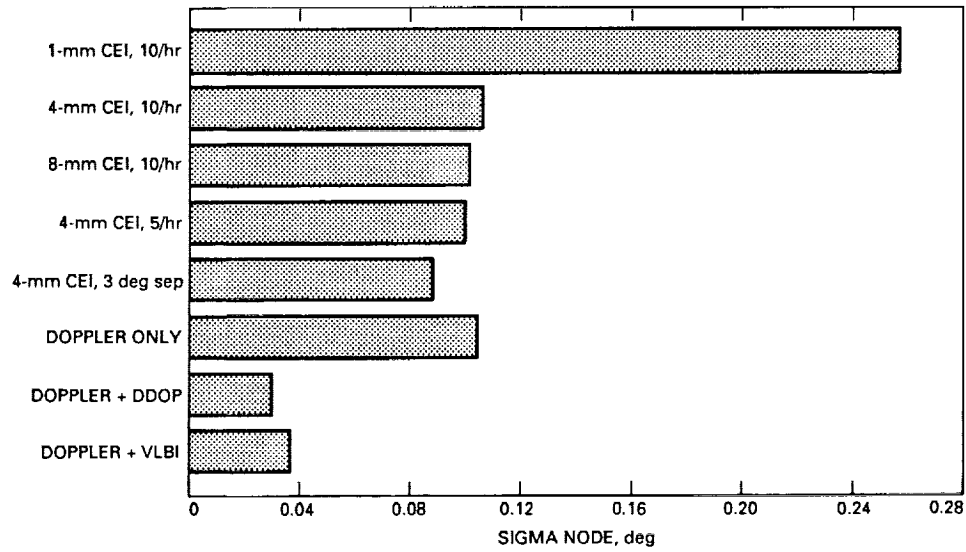


Fig. 1. Navigation performance obtained with different data strategies.

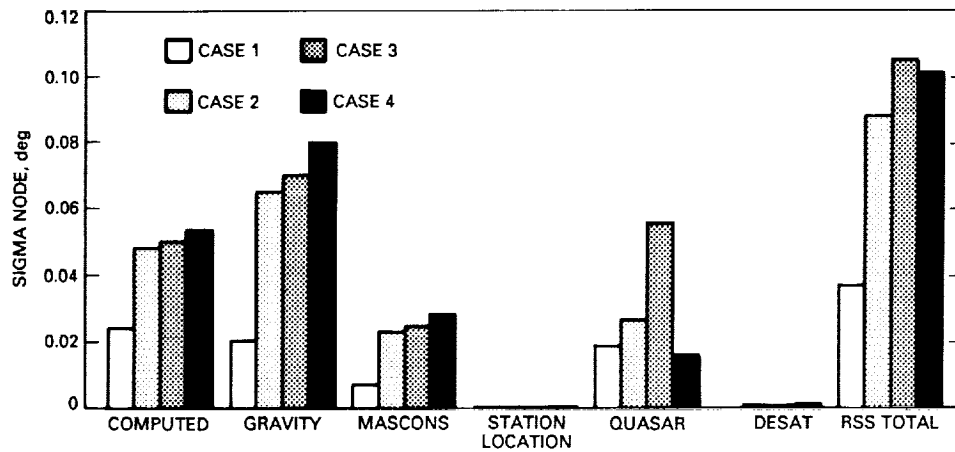


Fig. 2. RSS error breakdown for selected data strategies.

Voltage Source AC-to-DC Converters for High-Power Transmitters

R. Cormier

Radio Frequency and Microwave Subsystems Section

This work was done to optimize the design of the components used for the beam power supply, which is a component of the transmitters in the DSN. The major findings are: (1) the difference in regulation between a six-pulse and a twelve-pulse converter is at most 7 percent worse for the twelve-pulse converter; (2) the commutation overlap angle of a current source converter equals that of a voltage source converter with continuous line currents; (3) the sources of uncharacteristic harmonics are identified with SPICE simulation; (4) the use of an imperfect phase-shifting transformer for the twelve-pulse converter generates a harmonic at six times the line frequency; (5) the assumptions usually made in analyzing converters can be relaxed with SPICE simulation. The results demonstrate the suitability of using SPICE simulation to obtain detailed performance predictions of ac-to-dc converters.

I. Introduction

Large power converters (rated in megawatts) operate from three-phase power lines. Full-wave bridge rectifiers are utilized exclusively, resulting in six-pulse converters. By combining bridge rectifiers and an appropriate phase-shift transformer, multiples of six-pulse operation are obtained. The original transmitter beam supplies in the DSN are all six-pulse. Recently installed beam supplies, rated 75 kW and 2.25 MW, operate as twelve-pulse converters. These beam supplies are for klystron microwave amplifier tubes that require low values of ripple and good stability.

This investigation includes the results obtained for six-pulse and twelve-pulse operation. Two areas of particular interest for twelve-pulse operation are regulation and the effects of imperfections and approximations in the design and manufacture of the phase-shifting trans-

former on uncharacteristic harmonics. The twelve-pulse converter has been faulted for poor regulation. This, as will be shown, appears to be an exaggeration. Two types of harmonics are generated in converters. Characteristic harmonics result from operation under ideal conditions; the characteristic harmonics have been analyzed extensively, and numerous books have been published on the subject (e.g., [1]). Uncharacteristic harmonics are due to various imperfections in the generator and the converter. Measurements of the power supplies at Goldstone and in the laboratory yield significant values for uncharacteristic harmonics; the SPICE program has been used to identify their source. Current source converters provide a constant dc current to a load by the use of an inductor filter. Voltage source converters provide a constant dc voltage to a load by the use of a capacitor filter. The analysis found in the literature is limited to current source converters and voltage source converters with discontinuous line currents,

but the operation of voltage source converters with continuous line currents has not been analyzed previously. All the power supplies used in the DSN are voltage source converters with continuous line currents. The results of the SPICE simulation indicate that both types of converters have many similarities, but important differences do exist. The applicability of current source equations to voltage source converters is investigated in this article.

A. Assumptions

The mathematical analysis of any converter is based on assumptions that the transformer and the rectifiers meet the conditions known as regular connection and regular operation. For regular connection three conditions must be met. These are:

- (1) The total of all ampere-turns on each transformer leg is zero at any instant. The excitation current is excluded from this condition.
- (2) Line currents total zero at any instant. There are no zero- or negative-sequence components, and the excitation current is excluded.
- (3) Each successive interval of operation is identical to all other intervals.

Condition 2 for regular connections is never met in practice because power sources have distortion and practical transformers have stray capacitances. The distortion is equivalent to a negative-sequence component. The stray capacitances provide a path for the zero-sequence component.

For regular operation, three conditions must be met. These are:

- (1) The rectifiers are arranged in a regular connection, i.e., the dc provides a balanced dc line load.
- (2) For a current source converter, the load circuit contains an infinite inductance so that the current is constant. For a voltage source converter, the load circuit contains an infinite capacitance so that the voltage is constant.
- (3) The ac source has no impedance and the transformer has no leakage inductance. The individual commutations do not overlap, but they may occur simultaneously.

The conditions for regular operation are even more restrictive than those for regular connection. In practice, for a current converter the load inductance is not infinite; thus it follows that the direct current is not constant and

consequently contains ripple. Similarly, the load capacitance of a voltage source converter is not infinite, and the direct voltage is not constant. Practical generators and power distribution systems have inductances. Transformers have leakage inductance. The following analysis uses the SPICE program to analyze circuits that do not meet the conditions for regular operation.

B. Sequence Component Analysis

A standard method for the analysis of multiphase circuits is the use of the symmetric components (see [2]). The phase voltages are related to the symmetric components by the equation

$$\mathbf{V}_p = \mathbf{A}\mathbf{V}_s \quad (1)$$

where \mathbf{V}_p and \mathbf{V}_s are column vectors. The components of \mathbf{V}_p are V_a , V_b , and V_c , and those of \mathbf{V}_s are V_0 , V_1 , and V_2 . Given a set of three-phase voltages designated by the vector \mathbf{V}_p , these phase voltages can be resolved into a zero-, a positive-, and a negative-sequence set of components designated by the vector \mathbf{V}_s . For a three-phase source the matrix \mathbf{A} is

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{pmatrix} \quad (2)$$

and the inverse is

$$\mathbf{A}^{-1} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{pmatrix} \quad (3)$$

where $a = 1/\sqrt{3}$. Solving for the symmetric components

$$\begin{pmatrix} V_0 \\ V_1 \\ V_2 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & a & a^2 \\ 1 & a^2 & a \end{pmatrix} \begin{pmatrix} V_a \\ V_b \\ V_c \end{pmatrix} \quad (4)$$

From the above equations it is seen that the positive-sequence component (V_1) may be simulated by three voltage sources, one for each leg, and the negative-sequence component (V_2) by the addition of three series voltage sources of opposite phase rotation (see Fig. 1). The zero-sequence component (V_0) should be zero because the neutral wire is not connected to the load from the generator.

Any set of six-phase voltages can be represented by the addition of six symmetrical sequence components.

These are shown in Fig. 2. The zero- and the third-order-sequence components must have a neutral wire in order to exist. The positive- and second-order-sequence components rotate counterclockwise, while the negative- and fourth-order-sequence components rotate clockwise. Figure 3 shows the voltage vectors required for a twelve-pulse converter. The two sets of three-phase are phased 90 deg from each other, and in Fig. 3(a) the rotation is opposite, while in Fig. 3(b) the rotation is in the same direction. These are referred to respectively as counter rotation and synchronous rotation.

For a six-phase system, Eq. (1) can be written in terms of components

$$\begin{pmatrix} V_a \\ V_b \\ V_c \\ V_d \\ V_e \\ V_f \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -a & a^2 & -1 & a & -a^2 \\ 1 & a^2 & a & 1 & a^2 & a \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & a & a^2 & 1 & a & a^2 \\ 1 & -a^2 & a & -1 & a^2 & -a \end{pmatrix} \begin{pmatrix} V_0 \\ V_1 \\ V_3 \\ V_4 \\ V_5 \end{pmatrix} \quad (5)$$

where $a = 1/\sqrt{120^\circ}$.

To calculate the sequence components from the phase voltages, the inverse matrix A^{-1} is calculated

$$A^{-1} = \frac{1}{6} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -a^2 & a & -1 & a^2 & a^2 \\ 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & a^2 & a & 1 & a^2 & a \\ 1 & -a & a^2 & -1 & a & -a^2 \end{pmatrix} \quad (6)$$

and

$$\mathbf{V}_s = \mathbf{A}^{-1} \mathbf{V}_p \quad (7)$$

For the combination of two full-wave bridge rectifiers to form a twelve-pulse converter, a set of components that satisfies \mathbf{V}_p is

$$\mathbf{V}_p = (V_a \ V_b e^{j30} \ V_c e^{j120} \ V_d e^{j150} \ V_e e^{j240} \ V_f e^{j270})^t \quad (8)$$

This represents a synchronous rotating set of vectors. A counter rotating set, \mathbf{V}'_p , becomes

$$\mathbf{V}'_p = (V_a \ V_d e^{j150} \ V_c e^{j120} \ V_b e^{j30} \ V_e e^{j240} \ V_f e^{j270})^t \quad (9)$$

The symmetrical sequence components necessary to generate these phase voltages were calculated from Eq. (7). One solution for the synchronous rotating set is

$$\mathbf{V}_s = (0 \ 0 \ 0.2588e^{j75} \ 0 \ 0 \ 0.9659e^{-j15})^t \quad (10)$$

Another is

$$\mathbf{V}_s = (0 \ 0 \ 0.7071e^{j75} \ 0 \ 0 \ 0.7071e^{-j15})^t \quad (11)$$

An example of a counter-rotating set is

$$\mathbf{V}'_s = (0 \ 0.5e^{j90} \ 0.5 \ 0 \ 0.5e^{-j90} \ 0.5e)^t \quad (12)$$

C. Commutation Overlap

The commutation overlap angle μ is used to calculate the regulation, E_d/E_{do} , of a converter. The equation for the regulation of six-pulse converter is

$$E_d/E_{do} = 1/2 [\cos \alpha + \cos(\alpha + \mu)] \quad (13)$$

for $\mu < 60$ deg, and α is the angle introduced when phase control thyristors are used to control the output voltage. Since $\alpha = 0$ when the converter is implemented with rectifiers, this equation simplifies to

$$E_d/E_{do} = 1/2 [1 + \cos \mu] \quad (14)$$

This same equation applies to a twelve-pulse converter for μ less than or equal to 30 deg. For angles greater than 30 deg, a new phenomenon is observed in a twelve-pulse converter. The angle μ remains fixed at 30 deg and there is a delay in the start of commutation between the two converters. This phenomenon is described in [3]. The regulation equation derived for this mode is

$$E_d/E_{do} = \cos 15^\circ \cos(\alpha + 15^\circ) \quad (15)$$

The α in this equation is the same as that of the previous equation. The phase delay obtained with thyristors in a six-pulse converter appears as an automatic delay for a twelve-pulse converter, the result of interaction between the two six-pulse converters that make up the twelve-pulse operation. The two equations are identical, however, as can be seen from the trigonometric identity

$$\cos x \cos y = 1/2 [\cos(x + y) + \cos(x - y)] \quad (16)$$

Let $x = \mu/2$ and $y = (\alpha + \mu/2)$. Substituting these values for x and y in the trigonometric identity, Eq. (16), and

Eqs. (13) and (15) are identical. This implies that to calculate the regulation for a twelve-pulse converter with μ greater than 30 deg, first determine μ for the equivalent six-pulse converter and subtract 30 deg to find α . This is then used to calculate the regulation from either equation. The equations are valid for all values of $(\alpha + \mu) < 60$ deg. Equations (13) and (15) are derived in [3].

The conversion of ac to dc generates harmonics both on the dc side of the converter and on the ac lines. For a perfect system, only characteristic harmonics are generated; in practice, other harmonics are also generated. The characteristic harmonics are of the order np on the dc side and $np \pm 1$ on the ac lines, where n is the harmonic number and p is the converter pulse number. The current source converter has been analyzed and equations derived to calculate harmonics. The voltage source converter has not been analyzed. The applicability of the equations derived for the current source converter is limited and care must be exercised in making calculations with these equations. For example, the equation for the amplitude of the characteristic dc voltage harmonic for a current source converter is

$$C_n = \frac{2E_{do}}{(n^2 - 1)} \quad (17)$$

A similar equation is used to calculate the amplitudes of the characteristic current harmonics for a voltage source converter

$$C'_n = \frac{2I_{sc}}{(n^2 - 1)} \quad (18)$$

Equation (17) applies to an ideal converter operating at light load from a voltage source generator with zero line inductance. Equation (18) applies when the converter is operating into a short circuit from a voltage source generator with line inductances. With commutation overlap, the amplitude of the principal characteristic harmonic (C_q) increases with commutation angle. For a six-pulse current source converter the value may be calculated with the equation

$$C_q = \frac{E_{do}}{70} \sqrt{78 - 70 \cos 2\mu + 28 \cos 5\mu - 20 \cos 7\mu} \quad (19)$$

provided the commutation angle is less than 60 deg. (This equation is derived from [4].)

Equation (19) *cannot* be applied to a voltage source converter. The values of harmonics are a maximum at

short circuit and decrease with decreasing current. Unfortunately, no equation is available to calculate harmonics at operating load.

The operation of current and voltage source converters is similar even though not all equations can be applied to both. Large ac-to-dc power supplies are usually assumed to operate from a voltage source with negligible inductance. The outputs of the rectifiers are connected to a choke that provides constant dc current. However, it has been known for many years that a better solution is to operate multiphase converters into a capacitor in place of a choke. Power systems, either utility or auxiliary, have appreciable inductances. This results in the generation of voltage spikes at the output of the converter when operating into a choke. The problem of voltage spikes was experienced in the DSN when the high-power transmitters were first installed (see [5]). This caused the problem of initial failures of the filter chokes. The solution was to add a capacitor on the converter side of the choke. The generators had enough inductances so that line currents are continuous; otherwise the rms line currents become very large. The major advantage is the elimination of spike voltages and reduced ripple at the output of the converter.

If a current source generator were available, the converter would operate with no commutation overlap. (In fact, simulation of a voltage source converter was made and the output current was independent of the line inductance. For a discussion of the current versus voltage source converter operation, see [6].) This is not the case with a voltage source generator and line inductances. The commutation overlap angle is almost equal for both converters. Actually, the major portion of the inductance is located in the generator. A generator has three different values of inductances depending on the time of measurement after the initiation of the short circuit. The sustained value of the short circuit determines the synchronous reactance. The initial period is divided into two parts. The first cycle after initiation of the short is known as the subtransient reactance, followed by a longer period called the transient reactance. During converter operation, the commutation momentarily shorts the generator for a few degrees of a cycle. The response of the generator is based on the subtransient reactance, and this is the value used to determine the referenced short current. For the 400-Hz generators in the DSN, the subtransient inductance is roughly one-third that of the steady-state short-circuit inductance. The inductances should not be so large that the converter is operating in a higher mode (commutation overlap angle $\mu > 60$ deg), as such converter operation results in poor regulation. The short-circuit current is defined as that current which would be obtained with the

operational inductances in the ac lines. These are the sub-transient inductance of the generator plus the wiring and the leakage inductance of the transformers.

D. Other DC Ripple Harmonics

Other types of distortions, such as harmonics of the line frequency, are also present. These harmonics may be simulated by a harmonic voltage in series with each main generator's phase voltage. Low-frequency amplitude modulations may also be present; these are generally called subharmonics even though technically they are not necessarily frequency-related to the line frequency. In order to analyze the converter operation with different types of subharmonic distortion it is possible to introduce a set of positive- or negative-sequence components at various frequencies. However, the use of a single set of voltage sources does not represent the usual subharmonic distortion, which requires the insertion of two series sources to simulate amplitude modulation. As the armature rotates, the amplitude of the output line voltage varies, which results in amplitude modulation at a frequency below that of the line frequency. When the distortion is in the form of a low-frequency amplitude modulation, then the instantaneous voltage is

$$E_m(t) = V \sin w_c t + \frac{mV}{2} [\sin(w_c - w_s)t + \sin(w_c + w_s)t] \quad (20)$$

The first term represents the main generator voltage, and the second term represents the sidebands. The sidebands form the envelope that rides on top of the carrier. Although this equation is usually written in terms of cosines, the results remain unchanged when written in terms of sines. The use of sines is preferred because of the transients generated by the start of a cosine wave. Another type of distortion is frequency modulation, but this has not been investigated as it is not a usual distortion in power systems (see [7]).

II. Six-Pulse Converter Analysis

The circuit used for simulation consists of a three-phase wye-connected generator, a transformer with a delta primary, and an extended delta secondary. The output of the secondary is connected to a three-phase full-wave bridge diode converter. A capacitive filter and a resistive load are connected to the output of the rectifiers. For simulation purposes, two or three voltage sources are connected in series on each leg of the generator. One set generates the fundamental positive-sequence component, the other

sets introduce the various types of distortions being investigated. The results of the simulation are given in Tables 1 through 8. A listing of the basic circuit is given in Table 9; the listing was modified as appropriate for each test. Only steady-state conditions are of interest, therefore a delay in recording data was introduced so that initial turn-on transients did not contribute to the data. The capability of SPICE to analyze the converter operation is evaluated on the basis of a comparison of the values for the characteristic harmonics obtained from the simulation and those calculated. Various types of distortions are simulated and the resulting uncharacteristic harmonics identified.

A. DC-Side Ripple Harmonics

Table 1 lists the values of the characteristic harmonics obtained with the SPICE simulation for a voltage source converter. Column four lists the values for the current harmonics obtained; the next column contains the theoretical values calculated by Eq. (18). For example, when the principal harmonic at 2400 Hz is compared to the calculated theoretical value, a difference of only 0.06 dB was obtained. For a six-pulse converter, the characteristic harmonics are multiples of six times the line frequency; in our case the line frequency is 400 Hz. The significance of this result is twofold. First, there is close agreement of the SPICE simulation with calculated values. Second, the equations derived on the basis of E_{do} remain valid when I_{sc} is substituted for E_{do} , at least for this one equation. Note that the values of the characteristic harmonics are independent of the line inductances when the value is normalized to the dc short circuit current.

The excellent results obtained confirm the applicability of SPICE simulation for the analysis of the operation of converters. A second check of the capabilities of the SPICE program is the calculation of the uncharacteristic harmonics that are theoretically equal to zero. The values obtained are less than 60 dB from those for the characteristic harmonics, and this indicates an error of less than 0.1 percent for the capability of SPICE (see Table 2).

B. Commutation Overlap

Voltage converter harmonics are maximum at short circuit and decreased at rated load. Unfortunately, no formula is available to calculate the value of harmonics at reduced current. Operating at 1 MW, the simulation for the voltage source converter yields a commutation overlap angle of 28.8 deg, and the value of the principal harmonic is -36.86 dBI_{sc} . This is a decrease of 12 dB from the value at short circuit. At light load of 100 kW, the improvement was 15 dB (see Tables 3 and 4). The value of the principal harmonic for the voltage source converter referenced

to the rated current I_d in place of I_{sc} is $-24.86 + 18.35 - 12 = -18.51$ dB. (The -18.35 dB is the difference between short-circuit current and rated current.) This represents an increase of 6.35 dB or a doubling of ripple components in terms of currents. The klystron responds to voltage changes, so the current ripple must be converted to voltage ripple. This can be done provided the value of the filter capacitor is known. For example, using the value of $0.2 \mu\text{F}$, the principal harmonic that was calculated as equal to -18.51 dB becomes -34.12 dB when referenced to the dc output voltage E_d .

In order to compare the two types of converters, it is necessary to calculate the value of the principal harmonic for the current source converter operating with the same load. The SPICE listing was modified by replacing the capacitor with a choke filter. The commutation angle measured 30.4 deg. From Eq. (19) a value of -22.24 dB E_{do} was obtained. From Eq. (14), the output voltage is reduced from the light load value by 0.6 dB. Correcting for the change in output voltage, the ripple becomes $-22.24 + 0.6 = -21.64$ dB referenced to E_d . The difference of ripple values between the two converters, namely $-34.12 + 21.64 = -12.48$ dB, is lower for the voltage source converter, which is obtained at the expense of the addition of a $0.2 \mu\text{F}$ at the output of the converter. The voltage source converter with 0.6 mH of line inductance can operate from a maximum load of 1 MW to a light load of 100 kW before the load current becomes discontinuous. The choice of 0.6 mH is based on the value of subtransient reactance of the present generators that feed the DSN beam supply.

C. Negative-Sequence Component

Next, the effects of generator line distortion are investigated (see Table 5). It should be noted that the generators for the transmitters in the DSN are dedicated and no other loads contribute to the generation of line unbalance. For a capacitive filter, the harmonics are measured as currents and the value of the negative-sequence component is based on the short-circuit current in the same way as for the characteristic harmonics. For a negative-sequence component, the harmonic on the dc side has a frequency of twice the line frequency. As an example, a 1 percent negative-sequence component in a converter with a load of 1 MW and line inductances of 0.6 mH gives a -40.23 dB I_{sc} ripple component at 800 Hz. The generation of the 800 Hz appears to be the result of commutation angle variations. First, the angle is 18.86 deg, next 20.88 deg, and then 16.17 deg. This pattern repeats in a regular cyclic order. In terms of voltage, the component is -27.95 dB E_d . These simulations were performed with both sequence components in phase at startup. For the components out-of-phase, the output harmonic decreased

by approximately 3 dB. It should be noted that two out-of-phase sequence components can be resolved into in-phase components. From these results the generator negative-sequence component can be specified.

D. Harmonic Distortion

Simulations of the second, third, and fourth harmonics and both sequence components are made at a distortion level of 1 percent. The output ripple frequency depends on the input harmonic. The lowest output frequency of 400 Hz, which is the hardest to filter, is generated by the $+2$ and the -4 components. For the $+2$ component, a 1 percent (-40 dB) input harmonic level results in an output of -40.9 dB I_{sc} at a frequency of 400 Hz. In terms of voltage, the component value is -22.60 dB E_d . For the -4 component, the output is only one quarter as large. The results of the simulation are given in Table 6. Additional testing of other harmonics was not made because these have been analyzed and the results reported in [8] (Table 3.5).

E. Subharmonic Distortion

Another type of line voltage distortion generally present in practical generators is subharmonics of the line frequency. The subharmonics are not true harmonics at all but rather are low-frequency voltage modulations that appear on the lines. The simulated subharmonics are three-phase with either a positive or negative sequence. The first simulation was made with one series voltage source in each leg of the generator. This simulates single sideband distortion (see Table 7). Ripple frequencies appear as the sum or difference between subharmonic and the line frequencies. For a positive-sequence subharmonic, the ripple frequency is the difference frequency, while for a negative-sequence component, the ripple is the sum frequency. The subharmonic affects the commutation overlap angle, which appears to be the cause of the generation of these harmonics. The SPICE listing was changed to include two voltage sources to represent the two sidebands of an amplitude modulated signal. For the negative-sequence component, the ripple harmonics identified, although very weak, were detected at $2f_s$ and $3f_s$ (f_s = subharmonic frequency). When the positive-sequence component was tested, a stronger (-57 dB I_{sc}) single component was observed at the subharmonic frequency. This result, which is included at the bottom of Table 7, is consistent with measurements of transformer rectifier assemblies in service. A ripple of 20 Hz has been observed in the transmitters at Goldstone; this low a frequency is particularly troublesome because of the difficulty in filtering. By introducing a feedback component in the generator field supply, it is possible to reduce this component. The simulation iden-

tifies the subharmonic as an amplitude modulation of the output generator line voltage.

F. AC Line Harmonics

The ac line harmonics are found on the current and the voltage waveforms of the ac lines. For the usual current source converter, the line currents carry the harmonics, and the voltages are taken to be without any harmonics. This is changed for a voltage source converter. Ideally, the line currents are free of harmonics and the voltages carry all the harmonics. For an actual voltage source converter, both current and voltage harmonics are present. The theoretical values are calculated from knowing that the values of the ac harmonics are equal to $1/n$, where n is the harmonic number. For a six-pulse converter, n takes the values of 6 ± 1 for the first two harmonics. Values of -14 dB and -16.9 dB are calculated for the fifth and seventh harmonics, respectively (see [9]). The ac harmonics were simulated for various values of line inductances. The values were chosen to give commutation overlap angles from 7.5 deg to 60 deg. From the results, the current harmonics are very large at low values of line inductances. At 7.5 deg, the fifth harmonic current is almost half of the line current. As the commutation angle is increased, the current harmonics decrease and the voltage harmonics increase in value. At 60 deg, the fifth voltage harmonic equals one-fifth the line voltage. The voltage regulation increases as the commutation overlap angle increases (see Table 8, column 7). From these results, voltage converters operating from a voltage source must have line inductances and the operation should be with relatively large commutation overlap angles. The loss of voltage from commutation is nondissipative.

III. Twelve-Pulse Analysis

The generation of twelve-pulse requires a phase-shifting transformer to generate six-phase. If the six phases are connected to a single full-wave bridge, the ripple is not twelve-pulse but only six-pulse. The alternating voltages are pairwise equal and opposite and the result is the superpositioning of the pulses out of the converter. In order to obtain twelve pulses from a six-phase source, the phases must be broken up into two sets of three-phase arranged asymmetrically. This is done when a phase-shifting transformer with delta wye secondaries feeds two full-wave bridge rectifiers. An alternative is the use of two extended deltas, one phase-shifted 15 deg clockwise, the other phase-shifted 15 deg counterclockwise. For simulation purposes, a source with two symmetric components was developed which yielded the basis for comparison of all the other configurations.

A. Sequence Components

The phase-shifting transformer of a twelve-pulse converter can be replaced with a six-phase source. The two symmetric sequence components (V_2) and (V_5), based on Eq. (10), were placed in the SPICE model and the ripple was that of a twelve-pulse converter. To verify the model, an ideal converter and input power source were simulated. The ratio of the highest uncharacteristic harmonic to the principal characteristic harmonic at $12f$, where f is the line frequency, is greater than 80 dB. This represents a residual of less than 0.01 percent, more than adequate for our purposes.

Any distortion in either V_s or V'_s yields a harmonic at six times the input line frequency. The ratio of the amplitude of the two sequence components is 3.73 , and the angle between the two is 90 deg. When the amplitude of the larger component was decreased by 2.5 percent, the principal characteristic harmonic component of a six-pulse converter appears. The value referenced to the twelfth harmonic increased from practically zero (-95 dB) to -13 dB. Similar results are obtained when the phase is not 90 deg. The same results are obtained when the values of turns ratio or unbalance leakage inductance are simulated in a phase-shifting transformer used to supply a twelve-pulse converter. Other distortion of the three-phase power yields the same uncharacteristic harmonics with a twelve-pulse converter as with a six-pulse converter.

B. Commutation Overlap

For a six-pulse converter, normal operation is taken for μ less than 60 deg, and this is called Mode 1 operation. During Mode 1, the output voltage decreases linearly from 1 to 0.750 . A twelve-pulse converter consists of two out-of-phase six-pulse converters. Mode 1 operation exists for μ less than 30 deg only. At $\mu = 30$ deg, the two converters have the same value of regulation. For μ greater than 30 deg, the regulation is worse for the twelve-pulse converter. At $\mu = 60$ deg, the regulation for the twelve-pulse converter is 0.683 , a difference of 7 percent. Since the commutation overlap angle is usually made less than 60 deg, the difference in regulation between the two converters does not prevent the use of twelve-pulse converters.

IV. Conclusion

The use of SPICE simulation has made it possible to obtain detailed performance predictions for power converters. For the first time, an analysis of voltage source converters with continuous line current was obtained. The effects of line distortion such as generator harmonics and subharmonics on the output voltage ripple were deter-

mined. The effects of line voltage unbalances, both phase and amplitude, on converter operation were greatly simplified by the use of sequence components. The results indicate the ease of design of voltage source converters for medium- and high-power applications.

The use of twelve-pulse converters for the transmitters in the DSN has a number of advantages over six-pulse units. The ripple on the ac lines and on the dc side is re-

duced and the frequency is increased, resulting in reduced requirements for filtering. The uncharacteristic harmonics are equal for both types of converters. The only disadvantage is the poorer regulation for commutation overlap angles greater than 30 deg. At 60 deg, a penalty of 7 percent is found for the twelve-pulse converter over that of the six-pulse. If the design is limited to commutation of about 30 deg, both types of converters are equal in terms of regulation.

References

- [1] G. Seguier, *Power Electronic Converters*, New York: McGraw-Hill Book Company, 1986.
- [2] J. D. Glover and M. Sarma, *Power System Analysis and Design*, Boston: PWS Publishers, Chapter 3, 1987.
- [3] R. L. Witzke, J. V. Kresser, and J. K. Dillar, *Voltage Regulation of 12-Phase Double-Way Rectifiers*, AIEE Transactions, vol. 72, pt. I, pp. 244-253, July 1953.
- [4] J. Schaefer, *Rectifier Circuits: Theory and Design*, Chapter 15, New York: John Wiley & Sons, Inc., 1965.
- [5] E. J. Finnegan, "High-Power Transmitter High-Voltage Power Supply Ripple," *DSN Progress Report 42-31*, Jet Propulsion Laboratory, Pasadena, California, pp. 84-87, November and December 1975.
- [6] P. Wood, *Switching Power Converters*, New York: Van Nostrand Rheinhold Co., 1981.
- [7] B. P. Lathi, *Signals, Systems and Communication*, Chapter 11, New York: John Wiley & Sons, Inc., 1965.
- [8] J. Arrillaga et al., *Power System Harmonics*, New York: John Wiley & Sons, pp. 79-84, 1985.
- [9] Kimback, *Direct Current Transmission*, New York: John Wiley & Sons, pp. 308-309, 1971.

Table 1. Short circuit test (Fourier components: characteristic harmonics)

Harmonic No.	Frequency, Hz	Current, amp	Normalized, dBI _{sc}	Theoretical, dBr
6	2400	10.83	-24.80	-24.86
12	4800	2.551	-37.35	-37.09
18	7200	1.099	-44.70	-44.16
24	9600	0.6913	-48.70	-49.17
30	12000	0.4294	-52.83	-53.05
36	14400	0.2518	-57.47	-56.22
42	16800	0.2176	-58.74	-58.90
48	19200	0.1861	-60.09	-61.23
54	21600	0.1148	-64.29	-63.27

Note: A short circuit dc load of 1 ohm with line inductance of 0.6 mH is the basis for Table 1 data. The short circuit dc current is 188.1 A. The results are independent of line inductances. The values are identical to those obtained with a current source converter.

Table 2. Simulation capability test (uncharacteristic harmonics)

Harmonic No.	Frequency, Hz	Current, amp	Normalized, dBr
1	400	1.092E-3	-104.72
2	800	6.998E-5	-128.59
3	1200	2.355E-6	-158.05
4	1600	1.441E-5	-142.31
5	2000	7.835E-5	-127.61
7	2800	6.418E-5	-129.34
8	3200	1.501E-5	-141.96
9	3600	1.783E-6	-160.46

Note: A dc load of 1MW and line inductances of 0.6 mH are the basis for Table 2 data.

Table 3. Light load test (Fourier components: characteristic harmonics)

Harmonic No.	Frequency, Hz	Current, amp	Normalized, dBI _{sc}
6	2400	1.884	-39.99
12	4800	0.3057	-55.78
18	7200	0.1180	-64.05
24	9600	0.0616	-69.69
30	12000	0.0374	-74.03
36	14400	0.0248	-77.60
42	16800	0.0174	-80.67
48	19200	0.0128	-83.36
54	21600	0.0097	-85.74

Note: The converter is operating with a light load of 1.964 Adc. The power output is nominally 100 kW. Line inductance 0.6 mH.

Table 4. Full power test (Fourier components: characteristic harmonics)

Harmonic No.	Frequency, Hz	Current, amp	Normalized, dBI _{sc}
24	9600	0.1569	-61.58
30	12000	0.1062	-64.97
36	14400	0.0692	-68.68
42	16800	0.0540	-70.84
48	19200	0.0389	-73.69
54	21600	0.0326	-75.22

Note: The converter is operating with a rated load of 22.75 A dc. The power output is nominally 1 MW. Line inductance 0.6 mH.

Table 5. Voltage converter (negative-sequence component)

Harmonic No.	Frequency, Hz	Current, amp	Normalized, dBI _{sc}	AC Inductance, mH
2	800	4.471	-47.01	0.1
4	1600	1.358	-57.36	0.1
8	3200	0.08348	-67.06	0.1
2	800	4.639	-37.91	0.3
4	1600	0.8611	-52.54	0.3
8	3200	0.4399	-58.37	0.3
2	800	1.831	-40.23	0.6
4	1600	0.3048	-55.81	0.6
8	3200	0.1934	-59.76	0.6
2	800	1.025	-40.94	1
4	1600	0.159	-57.13	1
8	3200	0.1004	-61.12	1

Note: A dc load of 1 MW nominal, which is the rating of the supplies in the DSN, is the basis for Table 5 data. The negative-sequence component had an amplitude of 1 percent or -40 dB with respect to the fundamental component of line voltage.

Table 6. Voltage converter (harmonic distortion)

Input Harmonic No.	Output Frequency, Hz	Output DC Current, amp	Output Normalized, dBr
2	400	1.695	-40.9
-2	1200	0.9058	-46.35
3	800	1.404	-42.54
-3	800	0.4747	-51.96
-3	1600	0.6576	-49.13
-3	3200	0.05475	-70.72
4	1200	0.7734	-47.72
-4	400	0.4281	-52.86
-4	2000	0.4355	-52.71

Note: A dc load of 1 MW nominal, which is the rating of the supplies in the DSN, is the basis for Table 6 data.

Table 7. Subharmonic distortion

Input Frequency, Hz	Input Level, dB	Sequence Component	Output Frequency, Hz	Output Level, dBI _{dc}
Single-Side-Band Source				
300	1%	pos	100	-57.85
300	1%	neg	700	-38.37
200	1%	pos	200	-48.47
200	1%	neg	600	-36.44
100	1%	pos	300	-43.91
100	1%	neg	500	-36.36
40	1%	pos	360	-41.49
40	1%	neg	440	-38.22
Double-Side-Band Source				
440/360	1%	pos	40	-57.33

Note: A dc load of 1 MW nominal, which is the rating of the supplies in the DSN, is the basis for Table 7 data. The input level is referenced to the main 400-Hz component of the phase voltage.

Table 8. AC line harmonics

Line Ind., mHy	Voltage $f = 2000$, dB	Current $f = 2000$, dB	Voltage $f = 2800$, dB	Current $f = 2800$, dB	Overlap Angle, deg	Line Regulation, %
0.1	-25.55	-6.99	-29.55	-14.08	7.5	1.10
0.3	-20.51	-11.26	-28.83	-22.49	18.9	2.57
0.6	-17.26	-13.87	-23.26	-22.79	28.8	5.13
1	-15.36	-16.28	-19.53	-23.36	38.1	8.67
3	-14.03	-24.15	-16.92	-29.97	60	23.6

Note: The values are given in dB referenced to the level of the fundamental 400-Hz component of the phase voltage or the line current as appropriate. A dc load of 1 MW, which is the rating of the supplies in the DSN, is the basis for Table 8 data.

Table 9. Input listing (AC line harmonics, Ideal circuit)

```

TRANSFORMER ANALYSIS—2.25 MW, 400Hz ASSEMBLY TAP 8 (ONE PI)
.FOUR 4000 I(V4) V(52)
TRAN 1US 22.5MS 20MS
.WIDTH IN=133 OUT=133
.OPTIONS METHOD=GEAR LIMPTS=6000 NOMOD RELTOL=.03 ITL5=0 ITL4=50 LVTIM=1
V1 60 1 SIN(0 3000 400 0 0)
V2 61 1 SIN(0 3000 400 0.83333333MS 0)
V3 62 1 SIN(0 3000 400 1.66666667MS 0)
V4 63 60 SIN(0 0 0 0 0)
V5 64 61 SIN(0 0 0 0 0)
V6 65 62 SIN(0 0 0 0 0)
V10 51 63 SIN(0 0 0 0 0)
V11 54 64 SIN(0 0 0 0 0)
V12 57 65 SIN(0 0 0 0 0)
V7 1 0
L1 51 53 3M
L3 57 59 3M
L5 54 56 3M
R1 53 52 50M
R3 59 58 50M
R5 56 55 50M
RP10 52 71 21.96M
RP11 58 81 21.96M
RP12 55 91 21.96M
K123 LP11 LS13 .9999
K175 LP12 LS15 .9999
K148 LP13 LS18 .9999
K112 LSE11 LP11 .9999
K194 LSE19 LP13 .9999
K167 LSE16 LP12 .9999
K131 LS13 LSE11 .9999
K189 LS18 LSE19 .9999
K165 LSE16 LS15 .9999
LSE11 103 116 1.8H
LSE16 106 118 1.8H
LSE19 109 117 1.8H
LP11 91 58 .2H
LP12 81 52 .2H
LP13 71 55 .2H
LS13 116 115 5.408H
LS15 118 113 5.408H
LS18 117 114 5.408H
RS12 502 0 2K
C1 502 700 1UF
RC1 700 0 1
V8 501 502
RS13 116 113 630M
RS14 118 114 630M
RS15 117 115 630M
VD11 603 103
VD12 103 604
VD13 605 109
VD14 109 606
VD15 607 106
VD16 106 608
D11 501 603 DIODE
D12 604 0 DIODE
D13 501 605 DIODE
D14 606 0 DIODE
D15 501 607 DIODE
D16 608 0 DIODE
.MODEL DIODE D
.PRINT TRAN I(VD11) I(VD12) I(VD13) I(VD14) I(VD15) I(VD16) I(V4) V(52)
.END

```

Note: The listings were changed to obtain each of the conditions required for the different tables.

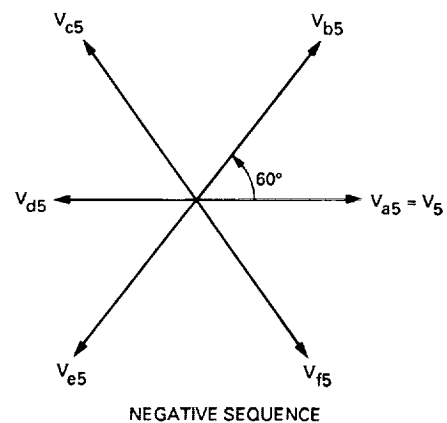
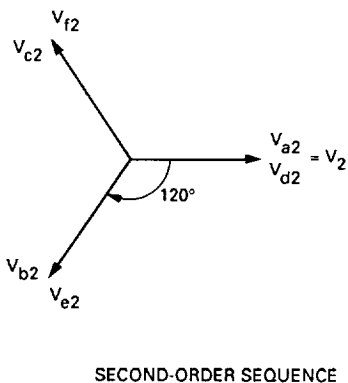
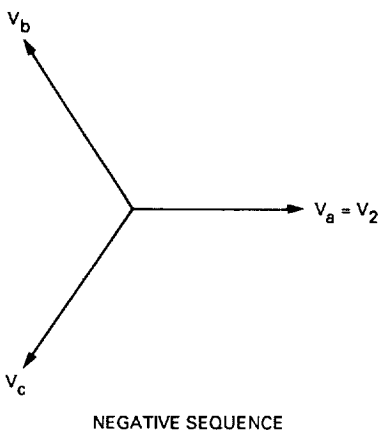
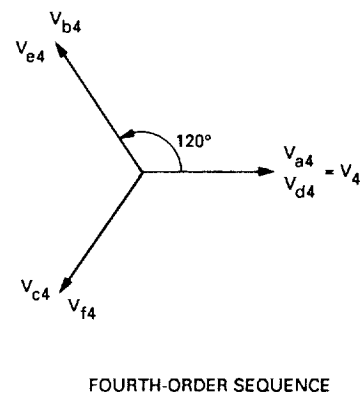
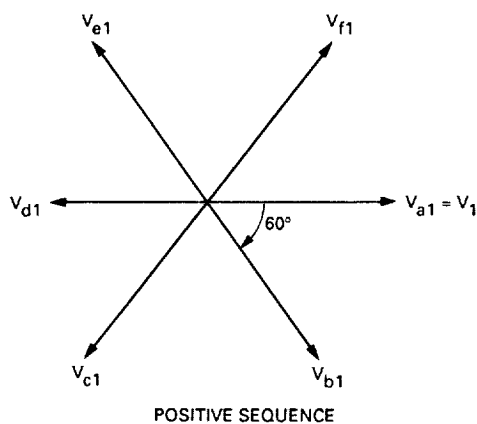
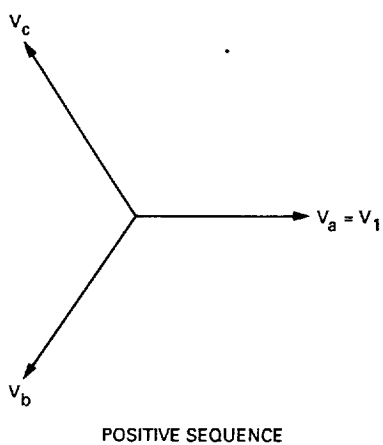
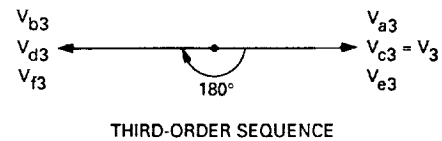
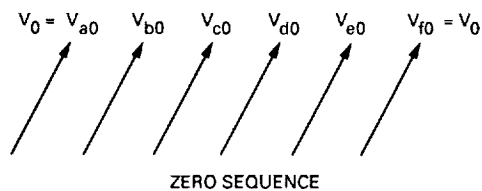
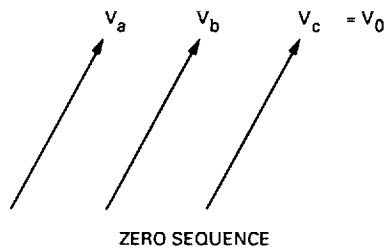


Fig. 1. Three-phase symmetric sequence components.

Fig. 2. Six-phase symmetric sequence components.

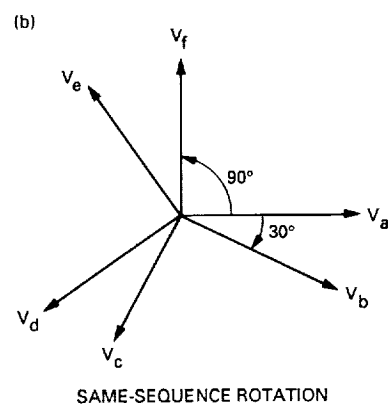
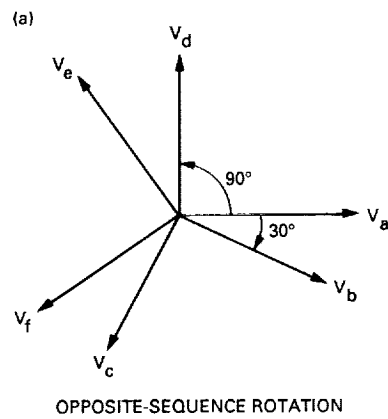


Fig. 3. Six-phase voltages required for twelve-pulse converter.

Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations

V. Galindo-Israel and W. Imbriale
Ground Antenna and Facilities Engineering Section

K. Shogen and R. Mittra
Electromagnetic Communications Laboratory
University of Illinois, Urbana, IL

In obtaining solutions to the first-order nonlinear partial differential equations (PDEs) for synthesizing offset dual-shaped reflectors, it is found that previously observed computational problems can be avoided if the integration of the PDEs is started from an inner projected perimeter and integrated outward rather than starting from an outer projected perimeter and integrating inward. This procedure, however, introduces a new parameter, the main reflector inner perimeter radius ρ_0 , when given a subreflector inner angle θ_0 . Furthermore, a desired outer projected perimeter (e.g., a circle) is no longer guaranteed. "Stability" of the integration is maintained if some of the initial parameters are determined first from an approximate solution to the PDEs. A one-, two-, or three-parameter optimization algorithm can then be used to obtain a "best" set of parameters yielding a close fit to the desired projected outer rim. Good low cross-polarization mapping functions are also obtained. These methods are illustrated by synthesis of a high-gain offset-shaped Cassegrainian antenna and a low-noise offset-shaped Gregorian antenna.

I. Introduction

The problem of synthesizing dual-shaped offset reflector antennas has received considerable attention in recent years, not only because it is theoretically challenging but also because it is of considerable practical interest. Galindo-Israel et al. [1] presented an exact algorithm for simultaneously synthesizing the shapes of the sub and main reflectors for an arbitrary feed pattern, aperture ampli-

tude, and phase distributions by solving a set of nonlinear first-order partial differential equations (PDEs). Jervase et al. [2] discussed an extension of the numerical method presented in [1] and gave some computed results for the reflector shapes. It has been observed that numerical difficulties may arise when applying these techniques, especially in the vicinity of the center region of the reflectors. This occurs when the PDEs' integration is initialized on the outer rims of the reflectors and proceeds "inward" [1].

In this article, we investigate the possibility of circumventing these difficulties by starting the synthesis procedure not from the periphery, but from inner or central rims of the reflectors with the integration then proceeding "outward" to the outer rims. The success of this procedure depends significantly on the initial conditions chosen at the inner rims. An approximate solution to the synthesis problem provides one good choice for the initial conditions. Numerical results for symmetric feed patterns and aperture distributions are included to illustrate the application of the synthesis procedure based on a "center" start. A uniform aperture distribution for a high-gain dual-offset shaped Cassegrainian-type reflector antenna is illustrated first. Next, an aperture distribution that is uniform over most of the aperture, but is given a Gaussian taper near the outer rim, is shown in a Gregorian-type configuration. This distribution is useful for a low-noise (but high-gain) large ground antenna where very low spillover is desired past the main reflector into a "hot" ground background.

Starting the integration in the central region of the reflectors and integrating outward does *not* guarantee a circular projected outer rim as integrating inward does (if that is desirable). It is possible to *optimize* one or more initial parameters to obtain a desired circular outer rim (or other desired results). This approach was investigated for obtaining a circular projected outer rim with very good results.

II. Numerical Method

We choose the specified feed power pattern $I(\theta)$ and the desired aperture power distribution $V(\rho)$ to be circularly symmetric. For the antenna system shown in Fig. 1, the differential equations for synthesizing the dual reflector system are given by [1]:

$$r_\theta = r_\theta(\theta, \phi | \rho, \psi, r) \quad (1)$$

$$r_\phi = r_\phi(\theta, \phi | \rho, \psi, r) \quad (2)$$

$$\rho_\theta = \rho_\theta(\theta, \phi | \rho, \psi, r; \rho_\phi, \psi_\phi) \quad (3)$$

$$\psi_\theta = \psi_\theta(\theta, \phi | \rho, \psi, r; \rho_\phi, \psi_\phi) \quad (4)$$

Equations (1) and (2) are obtained by applying Snell's Law to the subreflector while z is calculated by using the path length condition viz., $r + S - z = \text{constant}$. Equations (3) and (4) are obtained from the following three conditions:

- (A) Conservation of energy on a pointwise basis for geometrical optics:

$$V_c V(\rho) \rho [\rho_\theta \psi_\phi - \rho_\phi \psi_\theta] = I(\theta) \sin \theta \quad (5)$$

$$V_c = \int_0^{\theta_M} I(\theta) \sin \theta d\theta / \int_0^{\rho_M} V(\rho) \rho d\rho \quad (6)$$

- (B) Total differentiability at all points $(\theta, \phi, \rho, \psi)$

$$r_{\theta\phi} = r_{\phi\theta} \quad (7)$$

- (C) Snell's Law applied to the main reflector.

Equations (1) through (4) are solved by starting with some initial values for ρ, θ (very small), and ψ , and then increasing θ in small increments to obtain the main and the subreflector surfaces in a progressive manner. The procedure is identical to that in [1] except that the integration is outward (increasing θ and ρ).

III. Numerical Results for the Cassegrainian-Type System

Results are presented for the following system parameters: $a = 11$, $\Omega = 21^\circ$, $\theta_M = 16^\circ$, $\rho_M = 5$, $r_i = 7$ (see Fig. 1). The optical path length is assumed to be 22, the specified or object aperture power distribution is uniform, and the feed pattern is taken to be $\cos^n \theta$ with $n = 151$ (yielding a subreflector edge illumination level at -25.9 dB below the center illumination).

As mentioned above, the reflector shapes depend strongly on the initial values chosen for the solution of the differential Eqs. (1) through (4). For a given θ_0 (minimum θ rim value for the subreflector), a critical initial parameter is the value chosen for ρ_0 (minimum ρ rim value for the main reflector). As an example, Fig. 2 shows two results for the same system parameters but with different initial values of ρ_0 . It is observed from this figure that ρ_θ becomes very large near the center point ($\theta \approx 0$) for one of the ρ_0 choices (case B). This leads to unacceptable reflector shapes.

One approach to finding good initial values is to start with an approximate solution (see [3], for example). An approximate solution can be found by relaxing the total differentiability condition (i.e., Eq. 7), and then prescribing the mapping [1,3] in advance as $\psi = \phi$, $\psi_\phi = 1$, $\rho_\phi = 0$, and $\psi_\theta = 0$. Under these conditions the mapping from (θ, ϕ) to (ρ, ψ) is forced to be concentric. Thus the conservation of energy reduces to

$$V_c V(\rho) \rho \rho_\theta = I(\theta) \sin \theta \quad (8)$$

The above *approximate* solution [3] allows a value of ρ_0 to be found for a given θ_0 which then yields very good and stable numerical solutions by the exact method discussed earlier. While this value of ρ_0 is not *critical* for obtaining a “good” solution, it is critical if a perfectly circular (or other shape) is required for the *outer* rim of the main reflector. Alternate methods for choosing ρ_0 will be discussed in the next section.

Figure 3 presents the results for ψ_θ versus θ for two cases in order to illustrate the advantage of starting from a central rim instead of the outer edge. Note that the outward solution (solid line) exhibits a much better behavior than the corresponding inward solution (dashed line). Figure 4 shows the main and subreflector shapes derived by using the outward approach. Note that not only are the surfaces well-behaved, but also that the maximum value of ρ of the main reflector in this case does not deviate significantly from ρ_M , the value for maximum ρ , which was desired to be 5. Thus, the outer periphery condition, that the projected aperture be circular, is approximated very well in this solution.

Figure 5 shows a comparison between the radiation patterns in the offset plane of the shaped reflector (solid line) and a comparable Cassegrainian antenna. The diameter of the main reflectors of the two antennas is 80λ . The system parameters of the shaped reflector are $A = 2200$, $\Omega = 23^\circ$, $\theta_M = 8.5^\circ$, $\rho_M = 1000$, $r_i = 1000$ (see Fig. 1). The optical units of length can be converted to wavelengths by dividing by 25. The optical path length is assumed to be 4000. The specified main reflector aperture amplitude distribution was uniform and the feed pattern was taken to be $\cos^n \theta$, with $n = 151$. The radiation patterns are calculated by using geometrical optics for the subreflector and physical optics for the main reflector scattering. Some radiation characteristics of the two antennas are given in Table 1.

IV. Obtaining a Gregorian Design by Optimization of Parameters

In this section, we will discuss a special offset dual-shaped Gregorian design wherein the parameter ρ_0 and other initial parameters as well are set by an optimization procedure.

The synthesis software used herein can go through a complete integration in less than a half minute on a 386 20-MHz processor when approximately eight Fourier terms are used to represent *differences* in the functions

$$\rho(\phi) \text{ and } \psi(\phi) \quad (9)$$

between adjacent θ values. Since an FFT is used for representation of these functions, the increase in time with increasing number of Fourier terms is linear. Eight Fourier terms was sufficient for the case to be presented.

This very rapid two-dimensional integration of the nonlinear PDEs reasonably allows for the possibility of performing the integration *many times* in an optimization algorithm to “optimize” one or more initial parameters.

One parameter that can be optimized is the choice of ρ_0 for a given θ_0 value. Ideally, a value of ρ_0 can be chosen at the start of the optimization by solving for the approximate solution to the PDEs as discussed earlier.

The question immediately arises as to what object function the optimization algorithm should attempt to minimize. One useful object is a perfectly circular projected perimeter of the main reflector at $\theta = \theta_{MAX}$. Thus, the function

$$f = SQRT \left[\sum_{n=1}^N (\rho_{M_i} - \rho_M)^2 / N \right] \quad (10)$$

can be minimized, where

$$(\rho_{M_i} : i = 1, \dots, N) \quad (10a)$$

are the perimeter values obtained at the set

$$(\phi_i : i = 1, \dots, N) \text{ or } (\psi_i : i = 1, \dots, N) \quad (10b)$$

for any given integration of initial parameter values. The value ρ_M need not be a constant unless a circular perimeter is desired. The object function could be expanded to obtain a prescribed set of mapping functions

$$\rho(\theta, \phi), \psi(\theta, \phi) \quad (11)$$

as closely as possible in order to minimize cross-polarized currents in the aperture, for example.

In addition to the parameter ρ_0 , we have chosen two other initial parameters for optimization in a three-dimensional space. A second parameter permits *one* degree of freedom in the choice of the initial function ψ_{ϕ_0} (see [1]). For the Cassegrainian design discussed earlier, the value of ψ_{ϕ_0} was selected as unity. The parameter A_ψ is introduced such that

$$\psi_{\phi_0} \equiv 1 + A_\psi \left(\frac{\pi}{180} \right) \sin \phi \quad (12)$$

where A_ψ is given in degrees. The allowable field of A_ψ is bounded to prevent caustics from occurring where not desired. A higher order Fourier series could be used with more parameters available for optimization.

A third parameter used herein was the value of Ω (see Fig. 1 or Fig. 6).

Figure 6 presents the profiles of a three-dimensionally optimized dual-shaped Gregorian reflector pair in the off-set plane. A uniform phase in the main reflector aperture was required as indicated by the rays in Fig. 6. The object function was that in Eq. (10). The value θ_0 was chosen as 0.1 deg, while the optimum parameters were found as

$$\begin{aligned}\rho_0 &= 0.015636 \\ A_\psi &= 23.53905^\circ\end{aligned}\quad (13)$$

and

$$\Omega = -15.0858^\circ$$

The starting values for the above optimization were chosen from an approximate solution and "manual" optimization through several iterations. Table 2 presents an extract of the mapping functions obtained (see Eq. 11).

It is not clear whether this set of "optimum" parameters is global or local. Achieving a rim that is more circular and/or closer to the desired value of $\rho_M = 3$ with better than three significant figures may require greater control of the ψ_{ϕ_0} function (Eq. 12) than one parameter, A_ψ , can provide. Nevertheless, the mapping is approximately circular throughout, i.e.,

$$\rho(\theta, \phi) \approx \rho(\theta) \quad (14)$$

indicating that a low cross-polarization exists in the aperture. Another measure of this low cross-polarization mapping is the *nearly* point caustic illustrated between the Gregorian reflectors in Fig. 6.

Figure 7 illustrates the desired and the achieved aperture power distributions. The desired uniform phase was achieved exactly (see [1]). The small "hole" in the main reflector near $\rho = 0$ is determined by the "optimum" choice of ρ_0 . The difference between the geometrical optics (GO) computed power distribution and the desired power near $\rho \approx \rho_M = 3.0$ is at least in part due to the approximate GO analyses used. (The GO analyses utilized a *polar* grid wherein very wide ray "tubes" occur for larger ρ when

$\Delta\phi$ is fixed. As $\Delta\phi$ was decreased, it was found that the discrepancy grew smaller for the same set of reflectors.) Similarly, and for the same reason, the small discrepancy near $\rho = 0$ decreased as $\Delta\theta$ was decreased.

A two-dimensional optimization with Ω fixed at $\Omega = -15.0858$ deg, the value found with the three-dimensional optimization, which yields essentially the same reflector pair as illustrated in Fig. 6. In this case, however, the other values chosen for the start of the optimization were $\rho_0 = 0.0001$ and $A_\psi = 0.0$ deg.

Thus, we did not take advantage of the values of ρ_0 and A_ψ known from an approximate solution. The final values of the parameters are shown in Table 3.

It should be noted that in order to achieve the optimized values shown in Table 2 from very poor starting values, a careful construction of constraints had to be built into the optimization algorithm. However, the constraints were general and not limited to any particular geometry.

Table 4 shows the resultant mapping obtained from the two-dimensional algorithm. It is seen that the reflectors obtained in this way differ by very little from those obtained and illustrated in Table 2.

One-dimensional optimizations for an optimum value of ρ_0 , with both A_ψ and Ω fixed, are generally less successful unless a reasonably "good" value of A_ψ is chosen. Several cases wherein the starting value of ρ_0 is chosen as:

$$\rho_0(\text{start}) = 0.0001$$

are illustrated in Table 5. Values of

$$A_\psi = 0^\circ$$

$$A_\psi = 10^\circ$$

$$A_\psi = 20^\circ$$

were selected. When either $A_\psi = 0$ deg or $A_\psi = 10$ deg is used, a correct solution is *not* obtained for the PDEs. This is seen by the values of ρ_M obtained at $\theta = 16$ deg. However, when $A_\psi = 20$ deg is chosen, a good solution is obtained. This points up the utility of using the results for ρ_0 and A_ψ from an approximate solution as starting values in any optimization algorithm.

V. Numerical Results for the Gregorian-Type System

The synthesized Gregorian-type low-noise/high-gain dual-shaped offset reflector system was analyzed by GO

during the actual synthesis and with an independent GO computer program. The subreflector was further analyzed by the geometric theory of diffraction (GTD) and by physical optics (PO). The main reflector was analyzed by PO in all cases. Thus, the dual-reflector analysis was performed by GO/PO, GTD/PO, and PO/PO (indicating the sub-analysis/main analysis as shown). In addition, the object aperture distribution, Fig. 7, was integrated directly by PO for comparison.

The feed pattern used in the analysis was the same as that used in the synthesis. Principal polarization was chosen as right circular. A projection of the current distribution on the subreflector onto a plane normal to the axis of the feed is shown in Fig. 8. Note the taper of the feed power pattern to -20 dB at the subreflector edge at $\theta = 16$ deg. This sharp taper leads to essentially no spillover loss past the subreflector.

The *equivalent* current aperture distribution on the main reflector is shown in Fig. 9 when the subreflector is analyzed by GO. (The *equivalent* current is the actual current multiplied by the Jacobian determinant transformation of the actual main reflector surface to the aperture plane.) Note that the current distribution is uniform to the point where it then descends Gaussian to approximately -20 dB at the edge of the main reflector. This distribution is essentially the same as the object function in Fig. 7.

The main reflector ($D = 120\lambda$) equivalent current distribution obtained by a GTD subreflector analysis is shown in Fig. 10 (a PO subreflector analysis yields little observable difference). Note the shallow diffraction ripples over most of the aperture (where the current was uniform in Fig. 9). The current distribution still has a Gaussian taper as prescribed in Fig. 7. This will result in very little spillover loss past the main reflector edge and thus be ideal for a low-noise reflector system.

The far-field patterns (diameter of main reflector $= 120\lambda$) for each of these analysis cases is shown in Fig. 11. Note the very low loss in gain between the aperture integration of the object function (Fig. 7) and the various diffraction analyses. This illustrates that the chosen object pattern has achieved the goal of potentially very low noise due to spillover past the main reflector into a warm ground environment (typical of upward-looking ground-based antennas). This type of aperture distribution appears achievable only by dual-shaping methods.

VI. Conclusions

When integrating the nonlinear PDEs for offset dual-shaped reflector synthesis from a prescribed outer projected perimeter *inwards*, it was found earlier that computational problems do arise in the central region of the reflectors [1]. It is found herein that these computational problems are largely bypassed if the integration of the PDEs is started from a prescribed projected *inner* rim and integrated *outward*. In this case, however, a circular, or otherwise prescribed, *outer* rim is not guaranteed. In addition, a new parameter of the PDEs is introduced, the radius, ρ_0 , of the inner projected rim of the main reflector when the inner value of $\theta = \theta_0$ for the subreflector is prescribed. This parameter exists in addition to other initial parameters (e.g., Ω ; see Figs. 1 and 6) and the initial function $\psi_\phi(\phi)$ (see [1]).

A poorly selected inner radius parameter ρ_0 , or other poorly selected initial parameters will preclude the obtaining of useful solutions to the PDEs by the method discussed herein or in [1]. Two complementary methods for choosing the parameter ρ_0 and other parameters are discussed. The first method utilizes the initial values found from an approximate solution to the PDEs. A second method utilizes an optimization algorithm which searches for the parameters of the PDEs that minimize a prescribed object function. Such an object function may measure the deviation of the projected outer perimeter from circular (or some other desired shape). Alternatively, a minimum cross-polarization distribution in the aperture (optimum mapping function) may be chosen as the object function.

An offset dual-shaped Cassegrainian-type high-gain antenna is synthesized to illustrate the first method described above. The second method of optimization is illustrated by synthesizing a Gregorian low-noise/high-gain offset dual-shaped reflector antenna. One-parameter [ρ_0], two-parameter [ρ_0 and a parameter A_ψ describing $\psi_{\phi_0}(\phi)$], and three-parameter [ρ_0 , A_ψ , and Ω] optimizations are illustrated. Analyses of the synthesized reflectors by GO, GTD, and PO are shown to verify the utility of the results.

It should be noted that the method of using an approximate solution to the PDEs to find the initial parameters is complementary to utilization of the optimization algorithm since it provides an excellent set of starting values for the parameters in the optimization procedure.

Acknowledgments

The authors wish to acknowledge the significant assistance in this work of Prof. S. Rengarajan of California State University at Northridge and R. Hodges, JPL part-time employee and Ph.D. candidate at the University of California at Los Angeles.

References

- [1] V. Galindo-Israel, W. Imbriale, and R. Mittra, "On the Theory of Synthesis of Single and Dual Offset Shaped Reflector Antennas," *IEEE Trans. Antennas Propagat.*, vol. AP-35, no. 8, pp. 887-896, August 1987.
- [2] J. A. Jervase, R. Mittra, V. Galindo-Israel, and W. Imbriale, "Numerical Study of the Problem of Synthesis of Offset Dual Shaped Reflector Antennas," *URSI/IEEE AP-S Symposium*, pp. 1-2, Syracuse, New York, June 1988.
- [3] K. Shogen, V. Galindo-Israel, R. Mittra, and W. Imbriale, "A Numerical Approach for Synthesizing Dual-Shaped Offset Reflector Antennas," presented at *URSI/IEEE AP-S Symposium*, San Jose, California, June 1989.

Table 1. Comparison of the radiation characteristics of the shaped reflector antenna and the Cassegrainian antenna (see Fig. 5)

Main Reflector	Shaped Reflector	Cassegrainian Antenna
Edge illumination level	-4 dB	-16 dB
-3 dB beamwidth	0.77 deg	0.91 deg
1st sidelobe level	-10 dB	-32 dB

Table 2. Mapping function for three-dimensional optimization, values in Eq. (13)

Values of ρ at:			θ°
$\phi = -90^\circ$ $\psi = +90^\circ$	$\phi = 0^\circ$ $\psi \approx 180^\circ$	$\phi = +90^\circ$ $\psi = 180^\circ$	
0.156359E-01	0.156359E-01	0.156359E-01	0.100000E+00
0.345170E+00	0.353902E+00	0.364169E+00	0.100000E+01
0.691174E+00	0.699794E+00	0.709482E+00	0.200000E+01
0.101939E+01	0.102740E+01	0.103634E+01	0.300000E+01
0.132247E+01	0.132951E+01	0.133745E+01	0.400000E+01
0.159466E+01	0.160053E+01	0.160730E+01	0.500000E+01
0.183220E+01	0.183685E+01	0.184240E+01	0.600000E+01
0.203348E+01	0.203695E+01	0.204135E+01	0.700000E+01
0.219891E+01	0.220133E+01	0.220471E+01	0.800000E+01
0.233064E+01	0.233217E+01	0.233470E+01	0.900000E+01
0.243227E+01	0.243313E+01	0.243500E+01	0.100000E+02
0.251358E+01	0.251401E+01	0.251552E+01	0.110000E+02
0.258565E+01	0.258576E+01	0.258707E+01	0.120000E+02
0.265436E+01	0.265418E+01	0.265537E+01	0.130000E+02
0.272516E+01	0.272466E+01	0.272582E+01	0.140000E+02
0.280943E+01	0.280847E+01	0.280969E+01	0.150000E+02
$\rho_M = 0.301701E+01$	0.301343E+01	0.301529E+01	0.160000E+02

Table 3. Initial value parameters obtained by two-dimensional optimization: starting point: $\rho_0 = 0.0001$, $A_\psi = 0.0^\circ$

Parameters from Two-Dimensional Optimization	
Values Eq. (13)	Values Two-Dimensional Optimization
$\rho_0 = 0.015636$	$\rho_0 = 0.015734$
$A_\psi = 23.53905^\circ$	$A_\psi = 23.30467^\circ$

Table 4. Mapping function for two-dimensional optimization, values in Table 3

Values of ρ at:			θ°
$\phi = -90^\circ$ $\psi = +90^\circ$	$\phi = 0^\circ$ $\psi \approx 0^\circ$	$\phi = +90^\circ$ $\psi = 270^\circ$	
0.157339E-01	0.157339E-01	0.157339E-01	0.100000E+00
0.345189E+00	0.353831E+00	0.364174E+00	0.100000E+01
0.691235E+00	0.699701E+00	0.709547E+00	0.200000E+01
0.101948E+01	0.102728E+01	0.103644E+01	0.300000E+01
0.132259E+01	0.132937E+01	0.133758E+01	0.400000E+01
0.159479E+01	0.160038E+01	0.160745E+01	0.500000E+01
0.183235E+01	0.183667E+01	0.184258E+01	0.600000E+01
0.203364E+01	0.203676E+01	0.204154E+01	0.700000E+01
0.219908E+01	0.220113E+01	0.220491E+01	0.800000E+01
0.233082E+01	0.233197E+01	0.233491E+01	0.900000E+01
0.243246E+01	0.243292E+01	0.243522E+01	0.100000E+02
0.251379E+01	0.251378E+01	0.251576E+01	0.110000E+02
0.258588E+01	0.258550E+01	0.258733E+01	0.120000E+02
0.265463E+01	0.265387E+01	0.265568E+01	0.130000E+02
0.272548E+01	0.272429E+01	0.272618E+01	0.140000E+02
0.280984E+01	0.280797E+01	0.281016E+01	0.150000E+02
$\rho_M = 0.301779E+01$	0.301343E+01	0.301627E+01	0.160000E+02

Table 5. Parameters and perimeter values from one-dimensional optimization of ρ_0 with A_ψ fixed. Starting point for $\rho_0 = 0.0001$.

$\phi = -90^\circ$ $\psi = +90^\circ$	$\phi = 0^\circ$ $\psi \approx 180^\circ$	$\phi = +90^\circ$ $\psi = 270^\circ$	θ°
$A_\psi = 0.0^\circ$ Fixed			
$\rho_0 = 0.019594$ Final Value			
$\rho_0 = 0.195940E-01$	0.195940E-01	0.195940E-01	0.100000E+00
$\rho_M = 0.542957E+02$	0.499525E+02	0.459551E+02	0.160000E+02
$A_\psi = 10.0$ Fixed			
$\rho_0 = 0.005099$ Final Value			
$\rho_0 = 0.509900E-02$	0.509900E-02	0.509900E-02	0.100000E+00
$\rho_M = 0.815349E+00$	0.697426E+00	0.299195E+00	0.160000E+02
$A_\psi = 20.0$ Fixed			
$\rho_0 = 0.01747$ Final Value			
$\rho_0 = 0.174655E-01$	0.174655E-01	0.174655E-01	0.100000E+00
$\rho_M = 0.303909E+01$	0.300320E+01	0.302680E+01	0.160000E+02

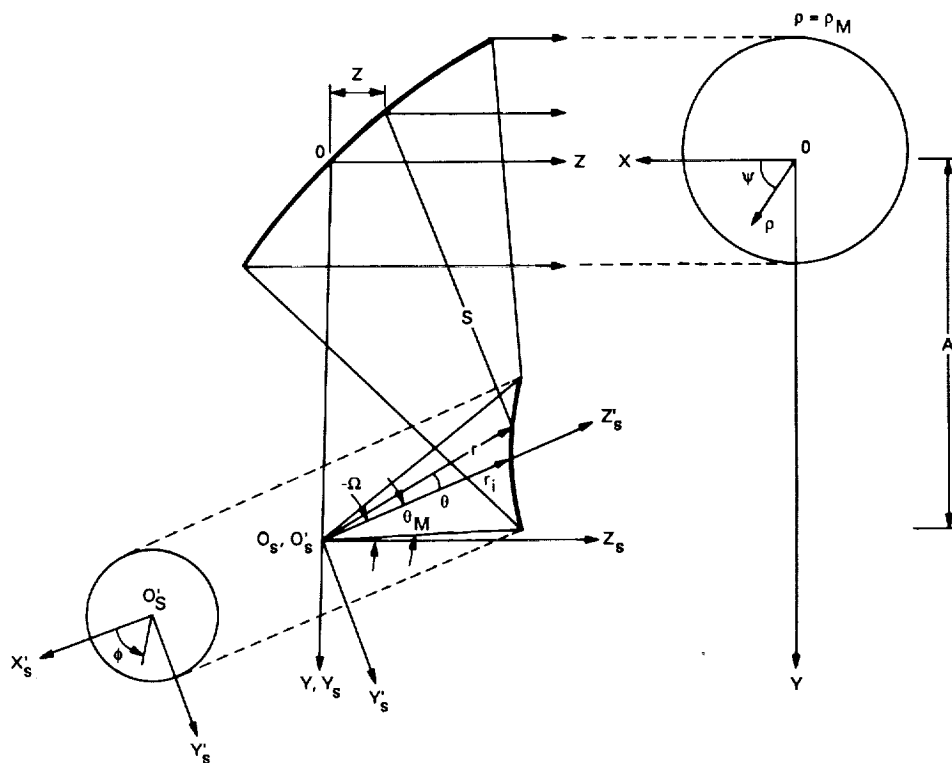


Fig. 1. Geometry of the Cassegrainian-type dual-shaped reflector antenna system.

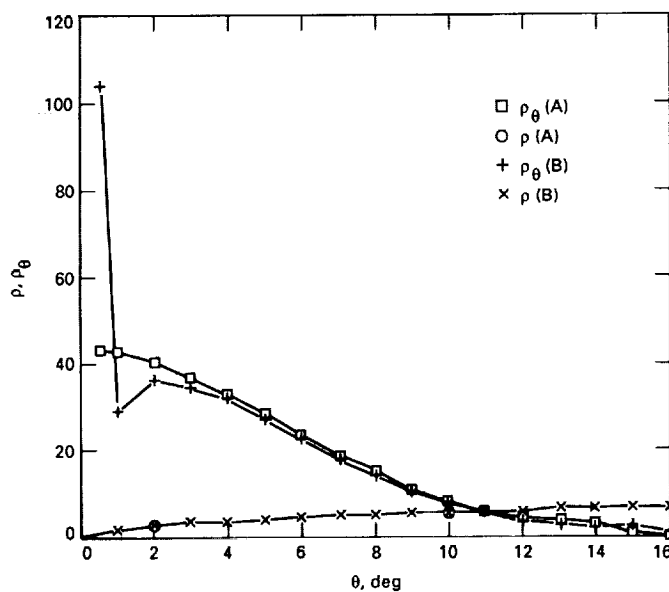


Fig. 2. Variation of ρ and ρ_θ with θ : (a) initial values of ρ_0 obtained from approximate solution and (b) arbitrarily chosen ρ_0 .

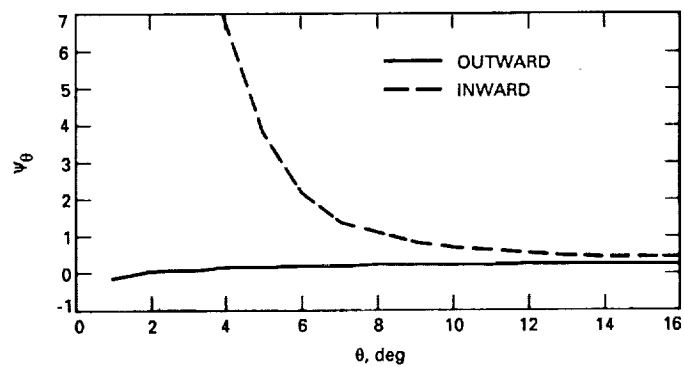


Fig. 3. Variation of ψ_θ with θ for an outward and an inward solution.

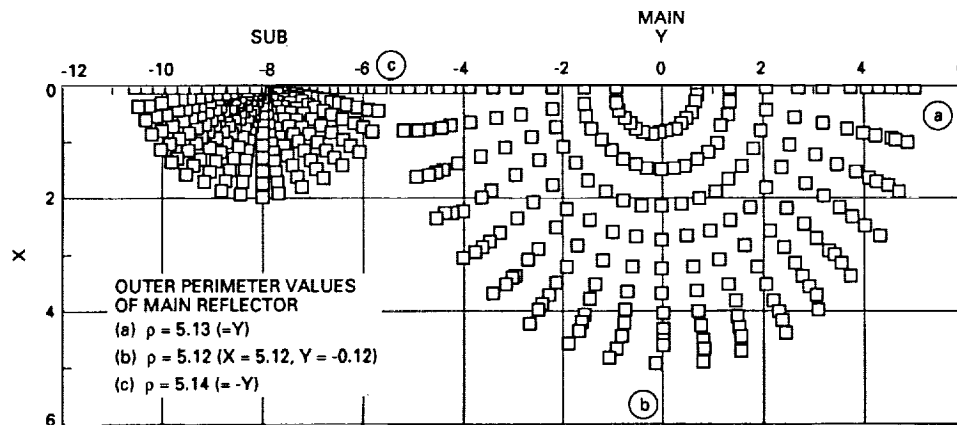


Fig. 4. Cassegrainian-type dual-reflector shape obtained from the synthesis procedure.

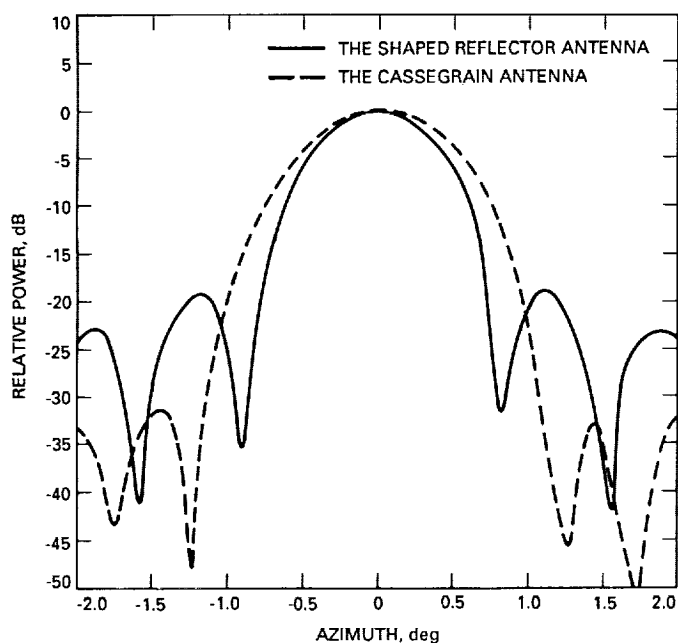


Fig. 5. Comparison of the radiation patterns—Cassegrainian-type case.

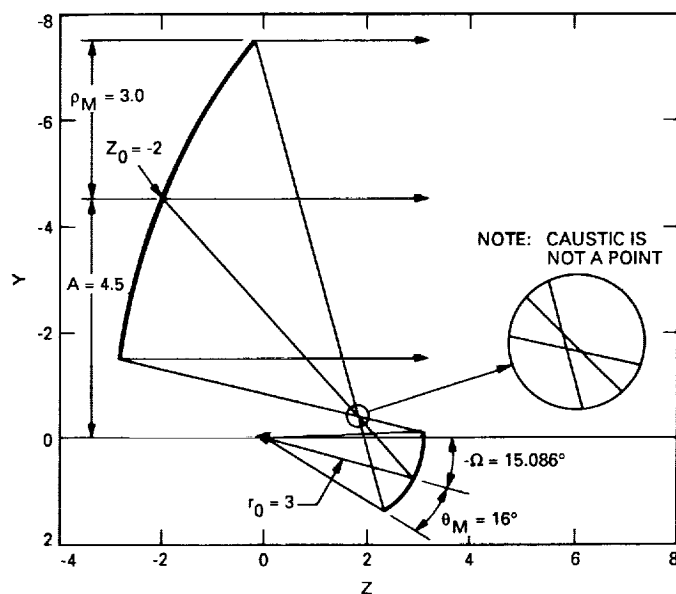


Fig. 6. Dual offset Gregorian reflectors with optimized parameters, ρ_0 , A_ψ , and Ω .

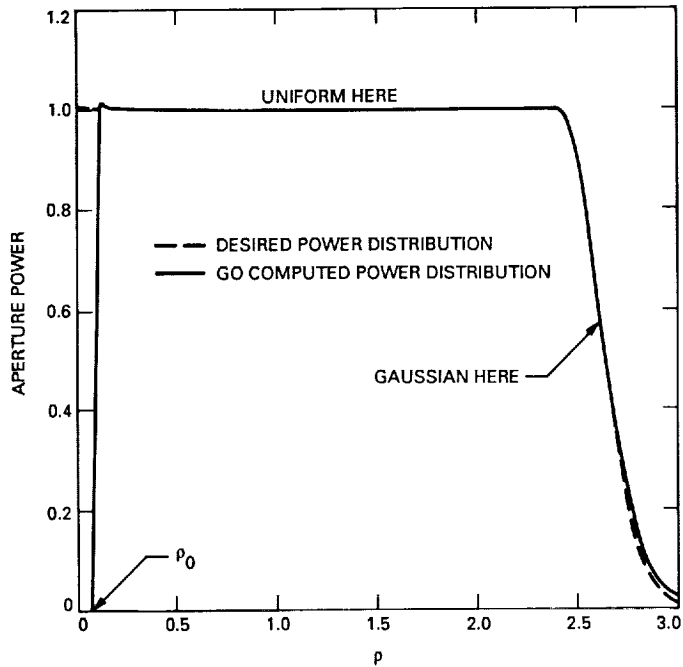


Fig. 7. Main reflector aperture power distributions—desired and computed.

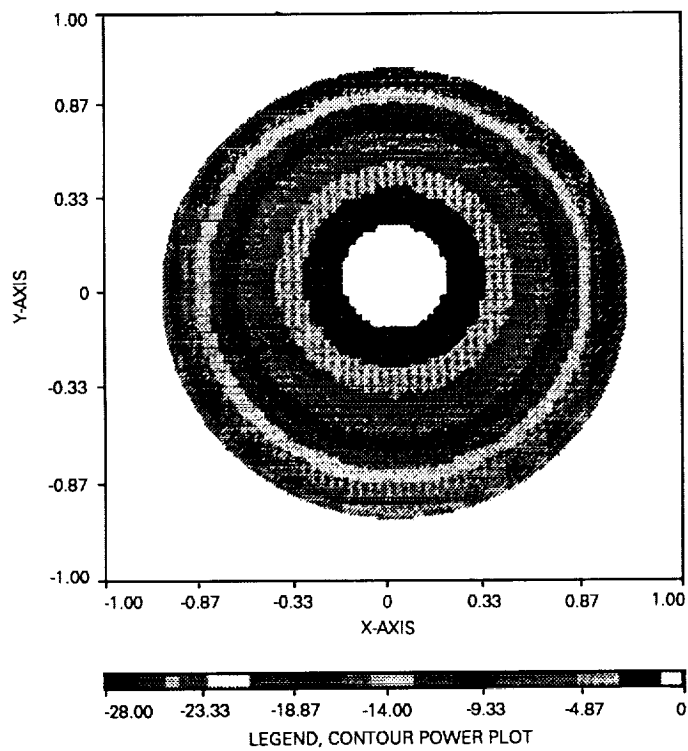


Fig. 8. Feed power pattern— $\cos^{116.6}\theta$, projected on $X_s - Y_s$ plane normal to axis of feed.

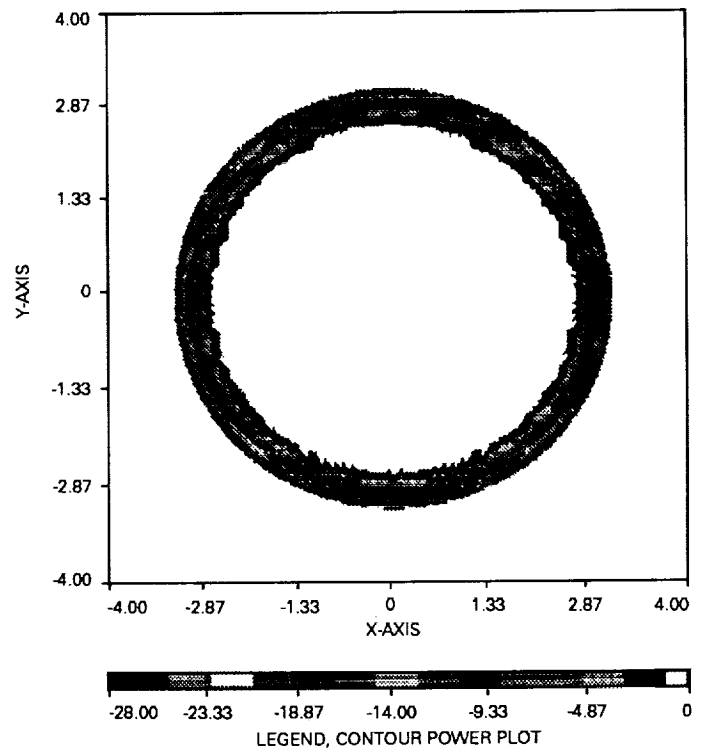


Fig. 9. Equivalent current aperture distribution. Subreflector analysis by GO (see Figs. 7 and 11).

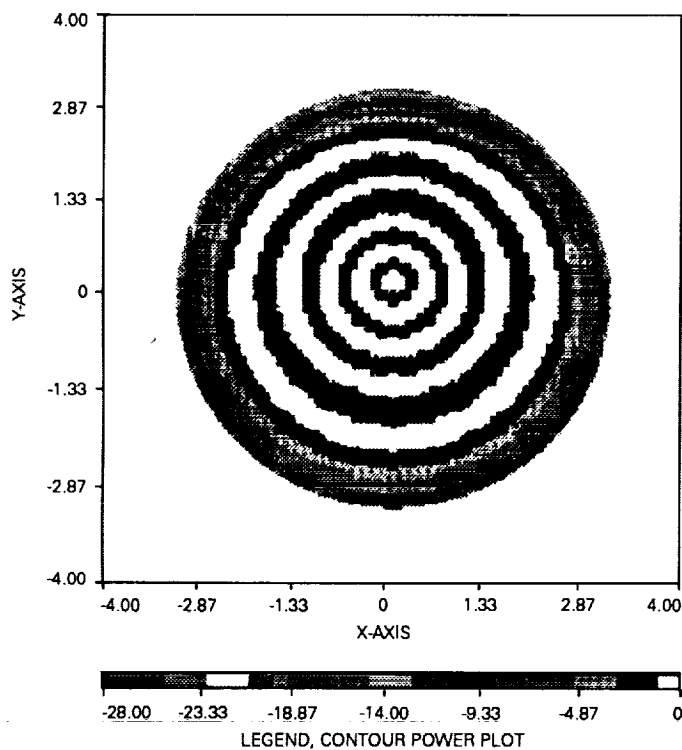


Fig. 10. Equivalent current aperture distribution. Subreflector analysis by GTD (PO analysis yields very similar result, see Fig. 11). Scale: 1 unit = 20λ ; radius = 60λ .

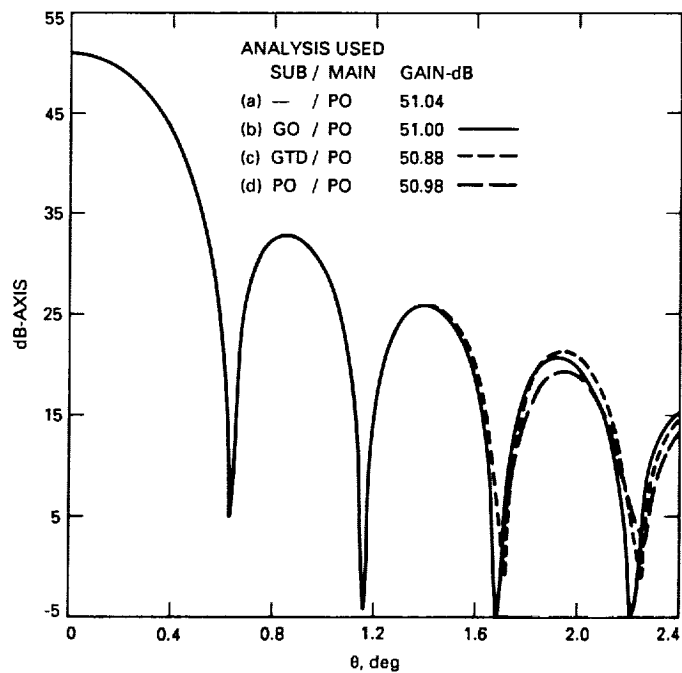


Fig. 11. Gregorian-type dual-reflector far-field power patterns, where $D = 120\lambda$ and feed $E_F = \cos^{58.3}\theta$ (-20 dB at edge of sub-reflector). Analysis by: (a) aperture PO integration (see Fig. 7); (b) GO for sub and PO for main (see Figs. 7 and 8); (c) GTD/PO (see Fig. 10); and (d) PO/PO (see Fig. 10).

N90-21888

A Real-Time Signal Combining System for Ka-Band Feed Arrays Using Maximum-Likelihood Weight Estimates

V. A. Vilnrotter and E. R. Rodemich
Communications Systems Research Section

A real-time digital signal combining system for use with Ka-band feed arrays is proposed. The combining system attempts to compensate for signal-to-noise ratio (SNR) loss resulting from antenna deformations induced by gravitational and atmospheric effects. The combining weights are obtained directly from the observed samples by using a "sliding-window" implementation of a vector maximum-likelihood parameter estimator. It is shown that with averaging times of about 0.1 second, combining loss for a seven-element array can be limited to about 0.1 dB in a realistic operational environment. This result suggests that the real-time combining system proposed here is capable of recovering virtually all of the signal power captured by the feed array, even in the presence of severe wind gusts and similar disturbances.

I. Introduction

There is considerable interest at the present time in operating the Deep Space Network at increasingly higher carrier frequencies for the purpose of enhancing its capabilities. To date, X-band (8.4 GHz) has been the highest carrier frequency employed for a deep-space mission due to the maturity of radio frequency (RF) components and technology at this frequency. Several recent studies concluded that Ka-band (32 GHz) is the proper next step in deep-space communications, providing 8- to 10-dB improvements in downlink telemetry capability [1]. These potential improvements can be attributed directly to the

higher Ka-band carrier frequencies, which yield greater antenna gains as well as reduced sensitivity to plasma effects and increased useful bandwidth. However, there are some disadvantages associated with the use of higher carrier frequencies, namely more stringent pointing requirements, increased losses due to weather effects, and greater sensitivity to imperfections of the reflecting surfaces of the antennas [2,3]. Such imperfections become particularly troublesome on large receiving antennas, which are subject to significant gravitational and thermal deformations, focussing and collimation problems, and mechanical and wind-induced vibrations. These imperfections and pointing errors translate directly to signal loss, resulting

in a corresponding degradation in telemetry performance. However, some of the signal can be recovered by means of a properly designed feed array. Our objective is to demonstrate a compensation technique that operates with negligible combining loss and hence recovers virtually all of the signal captured by the feed array.

A conceptual design for a possible real-time antenna compensation system is shown in Fig. 1. Here it is assumed that the received signal is in the form of a temporally modulated plane wave generated by a distant spacecraft. The Cassegrain receiving system consists of a large primary reflector, a secondary reflector, a focal-plane feed array, and signal processing and combining equipment designed to reconstruct the degraded signal. A possible feed-array configuration is also shown, using a maximally compact pattern of circular feed elements. This results in a hexagonal array geometry around the central feed, with each succeeding "ring" adding six more elements than the previous ring to the total; thus, one obtains 1, 7, 19, 37, etc., array elements as the number of rings is increased.

The ideal primary reflector surface and its associated signal power distribution over the focal plane are indicated by solid curves in Fig. 1, while the deformed surface with its power distribution is shown by dashed curves. Note that antenna imperfections always increase the effective spread of the signal power in the focal plane, thus reducing the maximum possible signal power coupled into any single feed element. However, the total signal power coupled into a properly designed feed array may not be greatly reduced, although its distribution over the feeds may change with time. Therefore, it is reasonable to assume that degradations in telemetry performance can be ameliorated by the use of a matched feed array in conjunction with appropriate real-time signal processing and signal-combining techniques.

Signal-processing and combining operations should be performed in a manner compatible with proposed DSN plans and equipment for future deep-space missions; the signal-combining system must interface properly with future DSN receivers. A block diagram of the proposed DSN "Advanced Receiver" front end is shown in Fig. 2(a). This receiver was designed to operate with a single feed, thus it has only one RF input and one complex baseband output (here represented by in-phase and quadrature components). The receiver front end consists of a low-noise amplifier, a downconverter chain that generates an intermediate frequency (IF) spectrum centered around 10 MHz, an automatic gain control (AGC) circuit to maintain constant average power, an 8-bit analog-to-digital (A/D) converter operating at 40 megasamples/sec, and in-phase and

quadrature phase reference samples obtained from a digital phase-locked loop, which downconverts the IF samples to baseband. The filtered outputs contain the entire modulated spectrum shifted to baseband.

As a first approximation, the front end of the proposed feed-array receiver is taken to be an array of receiver front ends (RFEs), each one associated with an element of the feed array. Under typical operating conditions, several array elements observe the signal simultaneously but with different amplitude and phase. Under favorable conditions, the distribution of signal power tends to change slowly with time as the antenna deforms due to gravitational and thermal loading. Antenna deformations and pointing errors introduced by wind gusts and turbulence-induced wavefront distortions tend to be rapidly varying effects, resulting in much less favorable conditions for reception.

A block diagram of the proposed feed-array combiner is shown in Fig. 2(b). It consists of K separate RFEs, each of which generates baseband I and Q signals that serve as inputs to a "Signal Combining System." Typically, $K = 1, 7, 19$, and so on. This system combines the K complex inputs and generates a single complex output. The receiver phase-locks to the residual carrier and provides an identical frequency reference to all RFEs simultaneously. Any phase difference between the reference signal and the received signal causes a measurable change in the complex output of the RFE. The signal combining system measures each complex output and uses these estimates to increase the effective signal level in the combined output.

II. NASA Telemetry Format

The telemetry format employed by deep-space vehicles has been, and presumably will continue to be, pulse code modulation/phase shift keying/phase modulation (PCM/PSK/PM) or PCM/PM [4]. In either case, the modulated carrier can be represented mathematically as [4]

$$s_T(t) = \sqrt{2P_T} \sin \left[\omega_c t + \sum_{i=1}^N \delta_i s^{(i)}(t) \right] \quad (1a)$$

where

$$s^{(i)}(t) = \begin{cases} d_i(t) & ; \text{PCM/PM} \\ d_i(t) \sin(\omega_{sc} t) & ; \text{PCM/PSK/PM} \end{cases} \quad (1b)$$

δ_i is the modulation index, and $d_i(t) = \pm 1, i = 1, 2, \dots, N$. Typically $N = 1$, although theoretically more than one data channel could be modulated simultaneously onto a carrier. For the case $N = 1$ let $\delta_i = \delta$, in which case the modulated carrier becomes

$$s_T(t) = \sqrt{2P_T}[\cos(\delta) \sin(\omega_c t) + s^{(1)}(t) \sin(\delta) \cos(\omega_c t)] \quad (2)$$

It is clear that the total received signal power is P_T watts, the power remaining in the carrier is $P_{RC} = P_T \cos^2(\delta)$ watts, and the power in the modulation sidebands is $P_M = P_T - P_{RC}$ watts. Since for $\delta = 90$ deg there is no power in the carrier component, this type of modulation is referred to as "suppressed-carrier modulation," while the term "residual-carrier modulation" is reserved for the case $0 < \delta < 90$ deg. To date, all DSN deep-space probes have employed the residual-carrier format [4]. Thus, residual-carrier modulation is assumed in this study, restricted to the case $N = 1$ for the sake of simplicity. The resulting signal spectrum consists of an impulsive term at the carrier frequency due to the residual carrier, and modulation sidebands centered around the fundamental subcarrier frequency and its harmonics.

A graphical representation of a typical received power spectrum is given in Fig. 3(a), showing spectral components around the fundamental frequency as well as around the higher-order harmonics of the square-wave subcarrier. A typical power spectrum of the complex baseband samples is shown in Fig. 3(b), consisting of a low-pass spectrum (the baseband version of the residual carrier) and the modulation spectra around the subcarrier fundamental and its harmonics. Observe that the dc component can be easily separated from the modulated subcarrier components by means of a narrow-band low-pass filter, provided the fundamental subcarrier frequency is significantly greater than the modulation bandwidth.

The effective bandwidth of this narrow low-pass filter will be denoted by B_A , whereas the effective bandwidth of the modulated subcarrier will be denoted by B_B in subsequent analysis. In particular, it will be convenient to consider digital filters that perform a finite averaging operation on the input sequence, in which case these bandwidths can be associated more precisely with the first zeros of the filter transfer function.

III. System Model

A block diagram of the signal combining system is shown in Fig. 4. The inputs to the combining system are

considered to be complex samples generated by the RFEs associated with each feed. The sample bandwidth is assumed to be large compared to the subcarrier frequency so that the signal modulation is not distorted by the sampling operation. Since both the amplifier and background radiation effectively add noise to the signal, independent zero-mean complex Gaussian noise samples are added to each channel, with variance determined by the strength of the total noise process and the sample integration time. The noise-contaminated samples are split into two streams: one of these (stream A) is input to the parameter estimation subsystem, while the other (stream B) serves as input to the combining subsystem. Stream A is filtered by a bank of low-pass filters that removes the modulation sidebands in order to simplify the estimator subsystem, while stream B enters the combiner unperturbed. Thus, the samples in stream B contain both the modulation sidebands and a complex constant due to the residual carrier.

It should be emphasized that the signal-combining system considered here consists of two separate subsystems: the parameter estimator and the channel combiner. The combiner structure does not depend on the form of the parameter estimator. Thus, various types of parameter estimators could be employed to determine the required complex weights, with varying degrees of complexity and performance. Here attention shall be restricted to maximum-likelihood parameter estimators, which yield the smallest estimation errors in the absence of a priori signal and noise statistics.

The received signal in the k th channel can be modeled as

$$s_k^{(1)}(t) = \sqrt{2}S_k [\cos(\delta) \sin(\omega_c t + \theta_k) + s^{(1)}(t) \sin(\delta) \cos(\omega_c t + \theta_k)] \quad (3)$$

where θ_k and $S_k = \sqrt{P_k}$ are the unknown phase and amplitude introduced by the antenna deformation, P_k is the signal power in the k th channel, δ is the modulation index, ω_c is the carrier radian frequency (nominally $2\pi \times 32$ Grad/sec) and

$$s^{(1)}(t) = d_1(t) \text{Sin}(\omega_{sc} t + \theta_{sc}) \quad (4)$$

is the data-modulated square-wave subcarrier at a fundamental frequency of $\omega_{sc}/2\pi$ Hz. Note that the total received signal power P_T is the sum of the individual signal powers, that is, $P_T = \sum_k P_k$. Both θ_k and S_k are assumed to be slowly varying functions of time, and may be considered constant over time intervals on the order of seconds to possibly minutes. The subcarrier phase θ_{sc} is taken to

be independent of the carrier phase θ_k , as it is typically generated by a separate oscillator on the spacecraft not coherent with the carrier.¹ Since $d(t)$ and $\text{Sin}(\omega_{sc}t + \theta_{sc})$ both take on the values ± 1 at all times, it follows that $s(t)$ is also restricted to the values ± 1 .

The received waveforms are downconverted to an IF frequency of about 10 MHz by means of a downconverter chain that mathematically corresponds to a single downconversion operation. It is assumed here that downconversion is performed with a "perfect" frequency reference that does not introduce any additional phase errors into the resulting IF signals. These signals can be represented as

$$\begin{aligned} s_{Ik}^{(1)}(t) &= \left[s_k^{(1)}(t) \sqrt{2} \cos((\omega_c - \omega_I)t) \right]_{LP} \\ &= S_k [\cos(\delta) \sin(\omega_I t + \theta_k) \\ &\quad + s^{(1)}(t) \sin(\delta) \cos(\omega_I t + \theta_k)] \end{aligned} \quad (5)$$

where "LP" refers to a low-pass filtering operation.

The IF waveforms are sampled by an A/D converter which generates 8-bit samples every T_0 seconds, or in the case of the Advanced Receiver, at a rate of 40 million samples per second. It is convenient to view these samples as the coefficients of an orthonormal expansion with basis functions the well-known "sampling functions"

$$\phi_i(t) = \frac{\sin[(\pi/T_0)(t - iT_0)]}{(\pi/T_0)(t - iT_0)} \quad (6)$$

which form a complete orthonormal set in the space of band-limited functions over the real line. In terms of these basis functions the expansion becomes

$$s_{Ik}^{(1)}(t) = \sum_{i=-\infty}^{\infty} s_{Ik}(i) \phi_i(t) \quad (7a)$$

with coefficients that represent samples of the waveform at the sampling instants " iT_0 ":

$$\begin{aligned} s_{Ik}(i) &= \int_{-\infty}^{\infty} s_{Ik}^{(1)}(t) \phi_i(t) dt \\ &= S_k [\cos(\delta) \sin(\omega_I iT_0 + \theta_k) \\ &\quad + s^{(1)}(iT_0) \sin(\delta) \cos(\omega_I iT_0 + \theta_k)] \end{aligned} \quad (7b)$$

Here it is assumed that the sampling time T_0 is small enough to avoid aliasing and that there are sufficient quantization levels to allow the representation of the quantized-amplitude samples by continuous-amplitude samples.

The final downconversion to baseband is accomplished using digital samples. Both in-phase and quadrature samples are generated simultaneously using samples of a perfect IF frequency reference, as indicated in Fig. 2(a). This operation is represented mathematically as

$$s_k(i) = [2s_{Ik}(i) \sin(\omega_I iT_0)]_{LP} \quad (8a)$$

$$c_k(i) = [2s_{Ik}(i) \cos(\omega_I iT_0)]_{LP} \quad (8b)$$

The sample products are filtered by digital low-pass filters that pass only the modulated subcarrier frequencies. The resulting baseband in-phase and quadrature samples in the k th channel at time iT_0 become

$$s_k(i) = S_k [\cos(\delta) \cos(\theta_k) - s(i) \sin(\delta) \sin(\theta_k)] \quad (9a)$$

and

$$c_k(i) = S_k [\cos(\delta) \sin(\theta_k) + s(i) \sin(\delta) \cos(\theta_k)] \quad (9b)$$

where

$$s(i) = d(iT_0) \text{Sin}(\omega_{sc} iT_0 + \theta_{sc}) \quad (9c)$$

These signals contain the unknown amplitude and phase that affect both the residual-carrier and the modulated-subcarrier signals.

Making use of the identities

$$\text{Re}\{e^{j\theta}[a + jb]\} = a \cos(\theta) - b \sin(\theta) \quad (10a)$$

$$\text{Im}\{e^{j\theta}[a + jb]\} = a \sin(\theta) + b \cos(\theta) \quad (10b)$$

and letting $\tilde{\cdot}$ denote a complex quantity, one can also write

$$s_k(i) = \text{Re}\{\tilde{s}_k(i)\} \quad (11a)$$

$$c_k(i) = \text{Im}\{\tilde{s}_k(i)\} \quad (11b)$$

where

$$\begin{aligned} \tilde{s}_k(i) &= S_k e^{j\theta_k} [\cos(\delta) + js(i) \sin(\delta)] \\ &= \tilde{V}_k [\cos(\delta) + js(i) \sin(\delta)] \end{aligned} \quad (12)$$

¹ P. Kinman, Telecommunications Systems Section, personal communication.

is the "complex envelope" of the baseband signal samples. The complex coefficient \tilde{V}_k contains both the amplitude and phase of the baseband sequences.

Next the representation of the additive noise is considered. Bandpass noise in the k th channel can be modeled as

$$n_k(t) = \sqrt{2}[n_{ck}(t) \cos(\omega_c t) + n_{sk}(t) \sin(\omega_c t)] \quad (13)$$

where $n_{ck}(t)$ and $n_{sk}(t)$ are independent zero-mean Gaussian baseband processes, each with spectral level $N_{0k}/2$.

Downconverting to IF yields

$$\begin{aligned} n_{Ik}(t) &= [n_k(t) \sqrt{2} \cos\{(\omega_c - \omega_I)t\}]_{LP} \\ &= n_{ck}(t) \cos(\omega_I t) + n_{sk}(t) \sin(\omega_I t) \end{aligned} \quad (14)$$

while the IF samples are again the coefficients of the orthonormal expansion

$$n_{Ik}(t) = \sum_{i=-\infty}^{\infty} n_{Ik}(i) \phi_i(t) \quad (15a)$$

$$\begin{aligned} n_{Ik}(i) &= \int_{-\infty}^{\infty} n_{Ik}(t) \phi_i(t) dt \\ &= n_{ck}(i) \cos(\omega_I i T_0) + n_{sk}(i) \sin(\omega_I i T_0) \end{aligned} \quad (15b)$$

Assuming that the noise processes are broadband compared to the sampling bandwidth, the correlation function of the noise samples becomes

$$\begin{aligned} E[n_{ck}(i) n_{ck}(\ell)] &= E[n_{sk}(i) n_{sk}(\ell)] \\ &= \frac{N_{0k}}{2T_0} \delta_{i\ell} \triangleq \sigma_k^2 \delta_{i\ell} \end{aligned} \quad (16)$$

In-phase and quadrature noise samples are generated along with the signal samples, with the result that

$${}_s n_k(i) = [2n_{Ik}(i) \sin(\omega_I i T_0)]_{LP} = n_{sk}(i) \quad (17a)$$

and

$${}_c n_k(i) = [2n_{Ik}(i) \cos(\omega_I i T_0)]_{LP} = n_{ck}(i) \quad (17b)$$

A complex representation for the noise samples is obtained by letting

$${}_s n_k(i) = \text{Re}\{\tilde{n}_k(i)\} \quad (18a)$$

$${}_c n_k(i) = \text{Im}\{\tilde{n}_k(i)\} \quad (18b)$$

where

$$\tilde{n}_k(i) = n_{sk}(i) + j n_{ck}(i) \quad (18c)$$

is a sequence of complex noise samples. Note that the variance of the complex noise samples is $2\sigma_k^2$, because the real and imaginary components are independent zero-mean processes, each with variance σ_k^2 . The use of complex notation will often allow considerable simplification in the derivations that follow.

Next, the problem is generalized somewhat by sampling of the primitive samples in streams A and B at different rates. The final samples are formed by averaging consecutive primitive samples, and hence represent an effective low-pass filtering operation. Since the samples in stream B must reflect the temporal variations in the modulated subcarrier accurately, typically only a few primitive samples can be averaged here. In stream A, however, enough samples must be averaged to ensure that the modulated subcarrier terms are filtered out, leaving only the residual carrier terms. Here one must be careful to make sure that the underlying processes of interest are not filtered out. Therefore, in stream A strict lower and upper bounds must be observed on the number of primitive samples averaged, while in stream B only the upper bound is relevant.

Suppose that M_A primitive samples are averaged in every channel of stream A, and M_B in stream B. This means that the time delay between samples is $T_A = M_A T_0$ in stream A and $T_B = M_B T_0$ in stream B. The situation is illustrated in Fig. 5, which shows the timing relationship between the samples in the two streams.

Applying the averaging operation of stream B to the primitive signal samples again yields Eq. (9), but evaluated at times $i_B = i M_B T_0$. The averaging operation in stream A effectively removes the subcarrier terms from the primitive samples, leaving the residual-carrier samples defined at times $i_A = i M_A T_0$:

$${}_s x_k(i_A) = [{}_s s_k(i)]_{LP} = S_k \cos(\delta) \cos(\theta_k) \quad (19a)$$

$${}_c x_k(i_A) = [{}_c s_k(i)]_{LP} = S_k \cos(\delta) \sin(\theta_k) \quad (19b)$$

Equation (19) can be expressed in complex form as

$$\begin{aligned} \tilde{x}_k(i_A) &= {}_s x_k(i_A) + j {}_c x_k(i_A) \\ &= S_k e^{j\theta_k} \cos(\delta) \\ &= \tilde{V}_k \cos(\delta) \triangleq \tilde{X}_k = X_{kR} + j X_{kI} \end{aligned} \quad (20)$$

Note that samples in both streams are obtained from the same sequence of primitive samples in the corresponding channels, hence the resulting sample variances in the corresponding channels are always related by the ratio of the number of primitive samples averaged in the two streams. Let the noise samples in stream A be denoted by $m_k(i_A)$, and in stream B by $n_k(i_B)$. The averaging operation reduces the noise-sample variances by factors of M_A and M_B , respectively, so that $\sigma_{A_k}^2 = \sigma_k^2/M_A$ and $\sigma_{B_k}^2 = \sigma_k^2/M_B$. The noise variances in the two streams are related by $\sigma_{B_k}^2 = \eta \sigma_{A_k}^2$, where $\eta \triangleq M_A/M_B$ is the ratio of the effective bandwidths of the two streams.

The parameter-estimation problem is greatly simplified if estimation is based only on the residual-carrier sequence as given by Eq. (20), because interference from the data modulation is not present. Since only a small fraction of the total signal power is in the residual carrier, one might expect substantial performance improvements by making use of the modulation sidebands as well. However, the more complicated estimation algorithms that make use of the full-spectrum sequence of Eq. (12) will not be addressed here. It is emphasized that while the averaging operation in stream B is optional, it is mandatory in stream A whenever modulation sidebands are present.

IV. The Combining Algorithm

The mathematical model for the combining system depicted in Fig. 4 can be summarized as follows: the noise-corrupted complex baseband samples in streams A and B are modeled as

$$\tilde{u}_k(i_A) = \tilde{x}_k(i_A) + \tilde{m}_k(i_A) \quad (21a)$$

$$\tilde{r}_k(i_B) = \tilde{s}_k(i_B) + \tilde{n}_k(i_B) \quad (21b)$$

Complex combining weights are determined by the estimator subsystem operating on stream A, and applied to the samples $\tilde{r}_k(i_B)$ in stream B to produce the combined output sequence $\tilde{z}(i_B)$:

$$\tilde{z}(i_B) = \sum_{k=1}^K \tilde{r}_k(i_B) \tilde{w}_k \quad (22)$$

The object of the combining algorithm is to maximize the signal-to-noise ratio of the combined output samples, defined as

$$\rho_z \triangleq \frac{|E[\tilde{z}(i_B)]|^2}{\text{var}[\tilde{z}(i_B)]} = \frac{|E[\sum_{k=1}^K \tilde{w}_k \tilde{r}_k(i_B)]|^2}{\text{var}[\sum_{k=1}^K \tilde{w}_k \tilde{r}_k(i_B)]} \quad (23)$$

Although in general ρ_z is a function of time, here it is assumed that all relevant processes are stationary, so that the time index can be ignored. Substituting Eq. (21b) for $\tilde{r}_k(i_B)$ yields

$$\rho_z = \frac{|\sum_{k=1}^K \tilde{w}_k \tilde{V}_k [\cos(\delta) + j \sin(\delta) s(i_B)]|^2}{\sum_{k=1}^K |\tilde{w}_k|^2 2\sigma_{Bk}^2} \quad (24)$$

Making use of the Schwarz inequality,

$$\begin{aligned} & \left| \sum_{k=1}^K (\sqrt{2} \tilde{w}_k \sigma_{Bk}) \left(\frac{\tilde{V}_k}{\sqrt{2} \sigma_{Bk}} [\cos(\delta) + j \sin(\delta) s(i_B)] \right) \right|^2 \\ & \leq \sum_{k=1}^K |\tilde{w}_k|^2 2\sigma_{Bk}^2 \sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} \end{aligned} \quad (25)$$

Dividing both sides of Eq. (25) by $\sum_k |\tilde{w}_k|^2 2\sigma_{Bk}^2$ yields the following upper bound on ρ_z :

$$\rho_z \leq \sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} \triangleq \rho \quad (26)$$

The choice of weights that achieves the value of the right-hand side of Eq. (26) is

$$\tilde{w}_k = \frac{\tilde{V}_k^*}{2\sigma_{Bk}^2} \quad (27a)$$

as can be seen by direct substitution

$$\begin{aligned} \rho_z &= \frac{|\sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} [\cos(\delta) + j \sin(\delta) s(i_B)]|^2}{\sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2}} \\ &= \sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} = \rho \end{aligned} \quad (27b)$$

Therefore, if the signal and noise parameters were known, then the combining operation defined by Eqs. (22) and (27a) would achieve the sample SNR ρ defined in Eq. (26). Note that with the optimum choice of weights, the mean value of the combined signal and the variance of the combined noise are both exactly equal to ρ . If all noise variances are equal, $\sigma_{Bk}^2 = \sigma_B^2$, then we have the interpretation $\rho = P_T/2\sigma_B^2$, where $P_T = \sum_k S_k^2$ is the total signal power collected by the array.

Since the required signal and noise parameters are generally not known, this upper bound on the sample SNR is not achieved in practice. When dealing with unknown parameters, a plausible approach is to use the best available estimates of the underlying parameters to estimate the combining weights. This is the approach adopted in this article; the estimated weights are obtained by applying the optimum weight formula (Eq. 27a) to the maximum-likelihood estimates of the underlying parameters $\{\tilde{V}_k\}$ and $\{\sigma_{Bk}^2\}$.

V. The Maximum-Likelihood Estimator

The combining algorithm requires an estimate of the complex signal coefficient \tilde{V}_k and of the noise variance σ_{Bk}^2 . These quantities can be estimated simultaneously from the observed samples, either in stream A or in stream B. Here attention is restricted to the residual-carrier sample stream (stream A), because of the relative simplicity of the problem when the data-modulated subcarrier is not present.

In the absence of a-priori statistics about the values of the unknown parameters \tilde{X}_k and σ_{Bk}^2 , the maximum-likelihood approach yields the best estimates of the embedded parameters, on the average. These estimates are obtained from the likelihood function, which is the conditional joint probability density of the observed samples, conditioned on the parameters \tilde{X}_k and σ_{Bk}^2 . The maximum-likelihood estimates are those values \hat{X}_k and $\hat{\sigma}_{Bk}^2$ that simultaneously maximize the likelihood function, or a monotonically increasing functional of it.

The log-likelihood function is obtained from the conditional probability density of the observables, conditioned on the parameters of interest. In stream A only the residual carrier is present, hence the observed complex samples in the k th channel at time " $i_A T_A$ " can be represented as

$$\tilde{u}_k(i_A) = \tilde{X}_k + \tilde{m}_k(i_A) \quad (28)$$

where $\tilde{X}_k = \tilde{V}_k \cos(\delta)$, and $\tilde{m}_k(i_A)$ is a zero-mean complex Gaussian random variable with variance $2\sigma_{A_k}^2$. It is assumed that all noise samples are independent random variables. If L consecutive samples are observed in each of K channels, the joint probability density of the entire $K \times L$ complex array becomes

$$p(\tilde{\mathbf{u}}|\tilde{\mathbf{X}}, \sigma_{\mathbf{A}}^2) = \prod_{k=1}^K \prod_{i_A=1}^L (2\pi\sigma_{A_k}^2)^{-1} e^{-|\tilde{u}_k(i_A) - \tilde{X}_k|^2 / 2\sigma_{A_k}^2} \quad (29)$$

where $\tilde{\mathbf{u}}$ is a $K \times L$ -dimensional complex matrix whose elements encompass all possible values of the indices " i_A " and " k ," while $\tilde{\mathbf{X}}$ and $\sigma_{\mathbf{A}}^2$ are K -dimensional complex vectors. Denoting the real and imaginary components of \tilde{X}_k by X_{kR} and X_{kI} , respectively, the natural logarithm of Eq. (29) becomes

$$\begin{aligned} \Lambda &\triangleq \ln \left[p(\tilde{\mathbf{u}}|\tilde{\mathbf{X}}, \sigma_{\mathbf{A}}^2) \right] \\ &= -L \sum_{k=1}^K \ln(2\pi\sigma_{A_k}^2) \\ &\quad - \sum_{k=1}^K \left(\frac{1}{2\sigma_{A_k}^2} \sum_{i_A=1}^L [(u_{kR}(i_A) - X_{kR})^2 \right. \\ &\quad \left. + (u_{kI}(i_A) - X_{kI})^2] \right) \end{aligned} \quad (30)$$

This is called the log-likelihood function. The maximum-likelihood estimates of X_{kR} , X_{kI} , and $\sigma_{A_k}^2$ are those values \hat{X}_{kR} , \hat{X}_{kI} , and $\hat{\sigma}_{A_k}^2$ that simultaneously maximize the log-likelihood function Λ . Differentiating Λ with respect to each X_{kR} and X_{kI} , setting the result equal to zero, and solving yields the maximum-likelihood estimates

$$\hat{X}_{kR} = \frac{1}{L} \sum_{i_A=1}^L u_{kR}(i_A) \quad (31a)$$

$$\hat{X}_{kI} = \frac{1}{L} \sum_{i_A=1}^L u_{kI}(i_A) \quad (31b)$$

and

$$\hat{\tilde{X}}_k = \hat{X}_{kR} + j\hat{X}_{kI} \quad k = 1, 2, \dots, K \quad (31c)$$

The estimate of the noise variance is obtained by writing the log-likelihood function, maximized over $\tilde{\mathbf{X}}$, as

$$\max_{\tilde{\mathbf{X}}} \Lambda = -L \sum_{k=1}^K \ln(2\pi\sigma_{A_k}^2) - \sum_{k=1}^K \left(\frac{G_k}{2\sigma_{A_k}^2} \right) \quad (32a)$$

$$G_k \triangleq \sum_{i_A=1}^L [(u_{kR}(i_A) - \hat{X}_{kR})^2 + (u_{kI}(i_A) - \hat{X}_{kI})^2] \quad (32b)$$

Differentiating with respect to $\sigma_{A_k}^2$, setting the result equal to zero, and solving yields

$$\hat{\sigma}_{A_k}^2 = \frac{1}{2L} G_k \quad (33)$$

Note that the maximum-likelihood signal estimates \hat{X}_{kR} and \hat{X}_{kI} are unbiased estimates of the mean values of the real sequences $u_{kR}(i_A)$ and $u_{kI}(i_A)$, also known in statistics as the "sample mean." However, the maximum-likelihood estimates of the noise variances are biased, and hence do not correspond exactly to the unbiased "sample variance" used in statistics. Here we use the unbiased variance estimate

$$\hat{\sigma}_{Ak,ub}^2 = \frac{1}{2(L-1)} G_k \quad (34)$$

For large L , the difference between the two estimates becomes negligibly small. The unbiased estimates shall be used in this analysis because each of these estimates can be modeled conveniently as the sum of the true variance plus a zero-mean error term.

The maximum-likelihood estimators defined above provide the best possible estimates of the desired parameters if a-priori statistics are not available, as long as the signal parameters and noise statistics do not change with time. These results can be extended to the slowly varying case by selecting the total observation time to be small compared to the characteristic time scale of the variations. Thus, over a short time interval the above solutions apply. However, the observation interval must now be continuously shifted in time to follow the parameter variations. This extension leads to the following "sliding-window" estimator structures:

$$\hat{X}_k(i_A) = \frac{1}{L} \sum_{\ell=i_A-L}^{i_A-1} \tilde{u}_k(\ell) \quad (35a)$$

$$\hat{\sigma}_{Ak,ub}^2(i_A) = \frac{1}{2(L-1)} \sum_{\ell=i_A-L}^{i_A-1} \left| \tilde{u}_k(\ell) - \hat{X}_k(i_A) \right|^2 \quad (35b)$$

Thus, the current estimates at time " $i_A T_A$ " are based on the previous L samples, up to and including the sample at time " $(i_A - 1)T_A$ ".

For large values of L and short sample durations, the computational burden required by these estimators may become quite severe. This difficulty can be ameliorated by observing that a recursive implementation is possible, where the current estimate is obtained in terms of the previous estimate, the current observable, and the "oldest" observable in the previous sum. The recursive forms are

$$\hat{X}_k(i_A) = \hat{X}_k(i_A - 1) + \frac{1}{L} [\tilde{u}_k(i_A - 1) - \tilde{u}_k(i_A - L - 1)] \quad (36a)$$

$$\begin{aligned} \hat{\sigma}_{Ak,ub}^2(i_A) = & \hat{\sigma}_{Ak,ub}^2(i_A - 1) + \frac{1}{2(L-1)} \left\{ |\tilde{u}_k(i_A - 1)|^2 \right. \\ & - |\tilde{u}_k(i_A - L - 1)|^2 \\ & \left. - L \left[\left| \hat{X}_k(i_A) \right|^2 - \left| \hat{X}_k(i_A - 1) \right|^2 \right] \right\} \quad (36b) \end{aligned}$$

Note that the corrections to the previous estimates can be obtained from a digital "first-in, first-out" stack, minimizing the number of computations required for each update.

It is well known that the maximum-likelihood estimate of the ratio of two parameters is equivalent to the ratio of the maximum-likelihood estimates [5]. It follows that the maximum-likelihood estimate of the k th weight can be obtained from the maximum-likelihood estimates of the complex signal estimates and the noise variances as

$$\hat{w}_k(i_A) = \frac{\hat{V}_k^*(i_A)}{2\eta\sigma_{Ak}^2(i_A)} = \frac{\hat{X}_k^*(i_A)}{2\cos(\delta)\eta\sigma_{Ak}^2(i_A)} \quad (37)$$

The maximum-likelihood estimate of the k th weight therefore depends on the signal parameters in each channel of stream A and on the corresponding noise variances.

An interesting special case occurs if the noise variance in each channel is the same. This situation could arise if the amplifier characteristics in the various channels were exceptionally well matched and all feeds observed the same background power levels, so that the noise temperatures in the various channels were essentially the same. In this case, the common variance does not have to be estimated, since it becomes a scaling factor applied to all of the weights. Because each signal and noise term is equally scaled, the SNR of the combined noisy signal is not affected if this common scaling factor is ignored.

VI. Combiner Performance Analysis

In this section, the performance of the signal combining system will be evaluated. Constant signal components will be assumed throughout, and it shall also be assumed that L is large, so that the biased and unbiased variance estimates can be considered equal. However, the unbiased variance estimates will be used in the analysis. Performance will be evaluated in the "steady-state," that is, assuming that L or more samples have already been observed.

Assuming the "sliding-window" configuration of Eq. (35), it is seen that the estimates at time $i_A T_A$ are

based on the previous L samples, up to and including the sample at time $(i_A - 1)T_A$. Since the estimates are unbiased, each can be expressed as the sum of the true parameter plus a zero-mean random variable representing the estimation error. This model yields

$$\hat{X}_k(i_A) = \tilde{X}_k + (\hat{X}_k(i_A) - \tilde{X}_k) \triangleq \tilde{X}_k + \tilde{\epsilon}_{X_k}(i_A) \quad (38a)$$

$$\hat{\sigma}_{Ak}^2(i_A) = \sigma_{Ak}^2 + (\hat{\sigma}_{Ak}^2(i_A) - \sigma_{Ak}^2) \triangleq \sigma_{Ak}^2 + \epsilon_{\sigma_{Ak}^2}(i_A) \quad (38b)$$

where

$$\tilde{\epsilon}_{X_k}(i_A) = \frac{1}{L} \sum_{\ell=i_A-L}^{i_A-1} (\tilde{u}_k(\ell) - \tilde{X}_k) \quad (39a)$$

$$\epsilon_{\sigma_{Ak}^2}(i_A) = \frac{1}{2(L-1)} \sum_{\ell=i_A-L}^{i_A-1} |\tilde{u}_k(\ell) - \hat{X}_k(i_A)|^2 - \sigma_{Ak}^2 \quad (39b)$$

Since the estimates are unbiased, it follows that the error terms are zero mean. The variance of the estimation errors can be obtained using well-known expressions for the variance of the sample mean and the variance of the sample variance, as follows:

$$\text{var} [\tilde{\epsilon}_{X_k}(i_A)] = \frac{E |\tilde{m}_k(i_A)|^2}{L} = \frac{2\sigma_{Ak}^2}{L} = \frac{2\sigma_{Bk}^2}{\eta L} \quad (40a)$$

$$\begin{aligned} \text{var} [\epsilon_{\sigma_{Ak}^2}(i_A)] &= \text{var} \left\{ \frac{1}{2(L-1)} \sum_{\ell=i_A-L}^{i_A-1} |\tilde{u}_k(\ell) - \hat{X}_k(i_A)|^2 \right\} \\ &= \frac{1}{2} \text{var} \left\{ \frac{1}{(L-1)} \sum_{\ell=i_A-L}^{i_A-1} [u_{kR}(\ell) - \hat{X}_{kR}(i_A)]^2 \right\} \\ &= \frac{1}{2L} \left(\mu_4 - \frac{L-3}{L-1} \sigma_{Ak}^4 \right) \\ &= \frac{\sigma_{Ak}^4}{L-1} \end{aligned} \quad (40b)$$

while

$$\text{var} [\epsilon_{\sigma_{Bk}^2}(i_A)] = \frac{\eta^2 \sigma_{Ak}^4}{L-1} = \frac{\sigma_{Bk}^4}{L-1} \quad (40c)$$

In deriving Eq. (40b) recall the fact that the fourth central moment of a Gaussian random variable is $\mu_4 =$

$3\sigma_{Ak}^4$. Since for a Gaussian population the sample mean and sample variance are independent random variables, it follows that $\tilde{\epsilon}_{X_k}(i)$ and $\epsilon_{\sigma_{Bk}^2}$ are independent as well. Since the i_A th estimate is based on L previous samples up to and including sample number $(i_A - 1)$, and since each sample is independent of all others, it also follows that the i_A th estimate is independent of the i_A th sample.

Making use of the model defined in Eq. (38), the maximum-likelihood estimate of the k th weight becomes

$$\hat{w}_k(i_A) = \frac{\tilde{X}_k^*}{2 \cos(\delta) \sigma_{Bk}^2} \left[\frac{(1 + \tilde{\epsilon}_{X_k}^*(i_A))}{(1 + \epsilon_{\sigma_{Bk}^2}(i_A))} \right] \quad (41a)$$

where

$$\tilde{\epsilon}_{X_k}^* = \tilde{\epsilon}_{X_k}^* / \tilde{X}_k^* \quad (41b)$$

and

$$\epsilon_{\sigma_{Bk}^2} = \epsilon_{\sigma_{Bk}^2} / \sigma_{Bk}^2 \quad (41c)$$

If the standard deviation of each estimation error is small compared to the estimate, then Eq. (41a) can be approximated as

$$\begin{aligned} \hat{w}_k(i_A) &= \frac{\tilde{X}_k^*}{2 \cos(\delta) \sigma_{Bk}^2} [1 + \tilde{\epsilon}_{X_k}^*(i_A)] \\ &\quad \times [1 - \epsilon_{\sigma_{Bk}^2}(i_A) + \epsilon_{\sigma_{Bk}^2}^2(i_A) - \dots] \\ &\simeq \frac{\tilde{X}_k^*}{2 \cos(\delta) \sigma_{Bk}^2} [1 + (\tilde{\epsilon}_{X_k}^*(i_A) - \epsilon_{\sigma_{Bk}^2}(i_A))] \\ &\triangleq \frac{\tilde{X}_k^*}{2 \cos(\delta) \sigma_{Bk}^2} [1 + \tilde{\epsilon}_{w_k}^*(i_A)] \\ &= \frac{\tilde{V}_k^*}{2 \sigma_{Bk}^2} [1 + \tilde{\epsilon}_{w_k}^*(i_A)] \end{aligned} \quad (42)$$

Therefore, the weight estimates may also be expressed approximately as the sum of the true weight plus an error term, provided the errors in the component estimates are

sufficiently small. As before, the error term is zero mean, with variance

$$\begin{aligned}\text{var} [\tilde{e}_{w_k}^*(i_A)] &= \text{var} [\tilde{e}_{X_k}^*(i_A)] + \text{var} [e_{\sigma_{Bk}^2}(i_A)] \\ &= \frac{2\sigma_{A_k}^2}{|\tilde{X}_k|^2 L} + \frac{1}{(L-1)} \\ &= \frac{2\sigma_{Bk}^2}{\eta |\tilde{V}_k|^2 L \cos^2(\delta)} + \frac{1}{(L-1)}\end{aligned}\quad (43)$$

Note that the current weight estimate is independent of the current sample, since it depends only on the previous L samples. Using the “small-error” approximation for the weighting factors, the combiner output sequence defined in Eq. (22) can be expressed as

$$\tilde{z}(i_B) \simeq \tilde{s}_c(i_B) + \tilde{n}_c(i_B) \quad (44)$$

where

$$\tilde{s}_c(i_B) \triangleq [\cos(\delta) + j \sin(\delta)s(i_B)] \sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} [1 + \tilde{e}_{w_k}^*(i_A)] \quad (45a)$$

and

$$\tilde{n}_c(i_B) \triangleq \sum_{k=1}^K \frac{\tilde{V}_k^* \tilde{n}_k(i_B)}{2\sigma_{Bk}^2} [1 + \tilde{e}_{w_k}^*(i_A)] \quad (45b)$$

In these equations, the index i_A refers to the last estimate in stream A that occurred prior to the current sample i_B in stream B.

The combiner output in Eq. (44) can be viewed as having two components: a signal term with random magnitude that takes on a new value each time the weight index changes, plus an equivalent additive-noise term that takes on a different value for each new sample of the combined signal. The magnitude of the mean value of the signal term is ρ , the sum of the channel SNRs that would be obtained if the weights were known without error:

$$\rho = \sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_{Bk}^2} \quad (46)$$

The variance of the signal term can be expressed in terms of ρ as

$$\text{var} [\tilde{s}_c(i_B)] = \frac{\rho}{\eta L \cos^2(\delta)} + \frac{1}{(L-1)} \sum_{k=1}^K \frac{|\tilde{V}_k|^4}{4\sigma_{Bk}^4} \quad (47)$$

Note the dependence of the signal variance on the sum of squares of the channel SNRs in Eq. (47). This term can be bounded as follows:

$$\frac{\rho^2}{K} \leq \sum_{k=1}^K \frac{|\tilde{V}_k|^4}{4\sigma_{Bk}^4} \leq \rho^2 \quad (48)$$

The left side follows from the Schwarz inequality: equality holds when all the channel SNRs are equal. The right side is trivial since one obtains this inequality by throwing away all of the cross terms in the square of the series for ρ . Equality holds when all the signal power is in a single channel. Useful bounds on ρ follow if the noise spectral levels in the various channels are bounded by known minimum and maximum values $N_{0 \min}$ and $N_{0 \max}$:

$$\left(\frac{P_T}{N_{0 \max}} \right) T_0 M_B \triangleq \rho_l \leq \rho \leq \rho_u \triangleq \left(\frac{P_T}{N_{0 \min}} \right) T_0 M_B \quad (49)$$

With the help of the above bounds, the variance of the signal term can be bounded as

$$\begin{aligned}\frac{\rho_l}{\eta L \cos^2(\delta)} + \frac{\rho_l^2}{K(L-1)} &\leq \text{var} [\tilde{s}_c(i_B)] \\ &\leq \frac{\rho_u}{\eta L \cos^2(\delta)} + \frac{\rho_u^2}{(L-1)}\end{aligned} \quad (50)$$

In addition to the known system parameters, these bounds involve only the total received signal power and the bounds on the noise spectral levels.

The equivalent noise term defined in Eq. (45b) consists of sums of random variables obtained from all K channels. Since the random processes in the various channels are independent, and since the current noise sample is independent of the current weight error, one can express the variance of the total complex noise as

$$\text{var} [\tilde{n}_c(i_B)] = \rho \left[1 + \frac{1}{L-1} + \frac{K}{\eta L \rho \cos^2(\delta)} \right] \quad (51)$$

Note that the variance of the combined noise is always at least as great as the variance of the ideally weighted noise terms, ρ , approaching that limit as L grows without

bound. The excess noise terms can be attributed directly to uncertainty in the weight estimates.

Perhaps the most direct measure of combiner performance is the extent to which the sample SNR of the combined signal approaches its maximum value, ρ , which would be obtained if the complex combining weights could be determined without error. A lower value here implies degraded performance. Analogous to the previous definition in Eq. (23), the sample SNR of the combined signal when the maximum-likelihood weight estimates are used is defined as

$$\rho_{ML} = \frac{|E[\tilde{z}(i_B)]|^2}{\text{var}[\tilde{z}(i_B)]} \quad (52)$$

The time index on ρ_{ML} is ignored since all relevant processes in this model are assumed to be stationary. The magnitude of the expected value of the combiner output is again ρ , while the variance is the sum of the signal and noise variances defined in Eqs. (47) and (51):

$$\begin{aligned} \text{var}[\tilde{z}(i_B)] = & \rho \left[1 + \frac{1}{L-1} + \frac{1}{\eta L \cos^2(\delta)} \left(1 + \frac{K}{\rho} \right) \right. \\ & \left. + \frac{1}{\rho(L-1)} \sum_{k=1}^K \frac{|\tilde{V}_k|^4}{4\sigma_{Bk}^4} \right] \end{aligned} \quad (53)$$

Substitution into Eq. (52) yields

$$\rho_{ML} = \frac{\rho}{\left[1 + \frac{1}{L-1} + \frac{1}{\eta L \cos^2(\delta)} \left(1 + \frac{K}{\rho} \right) + \frac{1}{\rho(L-1)} \sum_{k=1}^K \frac{|\tilde{V}_k|^4}{4\sigma_{Bk}^4} \right]} \quad (54)$$

Note that ρ_{ML} is always less than ρ , approaching that maximum value as L approaches infinity. This is reasonable, since the maximum-likelihood weight estimates approach the true weights as the number of observed samples increases. A direct measure of the loss in sample SNR due to the imperfect weight estimates is the quantity

$$\begin{aligned} \gamma & \triangleq \rho / \rho_{ML} \\ & = \gamma_s + \gamma_n \end{aligned} \quad (55a)$$

with components due to signal, γ_s , and noise, γ_n , defined as

$$\gamma_s = \frac{1}{\eta L \cos^2(\delta)} + \frac{1}{\rho(L-1)} \sum_{k=1}^K \frac{|\tilde{V}_k|^4}{4\sigma_{Bk}^4} \quad (55b)$$

$$\gamma_n = 1 + \frac{1}{L-1} + \frac{K}{\eta L \rho \cos^2(\delta)} \quad (55c)$$

Since these loss factors involve both sums of the individual sample SNRs and sums of their squares, it is useful to invoke the bounds of Eqs. (48) and (49) here. Thus, the component loss factors are bounded as

$$\begin{aligned} \epsilon \gamma_s & \triangleq \frac{1}{\eta L \cos^2(\delta)} + \frac{\rho_\epsilon}{K(L-1)} \leq \gamma_s \\ & \leq \frac{1}{\eta L \cos^2(\delta)} + \frac{\rho_u}{(L-1)} \triangleq \epsilon \gamma_s \end{aligned} \quad (56a)$$

$$\begin{aligned} \epsilon \gamma_n & \triangleq 1 + \frac{K}{\eta L \rho_u \cos^2(\delta)} + \frac{1}{(L-1)} \leq \gamma_n \\ & \leq 1 + \frac{K}{\eta L \rho_\epsilon \cos^2(\delta)} + \frac{1}{(L-1)} \triangleq \epsilon \gamma_n \end{aligned} \quad (56b)$$

while the total loss is loosely bounded by sums of the upper and lower bounds on its components:

$$\epsilon \gamma_s + \epsilon \gamma_n \triangleq \epsilon \gamma \leq \gamma \leq \epsilon \gamma \triangleq \epsilon \gamma_s + \epsilon \gamma_n \quad (56c)$$

Note that tighter bounds on the total combining loss exist, but these involve complicated formulas in different ranges of the parameters, and will not be considered here.

The equal noise-variance case can be treated as the general case considered above, but with considerable simplifications throughout. Although the noise variance is not estimated in this case, a common noise variance denoted by σ_B^2 shall be included in these equations for consistant results. Now the combining weights are given by

$$\hat{w}_k(i_A) = \frac{\hat{X}_k^*(i_A)}{2 \cos(\delta) \sigma_B^2} \quad (57)$$

where the denominator contains σ_B^2 , but not its estimate. The weight estimate can now be expressed as

$$\hat{w}_k(i_A) = \frac{\tilde{V}_k^*}{2\sigma_B^2} [1 + \tilde{e}_{Xk}^*] \quad (58)$$

and it is seen that there is no need to make a small-error approximation here. Therefore, these results will be valid

in general, not only where the small-error approximation holds. The variance of the combined signal and of the combined noise now becomes

$$\text{var}[\tilde{s}_c(i_B)] = \left(\sum_{k=1}^K \frac{|\tilde{V}_k|^2}{2\sigma_B^2} \right) \left(\frac{1}{\eta L \cos^2(\delta)} \right) \triangleq \frac{\rho}{\eta L \cos^2(\delta)} \quad (59a)$$

$$\text{var}[\tilde{n}_c(i_B)] = \rho \left[1 + \frac{K}{\eta L \rho \cos^2(\delta)} \right] \quad (59b)$$

while the variance of the total combined samples is their sum, as before:

$$\begin{aligned} \text{var}[\tilde{z}(i_B)] &= \text{var}[\tilde{s}_c(i_B)] + \text{var}[\tilde{n}_c(i_B)] \\ &= \rho \left[1 + \frac{\left(1 + \frac{K}{\rho}\right)}{\eta L \cos^2(\delta)} \right] \end{aligned} \quad (59c)$$

The combining loss for the equal-variance case also follows directly

$$e\gamma = e\gamma_n + e\gamma_s \quad (60a)$$

with components

$$e\gamma_n = 1 + \frac{K}{\eta L \rho \cos^2(\delta)} \quad (60b)$$

$$e\gamma_s = \frac{1}{\eta L \cos^2(\delta)} \quad (60c)$$

The behavior of the total combining loss and of its components will be examined in the next section.

VII. Numerical Results

The following example serves to illustrate the behavior of the combining loss in a realistic setting. Suppose the 70-m antenna at DSS 14 is receiving signals from a deep-space vehicle at Ka-band, and further suppose that the received power levels and noise temperatures are similar to those expected from the Galileo spacecraft during its encounter with Jupiter, that is, $P_T/N_0 = 55$ dB-Hz $= 3.16 \times 10^5$ Hz. Also assume that a standard NASA modulation format has been employed with $\delta = 80$ deg, subcarrier frequency of $f_{sc} = 2 \times 10^5$ Hz, and assume that we wish to recover the fifth subcarrier harmonic (the second and fourth harmonics are zero for a square wave) so that it is required that $f_B \geq 10^6$ Hz; let $f_B = 2 \times 10^6$

Hz. Assume a seven-element array, $K = 7$. As before, let $T_0 = 2.5 \times 10^{-8}$ sec, corresponding to a first-zero frequency of $f_0 = 1/T_0 = 4 \times 10^7$ Hz. Since the modulated subcarrier must be filtered out of stream A, the first zero of this effective filter should be at some frequency much smaller than the subcarrier fundamental, but large enough to pass the slowly-varying parameters generated by the antenna deformations. This requirement yields $f_A \leq 10^4$ Hz or so, implying that $M_A = f_0/f_A \geq 4 \times 10^3$. Observe that near the above parameter values, the "small-error" approximation upon which the general formulas are based remains valid, provided the inequality $L \gg 1$ is satisfied.

The bounds on the combining loss γ are shown in Fig. 6, for nominal P_T/N_0 of 45, 55, and 65 dB-Hz. Note that the combining-loss bounds, expressed in dB, are plotted on a logarithmic scale to show their relative behavior in regions where each combining-loss bound is very close to zero dB. The maximum and minimum values of P_T/N_0 are taken to be ± 2 dB above and below the nominal values presented in the graphs. Since the total averaging time used for each weight estimate T can be related to the number of observed samples L as $T = M_A T_0 L$, it is clear that the combining-loss bounds can be viewed either as functions of L or as functions of T . Therefore, the loss bounds are plotted in terms of both of these fundamental variables in Fig. 6. Because of the small-error approximation the bounds may not be accurate in the region $L < 10$ (or equivalently, $T < 10^{-3}$ sec), however, this region is of no interest in the present application, where the timescale of interest is typically on the order of a second. Note that a total combining loss of less than 0.1 dB can be guaranteed with observation times of less than 0.1 second, when operating at 55 dB-Hz. This result shows that during a typical encounter, this combining system can operate with negligible combining loss, even in the presence of severe wind gusts which can induce rapidly varying mechanical distortions into the antenna structure.

The components of the combining-loss bounds due to signal and noise are shown in Figs. 7(a) and 7(b), again expressed in dB, at a nominal P_T/N_0 of 65 dB-Hz, with all the other system parameters the same as above. Here the total combining-loss bounds and their components (in dB) are displayed on a linear scale in order to show the signal-loss bounds on the same graph. The loss components displayed in Fig. 7 indicate that for large L the greatest contribution to the total loss comes from the equivalent additive noise. The losses due to the random variations in the signal tend to be relatively insignificant for large L . This behavior is even more pronounced at lower P_T/N_0 , where the individual components of the combining loss are even more difficult to separate. However, the reader is

cautioned that the total combining loss defined here may not be an adequate indicator of performance in all applications. The relative contribution of these loss components to the performance degradation of specific systems may differ markedly and should be carefully examined for each application.

The combining loss for the equal-variance case can be determined exactly, without resorting to small-error approximations. These results are shown in Fig. 8, which confirms the previous conclusion that total combining losses can be limited to 0.1 dB with averaging times of less than 0.1 seconds when operating at 55 dB-Hz. The behavior of the loss components shown in Fig. 9 again indicates that the dominant contribution comes from the additive noise, so that losses due to random signal variations are relatively insignificant in this region.

VIII. Summary and Conclusions

A real-time signal combining system for use with array feeds has been proposed and evaluated. The combining system can be used to compensate for losses resulting from time-varying antenna distortions due to gravitational and atmospheric effects, provided the characteristic timescale of the variations is not too rapid — perhaps on the order of a second or so. It is ideally suited for recovering losses resulting from main-reflector gravity- and wind-induced deformations of large DSN antennas operating at Ka-band or higher frequencies.

Real-time compensation is achieved by forming a weighted combination of the array-feed signals observed in the presence of additive noise with combining weights continuously adjusted to maximize the signal-to-noise ratio of the combined signal. In the digital baseband model considered, the weights were complex numbers obtained from a “maximum-likelihood” weight estimator that obtained the required signal and noise parameters from observations of the array-feed outputs. A “sliding-window” implementa-

tion was proposed that effectively follows the variations in the signal and noise parameters, achieving automatic compensation in the combined output.

A mathematical model of the real-time compensation system was developed. Performance measures were defined and system performance evaluated under a “small-error” approximation, which is valid in the regions of interest. The model assumed constant signal and noise parameters, but remains valid for time-varying parameters, provided the variations are slow compared to the time scale of the observations. It was found that the combining operation produced two distinct effects: it introduced random variations in the combined signal and it increased the effective power of the additive noise. Both of these effects could be attributed to the uncertainty in the weight estimates due to the presence of additive noise in the system. It was shown that under receiving conditions similar to those expected during the Galileo encounter, the total combining loss could be limited to 0.1 dB, with averaging times of less than a hundred milliseconds for a seven-element array. This means that virtually all of the signal power collected by the feed array can be recovered, even in the presence of wind gusts that may produce changes in the signal parameters on the timescale of a second. For larger arrays, the combining loss tends to increase.

Although two different types of degradations have been defined and quantified, their effect on telemetry symbol detection, carrier tracking loop behavior, mechanical pointing-system behavior, etc., has not been completely determined here because the interaction of these components tends to be applications-dependent. These issues should be addressed in the future. Also, mathematical models should be developed to characterize the problem in a dynamic environment and the effect of imperfect phase reference on combiner performance should be addressed. In addition, the possibility of using the entire modulated spectrum for determining the combining weights should be addressed, since this approach will improve combiner performance in the low-SNR or highly dynamic environments.

Acknowledgment

The authors would like to thank Sam Dolinar for several helpful comments and useful discussions.

References

- [1] J. G. Smith, "Ka-Band (32 GHz) Downlink Capability for Deep Space Communications," *TDA Progress Report 42-88*, Jet Propulsion Laboratory, Pasadena, California, pp. 96-103, February 15, 1987.
- [2] W. A. Imbriale, A. M. Bhanji, S. Blank, V. B. Lobb, R. Levy, and S. A. Rocci, "Ka-Band (32 GHz) Performance of 70-Meter Antennas in the Deep Space Network," *TDA Progress Report 42-88*, Jet Propulsion Laboratory, Pasadena, California, pp. 126-134, February 15, 1987.
- [3] S. J. Blank and W. A. Imbriale, "Array Feed Synthesis for Correction of Reflector Distortion and Vernier Beam-steering," *TDA Progress Report 42-86*, Jet Propulsion Laboratory, Pasadena, California, pp. 43-55, August 15, 1986.
- [4] J. H. Yuen, *Deep Space Telecommunications Systems Engineering*, Chapter 5, New York: Plenum Press, 1983.
- [5] C. M. Thomas, *Maximum Likelihood Estimation of Signal-to-Noise Ratio*, Ph.D. dissertation, University of Southern California, 1967.

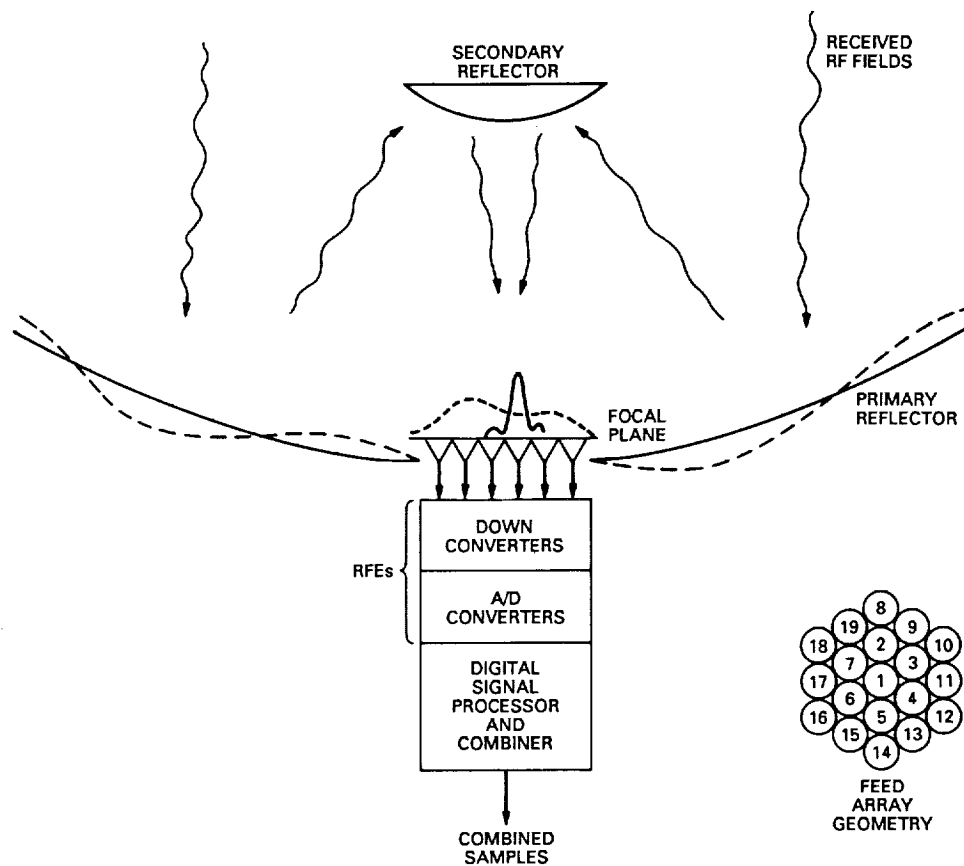


Fig. 1. Real-time antenna-compensation system conceptual design.

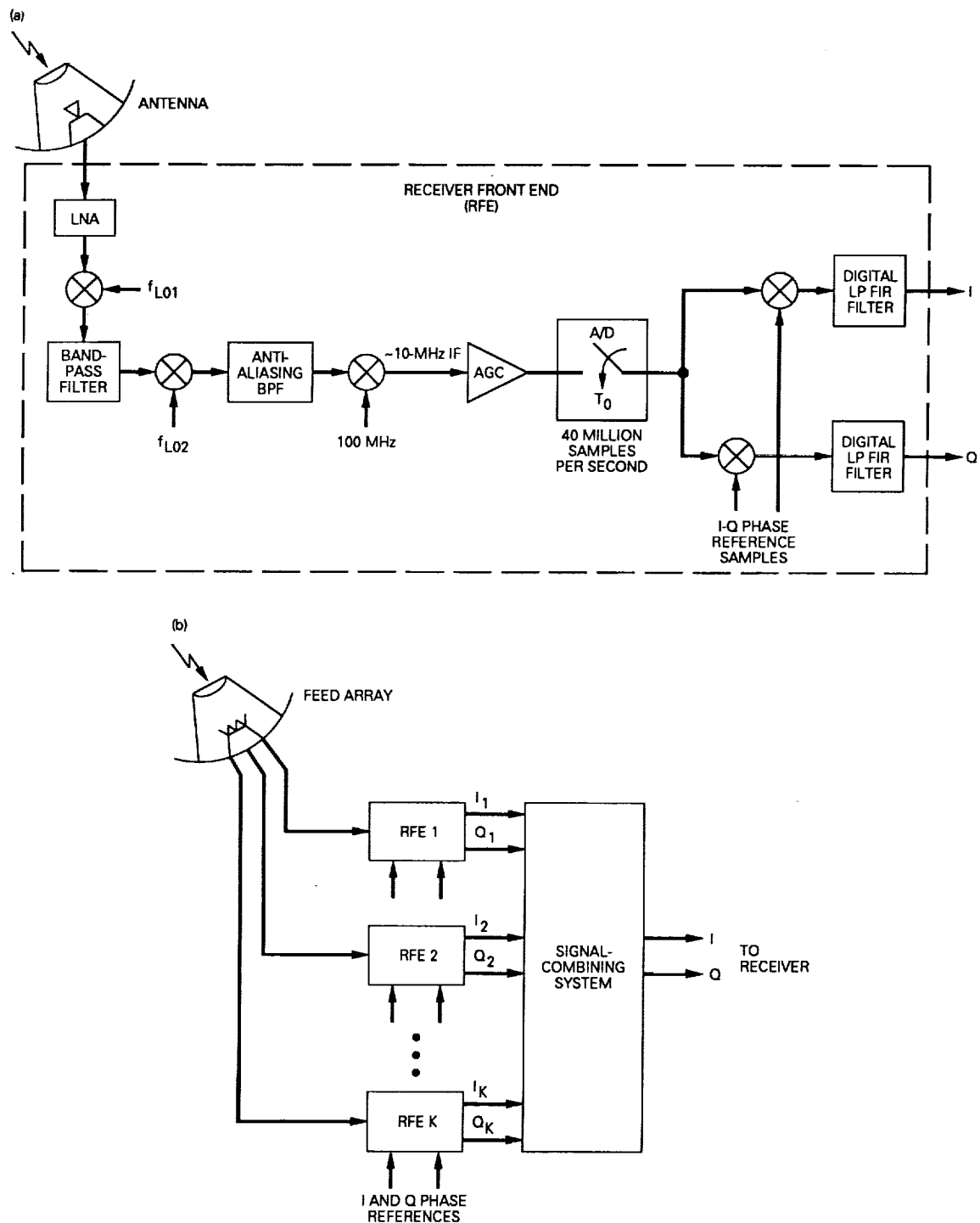


Fig. 2. System block diagrams: (a) proposed DSN Advanced Receiver front end and (b) feed-array signal-combining system.

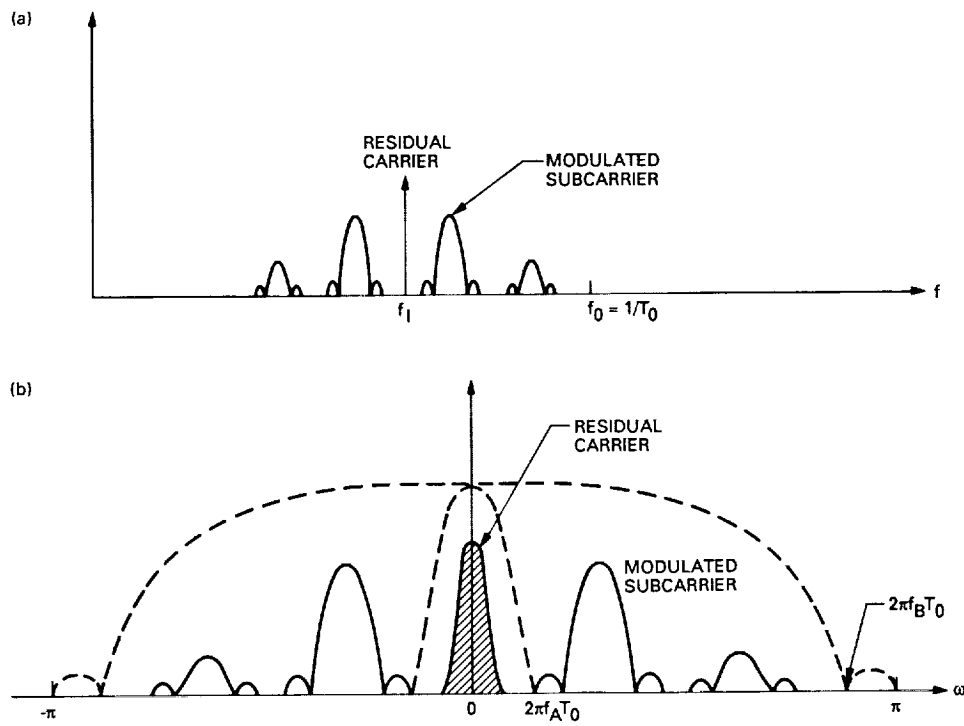


Fig. 3. NASA telemetry spectra: (a) signal spectrum after IF downconversion and (b) spectrum of baseband samples.

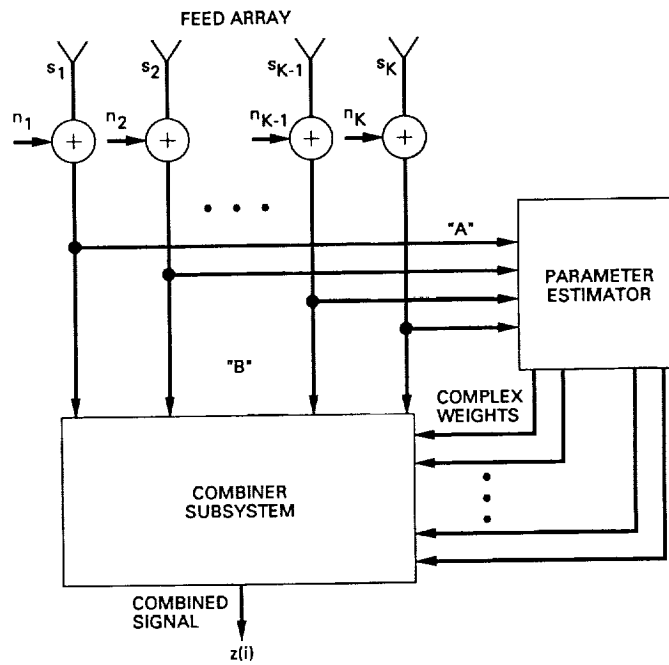


Fig. 4. Signal-combining system (baseband model) block diagram.

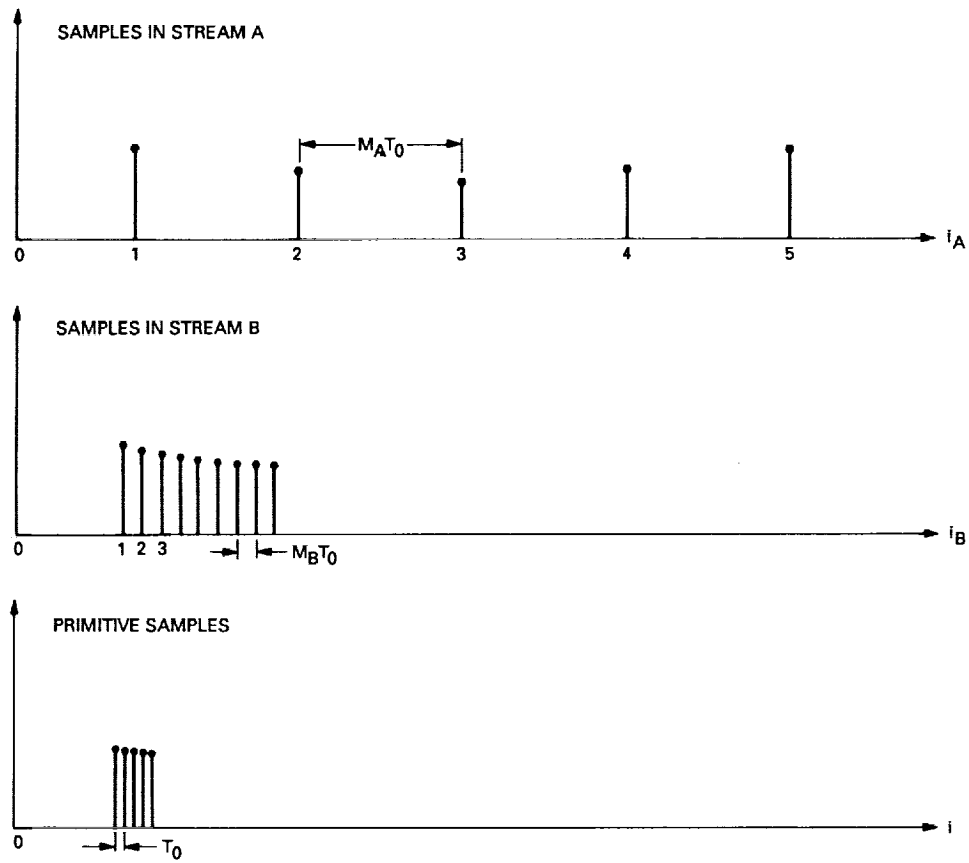


Fig. 5. Timing diagram.

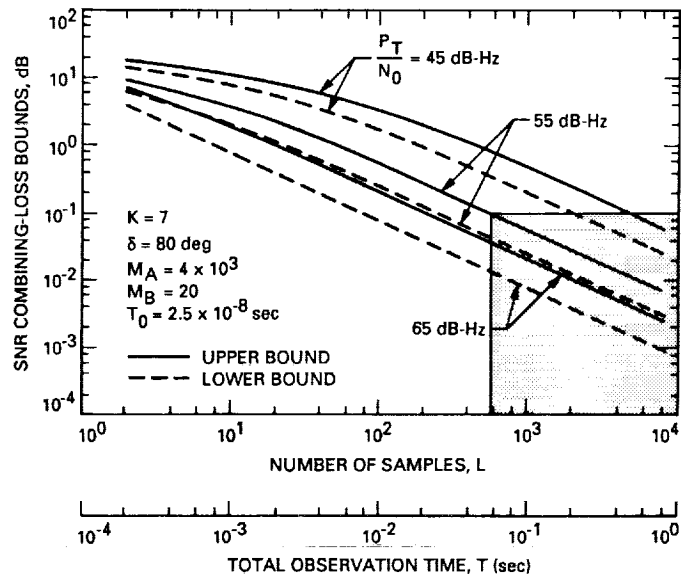


Fig. 6. SNR combining-loss bounds versus L and T .

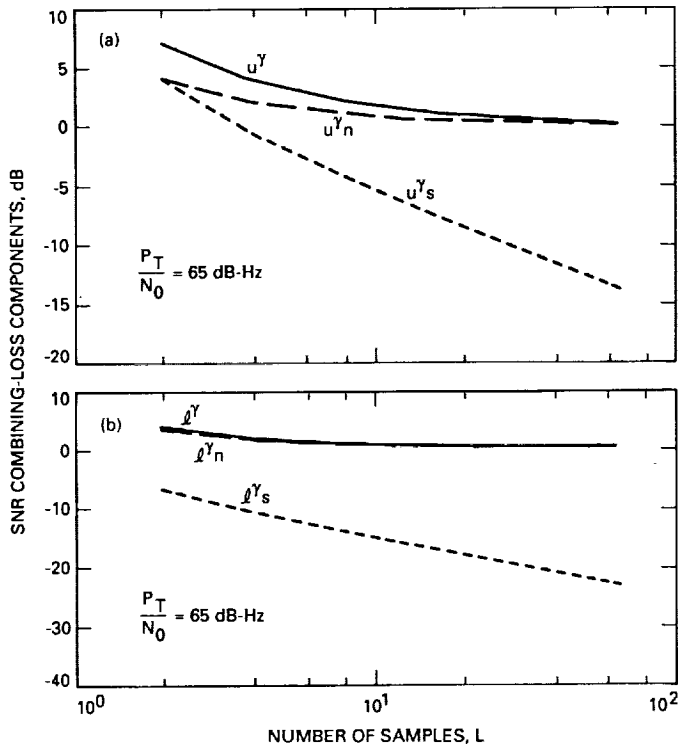


Fig. 7. SNR combining-loss components versus L (unequal-variance case): (a) upper bounds and (b) lower bounds.

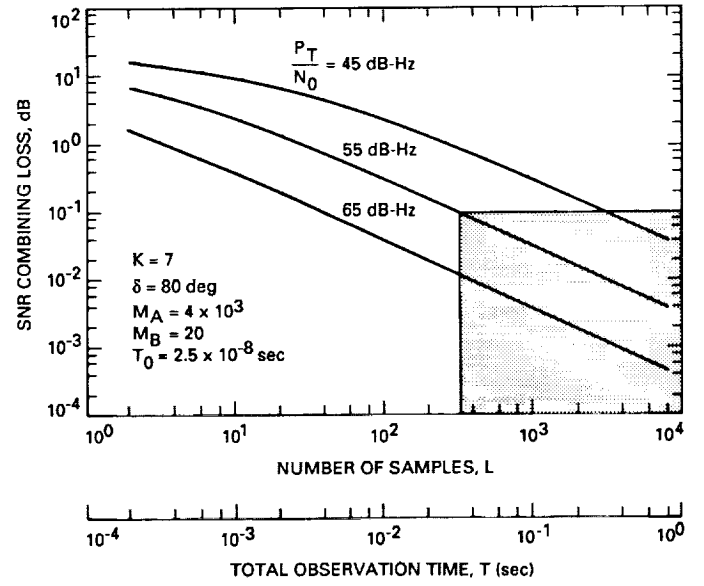


Fig. 8. SNR combining loss versus L and T (equal-variance case).

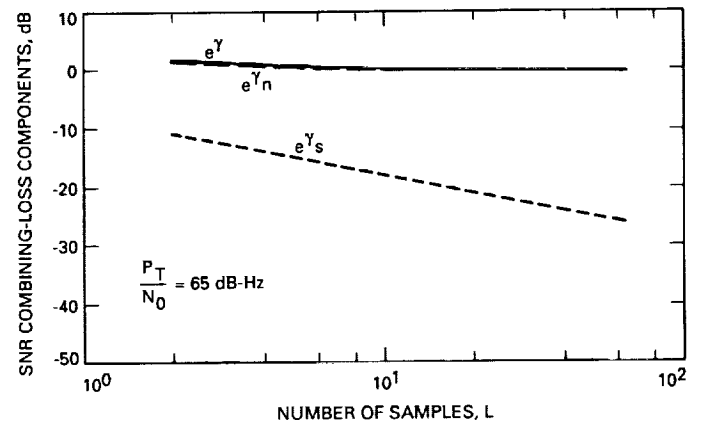


Fig. 9. SNR combining-loss components versus L (equal-variance case).

59-32
272 325

TDA Progress Report 42-100

118.

N90-21889

February 15, 1990

Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System

D. L. Johnson, S. M. Petty, J. J. Kovatch, and G. W. Glass
Radio Frequency and Microwave Subsystems Section

A total system noise temperature of 6.6 K has been demonstrated with an 8.4-GHz traveling wave maser and feedhorn operating in a cryogenic environment. Both the maser and feedhorn were inserted in the helium cryostat, with the maser operating in the 1.6-K liquid bath and the feedhorn cooled in the helium gas, with a temperature gradient along the horn ranging from the liquid bath temperature at its lower end to room temperature at its top. The ruby maser exhibited 43 dB of gain with a bandwidth of 76 MHz (-3 dB) centered at 8400 MHz. Discussions of the maser, cooled feedhorn, and cryostat designs are presented along with a discussion of the noise temperature measurements.

I. Introduction

The DSN currently uses low-noise maser amplifiers operating in 4.5-K closed-cycle refrigerators on the tipping structure of 70-m and 34-m Cassegrain antennas to receive microwave signals from deep space. A measure of the communications downlink performance for these antennas is the ratio of G/T , that is, the gain of the ground-based antenna divided by the system noise temperature. Improvements in G/T at a given frequency can be made by increasing the effective collecting area and efficiency of the antenna or lowering the system noise temperature.

The total system noise temperature of a typical 70-meter antenna with the 8450-MHz Block II-A maser operating in a receive-only mode is about 20 K at zenith on a clear, dry day. Of this 20 K, 4 K can be attributed to the antenna, 1.2 K to the dichroic plate, 6 K to the room-temperature feed components, and 3.8 K (nominal) to the maser, with the remaining noise temperature being attributed to the sky (cosmic background plus atmosphere). Thus, the room temperature feed components collectively are the largest single contributor to the antenna system noise temperature in the present X-band Deep Space Network (DSN) receiving systems.

New 34-m beam waveguide antennas are being implemented in the DSN that will provide, in contrast to present DSN antennas, a relatively large nontipping location for the maser and feed components. For these antennas, the use of a superfluid (below 2 K) liquid helium cryostat becomes very practical as a means to cool a maser and feed components, and achieve a very low system noise temperature. Operation of the maser near 1.6 K instead of 4.5 K dramatically improves its performance by both increasing its gain-bandwidth product and reducing its noise temperature from 3.8 K to about 1.0 K. The feedhorn and other feed components can be cooled as well, reducing their noise-temperature contributions to almost negligible levels. This reduction in noise temperature of the feed components and the maser can improve the total system-noise temperature by as much as 8.7 K and enable increased data rate capability for spacecraft missions such as Galileo.

This article describes a cryogenically cooled 8.4-GHz feedhorn and maser amplifier, enclosed in a 1.6-K superfluid helium dewar, that were developed and tested to demonstrate the above concept.

II. Description

The cryogenically cooled feedhorn-maser assembly described in this article is shown in Fig. 1. A block diagram of the microwave portion of the system is shown in Fig. 2. Major components include the dewar, feedhorn, waveguide directional input coupler, and the maser amplifier. The feedhorn-maser assembly is contained in a liquid-helium-filled dewar with the maser immersed in superfluid liquid helium and the feedhorn cooled by the discharging helium vapors. The 60-liter capacity dewar (International Cryogenics, Inc.) is 104-cm deep and has a 25.4-cm-diameter neck. The feedhorn (hereafter referred to as the horn) is attached by a flange at its aperture to the top plate of the dewar and suspended into the dewar. The maser is rigidly supported with stainless steel tubing attached to the lower end of the horn. A 7.6-cm-high, aluminum extension ring is set on top of the dewar to provide sufficient space under the horn aperture for microwave and electrical feedthroughs and for two helium vapor pumping ports. The maser output and pump waveguides are copper-plated, thin-walled stainless steel in the upper end of the dewar to maintain low radio frequency loss while minimizing heat leak to the helium bath. Mica waveguide windows and a Kapton¹ film horn aperture window were used to seal the dewar from the outside ambient air.

¹ Registered trademark of the DuPont Company.

The temperature of the liquid helium bath is reduced below 4.2 K by evacuating the helium vapors in the dewar. Helium vapor pumping is provided via two pumping ports at the top of the dewar connected to two helium-tight, single-stage Leybold-Hereaus S65B TRIVAC pumps. A combined pumping speed of 43 LPH (92 CFM) evacuates the dewar to a pressure of 800 Pa (6 torr). The helium flow rate was measured at the output of the vacuum pumps with a Teledyne-Hastings-Radist flow meter. A measured flow of 0.030 g/s (0.37 SCFM) corresponds to a 700-mW heat load into the bath.

A test signal port, required for maser tuning and calibration, is coupled to the maser input via a coaxial line (stainless-steel outer conductor/silver-coated beryllium copper inner conductor), a 5-dB attenuator, and a 35-dB waveguide directional coupler. The cooled attenuator and coupler reduce a 300-K ambient noise level at the test signal port to 0.03 K at the maser input. Two Gunn oscillators provide pump power to the maser amplifier.

A. Traveling-Wave Maser

The most economical and expedient means of providing a maser amplifier for this demonstration system was to use an available Block II-A maser structure [1], modifying it as necessary to accommodate operation at a lower physical temperature.

The Block II-A 8450-MHz maser amplifiers implemented in the DSN are installed in the vacuum environment of a closed-cycle helium refrigerator (CCR), and cooled by conduction to 4.5 K. The maser typically produces 40 to 45 dB of net gain over 100 to 115 MHz of bandwidth using 34 cm of total ruby length composed of four individual amplifier channels in cascade. Each amplifier channel, as shown disassembled in Fig. 3, is 8.5-cm long, with coupling probes at each end connected to the next section by short coaxial cables. The complete four-channel amplifier assembly is contained in a superconducting magnet, which provides the desired magnetic field for the ruby maser material and the yttrium-iron-garnet (YIG) isolators [2].

The input noise temperature of a maser amplifier assembly is approximately proportional to the physical temperature, while the electronic gain in decibels (excluding circuit losses) of a maser is approximately inversely proportional to the physical temperature. Therefore, while four maser channels are required at 4.5 K to provide 49 dB of electronic gain and 100 MHz of bandwidth, only one channel was needed to provide 47 dB of electronic gain with 76 MHz of bandwidth at 1.6 K.

The increased gain per unit length of the single maser channel, compared to the Block II-A maser, resulted in the need for increased reverse isolation per unit length. Yttrium-iron-garnet is distributed along the maser slow-wave structure to attenuate the signal propagation in the reverse direction, and prevent regenerative feedback. Our goal was to increase the reverse isolation from the Block II-A maser value of about 38 dB to about 100 dB per channel, while maintaining an adequate low level of forward loss (which adversely affects gain and noise temperature). A new isolator was designed, tested, and optimized in the single-channel, maser-amplifier structure operating at 4.5 K in a closed-cycle refrigerator. This isolator achieved approximately 74-dB reverse loss and 2.0-dB forward loss. The single-channel maser with this new isolator was integrated with the cryogenic feedhorn for system tests in the dewar. These system tests demonstrated that the isolator was marginally adequate at the lowest achievable bath temperature (1.6 K). Maser gain response, shown in Fig. 4, displayed some undesirable ripple (± 1 dB maximum) but was stable and satisfactory for system-noise temperature measurements. A sharp peak occurred in the maser gain at 8500 MHz due to inadequate isolator performance. The bandpass was tuned to a lower center frequency of 8400 MHz, instead of the typical 8450 MHz, to avoid this peak.

Table 1 lists the performance parameters of a typical Block II-A maser (operating at 4.5-K physical temperature), and the single-channel maser developed for this system, operating in a helium bath at 1.6 K.

B. Cooled Dual-Mode Horn

The goals for the cryogenic feedhorn design (as is the case with any cryogenic input transmission line) are to: (1) provide for a tolerable level of heat leak from ambient temperature to the helium bath, (2) provide minimum insertion loss at the signal frequency, and (3) reduce the feedhorn noise temperature contribution by cryogenically cooling as much of the horn as possible.

There are two high-quality X-band feedhorn designs that have been employed in the DSN: a wide bandwidth, corrugated horn that is presently used in the DSN, and a narrower bandwidth, smooth-walled, dual-mode horn. Neither horn design was suitable for cryogenic operation due to the thick aluminum wall construction. The coolable dual-mode horn used in the described system is a modified version of the DSN standard dual-mode horn, with the same overall dimensions, but built in two sections, the larger diameter section made of stainless steel and the smaller diameter section made of copper (Fig. 5).

The smaller diameter, 20.3-cm long, horn section was machined from copper to maintain its temperature isothermally with the bath temperature. This section of the horn accounts for nearly 60 percent of the microwave loss in the feedhorn, and maintaining it near the bath temperature virtually eliminates the noise contribution from it.

The larger diameter, 33.9-cm long, horn segment was machined from a solid billet of 304L stainless steel so that the flanges were an integral part of the horn. This was done to help maintain horn concentricity during and after fabrication. In addition to the top and bottom flanges, nine smaller flanges were machined onto the horn at 2.54-cm intervals. These flanges were used to attach perforated copper plates which serve to transfer heat from the feedhorn to the exiting helium vapor.

The perforated copper plates were 1.59-mm thick, perforated with 4.76-mm-diameter holes for a plate porosity of 33 percent. The outer diameters of the plates were kept within 1 mm of the neck diameter of the dewar to force the pumped vapors through the plate perforations. The operating temperatures of the plates are highly dependent on the helium vapor flow rate through the plates. The additional cooling provided by the vapor flow lowered the temperature along the stainless-steel horn section by up to 85 K, as compared to the calculated temperature profile of the horn with no helium vapor flow. The temperature profiles are shown for comparison in Fig. 6.

To minimize heat conduction from ambient temperature to the helium bath, the stainless steel horn wall was machined to a 1.0-mm thickness. This resulted in an acceptable heat leak of less than 0.25 W to the liquid bath. While the length and thickness of this horn section do not represent optimized dimensions for thermal considerations, they proved to be adequate for this demonstration.

The inside of the stainless horn section was copper plated to an average thickness of 2.3 microns (> 3 skin depths at 8400 MHz), which is sufficient to provide low microwave loss while not significantly adding to the conduction heat load to the bath.

Kapton polyimide film, 0.127-mm thick, was used at the horn aperture to serve as a vacuum window. This window contributes a noise temperature of approximately 0.1 K to the system noise temperature [3].

Low-density foam plugs (1.25-cm thick, dielectric constant of 1.04) were inserted in three places along the inside of the stainless horn section to serve as heat radiation blocks at intermediate temperatures. The top piece was

placed about 0.5 cm below the Kapton window and was just below room temperature, while the two lower plugs were placed about 10 and 20 cm below the Kapton window and were thus maintained at correspondingly lower temperatures. An estimated 0.04-K noise temperature is contributed by the top piece of foam, with smaller contributions from the colder foam plugs. The foam plugs also help prevent helium vapor circulation within the horn, thereby reducing the heat input due to convection.

III. System Tests

The assembled feedhorn-maser-dewar system and associated test equipment were moved to the roof of JPL's Building 238 to perform system-noise temperature measurements observing the sky. Figure 7 shows the experimental setup.

After filling the maser dewar with 4.2-K liquid helium, the maser's superconducting magnet and pump sources were energized and the maser gain response tuned to the desired frequency. Operating at an 8400-MHz center frequency, the gain of the maser at 4.2 K measured 13.5 dB.

The gain of the single-channel maser was measured at several different temperatures as the bath temperature was lowered to 1.6 K. A 7-dB increase in maser gain was measured (from 28 dB to 35 dB at 8400 MHz) upon cooling from 2.2 K to 2.17 K. This effect, shown in Fig. 8, is due to the onset of helium superfluidity at 2.17 K, which, with its very high effective thermal conductivity, insures that the ruby temperature is equal to the bath temperature. These test results suggest that above 2.17 K the ruby operates at a temperature about 0.5 K higher than that of the bath temperature.

The minimum bath temperature achieved was 1.6 K, which was determined by the helium boil-off rate and the size of the helium exhaust pumps. At 1.6 K, the maser exhibited a net gain of 43 dB, with a bandwidth of 76 MHz centered at 8400 MHz (shown in Fig. 4). The total system noise temperature measured 6.6 K with the feedhorn pointed at zenith. Of this 6.6 K, a calculated 2.8-K atmospheric noise contribution plus a 2.5-K "corrected" cosmic background noise temperature (see footnote in Table 2) implies that the input noise temperature of the feedhorn-maser system was 1.3 K (referenced to the feedhorn aperture). Table 2 summarizes these results.

Short-term gain and phase stability measured less than ± 0.02 dB and ± 0.3 deg, respectively, for 10-sec time intervals. During these measurements the temperature stability was better than ± 0.0002 K. No special attempts were made to increase the bath temperature stability.

IV. Conclusions

This demonstration has shown that an antenna feedhorn, waveguide feed components, and low-noise maser amplifier can be combined in a 1.6-K helium-cooled environment to achieve an extremely low input noise temperature—approximately 1.3 K. Use of such a system on a DSN antenna could reduce system noise temperatures by 2.4 dB, from 20 K (current 70-m Cassegrain feed, listen-only, zenith) to 11.5 K (future beam waveguide feed, listen-only, zenith). This system can best be utilized in an antenna environment that is nontipping and spacious enough to contain helium vacuum pumps and helium reliquefaction equipment. Such an environment will be available in the new beam waveguide antennas that the DSN will have in the 1990s.

Acknowledgments

The authors wish to thank the following people who helped with this project: Dan Hoppe for his calculations of the room-temperature horn noise temperature contributions; Mike Britcliffe for his assistance with the horn evaluations and for the use of his HEMT radiometer; Harry Reilly for his help with the Kapton window vacuum tests and valued discussions regarding feedhorns; and Steve Slobin for his calculations of the atmospheric noise temperature contributions.

References

- [1] D. L. Trowbridge, "Block IIA Traveling-Wave Maser," *TDA Progress Report 42-87*, Jet Propulsion Laboratory, Pasadena, California, pp. 158-164, November 15, 1986.
- [2] R. C. Clauss and R. B. Quinn, "Resonant Isolator for Maser Amplifiers," U. S. Patent 4,399,415, Issued 1983.
- [3] R. C. Clauss, H. F. Reilly, Jr., and M. S. Reid, "Low Noise Receivers: Microwave Maser Development," *JPL Space Programs Summary 37-62*, Jet Propulsion Laboratory, Pasadena, California, pp. 74-78, March 31, 1970.
- [4] C. T. Stelzreid, *The Deep Space Network—Noise Temperature Concepts, Measurements, and Performance*, JPL Publication 82-33, Jet Propulsion Laboratory, Pasadena, California, September 15, 1982.

Table 1. Performance of Block II-A maser and liquid-cooled maser compared

	Typical DSN Block II-A Maser	Maser Developed for This Demonstration
Center frequency, MHz	8450	8400
Maser structure temperature, K	4.5 ^a	1.6
No. of 8.5-cm ruby-filled channels	4	1
Electronic gain/unit length, dB/cm	1.4	5.6
Ruby absorption (A), dB	16.2	17
Inversion ratio ($I = E/A$)	3.0	2.8
Forward loss (copper and dielectric) (S), dB	7.5	2.5
Forward loss (due to isolator) Y, dB	1.5	2.5
Total forward loss ($F = S + Y$), dB	9	5.0
Total isolator reverse loss, dB	150	74
Calculated maser input noise temperature	3.4	1.0
Net gain ($G = E - F$), dB	40	43
Bandwidth (-3 dB) MHz	107	76

^aRuby bars cooled by conduction in a vacuum. Therefore, actual ruby temperature is assumed to be 5.0 K.

Table 2. Calculated feedhorn/maser noise temperatures

Center frequency	8400 MHz
Measured system noise temperature (A)	6.6 K
Measured atmospheric conditions (ground level):	
Altitude	0.4 km
Temperature	27.8 C
Pressure	0.0968 MPa
Relative humidity	30%
Calculated atmospheric noise contributions (from above data):	
Water vapor (B)	0.7 K
Oxygen (C)	2.1 K
Cosmic background (D)	2.5 K ^a
Resultant feedhorn/maser input noise temperature: (A-B-C-D)	1.3 K
Estimated component contributions to the above feedhorn/maser Input noise temperature:	
Feedhorn/window	0.15 K
Waveguide coupler, transitions, coax, VSWR loss	0.10 K
Test port noise leakage	0.03 K
Maser input noise temperature	1.00 K
Post amp noise contribution	0.02 K

^aThe "corrected" cosmic background noise temperature is 2.5 K (generally given as 2.7 K) as noted by Stelzreid in [4], which enables its use in simplified noise calculations that use the approximation for thermal noise $P_N = kTBG$.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

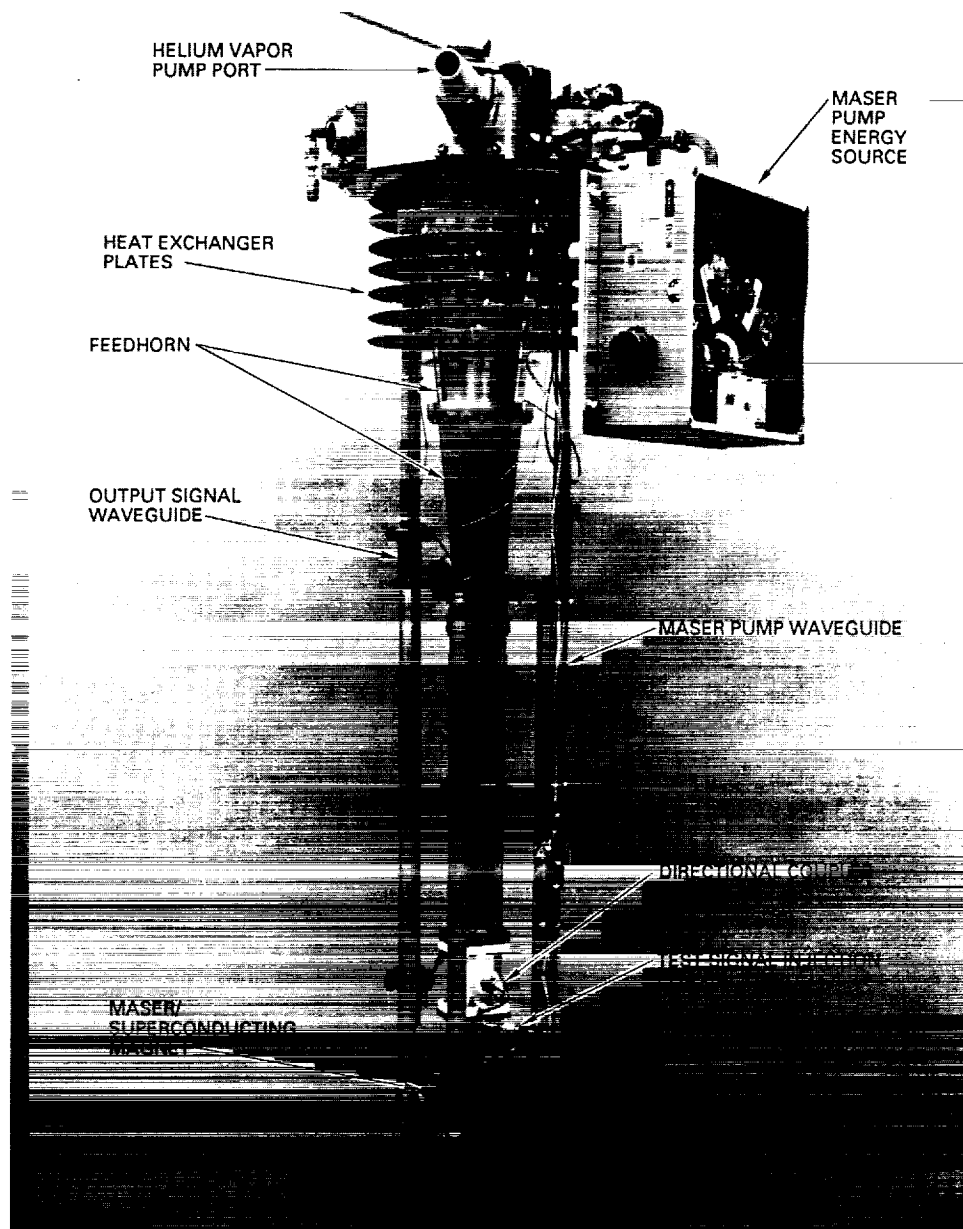


Fig. 1. Cryogenic feedhorn-maser assembly (shown removed from liquid helium dewar).

ORIGINAL PAGE IS
OF POOR QUALITY

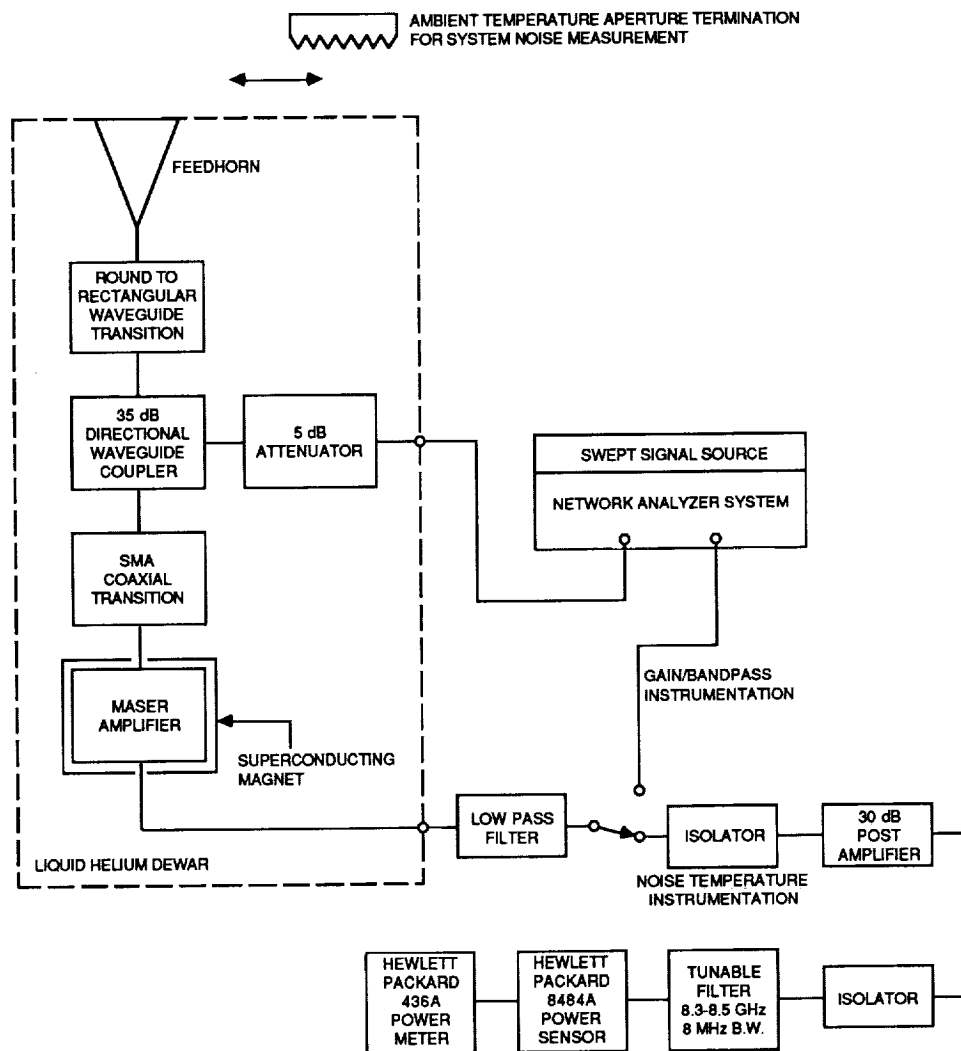


Fig. 2. Block diagram of cryogenic feedhorn-maser assembly and associated microwave instrumentation.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

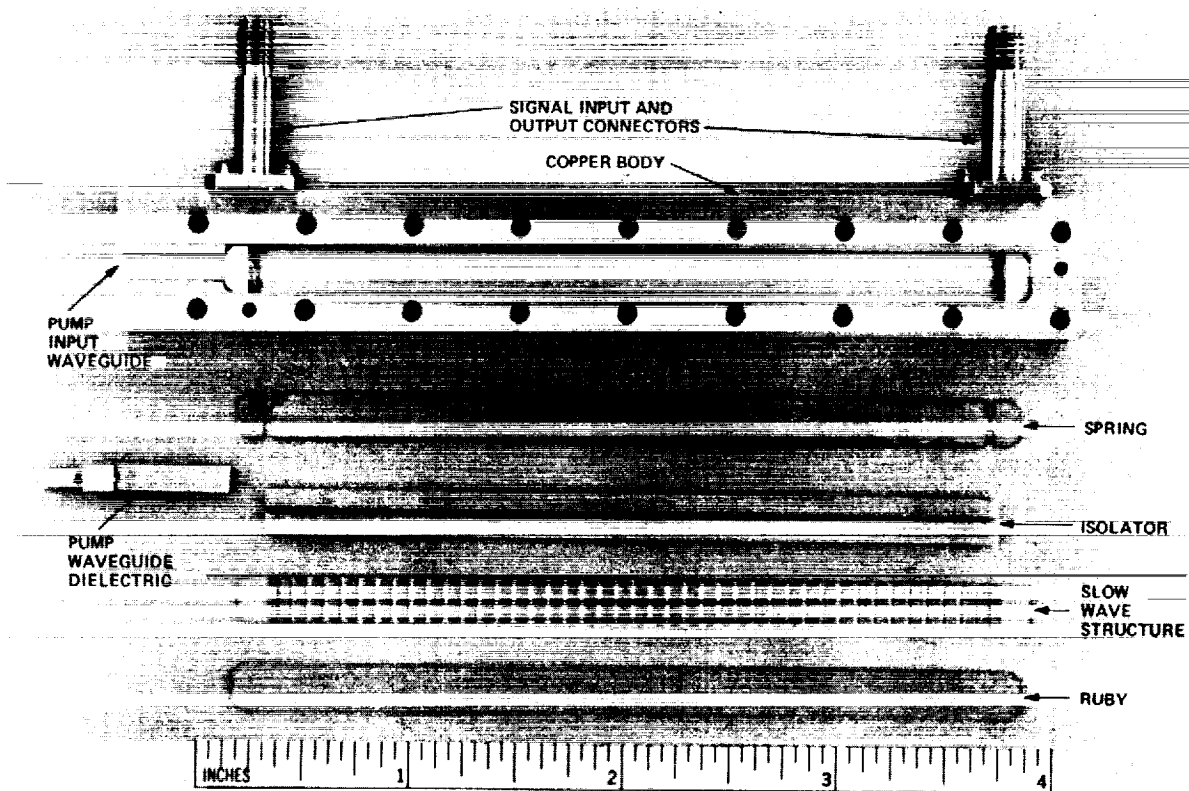


Fig. 3. 8.4-GHz traveling-wave maser structure (one of four channels).

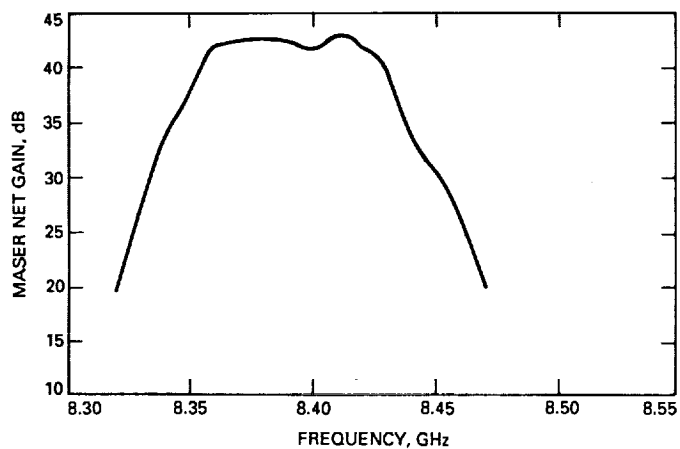


Fig. 4. Maser gain-bandpass at 1.64 K.

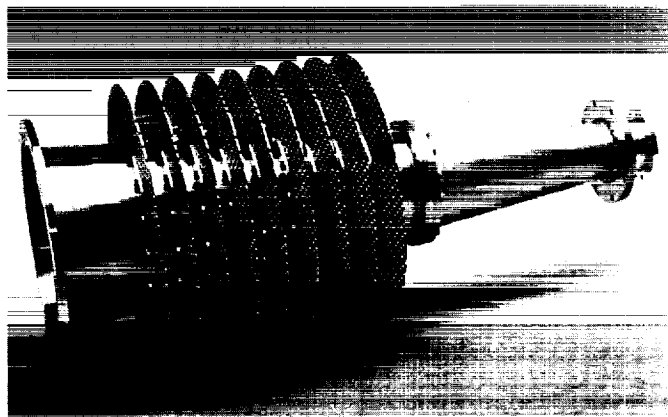


Fig. 5. Coolable dual-mode feedhorn.

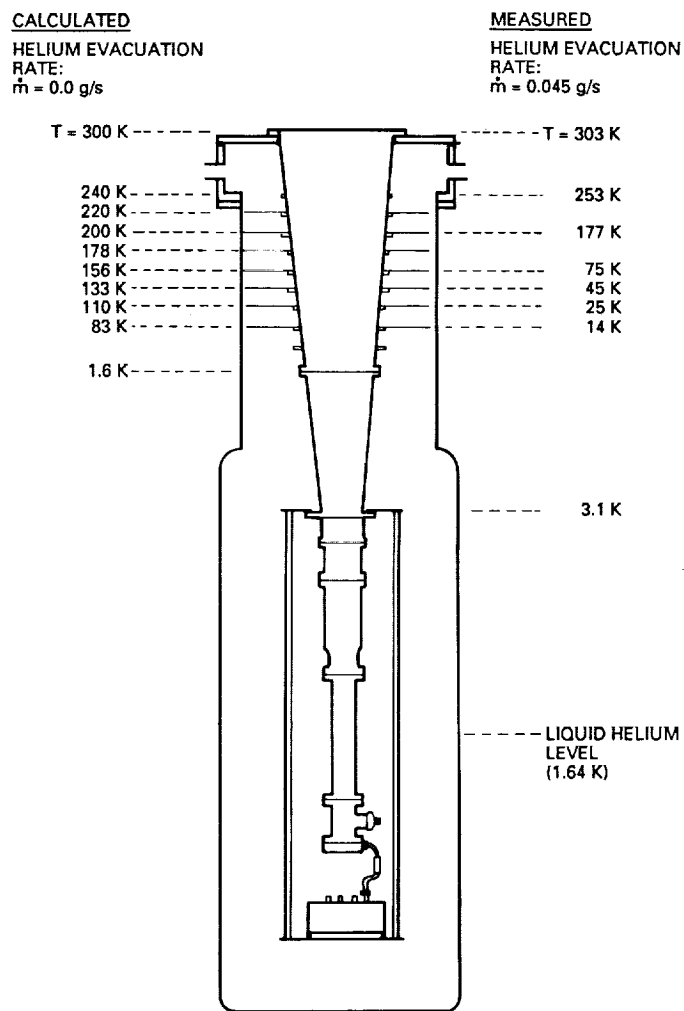


Fig. 6. Horn temperature profiles. The calculated values on the left represent the thermal gradient on the horn due to horn thermal conductivity only. The measured values on the right show the thermal gradient due to heat exchange with the exiting cold helium vapors.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

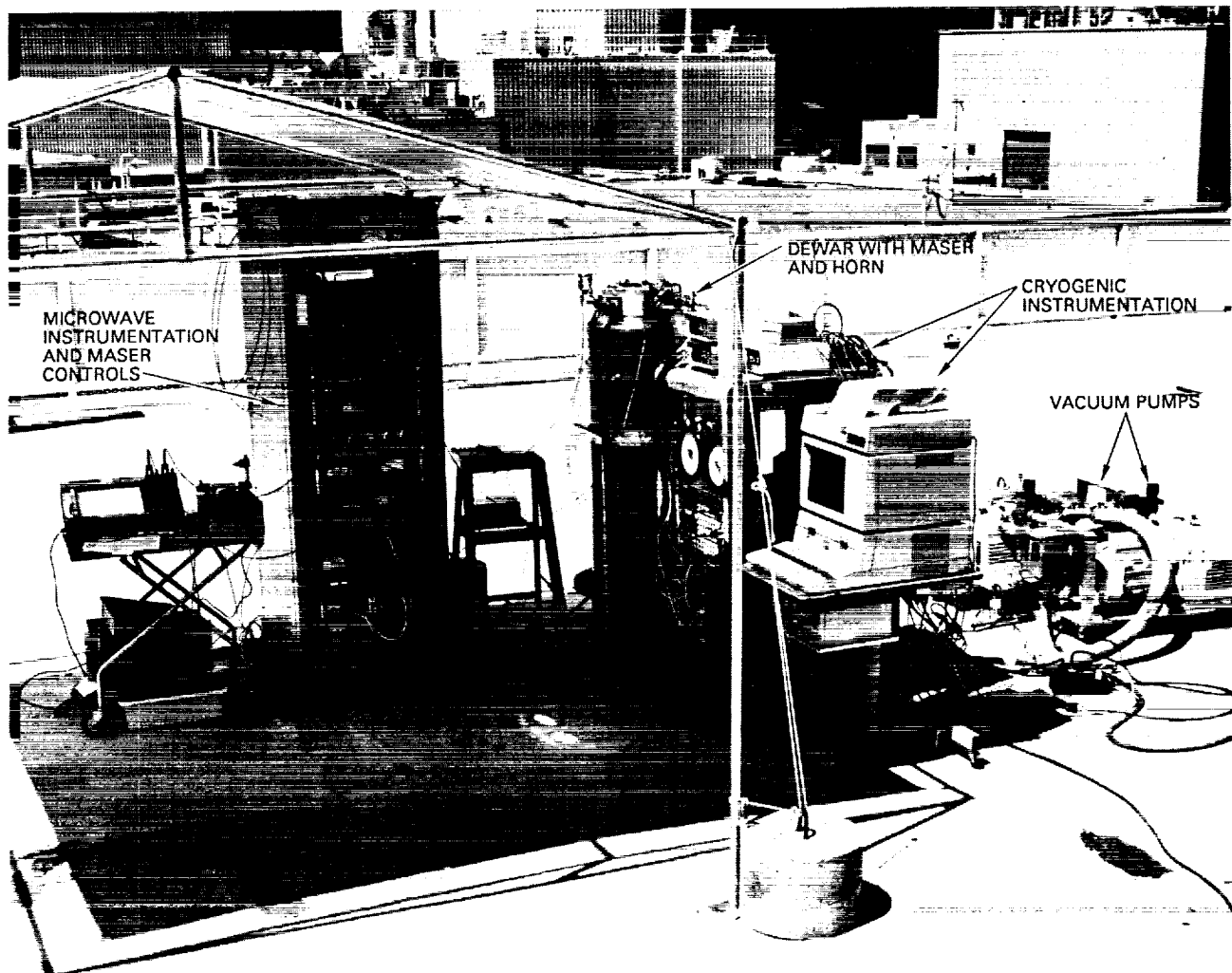


Fig. 7. Rooftop system test setup.

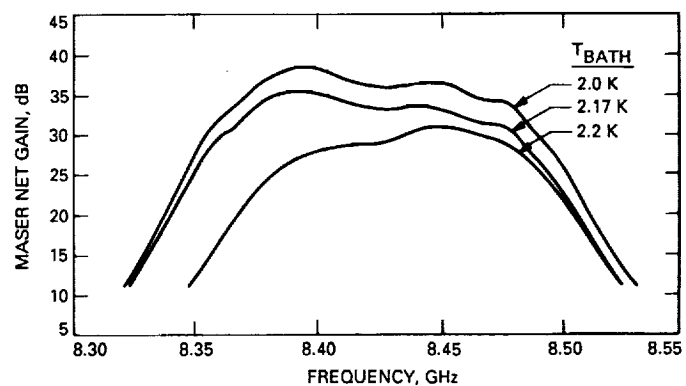


Fig. 8. Maser gain-bandpass as a function of temperature about the helium lambda point.

Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels

T. Y. Otoshi and M. M. Franco
Ground Antenna and Facilities Engineering Section

The results of X-band noise temperature tests on two types of antenna surface panels are presented. The first type tested was a solid antenna panel, while the second type was a perforated panel with 3/16-in.-diameter holes. Measurements were made at 8.45 GHz using an X-band radiometric system.

Included in this article are measured noise temperature contributions from (1) thermal diffusive white paint on solid and perforated panels and (2) water sprayed on both painted and unpainted perforated panels. For this article, experiments on perforated panels were restricted to the 3/16-in.-diameter hole panels formerly used on Deep Space Network 64-m antennas. Rigorous calibration equations, applicable to a variety of antenna panel and dichroic plate test configurations, are presented. It has been demonstrated that an accurate, stable radiometric measurement system of the type used for the results of this article makes it possible to obtain information that would be much more difficult to obtain using other techniques.

I. Introduction

This article presents results of noise temperature measurements made at 8.45 GHz on (1) painted and unpainted, solid-aluminum sheet reflectors and (2) painted and unpainted, perforated panels having 3/16-in.-diameter holes. The results of recent noise temperature tests on a painted 1/8-in.-hole panel were reported by R. Stevens

and R. Clauss.¹ Information concerning the dielectric constant and loss tangent of thermal diffusive white paint has been very difficult to obtain. To the authors' knowledge,

¹ R. Stevens and R. Clauss, "DSN Rain Effects Test Results and Recommendations," JPL Interoffice Memorandum RCC-89-019 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 26, 1989.

the only other sources of experimental data on thermal diffusive white paint used on Deep Space Network (DSN) antennas are (1) a 1961 Dalmo-Victor report from which paint data were extracted and used in a 1971 report by Otschi [1], and (2) an internal JPL report by A. Freiley,² who used an X-band cavity technique to obtain surface resistivity data.

Other radiometric measurements at 8.45 GHz on perforated panels have been made by S. Slobin and M. Franco in 1984.³ They performed measurements on a number of *unpainted* perforated panels having different hole diameters. The panels were tested in both wet and dry conditions using a test setup and test procedures that were almost identical to those used to obtain the results of this article. The primary difference in the two test setups was that the Slobin/Franco setup used an X-band maser, while the setup for this article used an X-band high-electron-mobility transistor (HEMT).

In Section II the test method will be described. Rigorous calibration equations, applicable to a general class of panel and dichroic plate testing configurations, are also presented. Then radiometric test data will be presented and discussed. The data presented in this article will be useful for assessing whether or not paint on antenna panels is a problem at X-band.

II. Test Procedure

A. Measurement System Description

Noise temperature measurements were made with the radiometric system located on the roof of the Telecommunications building at the Jet Propulsion Laboratory (Building 238). In this test setup the X-band horn used is a 22-dB corrugated horn, and the low-noise amplifier is a cryogenically cooled X-band HEMT. The HEMT is followed by about 30 ft of phase-stable X-band coaxial cable, a post-amplifier, and a preselect filter which is terminated by a Hewlett Packard (HP) Model 436A power meter sensor. The preselector filter is tuned to a center frequency of 8.45 GHz \pm 5 MHz. The test setup for this system was originally developed for purposes of measuring noise temperature increases due to rain on a DSN X-band dichroic

plate.⁴ When the horn is pointed at zenith sky, the operating noise temperature of the test setup at 8.45 GHz consists of the following noise temperature contributions:⁵

T_{cbr}	=	2.7 K
T_{atm}	=	3.3 K (clear sky, low humidity)
T_{horn}	=	1.5 K
T_{wg}	=	3.0 K
T_{hemt}	=	12.0 K
T_{fup}	=	4.0 K
T_{op}	=	<hr/> 26.5 K

Above, T denotes noise temperature while the subscripts *cbr*, *atm*, *horn*, *wg*, *hemt*, *fup*, and *op*, respectively denote cosmic background radiation, atmosphere, horn, waveguide, HEMT, followup receiver, and operating. In this article, the terms operating noise temperature and system temperature will be used interchangeably.

Calibration of this measurement system is done with a Y-factor method described in [2] and performed automatically by computer control. Updated antenna operating noise temperatures are displayed on a monitor every 5 seconds. Details of the computer-controlled system are presented in Appendix A.

It was originally thought that the polarizer for the test setup was set for circular polarization, but was inadvertently set for elliptical polarization. The incorrect setting caused the measured results of the panel losses to be slightly higher than they would be for a circular polarization configuration. The equations that apply to the elliptical polarization configuration actually used are given in Appendix B.

B. Panel Measurement

The test configurations for the panel tests are shown in Fig. 1. All panels tested were 36 in. \times 36 in. \times 0.072 in. and located 36 in. from the horn aperture and held at an angle of 45 deg so that, with the horn horizontal to the ground, the angle of incidence is 45 deg for a plane wave signal radiated from zenith sky (Fig. 1). Measurements of operating noise temperatures were made at 8.45 GHz first with a solid aluminum reference plate held at 45 deg with respect to the horizon. Then the panel under test was placed at the same angle and the operating noise temperature was again measured. If all other noise temperatures remained constant during the test, subtraction of the two

² A. Freiley, "RF Evaluation of Paint Samples," JPL Interoffice Memorandum 3331-73-008 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 9, 1973.

³ S. D. Slobin and M. M. Franco, "X- and Ka-band Noise Temperature Effects from Wet Perforated Aluminum Panels Simulating Rain on DSN Antenna Surfaces," JPL Interoffice Memorandum 3331-84-003 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 10, 1984.

⁴ Stevens and Clauss, *op. cit.*

⁵ M. Britcliffe, private communication.

measured operating noise temperatures will give the noise temperature difference between the panel under test and the reference plate.

To minimize measurement errors due to drift caused by changes in sky temperature versus time, the procedure used (to obtain the results for this article) was to take as many reference measurements as possible. A new reference measurement was taken after making measurements on two or three different test configurations.

The perforated panels were tested for the following three different "behind the panel" configurations: (1) horizon sky, (2) absorber, and (3) solid-aluminum plate.

For the perforated panel tests, the first test configuration (Fig. 1a) was to clamp the perforated panel to a wooden frame tilted at 45 deg. No additional reflecting plates or absorbers are placed behind the perforated panel. This test configuration is called the "horizon sky" behind-the-panel configuration because the signal that leaks through the panel will be absorbed primarily by the sky near the horizon. This configuration is convenient to use for quickly observing noise temperature differences. However, this setup is more susceptible to errors caused by changes in multiple reflections between fences, buildings, and other objects in close proximity to the test setup.

The second perforated panel test configuration was to place an absorber in back of the perforated panel to capture the leakage signal with an absorber at ambient temperature (see Fig. 1b). The difference between operating noise temperatures of this second configuration and that from the reference measurement provides information on leakage plus resistive panel losses of the perforated panel. For more details, refer to the equations presented in Appendix C.

The third configuration for testing perforated panels was to manually hold a solid-aluminum panel in back of the perforated panel at 45 deg with respect to level ground (see Fig. 1c). In this configuration, the signal that leaks through the perforated panel becomes reflected by the solid plate and becomes absorbed by zenith sky. As discussed in Appendix C, the difference between the operating noise temperature of this configuration and the reference measurement gives information on the resistive losses of the perforated panel.

Tests of wet test panels were made by spraying a fine mist of water on the panel. Due to the test panels being held at a 45-deg angle, most of the water ran off. Only

some fine droplets and those retained due to surface tension over the holes stayed on the panel.

Rigorous calibration equations for the three different test configurations are presented in Appendix C. These equations are useful for understanding the calibrations procedure, the contributions from perforated plate leakage, and potential errors caused by unwanted spillover. It should be pointed out that if one is interested in determining specific dissipative loss contributions, it is not just a simple matter of subtracting two measured operating noise temperatures. It becomes necessary to examine all noise temperature components (associated with reflection, leakage, and dissipation) causing the operating noise temperature values to change. It is shown in Appendix C that if two nearly similar panels are being compared, so that reflected and leakage components are nearly the same for the two panels, then subtraction of the two operating noise temperatures will produce the desired dissipative loss noise temperature relative to that of the other plate being compared.

To experimentally determine how much power was actually spilling over the edges of the 36-in. \times 36-in. test panel, a 24-in. \times 24-in. absorber was moved around the perimeter of the test panel. An increased contribution of about 2.0 K was measured when the absorber was placed above the top edge of the panel. This measured increase is consistent with a theoretically calculated value obtained with theoretical near-field pattern and brightness temperature data. The near-field horn patterns were obtained for the X-band horn located 36 in. away from a 36-in. \times 36-in. flat plate tilted at 45 deg. A separate report on this theoretical and experimental result is in preparation. As shown by the equations in Appendix C, the spillover loss does not affect the final test panel noise temperature values reported in this article, as long as the spillover-power to total-power ratio remains constant during measurement of the test plate and reference plate.

III. Experimental Results

The theoretical noise temperatures for an aluminum 6061 T6 solid panel (assuming a conductivity of 2.32×10^7 mhos/m) are 0.083 K and 0.165 K, respectively, for perpendicular and parallel polarizations for a 45-deg incidence angle at a frequency of 8.45 GHz. Then, for the elliptical polarization of the test setup (Appendix A), the theoretical effective noise temperature contribution due to this aluminum reference plate is calculated to be 0.153 K. If absolute values are of interest, the above calculated value can be applied to measurements that were made relative

to this reference plate. For this article, no corrections were made for the loss of the reference plate.

The measurements were started at 1:22 P.M. and completed at 2:04 P.M. on June 2, 1989. The total test time was 42 minutes and involved 29 different test configurations. In this test setup, the operating noise temperature value was displayed on a screen every 5 seconds. After a test configuration change was made, the values applicable to the test configuration of interest were written in a notebook. Table 1 shows the original recorded data, while Table 2 shows a tabulation of the reduced data. Table 1 may be useful as a guide for future test planning purposes. Table 2 shows the procedure for reducing the data from Table 1 for extracting particular noise temperature contributions of interest.

The information that can be extracted from the original data given in Table 1 are noise temperature increases due to (1) paint on a solid reflector, (2) paint on a perforated panel with 3/16-in.-diameter holes, (3) leakage through the holes of an unpainted perforated panel, (4) leakage through the holes of a painted perforated panel, (5) water spray on painted and unpainted perforated panels, (6) water spray on a painted solid panel, and (7) aluminum tape covering the holes of the unpainted perforated panel.

All test results apply only for test parameters with a 45-deg incidence angle, 8.45 GHz, and the elliptical polarization configuration described in Appendix A. Both the unpainted and painted perforated panels tested had 3/16-in.-diameter holes, 1/4-in. hole-to-hole spacings, equilateral triangle hole patterns, and a plate thickness of 0.072 in. The total thickness of the primer and Triangle Co. IR #6 thermal diffusive paint on the painted panels was measured with a micrometer and found to be about 0.002 in.

Although more data of interest exist in Table 2, an attempt will be made to summarize most of the highlights below:

(1) The increase of noise temperature due to paint on a solid-aluminum reflector was measured to be 0.1 K at 8.45 GHz for a 45-deg incidence angle. The repeatability of this measurement was ± 0.1 K. See Table 2, item B1.

(2) The increase of noise temperature due to resistive losses for the unpainted perforated panel was 0.6 K. See Table 2, Item B2(c).

(3) The noise temperature component associated with perforated plate leakage was determined to be 0.8 K. See Table 2, Item B2(d). This measured value compares fa-

vorably with a theoretical value of 0.64 K calculated from Chen's Program [3] for the described perforated panel when the incidence angle is 45 deg and an elliptically polarized incident wave with 85.4 percent of the total power in the parallel polarization component and 14.6 percent in the perpendicular polarization component.

(4) The increase of noise temperature due to paint on the perforated panels with 3/16-in.-diameter holes was measured to be 0.2 K. See Table 2, Item B4(b). It is of interest to compare this value to the 0.1 K value that was measured for the paint contribution on a solid-aluminum plate (Table 2, Item B1).

(5) The increase of noise temperature due to water spray on an unpainted perforated panel with 3/16-in.-diameter holes was measured to be 11.6 K. See Table 2, Item B6(b). The measured value is about three times smaller than a theoretical value calculated for this panel when the holes are completely filled with water. During the measurement, it could be seen that the water layer only covered some of the holes near the panel surface and did not penetrate deeply into the holes. Most of the water began to drain off due to gravity, but water in some holes was retained because of surface tension.

(6) The increase of noise temperature due to water spray on a painted perforated panel with 3/16-in.-diameter holes was measured to be 7.3 K. See Table 2, Item B8(b). From comparisons with the 11.6 K value obtained for the unpainted perforated panel, it can be concluded that less water is retained by the painted panel than the unpainted panel.

(7) The increase of noise temperature due to water on a painted solid plate was measured to be 1.2 K. See Table 2, Item B11. Due to an oversight, no measurements were made of the noise temperature increase due to water sprayed onto the reference plate (the unpainted solid-aluminum plate).

IV. Discussion of Errors

It should be mentioned that the reported noise temperature contributions due to paint and water are larger than will be the case for a circularly polarized wave. The reason for this is that an unintentional error in the setting of the radiometric system polarizer caused about 85 percent of the power to be in the parallel polarization component and only 15 percent in the perpendicular polarization component. For a circularly polarized wave, the power is equally divided between the two polarization components. For the type of tests that were done on the two types of antenna reflector surface panels, a theoretical study shows that at a 45-deg incidence angle, the noise temperatures

for parallel polarization are significantly larger than those for perpendicular polarization.

The differential measurements involving paint contributions on the solid plate are believed to be accurate to within ± 0.1 K. Errors due to drift (changing sky conditions) were small for the solid-plate measurements because the measurements were performed sequentially. For the perforated panel tests, there are several possible sources of errors in the reported data. Among these are uncorrectable operating system drifts due to changing sky conditions, an insufficiently large absorber for the absorber behind-the-perforated-panel tests, and an insufficiently large solid plate for the behind-the-panel test configuration. Water spray test data were difficult to repeat due to drainage and the uncontrollable amounts of water sprayed.

For a very accurate measurement of the resistive loss of a perforated panel, the solid plate in back of the perforated panel must be large enough to capture most (90 percent) of the leakage wave and be held at 45 deg, within ± 5 deg. However, for the test setup used, the solid plate was only 36-in. \times 36-in. and was held manually at an angle such that the solid plate was only approximately parallel to the perforated panel. The solid plate was only about 24 in. from the back side of the perforated panel (Fig. 1c). This distance was not large enough to keep multiple reflections from occurring between the solid panel and the back side of the perforated panel, but this fact was not known at the time of the measurements.

It is important to try to salvage the measured data concerning the resistive loss of the perforated plate because good experimental information concerning the resistive losses of perforated plates is very difficult to obtain. Resistive loss contributions in the past have been thought to be negligibly small and, therefore, seldom considered. For the perforated panel that was tested, the measured noise temperature contribution due to surface resistivity was 0.6 K with associated errors of about ± 0.2 K. This result indicates that the resistive loss of a perforated panel is not as negligibly small as was previously assumed.

Slobin and Franco⁶ have published experimental data that might be useful for determining the resistive losses of a number of perforated panels having different hole diameters. They used a larger (3 ft \times 6 ft) solid-aluminum panel in back of the perforated panels at a sufficiently large distance so that the leakage signal did not undergo multiple bounces between the back side of the perforated panel and the solid plate. A study of the experimental data pre-

sented in Slobin and Franco's report supports the findings that the noise temperature contributions due to resistive losses of unpainted perforated panels are much larger than previously believed. Theoretical values for resistive losses of perforated plates were calculated in the past through the use of a simple formula derived by Otoshi [4]. The previous formula now appears to be incorrect. As a result of recent experimental discoveries, a new formula has been derived for calculating perforated panel resistive losses. The validity of this new formula is currently being investigated and will be reported on separately.

V. Recommendations for Future Tests

The procedure to minimize errors due to sky noise temperature changes was to make as many reference measurements as might be practical. Then, differences were taken between the operating noise temperature of the configuration under test and the operating noise temperature of the nearest (in time) reference plate measurement. An improved procedure would have been to record the test times (hr, min) for each configuration, including the reference readings. Then, interpolation between reference plate readings could have been used to obtain a reference value that corresponded to the mid-time that measurements were made for the test configuration of interest. In this manner, the errors due to drift could have been reduced to a negligibly small value.

Figure 2 shows the magnitude of the drift relative to the operating noise temperature measured for all test configurations. The reference measurement drift is shown more clearly on a magnified scale in Fig. 3. The total drift in the reference operating noise temperature was 0.5 K in about 40 minutes of time. Had the actual test times of the configurations been recorded, then correction could have been made for the drift. Many reference readings were taken because thin clouds occasionally drifted over the main beam of the horn. For interest, the data points for the painted solid-plate measurements are also shown in Fig. 3. It can be seen that measurements on the painted solid-plate were done immediately after measurements on the reference plate. Therefore, drift error was minimal for determining the noise contributions of paint on the solid plate.

It should be mentioned that if one is interested in a particular noise temperature contribution such as paint on perforated panels, better test procedures than the ones described in this article could be developed. For example, if it had been the primary objective to determine the noise temperature contributions due to paint on a perfo-

⁶ Slobin and Franco, op. cit.

rated panel, then measurements of unpainted and painted perforated panels should be made sequentially rather than being taken relative to the reference plate.

A future improvement to the current computer-aided measurement system might be to have a computer program where the experimenter can type in a data file name, test configuration title, and the number of data points to be recorded. The computer then opens the specified data file and records the information into the data file and provides a prompt for start of the test. After the return key is hit, the specified number of operating noise temperature data points is measured. The computer averages the data, computes a standard deviation, records the computed data, records the mid-test time, and then closes the file. This feature would free the experimenter from having to rapidly write down numbers and avoid possible transcribing errors.

From the experience gained from these tests, it was found beneficial to have two people involved in performing the tests. One person is needed to devote full time to making configuration changes and preparing for the next change, while the second person is needed to write down the measured values. Speed is of the essence if the sky noise temperature is changing significantly such as would be caused by clouds drifting overhead.

VI. Concluding Remarks

A large amount of very useful data concerning paint contributions on solid and perforated panels have been obtained. In addition, new information was obtained on the resistive losses of perforated plates at X-band. The fact that reliable and repeatable data were obtained on 29 configurations (see Table 1) in about 40 minutes time is a tribute to the general usefulness of the noise temperature measurement system for performing these types of tests.

It has been demonstrated that an accurate, stable radiometric measurement system of the type used for the results of this article makes it possible to obtain information that would be much more difficult to obtain using other techniques. For example, the contributions of paint and resistive losses of perforated panels were easy to measure with the radiometric system that was used. A waveguide technique to determine perforated plate resistive losses requires ultraprecision and specially imaged waveguide samples [4]. Resistive loss on perforated panels cannot be determined with the cavity technique⁷ because the loss due to leakage through the perforated holes cannot be isolated from the loss due to surface resistivity.

⁷ Freiley, op. cit.

Acknowledgments

M. Britcliffe of the Radio Frequency and Microwave Subsystems Section (Section 333) developed the radiometric measurement system described in this article and gave permission to use the test setup for making the panel measurements. All panels for the noise temperature tests were furnished by V. Lobb of the Ground Antenna and Facilities Engineering Section.

References

- [1] T. Y. Otoshi, "Antenna Noise Temperature Contributions due to Ohmic and Leakage Losses of the DSS 14 64-m Antenna Reflector Surface," *DSN Progress Report for July and August 1971*, Technical Report 32-1526, vol. 26, no. V, pp. 115-119, Jet Propulsion Laboratory, Pasadena, California, October 15, 1971.
- [2] C. T. Stelzried, "Operating Noise Temperature Calibrations of Low Noise Receiving Systems," *Microwave Journal*, vol. 14, no. 6, pp. 41-48, June 1971.
- [3] C. C. Chen, "Transmission of Microwave Through Perforated Plates of Finite Thickness," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-21, pp. 1-6, January 1973.
- [4] T. Y. Otoshi, "Precision Reflectivity Loss Measurements of Perforated Plate Mesh Materials by a Waveguide Technique," *IEEE Transactions on Instrumentation and Measurements (Special Issue)*, vol. IM-21, no. 4, pp. 451-457, November 1972.

Table 1. Original data

Test. No.	Test Configuration	Readings of T_{op} , K	Average T_{op} , K	Comments
A1	Unpainted solid-aluminum sheet (unclean side)	27.00, 26.99 26.99	26.99	
A2	Reference reading unpainted aluminum sheet (clean side)	26.93, 26.87 26.85, 26.54 26.55, 26.60	26.72	Turned sheet to other side and wiped
A3	Painted solid-aluminum sheet	26.93, 26.97 26.95	26.95	Test A3-Test A2 values give paint loss
A4	Unpainted 3/16-in.-diameter hole panel, dry, horizon sky behind panel	27.63, 27.64 27.63	27.63	
A5	Unpainted 3/16-in.-diameter hole panel, dry, absorber behind panel	27.98, 27.99 28.00, 27.99	27.99	Test A5-Test A7 values give perforated panel dissipation and leakage loss
A6	Unpainted 3/16-in.-diameter hole panel, dry, solid plate behind panel	27.24, 27.25 27.23	27.24	
A7	Reference reading unpainted aluminum sheet (clean side)	26.67, 26.65 26.64	26.65	Compare to Test A2 reference value of 26.72 K
A8	Painted solid-aluminum sheet	26.81, 26.77 26.72, 26.77	26.78	Test A8-Test A7 values give paint loss
A9	Painted 3/16-in.-diameter hole panel, dry, horizon sky behind panel	27.75, 27.75 27.75	27.75	
A10	Painted 3/16-in.-diameter hole panel, dry, absorber behind panel	28.16, 28.19 28.18	28.18	Compare to Test A5 results
A11	Painted 3/16-in.-diameter hole panel, dry, solid plate behind panel	27.50, 27.52 27.48	27.50	
A12	Reference reading unpainted aluminum sheet (clean side)	26.61, 26.61 26.57	26.60	Compare to Test A7 reference value of 26.65 K
A13	Painted solid-aluminum sheet	26.77, 26.77 26.70	26.73	Test A13-Test A12 values give paint loss
A14	Unpainted 3/16-in.-diameter hole panel, wet, horizon sky behind panel	43.30, 41.70 41.09	42.04	Water draining off during test
A15	Unpainted 3/16-in.-diameter hole panel, wet, absorber behind panel	39.48, 39.26	39.37	Water draining off during test
A16	Unpainted 3/16-in.-diameter hole panel, wet, solid plate behind panel	38.60, 38.40 38.30	38.43	Water draining off during test
A17	Same as Test A16 except more water sprayed on panel, solid plate behind panel	40.97, 43.30 40.42	41.56	
A18	Reference reading unpainted aluminum sheet (clean side)	26.44, 26.42 26.43	26.43	Compare to Test A12 reference value of 26.60 K

Table 1 (contd)

Test. No.	Test Configuration	Readings of T_{op} , K	Average T_{op} , K	Comments
A19	Painted solid-aluminum sheet	26.57, 26.58 26.59	26.58	Test A19–Test A18 values give paint loss
A20	Painted 3/16-in.-diameter hole panel, wet, horizon sky behind panel	35.37, 36.18 35.60	35.72	Water draining off during test
A21	Painted 3/16-in.-diameter hole panel, wet, absorber behind panel	35.40, 35.37 35.25	35.34	Water draining off during test
A22	Painted 3/16-in.-diameter hole panel, wet, solid plate behind panel	34.97, 35.02 35.04	35.01	Water draining off during test
A23	Reference reading unpainted aluminum sheet (clean side)	26.42, 26.37 26.37	26.39	Compare to Test A18 reference value of 26.43 K
A24	Painted solid-aluminum sheet	26.49, 26.48 26.47	26.48	Test A24–Test A23 values give paint loss
A25	Painted solid-aluminum sheet, wet	30.17, 26.98 26.70, 26.67	27.63	Test A25–Test A24 values give water film loss on painted aluminum sheet
A26	Unpainted 3/16-in.-diameter hole panel, covered with 2-in.-wide aluminum tape (not too carefully done). Plate was not clamped to be as flat as in Test A4.	29.71, 29.74 29.71	29.72	Test A26–Test A28 values give aluminum tape loss (inconclusive)
A27	Same as Test A26 except wet	32.49, 35.89 30.94, 31.30	32.66	
A28	Reference reading unpainted aluminum sheet (clean side)	26.23, 26.21 26.22, 26.21	26.22	Compare to Test A23 reference value of 26.39 K
A29	Painted solid-aluminum sheet (dried out)	26.23, 26.25 26.24	26.24	Test A29–Test A28 values give paint loss

Table 2. Reduced test data from original data in Table 1

Item Id.	Item Being Tested and Test Configuration	Noise Temperature Contribution, K	Average K	Comments
B1	Increase due to white paint only on 6061 T6 aluminum sheet	(A# values are given in Table 1) A3 - A2 = 0.23 A8 - A7 = 0.13 A13 - A12 = 0.13 A19 - A18 = 0.15 A24 - A23 = 0.09 A29 - A28 = 0.02	0.1	Repeatability is ± 0.1 K
B2	Dry unpainted aluminum antenna panel with 3/16-in.-diameter holes	(from Table 1)		
	(a) With horizon sky behind panel	27.63 - 26.72 = 0.91 A4 A2	0.9	
	(b) With absorber behind panel	27.99 - 26.65 = 1.34 A5 A7	1.3	Approximate NT due to resistive and leakage losses
	(c) With solid plate behind panel	27.24 - 26.65 = 0.59 A6 A7	0.6	Approximate resistive loss NT for unpainted perforated panel
	(d) Calculations	(from this Table) 1.34 - 0.59 = 0.75 B2(b) B2(c)	0.8	Approximate NT due to leakage
B3	Dry painted aluminum antenna panel with 3/16-in.-diameter holes	(from Table 1)		
	(a) Horizon sky	27.75 - 26.65 = 1.10 A9 A7	1.1	
	(b) Absorber	28.18 - 26.60 = 1.58 A10 A12	1.6	Approximate NT due to resistive and leakage losses
	(c) Solid plate	27.50 - 26.60 = 0.90 A11 A12	0.9	Approximate NT due to resistive loss only
	(d) Calculations	(from this Table) 1.58 - 0.90 = 0.68 B3(b) B3(c)	0.7	Approximate NT due to leakage loss only
B4	Noise temperature increase due only to white paint on dry aluminum panels with 3/16-in.-diameter holes. Calculations based on unpainted and painted panel data with the same configurations behind the panels. It is interesting to compare the following results to the 0.13-K value for the painted solid panel.	(from this Table)		
	(a) Horizon sky	1.10 - 0.91 = 0.19 B3(a) B2(a)	0.2	Sensitive to changes in low elevation-angle sky temperature
	(b) Absorber	1.58 - 1.34 = 0.24 B3(b) B2(b)	0.2	Most accurate result
	(c) Solid plate	0.90 - 0.59 = 0.31 B3(c) B2(c)	0.3	Difficult to accurately position the solid plate

Table 2 (contd)

Item Id.	Item Being Tested and Test Configuration	Noise Temperature Contribution, K	Average K	Comments
B5	Unpainted aluminum antenna panel with 3/16-in.-diameter holes in wet condition after water sprayed on front side			
		(from Table 1)		
(a)	Horizon sky	$42.04 - 26.60 = 15.44$ A14 A12	15.4	
(b)	Absorber	$39.37 - 26.43 = 12.94$ A15 A18	12.9	Approximate resistive and leakage loss NT
(c)	Solid plate	$38.43 - 26.43 = 12.00$ A16 A18	12.0	Approximate resistive loss NT
(d)	Solid plate after more water sprayed	$41.56 - 26.43 = 15.13$ A17 A18	15.1	Approximate resistive loss NT
(e)	Calculations	(from this Table) $12.9 - 12.0 = 0.90$ B5(b) B5(c)	0.9	Approximate leakage loss NT for wet panel
B6	Noise temperature increase due to water on <i>unpainted</i> aluminum panel with 3/16-in.-diameter holes			
		(from this Table)		
(a)	Horizon sky	$15.44 - 0.91 = 14.53$ B5(a) B2(a)	14.5	
(b)	Absorber	$12.94 - 1.34 = 11.60$ B5(b) B2(b)	11.6	
(c)	Solid plate	$12.00 - 0.59 = 11.41$ B5(c) B2(c)	11.4	
(d)	Solid plate after more water sprayed	$15.13 - 0.59 = 14.54$ B5(d) B2(c)	14.5	
B7	Painted aluminum antenna panel with 3/16-in.-diameter holes in wet condition after water sprayed on front side			
		(from Table 1)		
(a)	Horizon sky	$35.72 - 26.43 = 9.29$ A20 A18	9.3	
(b)	Absorber	$35.34 - 26.43 = 8.91$ A21 A18	8.9	Approximate resistive and leakage loss NT
(c)	Solid plate	$35.01 - 26.39 = 8.62$ A22 A23	8.6	Approximate resistive loss NT
(a)	Calculations	(from this Table) $8.91 - 8.62 = 0.29$ B7(b) B7(c)	0.3	Approximate leakage loss NT
B8	Noise temperature increase due to water on <i>ainted</i> aluminum antenna panel with 3/16-in.-diameter holes			
		(from this Table)		
(a)	Horizon sky	$9.29 - 1.10 = 8.19$ B7(a) B3(a)	8.2	
(b)	Absorber	$8.91 - 1.58 = 7.33$ B7(b) B3(b)	7.3	Water draining
(c)	Solid plate	$8.62 - 0.90 = 7.72$ B7(c) B3(c)	7.7	

Table 2 (contd)

Item Id.	Item Being Tested and Test Configuration	Noise Temperature Contribution, K	Average K	Comments
B9	Differences of noise temperature increase due to water on <i>unpainted</i> and <i>painted</i> aluminum panel with 3/16-in.-diameter hole			
		(from this Table)		
	(a) Horizon sky	$15.44 - 9.29 = 6.15$ B5(a) B7(a)	6.2	
	(b) Absorber	$12.94 - 8.91 = 4.03$ B5(b) B7(b)	4.0	Water draining off
	(c) Solid plate	$12.00 - 8.62 = 3.38$ B5(c) B7(c)	3.4	
	Note: For all of the above test configurations, the unpainted panel with water has a higher noise temperature than the painted panel with water.			
B10	Water sprayed on white painted solid 6061 T6 aluminum sheet	(from Table 1) $27.63 - 26.39 = 1.24$ A25 A23	1.2	Relative to the reference plate
B11	Increase due only to water sprayed on white painted solid 6061 T6 aluminum sheet	(from this Table) $1.24 - 0.13 = 1.11$ B10 B1 (from Table 1) $27.63 - 26.48 = 1.15$ A25 A24	1.1 1.2	This value is more accurate
B12	Aluminum-tape-covered unpainted aluminum antenna panel with 3/16-in.-diameter holes	(from Table 1) $29.72 - 26.22 = 3.50$ A26 A28	3.5	Taped panel not clamped well to test frame
B13	Water sprayed on aluminum-tape-covered antenna panel with 3/16-in.-diameter holes	(from Table 1) $32.66 - 26.22 = 6.44$ A27 A28	6.4	
B14	Increase due only to water sprayed on aluminum-tape-covered antenna panel with 3/16-in.-diameter holes	(from this Table) $6.44 - 3.50 = 2.94$ B13 B12	2.9	

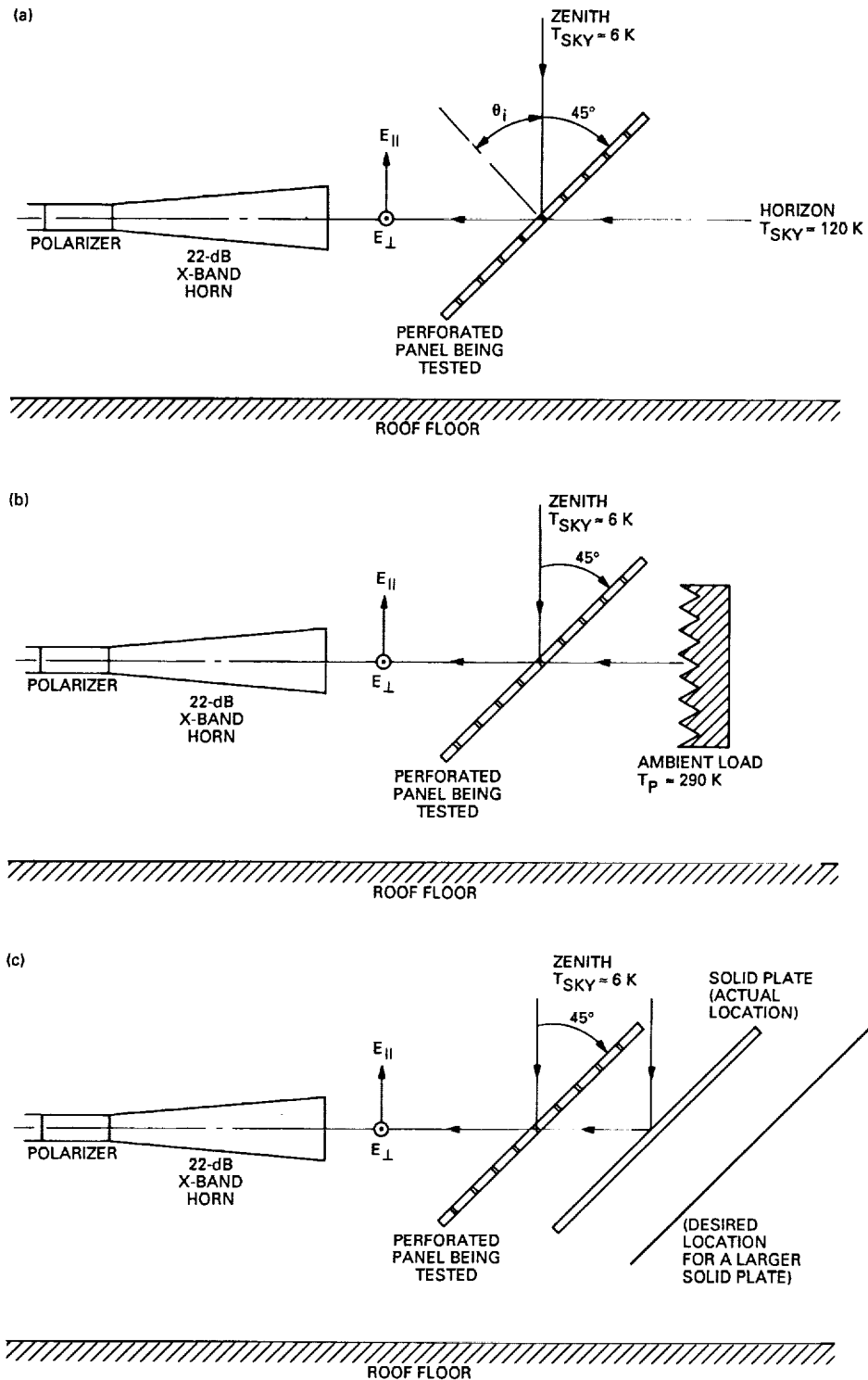


Fig. 1. Radiometric test setup showing the: (a) horizon sky, (b) absorber, and (c) solid plate behind-the-panel test configurations.

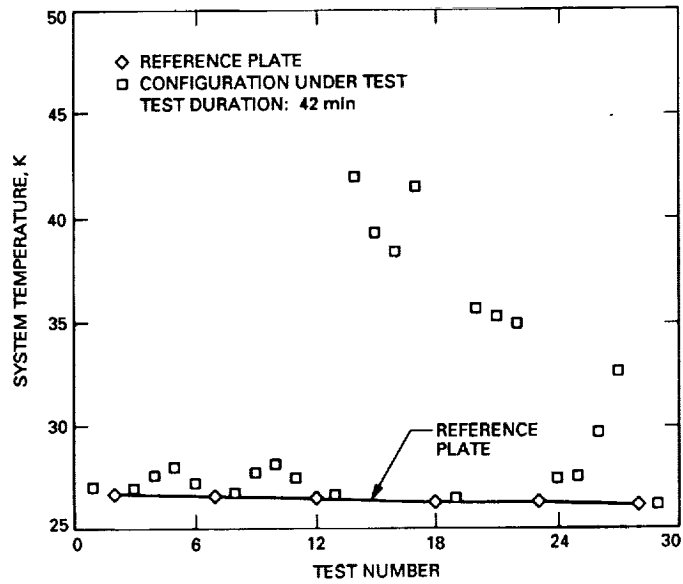


Fig. 2. Measured operating (system) noise temperatures versus test number. Descriptions of the tests are given in Table 1.

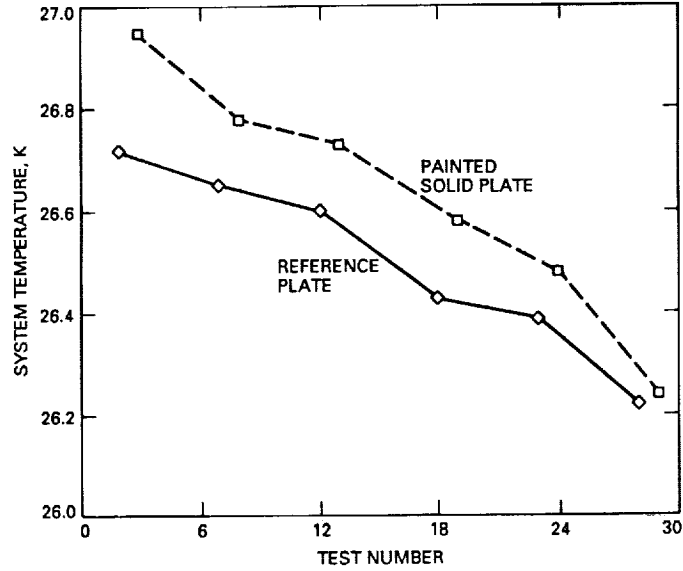


Fig. 3. Measured operating noise temperatures for the unpainted and painted solid-aluminum plate reflectors. The unpainted plate was used for all reference measurements.

Appendix A

Computer-Controlled Operating Noise Temperature Measurement System

A brief description of the special automatic features of the test setup will be given in this Appendix. For the test setup described in this article, it was not required that an aperture ambient load be placed over the horn in order to make Y-factor measurements for measuring the operating noise temperature. Instead the ambient load for this test setup is a waveguide load connected to a remote controllable waveguide switch.

In this automatic measurement system, a computer first commands the waveguide switch to switch to the "ambient load" position so that the waveguide ambient load is connected to the input to the HEMT. The ambient load physical temperature, furnished by a digital thermometer, is read by the computer. A command is then sent to rezero the HP digital power meter. Then, for a duration of about 5 sec, the average received power in the "ambient load" configuration is measured with the HP digital power meter and read by the computer.

The computer then commands the waveguide switch to switch to the "horn" position so that the horn is connected to the input of the HEMT. A measurement of the average received noise power in the "horn" position is then measured with the HP power meter for about 5 sec. Then, from the two measured noise power values (ambient load

and horn) and a Y-factor equation [2], the operating noise temperature for the horn system configuration is computed and displayed on the computer monitor. Then, the average received noise power for the horn configuration is again measured for 5 sec. The noise power measured previously for the ambient load configuration is assumed to be constant and used again to compute a new operating noise temperature for the current horn configuration. A new value of antenna operating noise temperature is obtained every 5 sec and displayed on the computer monitor.

After 15 min, the computer commands the waveguide switch to return to the "ambient load" position and the entire sequence is repeated. This procedure results in the system being recalibrated every 15 min. The assumption is made that the ambient load and receiver noise temperatures do not change during the 15-min period between calibration and recalibration. The assumption is a good one, provided that the ambient load temperature does not change more than about ± 3 K over the 15-min test period. The other assumption, that the receiver noise temperature be constant, is a good one because tests have shown that the HEMT noise temperature does not change more than ± 0.2 K over long periods (12 hr) of time. Test results from DSS 13 indicate that the X-band HEMT stability is also insensitive to changes in physical orientation (elevation angle).

Appendix B

Noise Temperature for Elliptical Polarization Configuration

Although it was originally thought that the polarizer for the test setup was set for circular polarization, it was found that installed polarizer was inadvertently set to 22.5 deg rather than 45 deg with respect to the E-field orientation of the rectangular waveguide. Thus, for the actual elliptical polarization configurations that existed for the tests, the following equations apply:

$$\left[\frac{\text{parallel pol. power}}{\text{total power}} \right] = \cos(22.5) \cos(22.5) = 0.854$$

$$\left[\frac{\text{perp. pol. power}}{\text{total power}} \right] = \sin(22.5) \sin(22.5) = 0.146$$

The measured noise temperature (NT) is then

$$NT = 0.854(NT)_{\parallel} + 0.146(NT)_{\perp}$$

where $(NT)_{\parallel}$ and $(NT)_{\perp}$ are the noise temperatures of the antenna panel that would be measured if the polarizer was set for receiving only parallel and perpendicular polarizations, respectively.

Appendix C

Noise Temperature Equations for Various Test Configurations

A. Operating Noise Temperature

For the test setups shown in Fig. 1, a general expression for the antenna system operating noise temperature is

$$T_{opa} = T_a + T_r \quad (C-1)$$

where T_a is the antenna noise temperature defined at the horn aperture and T_r is the receiver effective noise temperature defined at the horn aperture. The units are in kelvins.

The receiver consists of the horn, waveguide (polarizer, transition, and switch) connecting the horn to the HEMT, the HEMT, and the followup receiver. For purposes of this article, the receiver noise temperature is assumed to be constant and only the antenna noise temperature changes with the test configuration. Therefore, only expressions of T_a for the different test configurations will be provided.

B. Solid-Plate Test Configuration

Let $T_{a,sp}$ be the antenna noise temperature for the test configuration that involves the use of the solid aluminum plate test as the reference. Furthermore, let $T_{a',sp}$ be the general symbol for antenna noise temperature when the plate has a layer of paint or water. Then, if the solid plate replaces the perforated plate shown in the test setup of Fig. 1(a), it can be seen that

$$T_{a,sp} = A_{sp} [|(S11)_{sp}|^2 T_{sky,zen} + D_{sp} T_p] + (1 - A_{sp}) T_{b,sp} \quad (C-2)$$

$$T_{a',sp} = A'_{sp} [|(S11')_{sp}|^2 T_{sky,zen} + D'_{sp} T_p] + (1 - A'_{sp}) T_{b',sp} \quad (C-3)$$

where the A_{sp} is the fraction of total power captured by the solid reference plate, $T_{b,sp}$ is the average brightness temperature as seen by the spillover (uncaptured) power, and $|(S11)_{sp}|^2$ and D_{sp} are, respectively, the fractions of total incident power that are reflected and dissipated by the reference plate. The primed symbols have the same definitions as the corresponding unprimed symbols except that they apply to the solid plate having the paint or water layer. The symbol $T_{sky,zen}$ is the brightness temperature

as seen looking at the sky in the zenith direction and consists of noise contributions due to the atmosphere, cosmic background radiation, and the galaxy; T_p is the physical (or ambient) temperature in kelvins. It is assumed that the physical temperatures of the painted and unpainted (or wet) plates are the same during the test.

In Eqs. (C-2) and (C-3), it is of interest to note from conservation of energy considerations that

$$D_{sp} = 1 - |(S11)_{sp}|^2$$

$$D'_{sp} = 1 - |(S11')_{sp}|^2$$

and $|(S11)_{sp}|^2$ and $|(S11')_{sp}|^2$ can never be equal to unity in practice because of finite metal surface resistivity of the solid plate. Assuming that the spillover power ratio is the same in the two test configurations (i.e., $A'_{sp} = A_{sp}$), and $T_{b,sp} = T_{b',sp}$ then subtraction of the operating noise temperatures, using Eqs. (C-1), (C-2), and (C-3), leads to the result

$$T_{a',sp} - T_{a,sp} = A_{sp} [|(S11')_{sp}|^2 - |(S11)_{sp}|^2] T_{sky,zen} + (D'_{sp} - D_{sp}) T_p \quad (C-4)$$

For the solid plate with paint

$$|(S11')_{sp}|^2 \simeq |(S11)_{sp}|^2$$

and since for this test setup used at X-band, $T_{sky,zen} = 6$ K, the first term in Eq. (C-4) becomes very small compared to the other terms and may be dropped. A theoretical study showed that the same assumption can be made for the case of a thin layer of water on a solid reflector surface. Then it can be seen from Eq. (C-4) that subtraction of the two operating noise temperatures will lead to the differential noise temperature increase due to dissipative loss of the paint or water.

To account for polarization then, in Eqs. (C-2) and (C-3) one needs only to substitute the expressions

$$|(S11)_{sp}|^2 = |(S11)_{sp,\parallel}|^2 \cos^2 \theta_p + |(S11)_{sp,\perp}|^2 \sin^2 \theta_p \quad (C-5)$$

$$|(S11')_{sp}|^2 = |(S11')_{sp,\parallel}|^2 \cos^2 \theta_p + |(S11')_{sp,\perp}|^2 \sin^2 \theta_p \quad (C-6)$$

where $|(S11)_{sp,\parallel}|^2$ and $|(S11)_{sp,\perp}|^2$ are, respectively, the power reflection coefficient for parallel and perpendicular polarizations for the reference plate and primed symbols apply to the test plate configuration. The symbol θ_p is the polarizer angle defined such that when $\theta_p = 0$, all of the reflected power is in the parallel polarization component.

C. Perforated Panel in the Horizon Sky Behind-the-Panel Configuration

For the horizon sky behind-the-panel configuration shown in Fig. 1(a), let $T_{a,pp,hor}$ and $T_{a',pp,hor}$ be defined as the antenna noise temperatures, respectively, for the unpainted and painted (or wet) perforated panel test configurations. Then, from Fig. 1(a), it can be seen that

$$T_{a,pp,hor} = A_{pp}[|(S11)_{pp}|^2 T_{sky,zen} + D_{pp}T_p + |(S21)_{pp}|^2 T_{sky,hor}] + (1 - A_{pp})T_{b,pp} \quad (C-7)$$

$$T_{a',pp,hor} = A'_{pp}[|(S11')_{pp}|^2 T_{sky,zen} + D'_{pp}T_p + |(S21')_{pp}|^2 T_{sky,hor}] + (1 - A'_{pp})T_{b',pp} \quad (C-8)$$

where A_{pp} is the fraction of total power captured by the unpainted perforated reference plate and $T_{b,pp}$ is the average brightness temperature as seen by the spillover (uncaptured) power. The symbols $|(S11)_{pp}|^2$, $|(S21)_{pp}|^2$, and D_{pp} are, respectively, the fractions of total incident power reflected, transmitted, and dissipated by the unpainted perforated plate. The primed symbols have the same definitions as the corresponding unprimed symbols except that they apply to the perforated plate having a layer of paint or water. The symbol $T_{sky,hor}$ is the brightness temperature as seen looking at the sky in the horizon direction.

From conservation of energy considerations, in Eqs. (C-6) and (C-7), we could have let

$$D_{pp} = 1 - |(S11)_{pp}|^2 - |(S21)_{pp}|^2$$

$$D'_{pp} = 1 - |(S11')_{pp}|^2 - |(S21')_{pp}|^2$$

but it is clearer to explicitly show the dissipative components of interest as was done previously.

When the operating noise temperatures for the primed and unprimed expressions are differenced for this test configuration, this corresponds to subtracting Eq. (C-7) from (C-8). For most perforated panel testing, the differential term involving $T_{sky,zen}$ is typically small compared to other

terms and, therefore, can be dropped. The end result will be an expression for the differential noise contribution due to dissipation except for an error term involving $T_{sky,hor}$, which can also be neglected if the difference in leakage wave contributions for the two cases is small. The differential leakage wave contribution will be small if

$$|(S21')_{pp}|^2 - |(S21)_{pp}|^2 \approx 0$$

The above assumption may be valid for the tests involving comparisons of painted and unpainted panels, but not necessarily valid for tests involving comparisons of wet and dry panels.

Differencing the above primed and unprimed antenna temperatures, for the wet and dry panel tests, will give a net change due to changes in reflection, dissipation, and leakage. The measured change would still be of interest, but it cannot be stated that the entire change was caused by differences in dissipative loss alone.

For tests on the perforated panels, the polarization can be accounted for by substitutions of the following expressions into Eqs. (C-7) and (C-8):

$$|(S11)_{pp}|^2 = |(S11)_{pp,\parallel}|^2 \cos^2 \theta_p + |(S11)_{pp,\perp}|^2 \sin^2 \theta_p \quad (C-9)$$

$$|(S21)_{pp}|^2 = |(S21)_{pp,\parallel}|^2 \cos^2 \theta_p + |(S21)_{pp,\perp}|^2 \sin^2 \theta_p \quad (C-10)$$

$$|(S11')_{pp}|^2 = |(S11')_{pp,\parallel}|^2 \cos^2 \theta_p + |(S11')_{pp,\perp}|^2 \sin^2 \theta_p \quad (C-11)$$

$$|(S21')_{pp}|^2 = |(S21')_{pp,\parallel}|^2 \cos^2 \theta_p + |(S21')_{pp,\perp}|^2 \sin^2 \theta_p \quad (C-12)$$

where $|(S11)_{pp,\parallel}|^2$ and $|(S11)_{pp,\perp}|^2$ are, respectively, the power reflection coefficients of the unpainted perforated panels for parallel and perpendicular polarizations. Then, $|(S21)_{pp,\parallel}|^2$ and $|(S21)_{pp,\perp}|^2$ are, respectively, the power transmission coefficients of the unpainted perforated panels for parallel and perpendicular polarizations. The primed symbols have the same definitions except they apply to the perforated panel having a layer of paint or water.

D. Perforated Panel in the Absorber Behind-the-Panel Configuration

For the absorber behind-the-panel configuration shown in Fig. 1(b), let $T_{a,pp,abs}$ and $T_{a',pp,abs}$ be defined as the antenna noise temperatures, respectively, for the

unpainted and painted (or wet) perforated panel test configurations. Then, from Fig. 1(b), it can be seen that

$$\begin{aligned} T_{a,pp,abs} = & A_{pp} [| (S11)_{pp} |^2 T_{sky,zen} + D_{pp} T_p \\ & + A_b | (S21)_{pp} |^2 T_p \\ & + (1 - A_b) | (S21)_{pp} |^2 T_{b,b}] \\ & + (1 - A_{pp}) T_{b,pp} \end{aligned} \quad (C-13)$$

$$\begin{aligned} T_{a',pp,abs} = & A'_{pp} [| (S11')_{pp} |^2 T_{sky,zen} + D'_{pp} T_p \\ & + A_b | (S21')_{pp} |^2 T_p \\ & + (1 - A_b) | (S21')_{pp} |^2 T_{b,b}] \\ & + (1 - A'_{pp}) T_{b,pp} \end{aligned} \quad (C-14)$$

where A_b is the fraction of leakage wave power captured by the absorber and is assumed to be the same value for the above primed and unprimed measurements. Also, $T_{b,b}$ is the average brightness temperature as seen by the portion of perforated panel leakage wave that was not captured by the absorber, K . To account for polarization, it is necessary only to substitute Eqs. (C-9) and (C-10) into Eqs. (C-13), and Eqs. (C-11) and (C-12) into Eq. (C-14).

If the operating noise temperatures for the primed and unprimed cases are differenced, the result is equivalent to subtracting Eq. (C-13) from Eq. (C-14). Then, after the differencing operation is performed, the term involving $T_{sky,zen}$ may be dropped due to making an assumption that it is small compared to the other terms. The result is a noise temperature value that includes both differential panel dissipative loss and leakage loss. If A_b is not equal to unity, a small error term must also be considered. The same comments made for the wet plate case in the horizon sky configuration also apply to this "ambient load" configuration.

E. Perforated Panel in the Solid Plate Behind-the-Panel Configuration

For the solid plate behind-the-panel configuration shown in Fig. 1(c), let $T_{a,pp,sp}$ and $T_{a',pp,sp}$ be defined as the antenna noise temperatures, respectively, for the unpainted and painted (or wet) perforated panel test configurations. Then, from Fig. 1(c), it can be seen that

$$\begin{aligned} T_{a,pp,sp} = & A_{pp} \{ | (S11)_{pp} |^2 T_{sky,zen} + D_{pp} T_p \\ & + A_c | (S21)_{pp} |^2 [| (S11)_{sp} |^2 T_{sky,zen} + D_{sp} T_p] \\ & + (1 - A_c) | (S21)_{pp} |^2 T_{b,c} \} \\ & + (1 - A_{pp}) T_{b,pp} \end{aligned} \quad (C-15)$$

$$\begin{aligned} T_{a',pp,sp} = & A'_{pp} \{ | (S11')_{pp} |^2 T_{sky,zen} + D'_{pp} T_p \\ & + A_c | (S21')_{pp} |^2 [| (S11)_{sp} |^2 T_{sky,zen} + D_{sp} T_p] \\ & + (1 - A_c) | (S21')_{pp} |^2 T_{b,c} \} \\ & + (1 - A'_{pp}) T_{b,pp} \end{aligned} \quad (C-16)$$

where A_c is the fraction of power in the leakage wave captured by the solid plate behind the perforated panel and is assumed to be the same value for the above primed and unprimed measurements. Also, $T_{b,c}$ is the average brightness temperature as seen by the portion of perforated plate leakage signal that was not captured by the solid plate, K . To account for polarization, it is necessary only to substitute Eqs. (C-9) and (C-10) into Eq. (C-15) and Eqs. (C-11) and (C-12) into Eq. (C-16).

For this test configuration, after differencing the antenna noise temperatures for the primed and unprimed cases, the terms involving products with $T_{sky,zen}$ can be dropped. The term $D_{sp} T_p$ may also be dropped. The same comments made for the wet plate case in the horizon sky configuration also apply to this solid-plate configuration.

The presentations of equations for the various test configurations have now been completed. The equations are very useful for error analysis purposes and for developing methodologies for isolating particular noise temperature contributions of interest. It should be mentioned that if the dissipative loss contributions are very small, and it is desirable to determine them very accurately, then the dropping of terms may not be permissible.

F. Applications

An interesting application of the above presented equations is to separate the dissipative loss and leakage loss contributions of the perforated panel.

Subtraction of the operating noise temperatures for the perforated panel in the absorber behind-the-panel configuration and the solid reference plate [from Eqs. (C-1), (C-2), and (C-13)], results in the expression

$$\begin{aligned} T_{a,pp,abs} - T_{a,sp} = & [| (S11)_{pp} |^2 - | (S11)_{sp} |^2] T_{sky,zen} \\ & + (D_{pp} - D_{sp}) T_p + | (S21)_{pp} |^2 T_p \end{aligned} \quad (C-17)$$

where it was assumed in Eq. (C-2) that $A_{sp} = 1$ and in Eq. (C-13) $A_{pp} = 1$ and $A_b = 1$. The first, second, and third terms in the above equation, respectively correspond to differential contributions due to reflection, dissipation,

and leakage. If it can be assumed that the first term can be dropped because it is small compared to the other terms, then the remaining terms provide information concerning the contributions due to both dissipation and leakage.

Now, differencing the operating noise temperatures for the perforated panel in the solid-plate behind-the-panel configuration and the solid reference plate [from Eqs. (C-1), (C-2), and (C-15)] results in the expression

$$\begin{aligned} T_{a,pp,sp} - T_{a,sp} = & (|(S11)_{pp}|^2 - |(S11)_{sp}|^2) T_{sky,zen} \\ & + (D_{pp} - D_{sp}) T_p \\ & + |(S21)_{pp}|^2 |(S11)_{sp}|^2 T_{sky,zen} + D_{sp} T_p \end{aligned} \quad (C-18)$$

where it has been assumed in Eq. (C-2) that $A_{sp} = 1$ and in Eq. (C-15) that $A_{pp} = 1$ and $A_c = 1$. In Eq. (C-18), if it is assumed that the first and third terms can be dropped,

then the results are just the differential dissipative loss noise temperature contribution of the perforated plate.

In order to obtain the leakage wave contribution, one can difference the operating noise temperatures of the perforated panel in the absorber and solid plate behind the panels. Then, performing this differencing through the use of Eqs. (C-1), (C-13), and (C-15), one obtains

$$\begin{aligned} T_{a,pp,abs} - T_{a,pp,sp} = & |(S21)_{pp}|^2 [T_p - |(S11)_{sp}|^2 (T_{sky,zen} + D_{sp} T_p)] \end{aligned} \quad (C-19a)$$

where it has been assumed that $A_{pp} = 1$, $A_b = 1$, and $A_c = 1$. If we further assume that $|(S11)_{sp}|^2 = 1$ and $D_{sp} T_p$ is small compared to $T_{sky,zen}$, then differencing the operating noise temperatures will result in an approximate leakage wave noise temperature contribution of

$$T_{a,pp,abs} - T_{a,pp,sp} = |(S21)_{pp}|^2 (T_p - T_{sky,zen}) \quad (C-19b)$$

A Functional Description of the Advanced Receiver

S. Hinedi

Communications Systems Research Section

This article describes the breadboard Advanced Receiver II (ARX II) that is currently being built for future use in NASA's Deep Space Network (DSN). The hybrid analog/digital receiver performs multiple functions including carrier, subcarrier, and symbol synchronization. Tracking can be achieved for residual, suppressed, or hybrid carriers and for both sinusoidal and square-wave subcarriers. Other functions such as time-tagged Doppler extraction and monitor/control are also discussed, including acquisition algorithms and lock-detection schemes. System requirements are specified and a functional description of the ARX II is presented. The various digital signal-processing algorithms used are also discussed and illustrated with block diagrams.

I. Introduction

The Advanced Receiver (ARX) [1] is a system that has been under development for future use in the Deep Space Network (DSN) [2]. It is intended to replace various receivers currently in use to demodulate and process signals from deep-space spacecraft. The ARX has been under development for several years and has undergone several tests [3-6] and modifications to improve its performance. It has also been used as a testbed for high-dynamic frequency tracking for Global Positioning System (GPS) applications [7-9].

Two versions of the ARX have been built, ARX I and ARX II; they differ mainly in the processing strategy used. In ARX I, the sampling clock was driven by the symbol-synchronization loop and hence, an integer number of samples per symbol was obtained. Since the sampling period was not fixed, the time base varied. As a result, time-tagged digital Doppler extraction was unfeasible. In ARX II, the sampling clock was fixed, resulting in the possibility

of a noninteger number of samples per symbol. However, there is a fixed time base and the system can be easily synchronized to a station clock. This article concerns itself only with ARX II, as its architecture was chosen as the final candidate for a possible future implementation of a Block V receiver. The system requirements and a functional description of the operating principles of the receiver are given. Performance trade-off issues as a function of complexity are discussed and design decisions are clarified.

Before discussing the design of the ARX II, it is important to understand the different requirements of the system and the constraints they might impose on its architecture. First, functional and performance requirements will be discussed. Before doing that however, it is worthwhile to make sure that the reader understands what is meant here by the term "receiver." The received signal, which may be at L (1628-1708 MHz), S (2200-2300 MHz) or X (8400-8500 MHz) band, gets downconverted in an open-loop fashion to an intermediate frequency (IF) in the

200–400-MHz range. Following that downconversion, all operations including further downconversion to more appropriate frequencies, carrier demodulation, and symbol detection are all included in what is referred to here as the “receiver.” Therefore, the input to the receiver is a signal in the 200–400-MHz range and one of its outputs consists of the detected soft symbols. These symbols are later decoded in several stages (Reed-Solomon as well as convolutional or Viterbi decoding) to produce bits. That decoding is not part of what is referred to here as the “receiver.” Other decoding schemes for source decompression might also be present but are not considered as part of the ARX II. Another important output of the ARX II is a time-tagged received phase cycle count, which is eventually used in other subsystems to perform ranging and navigation functions.

This article is divided into six sections. Section II discusses system requirements, both functional and technical. The architecture of the receiver is the topic of Section III, along with the interface issues that affect and limit the ARX II design. Section IV deals with the tracking algorithms and the other signal-processing schemes employed to perform the various tracking tasks. Implementation versus performance trade-off issues are discussed in Section V and the conclusions are presented in Section VI.

II. Receiver Requirements

In order to assess the performance of the system, a set of performance requirements is needed to serve as a reference against which a measured set of data can be compared. Secondly, a more basic set of functional requirements is needed to define the operations required from the receiver: What is it supposed to do? What are its inputs and outputs? Finally, interface issues need to be defined for the receiver to operate with other subsystems. The latter issue is probably the most important because it typically dictates certain features in the receiver. These issues, as well as other matters such as size, power consumption, etc., constitute the set of requirements and may result in constraints that the system engineer should be aware of prior to starting the design.

Functionally, the receiver must perform two fundamental tasks: Doppler extraction and telemetry processing. Doppler extraction involves tracking a pure tone in the presence of time-varying Doppler and recording its phase, time tagged for further off-line processing. Telemetry requires demodulating a signal whose spectrum is data-rate dependent and producing soft detected symbols to be used by other subsystems (for example, a Reed-Solomon decoder). Note that while Doppler extraction

requires narrow bandwidth, the telemetry task needs wide bandwidth to accommodate the highest data rate foreseen. The latter task involves demodulating the symbols of a received signal that can be either binary phase shift keyed (BPSK), quadrature phase shift keyed (QPSK), offset QPSK (OQPSK), or minimum shift keyed (MSK). In OQPSK, the in-phase and quadrature baseband pulses are shifted with respect to each other by half a symbol period while in MSK, sinusoidal pulse shaping is employed instead of rectangular pulses. For BPSK, the signal might have a residual carrier component, in which case the data spectrum will be shifted away using a square or a sinusoidal subcarrier. In the case of a fully suppressed BPSK carrier, the subcarrier may or may not exist. The data itself can be either non-return-to-zero (NRZ) or Manchester (bi-phase) encoded. In communication terms, the receiver needs to perform carrier tracking, whether residual or suppressed, subcarrier tracking, and symbol synchronization for the different modulation formats listed above. Various lock detectors are required to monitor the loops and to help automate the receiver's operation. Furthermore, a friendly receiver/user interface is highly desirable to facilitate operations.

Now that the system's functionality has been defined, performance-related issues need to be addressed to specify the receiver's requirements. In order to support current and future (near-term) missions, it has been decided that a predetection bandwidth of 16 MHz (3-dB bandwidth) is more than adequate. The latter bounds the highest data rate that can be processed and is roughly half the processing (or effective sampling) rate that can be accommodated with current technology in a digital implementation.

The dynamic range of the receiver should be about 73 dB to handle carrier-to-noise ratios (CNRs) from 0 to 73 dB-Hz (CNR is defined as the ratio of carrier power P_C to the one-sided noise spectral density level N_0). Moreover, the data-rate capability should be from 8 symbols-per-second (SPS) to 6.6 MSPS with direct BPSK modulation and from 8 SPS to 700 kSPS in the presence of a subcarrier, whose frequency lies in the 100-Hz to 2-MHz range. The modulation index, the parameter that delegates power to the data, varies between 0 deg and 90 deg, the first corresponding to a pure tone, the latter to a suppressed carrier. Radio (symbol signal-to-noise ratio) loss due to the carrier loop should be less than 0.3 dB over data rates from 8 SPS to 6.6 MSPS. Similar bounds exist on the loss due to subcarrier and symbol-synchronization loops over the data rates from 8 SPS up to 1 MSPS.

These bounds, from a designer's point of view, determine the range of loop bandwidths that need to be han-

dled. The receiver should be able to acquire the carrier within 10 sec at 8 dB above loop threshold and acquire the subcarrier frequency and symbol epoch within 20 sec for symbol rates above 200 SPS and symbol signal-to-noise ratios (SNR) greater than -2.5 dB; otherwise, acquisition should occur within 6000 symbols. The predict uncertainty is about ± 10 kHz in carrier frequency, ± 1 kHz in subcarrier frequency, and ± 20 SPS for data rates above 20 kSPS or $\pm (\text{symbol rate}/2000)$ for data rates below 20 kSPS.

Furthermore, the receiver should monitor its status by measuring and reporting E_s/N_0 (symbol energy-to-noise ratio) to within ± 0.1 dB, the carrier and subcarrier frequencies to within ± 0.1 Hz, the symbol rate to within ± 0.1 SPS, and the carrier static phase error to within ± 1 deg. Moreover, the receiver should report in-lock/out-of-lock status for all loops; perform sideband aiding, Doppler-rate aiding, fast Fourier transforms (FFTs) for acquisition as well as for monitor/control, and real-time loop bandwidth optimization; and implement loop types I, II, and III for all the various loops with different update rates that range from 100 Hz to 1 kHz.

The last set of requirements to be discussed is related to the interface issues that need to be met for the new system to operate with existing equipment. For example, the receiver will be supplied only 1-MHz, 5-MHz, and 10-MHz references from the Frequency and Timing Subsystem (FTS). Any additional references that are needed by the receiver must be generated internally. Furthermore, to supply time-tagged carrier phase cycle count, a time code translator (TCT) is needed to read the time from the station's time code generator (TCG); this ensures that the receiver's time is synchronized with the station time. Also, the receiver needs to interface with the sequential ranging assembly (SRA) subsystem by providing it with an analog 10-MHz locked IF signal. The SRA itself will undergo upgrading in the near future and it will then only accept digital baseband signals. Therefore, the receiver should be able to provide both analog IF and digital baseband signals to interface with existing and future SRAs.

Only the key elements of the requirements that directly influence the design have been discussed above since a complete specification would be lengthy and inappropriate for the purpose of this article.

III. The Architecture of the Advanced Receiver

Both ARXs were implemented in a hybrid analog/digital fashion to offer flexibility and low cost. The front end, which performs mainly filtering and downconversion

operations, was implemented using analog hardware, while the "signal-processing part" was built using the latest commercially available digital hardware.

Figure 1 depicts a top-level block diagram of the receiver with digital loop closure. The first stage of conversion from L-, S-, or X-band to the 200–400-MHz range is accomplished by an external subsystem (IF Distribution Subsystem) and is not part of the receiver. Its output, which constitutes the input to the ARX II, is further down-converted to a more appropriate frequency (70 MHz) for fine-tuned filtering and then to a different lower frequency (10 MHz) for sampling. Because of the high data-rate requirement of 6.6 MSPS, the 10-MHz IF signal is sampled at roughly 40 MHz, digitally in-phase and quadrature (I and Q) mixed to baseband, and filtered with the half-band filters (HBFs) to remove the double-frequency terms. The baseband I and Q signals are then processed to digitally perform carrier, subcarrier, and symbol synchronization, respectively. Note that the additional downconversion to 10 MHz could be avoided by undersampling the 70-MHz IF signal at 40 MHz. This would not violate Nyquist sampling because the signal is band-limited to 20 MHz. However, this would require higher quality analog filters (than their counterparts at 10 MHz) to reduce the potential aliasing.

Both analog and digital loop closures were implemented to satisfy the interface to the current and future SRAs. The closures are depicted in Fig. 2 along with the appropriate frequencies. In the analog loop closure, the downconverted 70-MHz signal is mixed with an 80-MHz phase-locked signal to produce the 10-MHz signal required by the current SRA. The analog-to-digital (A/D) converter operates on the 10-MHz IF locked signal to produce 8-bit samples at 40 MHz. The 10-MHz IF is then digitally removed using a look-up table. After filtering to remove the sum frequency term, the resulting signal has a 10-MHz effective bandwidth and a 40-MHz processing rate. A decimation by two is performed to reduce the processing rate to the required 20-MHz rate specified by Nyquist sampling. The Q samples are then accumulated to further reduce the rate to the residual-carrier loop update rate to enable a software implementation of the loop filter. The filter outputs a frequency-error estimate at the loop update rate and adjusts the nominal frequency (4 MHz) of the numerically controlled oscillator (NCO). Note that the NCO's phase is actually changing at the much higher rate of 20 MHz. This feeds the digital-to-analog (D/A) converter whose output is a 4-MHz sine wave. The latter is subsequently mixed with a fixed 84-MHz signal to produce the 80-MHz analog signal that closes the loop.

The operating principle of the digital loop closure is identical except that the 80-MHz signal is now a fixed refer-

ence and an NCO running at 40 MHz with a 10-MHz nominal frequency is now used instead of a look-up table. In either case, A/D conversion is performed at the 10-MHz IF signal rather than at baseband to avoid the potential problem of DC bias generated by the A/D converter. Moreover, the sampling frequency can be changed to 39.8 MHz to avoid any subharmonics of the sampling clock that might be generated by the A/D in the signal band of interest. A software command can switch the system between analog and digital closure without losing carrier lock. However, the sampling frequency cannot be changed on the fly.

A. Signal Model

A functional block diagram of the ARX II is shown in Fig. 3 for the digital-loop closure case. Depicted are the residual carrier tracking phase-locked loop (PLL), the suppressed carrier tracking loops for both BPSK and QPSK (Costas loop and cross-over Costas loop respectively), the subcarrier loop (also a Costas loop), the symbol-synchronization loop (data-transition tracking loop [DTTL]), lock detectors for the residual carrier loop and for the various Costas loops, sideband aiding, a symbol SNR estimator (split-symbol moments estimator [SSME]), and weighted integrate-and-dump filters (WIDFs) for symbol detection. A total-power automatic gain control (AGC) circuit is present to ensure that the signal is within the dynamic range of the A/D converter. The basic operating principles of each loop will be discussed briefly and the reader is referred to the appropriate references for additional information.

The received signal at the output of the AGC is given by

$$r(t) = \sqrt{2P} \sin[\omega_i t + \Delta D(t) + \theta_c] + n(t) \quad (1)$$

where P is the average signal power, ω_i the IF frequency (70 MHz) in rad/sec, θ_c the carrier phase in rad, Δ the modulation index, $D(t) = d(t)\text{Sin}(\omega_{sc}t + \theta_{sc})$ with $\text{Sin}(x) = \text{sgn}(\sin(x))$ for a square wave ($\text{sgn}(x)$ denotes the "signum" function) and $\text{Sin}(x) = \sin(x)$ for a sine-wave subcarrier, ω_{sc} the subcarrier frequency in rad/sec, θ_{sc} the subcarrier phase in rad, and $d(t)$ the data modulation, i.e.,

$$d(t) = \sum_{k=-\infty}^{+\infty} a_k p(t - kT) \quad (2)$$

with $a_k = \pm 1$ equally likely and $p(t)$ the baseband NRZ or Manchester pulse limited to T seconds. The narrow-band noise $n(t)$ can be written as

$$n(t) = \sqrt{2}n_c(t) \cos(\omega_i t + \theta_c) - \sqrt{2}n_s(t) \sin(\omega_i t + \theta_c) \quad (3)$$

where $n_c(t)$ and $n_s(t)$ are statistically independent stationary band-limited white Gaussian noise processes with one-sided spectral density N_0 watts/Hz and one-sided bandwidth W Hz (roughly 20 MHz in this case). The 70-MHz IF signal is downconverted to 10 MHz and sampled at 40 MHz.

B. Residual Carrier Tracking

In-phase and quadrature digital mixing are employed to convert the 10-MHz digital IF signal to the following baseband samples [10]

$$I_{cr}(n) = \sqrt{P} [D(n) \sin \Delta \sin \phi_c(n) + \cos \Delta \cos \phi_c(n)] + n_I(n) \quad (4)$$

$$Q_{cr}(n) = \sqrt{P} [D(n) \sin \Delta \cos \phi_c(n) + \cos \Delta \sin \phi_c(n)] + n_Q(n) \quad (5)$$

where $n_I(n)$ and $n_Q(n)$ are independent Gaussian random variables with variances $N_0/2T_s$ and $\phi_c(n)$ is the carrier phase error at time nT_s where T_s denotes sampling interval. Note that the power available for residual carrier tracking is $P_C = P \cos^2 \Delta$ due to the nonzero modulation index. The subscripts "cr" denote the carrier (c) residual (r) component. Several other subscripts will be used: "sc" to denote subcarrier, "cs" to indicate the carrier suppressed component and "sy" to denote symbol-synchronization loop related terms.

In residual phase tracking, the Q_{cr} s are accumulated over N_1 samples to reduce the processing rate from 20 MHz to the 100-Hz–1-kHz range to enable a software implementation of the loop filter $F(z)$. Due to the averaging (accumulation) operation, the first component of Q_{cr} becomes zero due to summing over several cycles of the subcarrier. In the ARX II, the loop filters are all given by

$$F(z) = G_1 + \frac{G_2}{1 - z^{-1}} + \frac{G_3}{(1 - z^{-1})^2} \quad (6)$$

where

$$G_1 = rd/T_u \quad (7)$$

$$G_2 = rd^2/T_u \quad (8)$$

$$G_3 = kr d^3/T_u \quad (9)$$

and

$$d = 4B_L T_u (r - k)/r(r - k + 1) \quad (10)$$

T_u denotes the loop update time, B_L the design loop bandwidth in Hz; r is typically 2 or 4 and is equal to 4ξ where ξ is the damping ratio, and k is a type-III loop gain parameter ($k = 0$ for type-II loop) with typical values ranging from $1/4$ to $1/2$. The filter of Eq. (6) was derived from an equivalent analog filter using the impulse-invariant transformation (IIT). The actual loop bandwidth B_L^a might be larger than the designed B_L depending on the product $B_L T_u$. Generally, for $B_L T_u < 0.05$, the actual loop bandwidth is very close to the design loop bandwidth B_L . The infinite impulse response (IIR) digital filter described above was found to provide the "best" performance for small values of $B_L T_u$ in terms of gain margins, transient response, and steady-state error performances [11]. Other filters derived from optimum estimation theory were also considered [12–16] but were not implemented due to their additional complexity and relative performance.

In the case of the PLL, the filter provides a phase-error estimate $\hat{\phi}_c$ which adjusts the frequency of the NCO every update period. An averaged version of the in-phase component of the residual carrier, I_{cr} , is used to provide an indicator for lock detection. Assuming that the residual carrier loop is in lock, the averaged in-phase samples are roughly unity (assuming proper normalization) since they are proportional to $\cos \phi_c$. In fact, that same quantity can be used to verify in real time that the loop gain is actually what it is supposed to be and thus that the operating bandwidth is actually B_L as designed. The normalization involved depends on the actual implementation of the hardware and will not be discussed in this article. A cycle count of the received phase is performed and time tagged using station time to enable digital Doppler extraction.

C. Subcarrier Tracking

In an ideal carrier-tracking situation, $\phi_c(n) = 0 \forall n$ and $Q_{cr}(n)$ of Eq. (5) becomes proportional to the subcarrier waveform $D(n)$. Hence, the input to the subcarrier tracking loop is the quadrature component of the residual carrier. Mixing $Q_{cr}(n)$ with the subcarrier in-phase and quadrature references and ignoring the double-frequency terms, one obtains

$$I_{sc}(n) = \sqrt{P_D} d(n) F_I(\phi_{sc}(n)) \cos(\phi_c(n)) + N_I(n) \quad (11)$$

and

$$Q_{sc}(n) = \sqrt{P_D} d(n) F_Q(\phi_{sc}(n)) \cos(\phi_c(n)) + N_Q(n) \quad (12)$$

where $P_D = P \sin^2 \Delta$ is the data power, ϕ_{sc} is the subcarrier phase error in rad, and $N_I(n)$ and $N_Q(n)$ are independent zero-mean Gaussian random variables with vari-

ances $N_0/2T_s$ and $N_0 W_{sc}/2T_s$, respectively. W_{sc} denotes the width of the window in fraction of cycles that might be used in the quadrature arm to improve the performance for square-wave tracking ($W_{sc} \leq 1$ and $W_{sc} = 1$ corresponds to no window).

For a square-wave subcarrier with window W_{sc} , the I and Q phase functions $F_I(\phi)$ and $F_Q(\phi)$ are given by [17]

$$F_I(\phi) = 1 - \frac{2}{\pi} |\phi| \quad |\phi| \leq \pi \quad (13)$$

and

$$F_Q(\phi) = \begin{cases} (2/\pi)\phi & |\phi| \leq (\pi/2)W_{sc} \\ \text{sgn}(\phi)W_{sc} & (\pi/2)W_{sc} \leq |\phi| \leq \pi(1 - W_{sc}/2) \\ 2\text{sgn}(\phi) - (2/\pi)\phi & \pi(1 - W_{sc}/2) \leq |\phi| \leq \pi \end{cases} \quad (14)$$

whereas in the case of a sine wave, they become

$$F_I(\phi) = \cos \phi \quad F_Q(\phi) = \sin \phi \quad (15)$$

The signals $I_{sc}(n)$ and $Q_{sc}(n)$ are then accumulated over a symbol duration to produce

$$I_{sc,mf}(k) = \sqrt{P_D} a_k F_I(\phi_{sc}(k)) \cos(\phi_c(k)) + N_{I,mf}(k) \quad (16)$$

and

$$Q_{sc,mf}(k) = \sqrt{P_D} a_k F_Q(\phi_{sc}(k)) \cos(\phi_c(k)) + N_{Q,mf}(k) \quad (17)$$

where the subscript "mf" indicates the output of the matched filter. The control signal for the matched filters is derived from the symbol synchronization loop and will be discussed later. The noises $N_{I,mf}(k)$ and $N_{Q,mf}(k)$ are still independent with respective variances $N_0/2T$ and $N_0 W_{sc}/2T$, where T denotes symbol duration and k the discrete time kT .

The error signal of the subcarrier loop is obtained by forming the product of $N_{I,mf}(k)$ and $N_{Q,mf}(k)$ to wipe out the data, accumulating over several symbols to reduce the processing from the symbol rate to a more appropriate subcarrier loop update rate, and feeding the average to the loop filter. The in-phase and quadrature subcarrier signals

are also individually squared and subsequently subtracted from one another to provide a lock indicator signal to the operator [18].

D. Suppressed Carrier Tracking and Sideband Aiding

For suppressed carrier tracking, the subcarrier in-phase component, $I_{sc,mf}(k)$ given by Eq. (16), is also the suppressed carrier in-phase component since it is already proportional to $\cos \phi_c(k)$ and to $F_I(\phi_{sc})$. In order to generate the quadrature component of the suppressed carrier, the in-phase component of the residual carrier ($I_{cr}(n)$ defined by Eq. 4) is used since it is proportional to $\sin \phi_c$. It is demodulated with the same subcarrier reference used to get $I_{sc}(n)$ and accumulated over a symbol period. Hence, the suppressed carrier signals become

$$I_{cs}(k) = \sqrt{P_D} a_k F_I(\phi_{sc}(k)) \cos \phi_c(k) + N_{I,mf}(k) \quad (18)$$

and

$$Q_{cs}(k) = \sqrt{P_D} a_k F_I(\phi_{sc}(k)) \sin \phi_c(k) + N_{Q,mf}^c(k) \quad (19)$$

where $\phi_c(k)$ is the average carrier phase error over the k th symbol and $N_{Q,mf}^c(k)$ is the noise quadrature component with variance $N_0/2T$. The phase-error estimate from the BPSK suppressed carrier tracking loop can be combined with its counterpart from the residual carrier tracking loop, with the appropriate weights, to form a combined carrier phase-error estimate, which ultimately adjusts the frequency of the carrier NCO. This is referred to as *sideband aiding* since the power in the data or sidebands is also used to track the carrier along with the power of the residual carrier. Depending on the modulation format (BPSK or QPSK), the samples $I_{cs}(k)$ and $Q_{cs}(k)$ undergo different processing to produce a phase-error estimate; for BPSK, the error signal is proportional to their product whereas for QPSK, the feedback signal is derived using $I_{cs} \text{sgn}(Q_{cs}) - Q_{cs} \text{sgn}(I_{cs})$ (where $\text{sgn}(x)$ denotes the "signum" function) as described in [19].

E. Symbol Synchronization

Obviously, symbol synchronization needs to be maintained to enable subcarrier and suppressed-carrier tracking. This is accomplished by the digital data-transition tracking loop (DTTL) [20], which uses the in-phase signal of the subcarrier, $I_{sc}(n)$, as its input. Two integrate-and-dump filters are used, one to accumulate over a symbol and the second over a symbol transition with possible windowing. The first, $I_{sc}(k)$, is already performed by both the subcarrier and suppressed carrier loops and can thus

be used in the symbol transition detector. From Fig. 3, the second signal is obtained at the output of the symbol transition WIDF, then multiplied by the first and accumulated to form the symbol phase-error estimate. A lock indicator on the symbol synchronization loop can be obtained by the SSME output which provides a symbol SNR estimate to the operator. The quality of those estimates is a function of both the performance of the symbol synchronization loop and the number of samples per symbol, as will be discussed later.

IV. Algorithm Descriptions

It is very important from an operational point of view to understand the limitations of the various algorithms involved to decide whether the system itself is not performing satisfactorily, or whether the system's status-reporting method is generating misleading data. This requires not only an intuitive understanding of each algorithm, but also an understanding of the possible correlation between the various parameters and coupling among the different loops. In this section, the loops are discussed and their respective performances indicated, but not derived. Moreover, the limitations of the various analyses are clarified and the underlying assumptions are identified.

A. Performance of the BPSK Carrier Tracking Loops

First consider the performance of the simplest loop, the digital PLL. It is well known that the phase-error variance is given by [21]

$$\sigma_{\phi_{cr}}^2 = \frac{N_0 B_{L,cr}^a}{P_C} \quad (20)$$

where $B_{L,cr}^a$ is the actual operating bandwidth of the residual carrier loop. The latter can be different from the design bandwidth $B_{L,cr}$ and is dictated by the difference between the predicted P_C/N_0 , denoted $(P_C/N_0)_p$, and the actual P_C/N_0 . Typically, the loop bandwidth is dependent on the loop gain, which depends on the amplitude of the incoming signal. But because of the total-power AGC present in the IF stage, the amplitude of the signal at the output of the A/D converter in units of the least significant A/D levels becomes dependent on the incoming P_C/N_0 . During system initialization, the loop gain is normalized so as to obtain the desired loop bandwidth $B_{L,cr}$ based on the available predict $(P_C/N_0)_p$. The equations describing the various normalizations are beyond the scope of this article.

Note that when the actual P_C/N_0 is different from its predicted counterpart, the amplitude of the signal will be different from that predicted at the A/D output and

the normalization performed will not provide the designed loop gain, and hence, the desired loop bandwidth. However, the loop bandwidth can be monitored in real time by averaging the in-phase residual carrier component over an appropriate length of time and comparing it with the predicted average, which for convenience can be normalized to unity. If the incoming amplitude is different from its predict, a correction can be made to renormalize it properly. For the averaging process to be reliable, it needs to be performed over a period during which the PLL has approximately zero steady-state phase error due to dynamics. Otherwise, the averaged quantity will reflect more than the difference in loop gain.

A steady-state phase-error estimate can be obtained in real time by computing

$$\widehat{\phi}_{ss}(j) = \tan^{-1} \left(\frac{\sum_{n=j}^{n=j+L+1} Q_{cr}(n)}{\sum_{n=j}^{n=j+L+1} I_{cr}(n)} \right) \quad (21)$$

where L controls the estimation period. In the ARX II, $\widehat{\phi}_{ss}(j)$ can be sent to the monitor's screen at different rates ranging from once every second to once every minute. Another very useful parameter is the incoming P_C/N_0 , which is also estimated by the receiver according to

$$\frac{\widehat{P_C}}{N_0}(j) = \frac{\left(\sum_{n=j}^{n=j+L+1} I_{cr}(n) \right)^2}{2T_{u,cr} \sum_{n=j}^{n=j+L+1} Q_{cr}^2(n)} \quad (22)$$

It can also be sent to the screen at the estimation rate ($T_{u,cr}$ is the residual carrier loop update period). The main disadvantages of that estimator are that first, it requires residual carrier lock, and second, the phase jitter needs to be "small," otherwise, a degraded estimate results, the amount of which highly depends on the residual-carrier loop SNR.

A residual carrier lock indicator can be mechanized by averaging the $I_{cr}(n)$ samples of Eq. (4) over a sufficient period and comparing the result to a threshold to obtain a decision on the lock status. If in-lock is reported, the receiver continues its regular tasks. However, if an out-of-lock status is indicated, an additional verification of the status is performed to lower the probability of a false alarm or indication. This process is well documented elsewhere [2] and is referred to as *sequential detection*. In the case of a final out-of-lock decision, the loop is disabled and an FFT is performed on the complex in-phase and quadrature samples using a separate FFT channel to measure the frequency difference between the signal and the NCO. Since a separate channel is used, the FFT can be performed at

a different rate than the residual carrier update rate and can involve several FFT sweeps which are eventually non-coherently averaged to reduce the effect of the noise; this is particularly useful for very weak signals. A peak detection is then performed on the averaged power spectrum, the NCO is shifted accordingly, and the loop is enabled. FFTs can still be performed once the loop is enabled to verify that frequency acquisition and tracking are actually in progress. Other strategies such as Adaptive Least Squares [22] or Frequency Sweeping [20, 21] have been considered for automating this process so that the receiver can locate the signal, acquire it, track it, verify that the right signal is being tracked by comparing predicts and measurements, monitor its status, and proceed accordingly depending on the lock indicators.

Depending on the value of the modulation index Δ , the carrier might become fully suppressed and thus, needs to be tracked with a Costas loop. The latter forms a phase discriminator by forming the product of $I_{cs}(k)$ and $Q_{cs}(k)$ at the symbol rate and accumulating the result over M_4 symbols to average the phase error. Obviously, the time constant of the phase-error process is much larger than the symbol duration and the averaging does indeed reduce the noise. The performance of the loop in terms of phase-error variance is easily derived and is given by [10]

$$\sigma_{\phi_{cs}}^2 = \frac{N_0 B_{L,cs}^a}{P_D S_{L,BPSK}} \quad (23)$$

where $B_{L,cs}^a$ is the actual loop bandwidth of the suppressed carrier loop and E_s/N_0 is the symbol energy-to-noise ratio equal to $P_D T/N_0$; T is the symbol period defined earlier and $S_{L,BPSK}$ is the Costas loop "squaring loss" given by

$$S_{L,BPSK} = \frac{1}{1 + \frac{1}{2E_s/N_0}} \quad (24)$$

Here again, depending on the difference between the incoming P_D/N_0 and the predicted P_D/N_0 , $(P_D/N_0)_p$, the actual loop bandwidth $B_{L,cs}^a$ might be different from its design counterpart. However, unlike the residual carrier loop, monitoring the loop bandwidth in real time cannot be performed by the loop itself due to the presence of the data modulation but can be accomplished by a separate P_D/N_0 estimator.

A "square law"-type lock detector is used to monitor the BPSK suppressed carrier loop status. This detector is not shown in Fig. 3 on the Costas loop but is shown for the subcarrier loop. The detector generates a signal proportional to $\cos 2\phi_{cs}(k)$ by squaring the samples $I_{cs}(k)$ and $Q_{cs}(k)$, subtracting them, averaging over many sym-

bols and thresholding the result. A sequential detection scheme can also be used to improve the detector's performance. This detector was chosen over the "absolute"-type detector because it requires 1 dB less in symbol energy-to-noise ratio to achieve the same lock-detection probability for a given false-indication rate [18].

B. Sideband-Aiding Performance

In most instances, the incoming carrier is neither a residual nor a fully suppressed carrier, but rather a mixture of the two. In those cases, both the carrier and the data power can be used to perform carrier synchronization and should in principle jointly outperform any individual tracking scenario. This is referred to as *sideband aiding* since the power in the sidebands (or data) is used with the carrier power to perform carrier tracking. In this case, both loop update rates are identical and the joint carrier phase-error estimate, $\hat{\phi}_c$, that ultimately adjusts the frequency of the carrier NCO is given by the weighted sum of the individual estimates as

$$\hat{\phi}_c = \alpha \hat{\phi}_{cr} + (1 - \alpha) \hat{\phi}_{cs} \quad (25)$$

where $\hat{\phi}_{cr}$, $\hat{\phi}_{cs}$ are the individual phase-error estimates by the residual and suppressed carrier loops respectively and α is the optimum weighting factor [10] given by

$$\alpha = \frac{\cos \Delta \sqrt{P_D + N_0/2T}}{\sin \Delta + \cos \Delta \sqrt{P_D + N_0/2T}} \quad (26)$$

Note that the optimum weight α requires knowledge of both P_D and N_0 individually and not just the ratio. When the carrier is fully suppressed (i.e., $\Delta = 90$ deg), $\alpha = 0$ and $\hat{\phi}_c = \hat{\phi}_{cs}$. On the other hand, when the carrier is a pure tone (i.e., $\Delta = 0$ deg), $\alpha = 1$ and $\hat{\phi}_c = \hat{\phi}_{cr}$ as it should. When using the optimum weight α , the performance of the sideband-aided loop becomes

$$\sigma_{\phi_c}^2 = \frac{1}{1/\sigma_{\phi_{cr}}^2 + 1/\sigma_{\phi_{cs}}^2} \quad (27)$$

where $\sigma_{\phi_{cr}}^2$, $\sigma_{\phi_{cs}}^2$ are the variances of the individual tracking loops given by Eqs. (20) and (23), respectively.

C. Performance of the Costas Cross-over Loop

The remaining carrier tracking loop to be discussed is the QPSK loop, which also processes the $I_{cs}(k)$ and $Q_{cs}(k)$

samples to provide a phase-error estimate. The received signal in this case is given by

$$r(t) = \sqrt{P} [d_1(t) \sin(\omega_c t + \theta) + d_2(t) \cos(\omega_c t + \theta)] + n(t) \quad (28)$$

where $d_1(t)$, $d_2(t)$ are the in-phase and quadrature modulations similar to Eq. (2) and T the QPSK symbol duration. The implemented loop is the so-called Costas cross-over loop, which is an approximation to the maximum-likelihood (ML) estimator at high signal-to-noise ratios.

The error signal that forms the input to the loop filter is obtained by averaging M_3 samples of the form

$$e(k) = I_{cs}(k) \text{sgn}(Q_{cs}(k)) - Q_{cs}(k) \text{sgn}(I_{cs}(k)) \quad (29)$$

The average can be shown to be proportional to the phase error ($\phi_{QPSK}(k)$), assuming a linearized model. The performance of the loop has been derived in various places and is given by [19]

$$\sigma_{\phi_{QPSK}}^2 = \frac{1}{\rho S_{L,QPSK}} \quad (30)$$

where $\rho = P/N_0 B_L$ is the loop signal-to-noise ratio of a "classical" PLL and $S_{L,QPSK}$ is the QPSK loop "squaring" loss. It can be shown that [19]

$$S_{L,QPSK} = \frac{[\Phi(\sqrt{R_d/2}) - \sqrt{2R_d/\pi} e^{-R_d/2}]^2}{1 + R_d - 2 \left[\frac{1}{\sqrt{\pi}} e^{-R_d/2} + \sqrt{R_d/2} \Phi(\sqrt{R_d/2}) \right]^2} \quad (31)$$

where $R_d = PT/N_0 = E_s/N_0$ is the QPSK symbol energy-to-noise ratio and $\Phi(x)$ is the error function given by

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (32)$$

This loop was chosen for simplicity of implementation.

The other two candidates, the ML loop and its low SNR approximation, require more hardware to implement and do not offer significant advantages. The ML loop requires implementation of the hyperbolic tangent function, which can be achieved with read-only memory (ROM) chips, but the input to the ROM chips has to be scaled by the signal amplitude and the noise spectral level. These are not known to the receiver and therefore must be esti-

mated. The sensitivity of the performance of the ML loop with respect to those parameters is not known and needs to be determined before the loop can be chosen as an implementable candidate. The generalized Costas loop (which is a low SNR approximation to the ML loop) requires implementing squaring operations, and thus scaling is also required to handle the wide dynamic range of signal-to-noise ratios and symbol rates involved. The Costas cross-over loop is the easiest to implement since the signum functions use only the sign bit of the input signals and are thus very straightforward to implement digitally. The remaining operations involve multiplications, additions, and accumulations, which are easily done. Furthermore, the Costas cross-over loop does not suffer from a significant degradation in performance [19] in the region of interest.

D. Subcarrier Loop Performance

Once carrier acquisition and tracking are achieved, subcarrier (if present) and symbol synchronization are still required to eventually detect the received symbols. Note that a subcarrier will not be used with QPSK, but only with nonsuppressed BPSK (i.e., $\Delta \neq 90$ deg), in which case a residual carrier component is present. The subcarrier can be either a square wave or a sine wave and is used to shift the data spectrum away from the residual component. In fact, bi-phase modulation can be thought of as a subcarrier which shifts the data away from the carrier by $1/T$ Hz. Often, one would like to shift the data by more than $1/T$ Hz so that enough guard band is included to reduce the interference from the tails of the data spectrum.

The Costas loop described earlier can be used to track both sine-wave and square-wave subcarriers. In this case, an averaged, filtered version of the product of $I_{sc,mf}(k)$ and $Q_{sc,mf}(k)$ is used to drive the subcarrier NCO. For square-wave subcarriers, a windowing operation can be performed on the Q-channel. This results in improved tracking performance, but smaller acquisition pull-in range [17]. In that case, the phase jitter is given by

$$\sigma_{\phi_{sc}}^2 = \left(\frac{\pi}{2}\right)^2 W_{sc} \left(\frac{N_0 B_{sc}^a}{P_D}\right) (1 + N_0/2E_s) \quad (33)$$

while for a sine-wave subcarrier, one has

$$\sigma_{\phi_{sc}}^2 = \left(\frac{N_0 B_{sc}^a}{P_D}\right) (1 + N_0/2E_s) \quad (34)$$

where B_{sc}^a is the actual subcarrier loop bandwidth and W_{sc} the subcarrier window. A square-law detector is used to indicate subcarrier lock to the operator and to facilitate system automation.

E. Symbol Synchronization Performance

For both the suppressed carrier and subcarrier loops to function properly, symbol epoch tracking has to be maintained so that the loop arm-filter outputs are actually those given by Eqs. (16) and (17). The task of symbol synchronization is accomplished by the data-transition tracking loop (DTTL), which uses the $I_{sc}(n)$ samples of Eq. (11) as its input. The DTTL was chosen because of its simplicity and improved performance when using a window. The other two candidates, the "Absolute Value Type of Early-Late Gate Symbol Synchronizer" (AVTS) and the "Difference-of-Squares Loop" (DSL) [20], are harder to implement and do not offer significant advantages.

In the DTTL, two accumulators are used, one to detect the received symbols and the other to accumulate over symbol transitions. The first is typically referred to as the "in-phase" filter and the other as the "mid-phase" filter (WIDF with symbol transition window in Fig. 3). Following the in-phase operation, a transition detector outputs a 0 or ± 1 to indicate a no-transition, a $+1$ to -1 , or a -1 to $+1$ transition respectively. The output is then multiplied by the mid-phase accumulator to wipe out the effect of the data and accumulated over several symbols to reduce the rate from the symbol to the loop update rate. Note that a windowing operation can be performed during the mid-phase accumulation and that results in a decrease in the loop phase jitter.

The performance of the loop when the arm filters are analog devices has been evaluated elsewhere [20] and the phase-error variance is given approximately by

$$\sigma_{\phi_{sy}}^2 \simeq 2\pi^2 W_{sy} \left(\frac{N_0}{E_s}\right) \left(\frac{B_{sy}^a}{R_s}\right) \frac{1}{\Phi^2(\sqrt{E_s/N_0})} \quad (35)$$

where B_{sy}^a is the actual symbol synchronization loop bandwidth, W_{sy} the symbol transition window, and R_s the symbol rate. In the ARX II, the "typical" analog arm filters are digital accumulators and the overall loop performance is still given by Eq. (35) as long as the number of samples per symbol (denoted by β) remains "large." However, the number of samples per symbol is not always "large" since at the high data-rate goal of 6.6 Msymbols/sec, only 3 to 4 samples per symbol are available (because of the 20-MHz processing rate). Furthermore, since the sampling clock is fixed and not driven by the symbol loop, β is not in general an integer even though it is a real number.

This problem was examined in a separate study [25]. It was concluded that the loop phase jitter will remain

small as long as "some" Doppler rate is present in the communication link. In the deep-space applications of interest, Doppler rate due to Earth rotation is enough to help the DTTL achieve sufficiently small phase jitter. The main problem for the DTTL occurs when the number of samples per symbol is exactly a small integer number, for example, $\beta_I = 4$. (When β is an integer, it will be denoted by β_I). In this case, the phase error increases dramatically due to the resolution offered by exactly 4 samples/symbol and Eq. (35) is no longer valid due to the additional error of the "self-noise," which was not accounted for. That error is equal to the variance of a uniformly distributed random variable over $(-1/2\beta_I, 1/2\beta_I)$, which for $\beta_I = 4$ gives a 73 mcycle root-mean-squared (rms) error! (In general, the rms error due to the self-noise of an integer number of samples per symbol is given by $1/(2\sqrt{3}\beta_I)$.)

Note that the true symbol phase jitter is not the sum of the white-noise and the self-noise jitters. Rather, it is related to both through a nonlinear equation involving the integral of the density function of the phase-error process. Only in special circumstances is the sum assumption a "good" approximation, but it can be used throughout as a rough rule of thumb. In real communication links, nonzero Doppler rate is typically present and is enough to guarantee either that β will not be an integer or that it will be an integer for a "short" period of time only. In the latter case, the symbol synchronization loop will not be affected due to the long averaging performed with the "typical" small loop bandwidth (1-100 mHz) it operates with.

F. Symbol SNR Estimator Performance

One way to monitor the symbol-synchronization loop performance in real time is through the use of a symbol SNR estimator. This is different from other loops where a binary decision is provided on the status of the loop. In this case, the estimator will continuously provide real-time estimates of the symbol SNR, which are compared to the expected symbol SNR. If the numbers are "close," symbol phase lock is assumed and system operation is resumed. If the numbers are different, then three cases are possible: either the loop is out of lock, the received symbol SNR is different from that expected, or, the estimator itself is providing misleading data due to its own limitation. From a user's point of view, these possible scenarios are too confusing and an ultimate binary decision is highly desirable. The most disturbing of the three outcomes is the case involving misleading data from the SNR estimator. This is obviously highly dependent on the estimator structure used and thus an understanding of the algorithm is required.

The algorithm used in the ARX II is the so-called "Split Symbol Moment Symbol SNR Estimator" (SSME) [26], which was originally designed to operate in an additive white Gaussian noise (AWGN) channel. A table look-up technique [27] could also be used, but the SSME was chosen due to its simplicity.

The input to the estimator is the data stream after it has been demodulated, i.e., the sequence of $I_{s,c}(n)$ samples. Assuming that perfect carrier and subcarrier lock has been achieved, the samples become

$$I_{s,c}(n) \simeq \sqrt{P_D}d(n) + N_I(n) \quad (36)$$

These samples are accumulated over a symbol duration by the matched filters in the various loops, where the accumulation start and stop times are provided by the DTTL. In fact, the carry-out bit of the symbol NCO can be used to clock all the accumulators that operate over a symbol duration. The SSME uses those samples and accumulates them over the first and second halves of a symbol period using a control signal derived from the symbol NCO phase. These are shown in Fig. 3 as "First-Half WIDF" and "Second-Half WIDF." By further processing these outputs, an averaged SNR estimate can be obtained over many symbols.

The principle of operation of the algorithm is as follows. A total (signal plus noise)-power estimate is obtained by summing, squaring, and averaging the outputs of the half WIDFs. Simultaneously, a signal-power estimate is obtained by forming the product of the half-WIDF outputs and then averaging over many symbols. A noise-power estimate is then easily obtained by subtracting the previous estimates with the appropriate weighting. Finally, a symbol-SNR estimate is derived from the separate signal- and noise-power estimates by taking the ratio. It was shown [28] that the algorithm works well when operating at low symbol rates and in the presence of additive white Gaussian noise (AWGN). However, at high symbol rates, filtering of the data becomes significant and does result in erroneous estimates. These estimates depend on β (number of samples per symbol), the data rate, and the effective filtering in the data path. This is because the SSME is operating in a region that it was not designed for. The user should be aware of this limitation of the SSME and act accordingly.

G. Weighted Integrate-and-Dump Filters

Another algorithm requiring description is the weighted integrate-and-dump filter (WIDF). It is well known in detection theory that the optimum detector in an AWGN channel is the matched filter followed by a sampler.

The filter's impulse response is matched to the incoming signal such that at the sampling instant maximum symbol SNR is achieved.

In digital circuits, one way to detect the symbols is to accumulate the samples over a symbol duration. By accumulation is meant a straight summation. This will maximize the symbol SNR as long as all the samples have the same SNR. However, this is rarely the case because the first and last samples of a symbol are affected by the rise and fall times of the pulse. Consequently, they will have different pulse amplitudes and thus, different sample SNRs.

When "many" samples per symbol are available, the first and last samples form a minority of all the samples and a straight accumulation will probably provide the maximum symbol SNR that can be achieved. However, when only three or four samples per symbol are taken, the first and last samples form a substantial percentage of the samples and thus need to be accounted for. This is accomplished by weighting the samples while accumulating them to improve the symbol SNR. This is easily implemented using a multiply-and-accumulate chip (MAC) with the weights being read from a memory chip.

The optimum weights depend on the filtering involved, on the number of samples per symbol, and on where the first sample occurs within the symbol (typically referred to as sampling offset); the latter is estimated by the symbol-synchronization loop. Assuming perfect knowledge of the sampling offset, it has been shown that weighting provides an average 0.2-dB improvement in symbol SNR over the straight accumulation (in DSN applications, this is a significant improvement). Moreover, when the sampling offset is exactly zero, the improvement can be as large as 1.5 dB in symbol SNR [28, 29].

In a practical system, less gain would be achieved due to the jitter in estimating the sampling offset and due to not knowing the exact filtering taking place. Nevertheless, some gain is expected in telemetry SNR since the weights are not assumed to be very sensitive to those parameters. The effect of filter distortion on telemetry SNR has also been examined [31] and it was determined that an amplitude ripple of less than 0.42 dB would guarantee less than 0.01-dB loss in symbol SNR.

V. Implementation Trade-off Issues

The previous section provided a description of the various algorithms employed to perform the necessary tasks to demodulate the received signal and monitor the receiver's performance. Implementation issues were pur-

posely avoided to give the reader a clear understanding of each algorithm's performance based on theoretical results. In this section, some of the characteristic losses inherent to our implementation are discussed and clarified.

The starting point of the digital hardware is the A/D operation, which operates at 40 MHz using eight bits. As far as SNR losses due to signal quantization are concerned, the use of four or more bits has negligible effect on sample SNR. In fact, for four bits, the loss is about 0.04 dB and is practically zero for more bits. Eight bits were used in the ARX II to provide some protection against the presence of relatively strong interference. For example, for an interference-to-noise ratio (at the input to the A/D converter) of 32.6 dB, 8 and 4 bits result in about 0.01-dB and 12.36-dB losses respectively. The loss assumes that the interference contributed only to quantization noise. A more realistic scenario would be to consider the effect of some spectral components due to interference in the signal frequency band, as this would create additional in-band noise.

Following the digital conversion, carrier demodulation is performed followed by low-pass filtering. At the output of the low-pass filter, the signal spectrum extends up to 10 MHz only and thus a sampling rate of 20 MHz is sufficient to perform further signal processing. The combination of low-pass filtering at 40 MHz and sample decimation by two can be implemented in an equivalent processing at 20 MHz [32] using half-band FIR filters. The filter response is shown in Figs. 4 (a) and (b) along with its finite-bit implementation. It is clear that a 9-bit implementation would not result in significant changes in filter response whereas an 8-bit version would contribute more pass-band ripple and higher sidebands.

Because of the finite-bit implementation, scaling is required throughout the receiver to accommodate the wide range of signal-to-noise ratios and symbol rates anticipated. As an example, most accumulators accept 16-bit inputs and output a 16-bit result. The accumulation is usually performed using 32 bits but only 16 of those can be accessed at the maximum chip speed of 20 MHz. Typically, the most significant 16 bits are hardwired to the next stage and software-controlled scaling is used to make sure that the result of the accumulation does indeed lie in that range. As mentioned previously, all loop filters are implemented in software and their outputs are the various frequency errors that need to be compensated for by the respective NCOs.

In the digital loop closure, the fractional phase is accumulated internally using 32 bits but only the most significant 12 bits are used in the look-up table, resulting in a

0.24-mcycle resolution. The phase is recorded using 48 bits for the integer part and 12 bits for the fractional part. The NCO is running at 10 MHz with a 40-MHz clock, resulting in 9.3-mHz ($4 \times 10^7/2^{32}$) frequency resolution.

In the analog loop closure, the fractional phase is recorded using 14 bits (resulting in an improved 0.06-mcycle resolution), two of which are accumulated externally using the carry-out bit of the NCO. The NCO is actually running at 16 MHz with a 20-MHz clock, but it effectively produces a 4-MHz signal because of the additional bits. The latter also improves the frequency resolution to 1.16 mHz ($2 \times 10^7/2^{34}$), but limits the maximum nominal frequency to one fourth of the clock rate (the practical limit of an NCO is roughly 40 percent of the clock). That in return limits the dynamic excursion of the frequency to ± 1 MHz in the analog loop closure, whereas it is about ± 5 MHz in the digital closure. Time tagging is performed using ± 1 μ sec accuracy with a ± 1 nsec stability. The logging rate can be chosen by software from 1, 2, 5, 10, or 20 Hz.

Note that in both analog and digital loop closures, 16-bit sine and cosine numbers are generated to reduce the in-phase and quadrature DC offsets [32] to an acceptable

level. Since the subcarrier frequency and symbol rate are always less than the IF carrier frequency, the frequency resolution of their respective NCOs will be much higher because they also use 32-bit commands. From the symbol and the subcarrier NCOs, only the most significant 8 bits of fractional phase (3.9-mcycle resolution) are used to generate the subcarrier waveform and to clock all the WIDFs. This will limit the number of windows that can be used to eight, with W_{sc} or W_{sy} equal to 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\frac{1}{16}$, $\frac{1}{32}$, $\frac{1}{64}$, or $\frac{1}{128}$ (7.81 mcycle) only. This limit does not affect the system performance in any way, and the number of available windows can be increased at the cost of additional hardware.

VI. Conclusion

This article provided a functional description of the Advanced Receiver currently under development for future use in NASA's DSN. The requirements were specified and the receiver's architecture described. Moreover, the various signal-processing schemes were briefly discussed and their limitations clarified. The receiver incorporates functions currently available in a variety of receivers and does so digitally and with improved performance and options for spacecraft designers.

Acknowledgment

I thank the ARX II design team, which includes Mr. J. Statman, Mr. Hector Del Castillo, Mr. Roland Bevan, and Mr. Remi Labelle.

References

- [1] D. H. Brown and W. J. Hurd, "DSN Advanced Receiver: Breadboard Description and Test Results," *TDA Progress Report 42-89*, vol. January–March 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 48–66, May 15, 1987.
- [2] J. H. Yuen, *Deep Space Telecommunication Systems Engineering*, New York: Plenum Press, 1983.
- [3] D. H. Brown, W. J. Hurd, V. A. Vilnrotter, and J. Wiggins, "Advanced Receiver Tracking of Voyager II Near Solar Conjunction," *TDA Progress Report 42-93*, vol. January–March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 75–82, May 15, 1988.
- [4] W. J. Hurd, D. H. Brown, V. A. Vilnrotter, and J. D. Wiggins, "Telemetry SNR Improvement Using the DSN Advanced Receiver with Results From Pioneer 10," *TDA Progress Report 42-93*, vol. January–March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 64–82, May 15, 1988.
- [5] V. A. Vilnrotter, W. J. Hurd, and D. H. Brown, "Optimized Tracking of RF Carriers with Phase Noise, Including Pioneer 10 Results," *TDA Progress Report 42-91*, vol. July–September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 141–157, November 15, 1987.
- [6] V. A. Vilnrotter, D. H. Brown, and W. J. Hurd, "Spectral Estimation of Received Phase in the Presence of Amplitude Scintillation," *TDA Progress Report 42-93*, vol. January–March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 8–17, May 15, 1988.
- [7] V. Vilnrotter, S. Hinedi, and R. Kumar, "Frequency Estimation Techniques for High Dynamic Trajectories," *IEEE Trans. on AES*, vol. AES-25, no. 4, pp. 559–577, July 1989.
- [8] S. Aguirre and S. Hinedi, "Two Novel Automatic Frequency Tracking Loops," *IEEE Trans. on AES*, vol. AES-25, no. 5, pp. 749–760, September 1989.
- [9] S. Hinedi and J. Statman, "Digital Accumulators in Phase and Frequency Tracking Loops," *IEEE Trans. on AES*, vol. AES-26, no. 1, pp. 1–13, January 1990.
- [10] R. Sfeir, S. Aguirre, and W. J. Hurd, "Coherent Digital Demodulation of a Residual Carrier Signal Using IF Sampling," *TDA Progress Report 42-78*, vol. April–June 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 135–142, August 15, 1984.
- [11] S. Aguirre, W. J. Hurd, R. Kumar, and J. Statman, "A Comparison of Methods for DPLL Loop Filter Design," *TDA Progress Report 42-87*, vol. July–September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 114–124, November 15, 1986.
- [12] C. A. Pomalaza-Raez and W. J. Hurd, "Improved Carrier Tracking by Smoothing Estimators," *TDA Progress Report 42-79*, vol. July–September 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 96–106, November 15, 1984.
- [13] C. A. Pomalaza-Raez and W. J. Hurd, "Carrier Tracking by Smoothing Filter Can Improve Symbol SNR," *TDA Progress Report 42-83*, vol. July–September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 57–62, November 15, 1985.

- [14] R. Kumar and W. J. Hurd, "A Class of Optimum Digital Phase Locked Loops for the DSN Advanced Receiver," *TDA Progress Report 42-83*, vol. July–September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 63–80, November 15, 1985.
- [15] J. I. Statman, "A Recursive Solution for a Fading Memory Filter Derived From Kalman Filter Theory," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 70–76, August 15, 1986.
- [16] J. I. Statman and W. J. Hurd, "An Estimator-Predictor Approach to PLL Loop Filter Design," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 77–89, August 15, 1986.
- [17] W. J. Hurd and S. Aguirre, "A Method to Dramatically Improve Subcarrier Tracking," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 103–110, August 15, 1986.
- [18] A. Mileant and S. Hinedi, "Costas Loop Lock Detection in the Advanced Receiver," *TDA Progress Report 42-99*, vol. July–September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 72–89, November 15, 1989.
- [19] M. K. Simon, "On the Optimality of the MAP Estimation Loop for Carrier Phase Tracking BPSK and QPSK Signals," *IEEE Trans. on Comm.*, vol. COM-27, no. 1, pp. 158–165, January 1979.
- [20] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, New Jersey: Prentice-Hall Inc., 1973.
- [21] S. Aguirre and W. J. Hurd, "Design and Performance of Sampled Data Loops for Subcarrier and Carrier Tracking," *TDA Progress Report 42-79*, vol. July–September 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 81–95, November 15, 1984.
- [22] R. Kumar, "Fast Frequency Acquisition via Adaptive Least Squares Algorithm," *TDA Progress Report 42-85*, vol. January–March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 53–60, May 15, 1986.
- [23] S. Aguirre, D. H. Brown, and W. J. Hurd, "Phase Lock Acquisition for Sampled Data PLL's Using the Sweep Technique," *TDA Progress Report 42-86*, vol. April–June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 95–102, August 15, 1986.
- [24] S. Aguirre, "Acquisition Times of Carrier Tracking Sampled Data Loops," *TDA Progress Report 42-84*, vol. October–December 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 88–93, February 15, 1985.
- [25] U. Cheng and S. Hinedi, "Performance of the All-Digital Data Transition Tracking Loop in the Advanced Receiver," *TDA Progress Report 42-99*, vol. July–September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 60–71, November 15, 1989.
- [26] M. K. Simon and A. Mileant, "SNR Estimation for the Baseband Assembly," *TDA Progress Report 42-85*, vol. January–March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 118–126, May 15, 1986.
- [27] V. A. Vilnrotter and E. R. Rodemich, "Table Look-Up Estimation of Signal and Noise Parameters From Quantized Observables," *TDA Progress Report 42-87*, vol. July–September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 104–113, November 15, 1986.

- [28] B. Shah and S. Hinedi, "Performance of the Split Symbol Moments SNR Estimator in the Presence of Inter-Symbol Interference," *TDA Progress Report 42-98*, vol. April-June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 157-173, July 15, 1989.
- [29] R. Sadr and W. J. Hurd, "Detection of Signals by the Digital Integrate-and-Dump Filter With Offset Sampling," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 158-173, November 15, 1987.
- [30] R. Sadr, "Detection of Signals by Weighted Integrate-and-Dump Filter," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 174-185, November 15, 1987.
- [31] R. Sadr and W. J. Hurd, "Filter Distortion Effect on Telemetry SNR," *TDA Progress Report 42-88*, vol. October-December 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 59-66, February 15, 1986.
- [32] R. Sadr and W. J. Hurd, "Digital Carrier Demodulation for the DSN Advanced Receiver," *TDA Progress Report 42-93*, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 45-63, May 15, 1988.

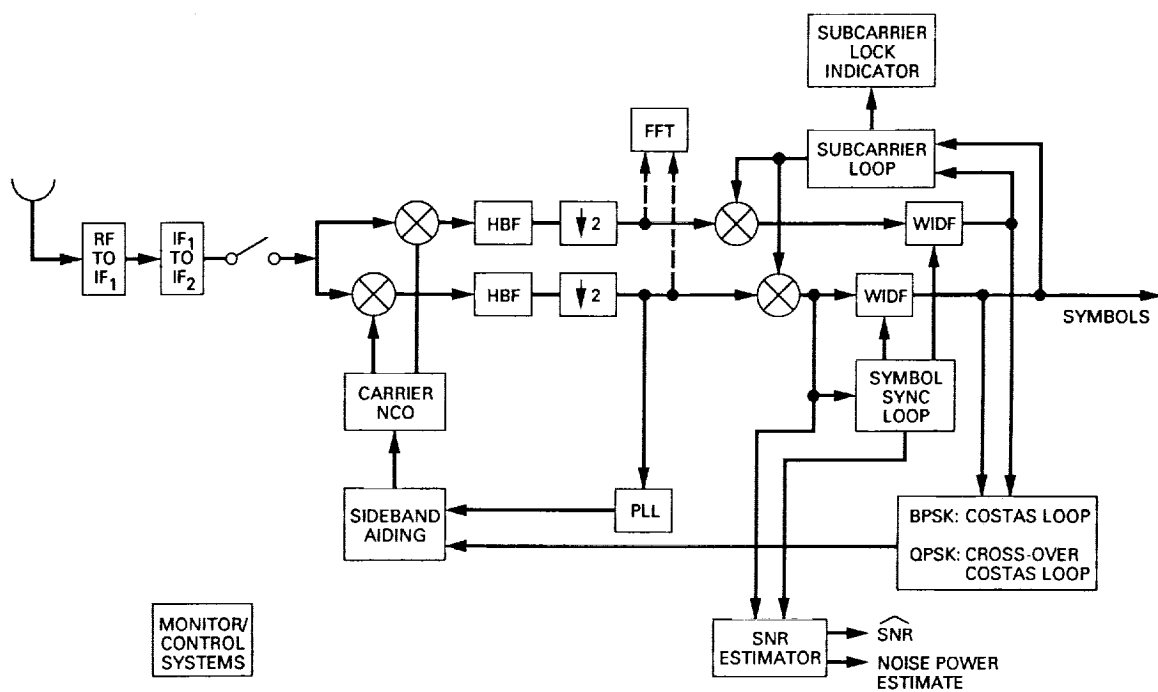


Fig. 1. Top-level block diagram of the ARX II.

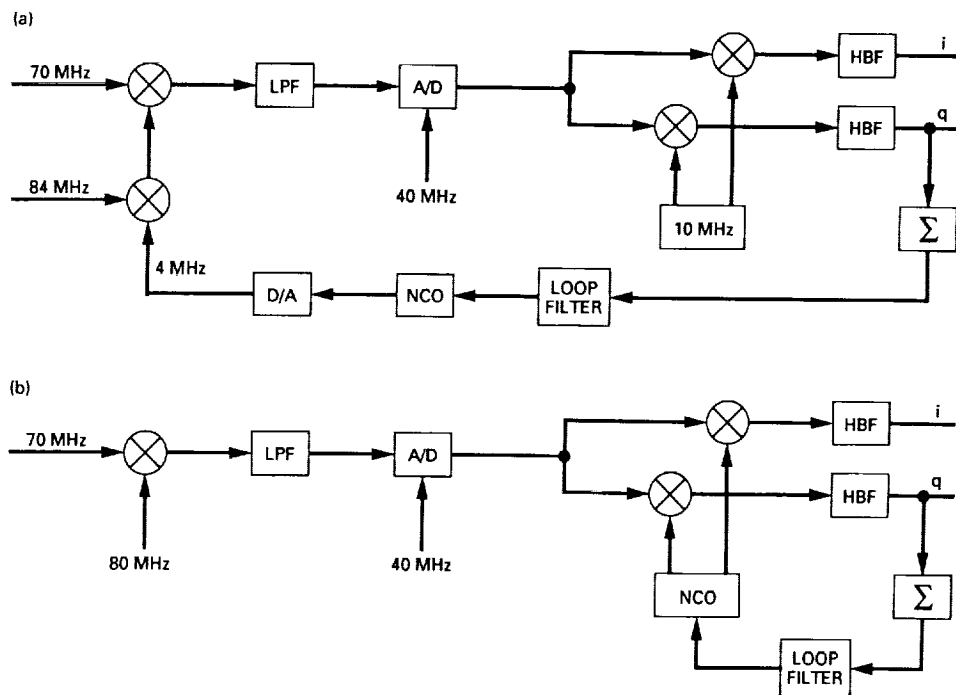
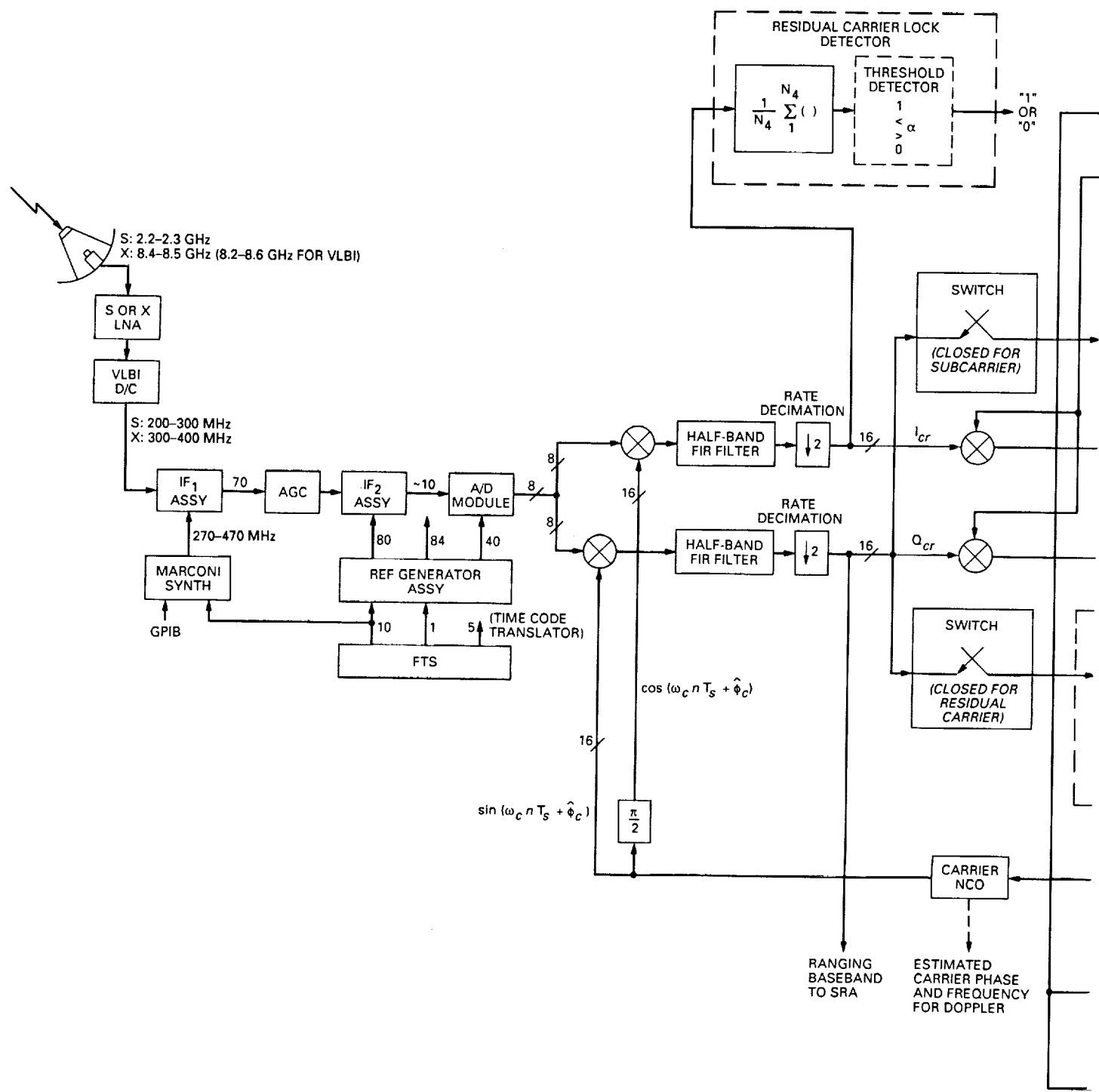


Fig. 2. Loop closures: (a) analog and (b) digital.

FOLDOUT FRAME /



FOLDOUT FRAME 2

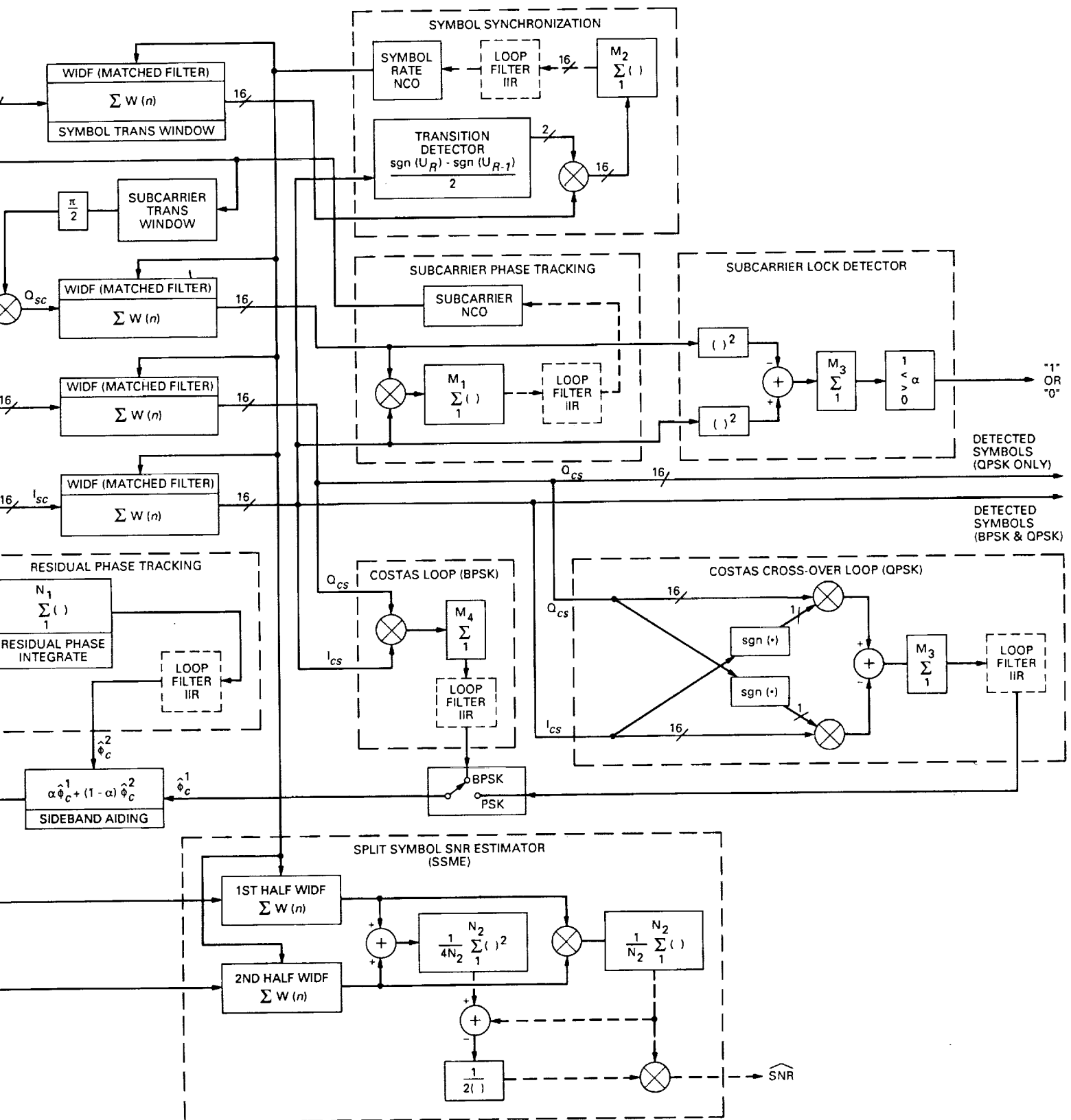


Fig. 3. ARX II functional block diagram.

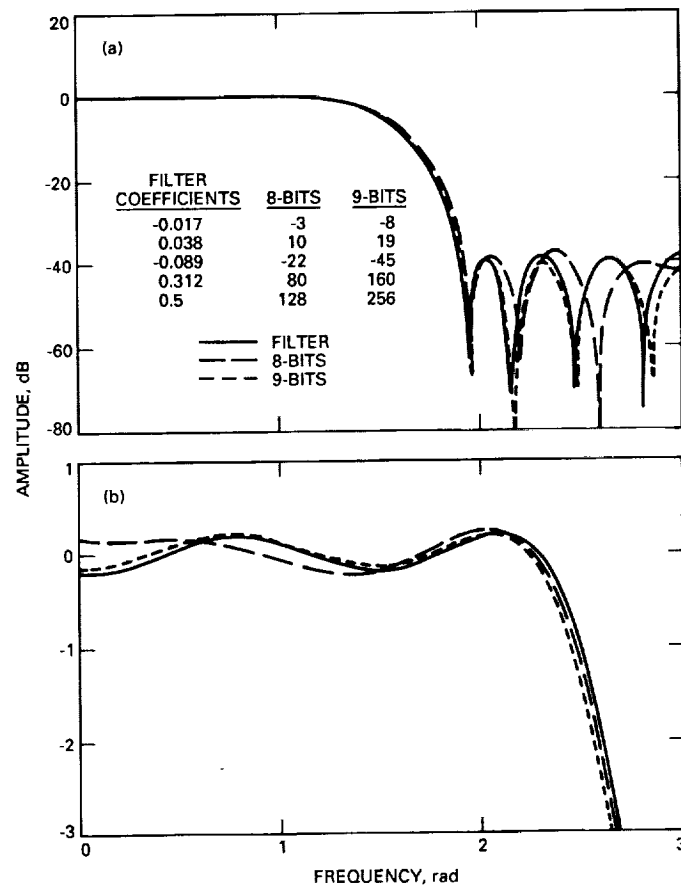


Fig. 4. Half-band filter: (a) amplitude response and (b) amplitude ripple.

QPSK Carrier-Acquisition Performance in the Advanced Receiver II

S. Hinedi

Communications Systems Research Section

B. Shah

Radio Frequency and Microwave Subsystems Section

This article describes the frequency-acquisition performance of the Costas cross-over loop which is used in the Advanced Receiver II (ARX II) to perform Quadrature Phase Shift Keying (QPSK) carrier tracking. The performance of the Costas cross-over loop is compared to two other QPSK carrier tracking loops: the MAP estimation loop and the generalized Costas loop. Acquisition times and probabilities of acquisition as functions of both loop signal-to-noise ratio and frequency-offset to loop-bandwidth ratio are obtained using computer simulations for both type-II and type-III loops. It is shown that even though the MAP loop results in the smallest squaring loss for all signal-to-noise ratios, the MAP loop is sometimes outperformed by the other two loops in terms of acquisition time and probability.

I. Introduction

Quadrature Phase Shift Keying (QPSK) is an efficient modulation scheme which has been used extensively in communications systems. It offers a significant advantage over Binary Phase Shift Keying (BPSK) since it permits the transmission of twice the data rate in the same channel bandwidth [1], as both the in-phase and the quadrature components of the carrier are modulated. An even better use of bandwidth can be obtained by going to higher-order phase shift keying, but that would require a higher signal-to-noise ratio to achieve a required bit-error probability.

Depending on the application, derivatives of QPSK can also be used; these include Staggered QPSK (SQPSK), which is better suited for satellite channels with Travelling Wave Tube (TWT) amplifiers and Minimum Shift Keying (MSK), which can be thought of as SQPSK with sinusoidal pulse shapes.

One of the functional requirements of the Advanced Receiver II (ARX II) is that it acquire and track QPSK signals. The most likely candidate algorithm for the ARX II is the Costas cross-over loop because of its ease of implementation and relative performance. Carrier synchroniza-

tion for QPSK has been addressed by many authors in the literature and several loops have been introduced and analyzed in terms of their respective phase-error variance [2-4]. However, very little has been reported on the acquisition performance of these loops, mainly because the problem becomes highly nonlinear and precludes any analysis that provides insight into loop behavior. The suppressed carrier tracking loops are typically derived from Maximum A Posteriori (MAP) estimation theory; the derivative of the likelihood function is used as an error signal to provide a closed-loop implementation commonly referred to as *MAP estimation loops* [2].

Three QPSK loops are considered in this article in terms of their acquisition performance: the MAP estimation loop, the Costas cross-over loop (also known as the polarity-type Costas loop), and the generalized Costas loop (sometimes referred to as the conventional Costas loop). Mean acquisition times and probabilities of acquisition as functions of both loop signal-to-noise ratio (SNR) and frequency-offset-to-loop-bandwidth ratio are investigated using computer simulations for type-II and type-III loops.

The loops are all shown in Fig. 1 where different processing of the in-phase and quadrature channels identifies which loop is being implemented. The MAP estimation loop produces an error signal given by

$$Z_0 = Z_c \tanh(RZ_s) - Z_s \tanh(RZ_c) \quad (1)$$

where R is the symbol SNR and Z_c , Z_s are the in-phase and quadrature signals respectively. This loop is "difficult" to build because it requires a priori knowledge of the symbol SNR and implementation of the hyperbolic tangent operations. The latter is hard to achieve in an analog fashion, but is easily done digitally with current technology using read-only memory (ROM) chips. To reduce the complexity of the MAP estimation loop, approximations to the hyperbolic tangent function can be used at low and high symbol SNRs. For instance, at high symbol SNRs, the hyperbolic tangent can be approximated by the signum function ($\tanh(x) \simeq \text{sgn}(x)$) to yield the Costas cross-over loop with error signal

$$Z_0 = Z_c \text{sgn}(Z_s) - Z_s \text{sgn}(Z_c) \quad (2)$$

At low symbol SNRs, the nonlinearity hyperbolic tangent can be expanded using a Taylor series to give $\tanh(x) \simeq x - x^3/3$, which results in the generalized Costas loop with Z_0 given by

$$Z_0 = Z_c Z_s (Z_c^2 - Z_s^2) \quad (3)$$

The two approximate loops are straightforward to implement and do not require knowledge of the symbol SNR. However, all three loops require an estimate of the incoming signal amplitude in order to operate the loop at the desired bandwidth.

Section II of this article develops a mathematical model for the various signals in Fig. 1. Simulation results are presented and discussed in Section III followed by the conclusion in Section IV.

II. Mathematical Model

In a typical communication system, the received signal is first downconverted to an appropriate intermediate frequency (IF) for further processing. At that point, the QPSK signal can be modeled as

$$r(t) = \sqrt{S}[d_I(t)\sin(\omega_i t + \theta) + d_Q(t)\cos(\omega_i t + \theta)] + n(t) \quad (4)$$

where S is the received power (watts), ω_i the IF radian frequency (rad/sec), θ the signal phase (rad), and $d_I(t)$, $d_Q(t)$ the in-phase and quadrature data streams given by

$$d_i(t) = \sum_{k=-\infty}^{+\infty} a_{k,i} p(t - kT) \quad i = I, Q \quad (5)$$

where $p(t)$ is the baseband Non-Return-To-Zero (NRZ) pulse limited to T seconds and $a_{k,i}$ the equally likely ± 1 binary symbols. For SQPSK, the in-phase and quadrature baseband pulses will be offset by half a symbol period and for MSK, the pulses will be sinusoidal. The narrow-band noise process $n(t)$ can be expressed as

$$n(t) = \sqrt{2}n_c(t)\cos(\omega_i t + \theta) - \sqrt{2}n_s(t)\sin(\omega_i t + \theta) \quad (6)$$

where $n_c(t)$ and $n_s(t)$ are statistically independent stationary band-limited white Gaussian noise processes with one-sided spectral density N_0 (watts/Hz) and one-sided bandwidth W (Hz). Assuming a digital implementation, the signal is first digitized at a high rate of R_s samples per sec ($T_s = 1/R_s$ is the sampling time) and subsequently demodulated using the in-phase and quadrature references (see Fig. 1). The resulting samples are then accumulated over a symbol duration and normalized to produce

$$Z_s(iT) = a_{i,I}\cos\phi - a_{i,Q}\sin\phi - N_I\cos\phi - N_Q\sin\phi \quad (7)$$

$$Z_c(iT) = a_{i,I}\sin\phi + a_{i,Q}\cos\phi - N_I\sin\phi + N_Q\cos\phi \quad (8)$$

where $\phi = \theta - \hat{\theta}$ is the phase error and N_I, N_Q are zero-mean Gaussian random variables with variances N_0/TS . It is assumed that perfect symbol synchronization has been achieved and no attempt is made to quantify the results of timing jitter on the carrier-acquisition process in this preliminary study. However, the acquisition performance with unknown symbol epoch is of prime interest as it reflects the carrier-acquisition process in a practical system. Moreover, the various loops might react differently in the absence of symbol synchronization and a complete study incorporating those effects is necessary. Depending on the loop, the error signal Z_0 is computed from Z_c and Z_s using the appropriate equations as discussed earlier. Further averaging over N symbols can be performed before updating the loop filter, which in turn controls the frequency of the numerically controlled oscillator (NCO).

The tracking performances of these loops have been derived elsewhere and are repeated below for convenience. In all cases, the variance of the phase-error process is given by

$$\sigma_\phi^2 = \frac{1}{\rho_c S_L} = \frac{1}{\rho} \quad (9)$$

where ρ is the loop SNR of the QPSK loop, $\rho_c = S/N_0 B_L$ the loop SNR of a "classical" phase-locked loop [1], and S_L is the "squaring" loss which for the MAP estimation loop is given by [2] as

$$S_L =$$

$$\frac{[\tanh(R - \sqrt{R}x) - R \operatorname{sech}^2(R - \sqrt{R}x)]^2}{(1 + R)\tanh^2(R - \sqrt{R}x) - [(x - \sqrt{R})\tanh(R - \sqrt{R}x)]^2} \quad (10)$$

where x is a zero-mean, unit-variance Gaussian random variable and the overbar denotes statistical expectation. For the Costas cross-over loop [3] ($\operatorname{erf}(x) \triangleq (2/\sqrt{\pi}) \int_0^x e^{-t^2} dt$ denotes the error function), S_L becomes

$$S_L = \frac{[\operatorname{erf}(\sqrt{R/2}) - \sqrt{2R/\pi} e^{-R/2}]^2}{1 + R - [\sqrt{2/\pi} e^{-R/2} + \sqrt{R} \operatorname{erf}(\sqrt{R/2})]^2} \quad (11)$$

and

$$S_L = \frac{1}{1 + \frac{9}{2R} + \frac{6}{R^2} + \frac{3}{2R^3}} \quad (12)$$

for the conventional Costas loop [4]. In all cases, S_L is a function of only one parameter, the symbol SNR R defined as ST/N_0 . Figure 2 depicts the "squaring" losses for the various loops; it is clear that the MAP estimation

loop outperforms the other two loops for all values of R . For $R \geq 8$ dB, the performance of the Costas cross-over loop approaches that of the MAP loop. Similar arguments apply for the generalized loop at low SNRs ($R \leq -4$ dB). The analytical results were reproduced through computer simulations to obtain greater confidence in the programs before generating new results on acquisition.

III. Simulation Model, Results, and Discussion

The loops in Fig. 1 are simulated by generating the in-phase and quadrature signals Z_s and Z_c according to Eqs. (7) and (8). Depending on the loop under consideration, the instantaneous error signal Z_0 is computed using Eqs. (1), (2), or (3). The simulations were carried out assuming $N = 1$, i.e., the loop update rate is identical to the symbol rate. In this case, the NCO transfer function is given by

$$N(z) = \frac{T_u z + 1}{2z - 1} \quad (13)$$

(where T_u denotes the loop update time) and it relates the filter output to the incoming phase estimate in the z -domain. In some digital implementations, the NCO might have an additional delay, in which case the loop performance will be slightly worse than the one considered here. Unlike tracking, which depends solely on the loop bandwidth, acquisition behavior is a function of the loop filter itself and not just the loop bandwidth. The simulations were carried out with the following filter

$$F(z) = G_1 + \frac{G_2}{1 - z^{-1}} + \frac{G_3}{(1 - z^{-1})^2} \quad (14)$$

where

$$G_1 = rd/T_u \quad (15)$$

$$G_2 = rd^2/T_u \quad (16)$$

$$G_3 = krd^3/T_u \quad (17)$$

and

$$d = 4B_L T_u (r - k) / (r(r - k + 1)) \quad (18)$$

B_L denotes the design-loop bandwidth in Hz, r is typically 2 or 4 and is equal to 4ξ where ξ is the damping ratio, and k is a type-III loop gain parameter ($k = 0$ for a type-II loop) with typical values ranging from 1/4 to 1/2. The filter of Eq. (14) was derived from an equivalent analog filter using the Impulse Invariant Transformation (IIT) [5]. The actual loop bandwidth might be larger than the designed B_L depending on the product $B_L T_u$. Generally, for

$B_L T_u < 0.05$, the actual loop bandwidth is very close to the design loop bandwidth B_L .

Each simulation point was obtained using 5000 runs for each combination of loop SNR and frequency offset. Each run is for a maximum of $100/B_L$ seconds and the loop is declared to have acquired the incoming signal if the magnitude of the residual phase error becomes less than $\pi/8$ radians and remains there for at least $10/B_L$ seconds. By residual phase error is meant the phase error modulo $\pi/2$ since the stable lock points are integer multiples of $\pi/2$. The "acquisition" definition used in this article might be too strict for some applications, especially at low loop SNRs. This will be discussed later in more detail.

Since the acquisition results are based solely on computer simulations, it is worthwhile at this point to discuss their confidence level. Note that the outcome of each run is either "loop acquired" or "loop did not acquire," i.e., a binary decision. Let \mathcal{X} be a random variable denoting the number of "loop acquired" decisions in n runs. Then, \mathcal{X} can be modeled by a binomial distribution with mean np and variance $np(1-p)$ where p denotes the expected probability of acquisition in a single run. Hence, the error level (ratio of the standard deviation to the mean) becomes $[(1-p)/np]^{0.5}$ (accuracy level is $1 - [(1-p)/np]^{0.5}$). As an example, when simulating the loops with an expected probability of acquisition of 10 percent using 5000 runs, the results will be correct with a 95.4-percent accuracy level. On the other hand, for a 90-percent probability of acquisition, the accuracy increases to 99.53 percent.

Probability of acquisition as a function of time (normalized by loop bandwidth) is depicted in Fig. 3 for all type-III loops and in Fig. 4 for all type II loops. It is clear that for a given acquisition time, the acquisition probability increases with decreasing frequency-offset-to-loop-bandwidth ratio $\Delta f/B_L$, as expected. Furthermore, type-II loops seem to outperform their respective type-III loops in terms of acquisition time for $\Delta f/B_L < 1$, but type-III loops are comparable for $\Delta f/B_L > 1$ for "high" acquisition probabilities. As an example, a type-II generalized Costas loop requires about $18/B_L$ seconds to acquire a $0.5B_L$ -Hz frequency offset with a 60-percent probability, whereas the corresponding type-III loop requires about $30/B_L$ seconds for the same probability. On the other hand, when the frequency offset is about $1.25B_L$, then the type-II and type-III loops require about $90/B_L$ and $40/B_L$ seconds respectively for a 20-percent acquisition probability. This is better shown in Fig. 5 for the type-II and type-III Costas cross-over loops where either type can dominate depending on the ratio $\Delta f/B_L$. Note that at high probabilities of acquisition (> 90 percent), the

performance of type-II and type-III loops will become comparable as shown by the trends of Fig. 5 for $\Delta f/B_L = 0.75$ and 1.25, with type-II loops seemingly in the lead.

When comparing the different loop structures, the MAP estimation loop is sometimes outperformed by one or both of the other two loops; for example, at 50 percent acquisition probability with a frequency offset of $1.0B_L$ Hz, the MAP estimation loop and the Costas cross-over loop require about $73/B_L$ and $63/B_L$ seconds, respectively. This is an unexpected result but not very surprising as the MAP loop is the optimum closed-loop estimator for an unknown phase but not necessarily for an unknown frequency. Figures 6 and 7 compare the acquisition performance of the three different loops structures at 15- and 25-dB loop SNRs for both type-II and type-III loops. It seems that for type-II loops, the MAP loop is outperformed by both the Costas cross-over loop and the generalized Costas loop, depending on the loop SNR (Fig. 6); however, for type-III loops, the performance of the MAP loop is almost identical to the best of the other two loops (Fig. 7) for all practical purposes.

Another parameter of interest is the probability of acquisition assuming an unlimited acquisition time (simulations were actually carried out with $100/B_L$ seconds). This differs from the cumulative probability of acquisition plotted earlier which employs a fixed acquisition time. The probability of acquisition as a function of frequency-offset-to-loop-bandwidth ratio for a given loop SNR is plotted in Fig. 8 for type-II loops and in Fig. 9 for type-III loops. As expected, the probability increases with increasing loop SNR but decreases with increasing $\Delta f/B_L$. It is interesting to note that for both type-II and type-III loops, the MAP estimation loop for all practical purposes achieves the highest probability of acquisition for a given loop SNR and a fixed $\Delta f/B_L$. However, from the previous discussion, the MAP loop does not necessarily have the highest probability of acquisition for a fixed acquisition time.

Finally, Fig. 10 depicts the effect of $B_L T_u$ on the cumulative probability for the Costas cross-over loop. As expected, better performance is obtained by decreasing $B_L T_u$, which results in a faster loop update rate for a given loop bandwidth.

As mentioned earlier, the acquisition definition invoked so far might be too strict at low loop SNRs. This is because even though the loop can acquire the incoming phase (i.e., the residual phase error becomes less than $\pi/8$ as required), it cannot remain there for $10/B_L$ seconds because the mean time to lose lock at low loop SNR becomes "small" and affects the probability of acquisition. This is

illustrated in Fig. 11, where the cumulative probability of acquisition is plotted against $B_L t$ for different lock criteria for the Costas cross-over loop at both 15- and 20-dB loop SNRs. For a 20-dB loop SNR, the cumulative probability of acquisition is for all practical purposes independent of the lock criterion used because the mean time to lose lock is larger than the $10/B_L$ seconds required. However, for a 15-dB loop SNR, the mean time to lose lock is small compared to $10/B_L$ seconds and the performance varies with each lock criterion as expected. In that sense, the lock

criterion of $10/B_L$ seconds might not be fair for low loop SNRs.

IV. Conclusion

Three QPSK carrier tracking loops were compared in this article in terms of their acquisition performance. It was determined that the Costas cross-over loop is the most likely candidate for implementation in the ARX II due to its simplicity and overall relative performance.

References

- [1] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Inglewood Cliffs, New Jersey: Prentice-Hall Inc., 1973.
- [2] M. K. Simon, "On the Optimality of the MAP Estimation Loop for Carrier Phase Tracking BPSK and QPSK Signals," *IEEE Trans. on Communications*, vol. COM-27, no. 1, pp. 158-165, January 1979.
- [3] W. R. Braun and W. C. Lindsey, "Carrier Synchronization Techniques for Unbalanced QPSK Signals—Part I," *IEEE Trans on Communications*, vol. COM-26, no. 9, pp. 1325-1333, September 1978.
- [4] J. J. Stiffler, *Theory of Synchronous Communications*, Inglewood Cliffs, New Jersey: Prentice-Hall Inc., 1971.
- [5] S. Aguirre, W. Hurd, R. Kumar, and J. Statman, "A Comparison of Methods for DPLL Loop Filter Design," *TDA Progress Report 42-87*, vol. July-September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 114-124, November 15, 1986.

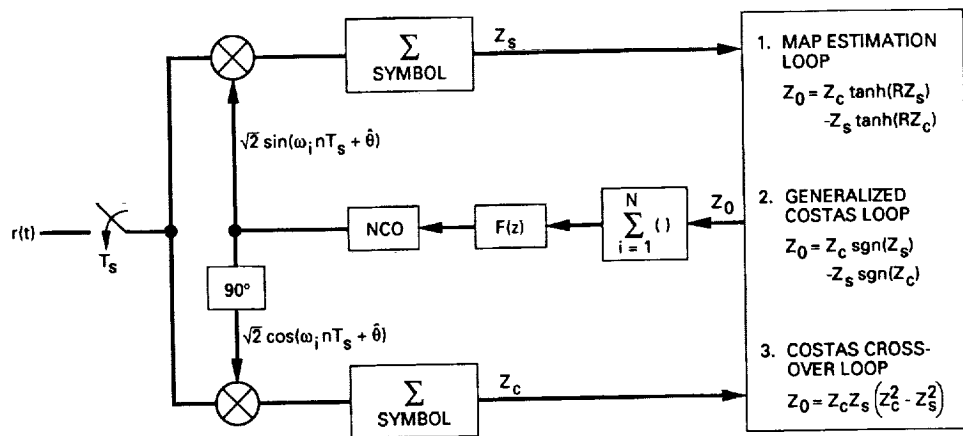


Fig. 1. The QPSK carrier-tracking loops.

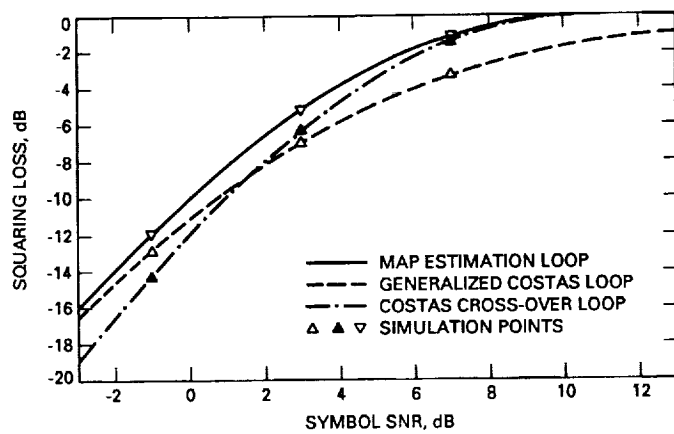


Fig. 2. Squaring losses of the QPSK loops.

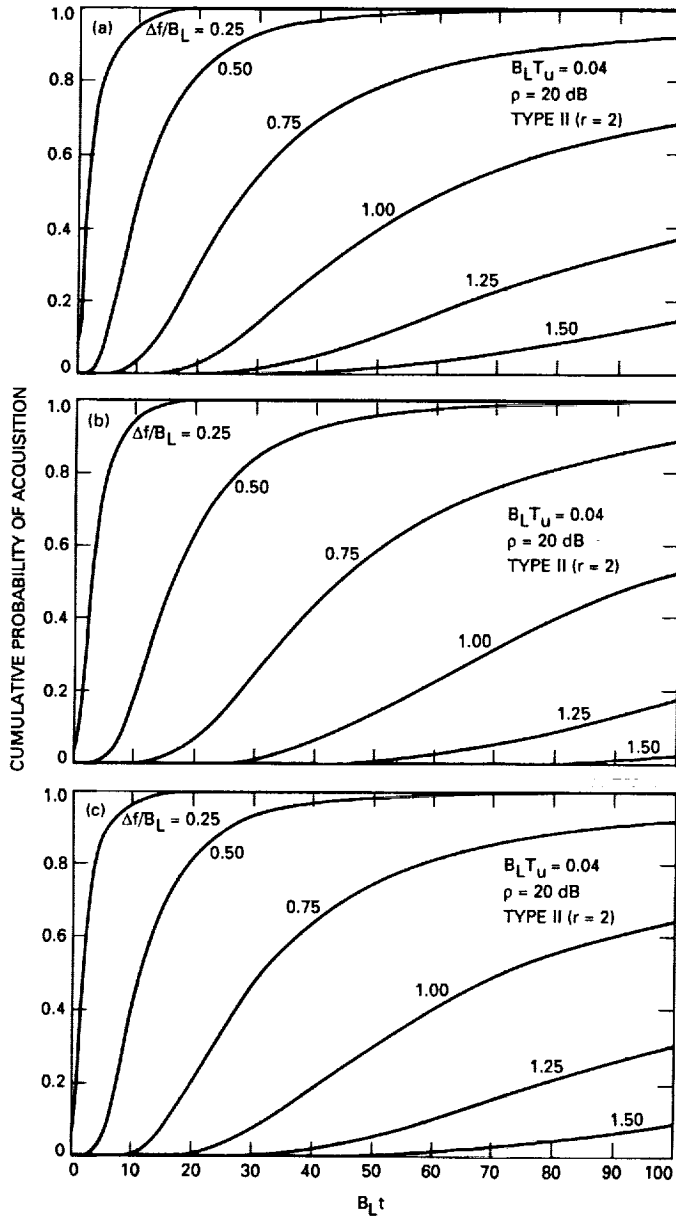


Fig. 3. Cumulative probability of acquisition versus $B_L t$ for type-II loops: (a) Costas cross-over loop, (b) generalized Costas loop, and (c) MAP estimation loop.

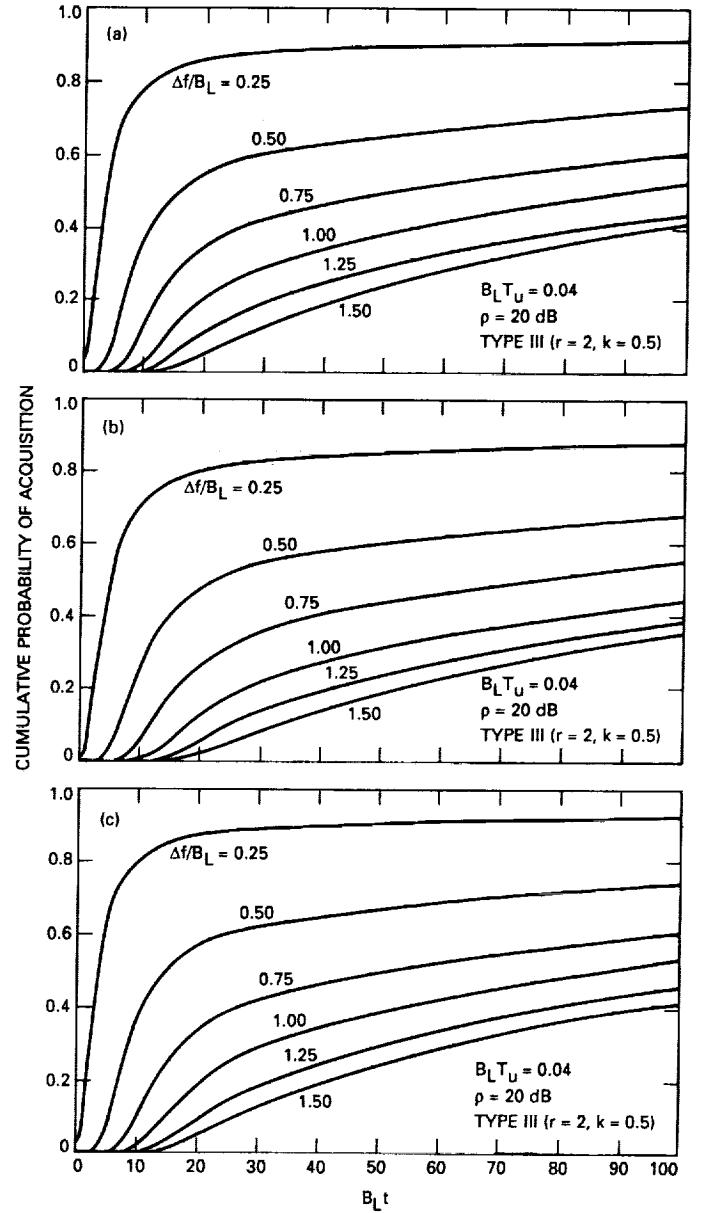


Fig. 4. Cumulative probability of acquisition versus $B_L t$ for type-III loops: (a) Costas cross-over loop, (b) generalized Costas loop, and (c) MAP estimation loop.

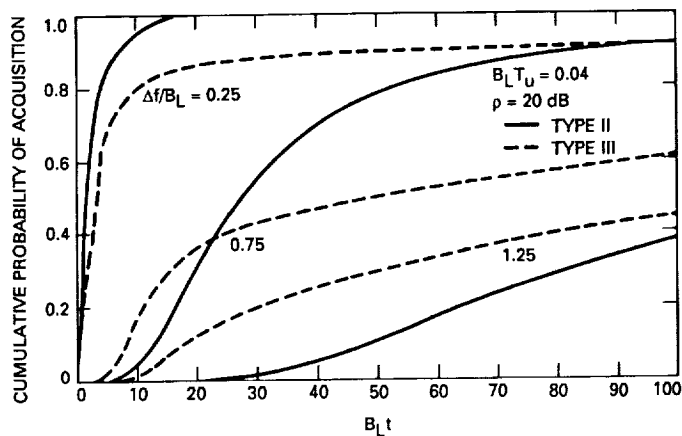


Fig. 5. Cumulative probability of acquisition versus $B_L t$ for the type-II and -III Costas cross-over loops.

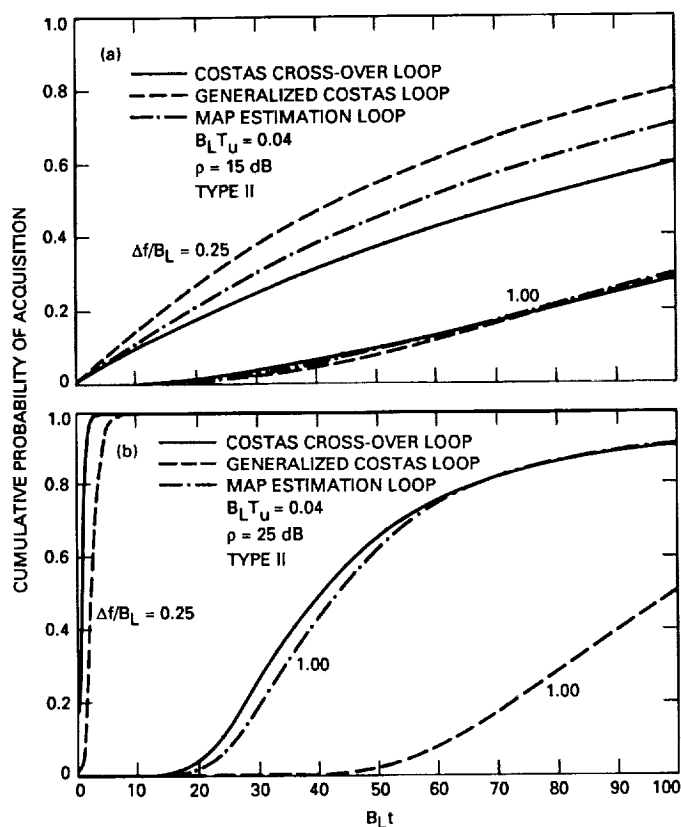


Fig. 6. Cumulative probability of acquisition versus $B_L t$ for all three type-II loops: (a) $\rho = 15$ dB and (b) $\rho = 25$ dB.

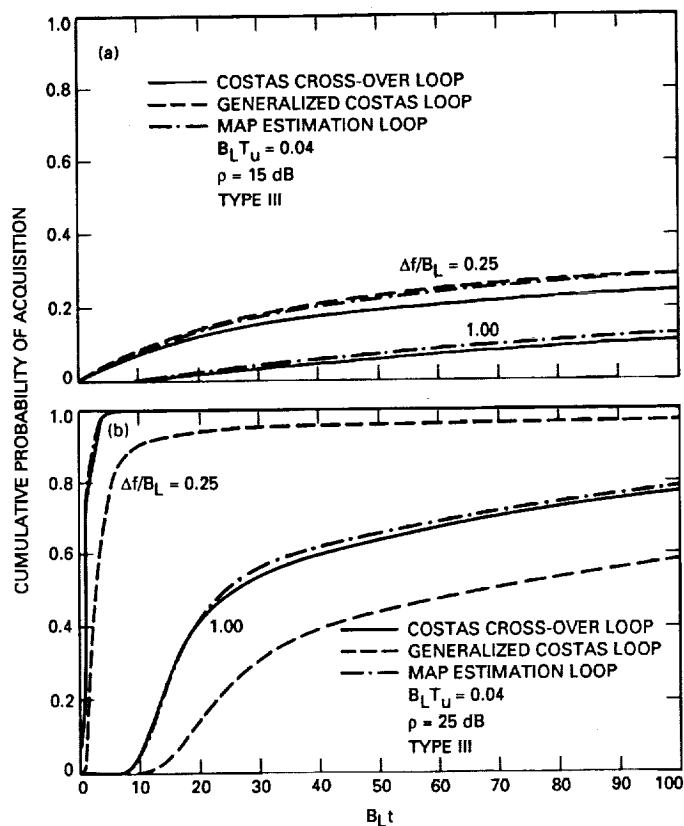


Fig. 7. Cumulative probability of acquisition versus $B_L t$ for all three type-III loops: (a) $\rho = 15$ dB and (b) $\rho = 25$ dB.

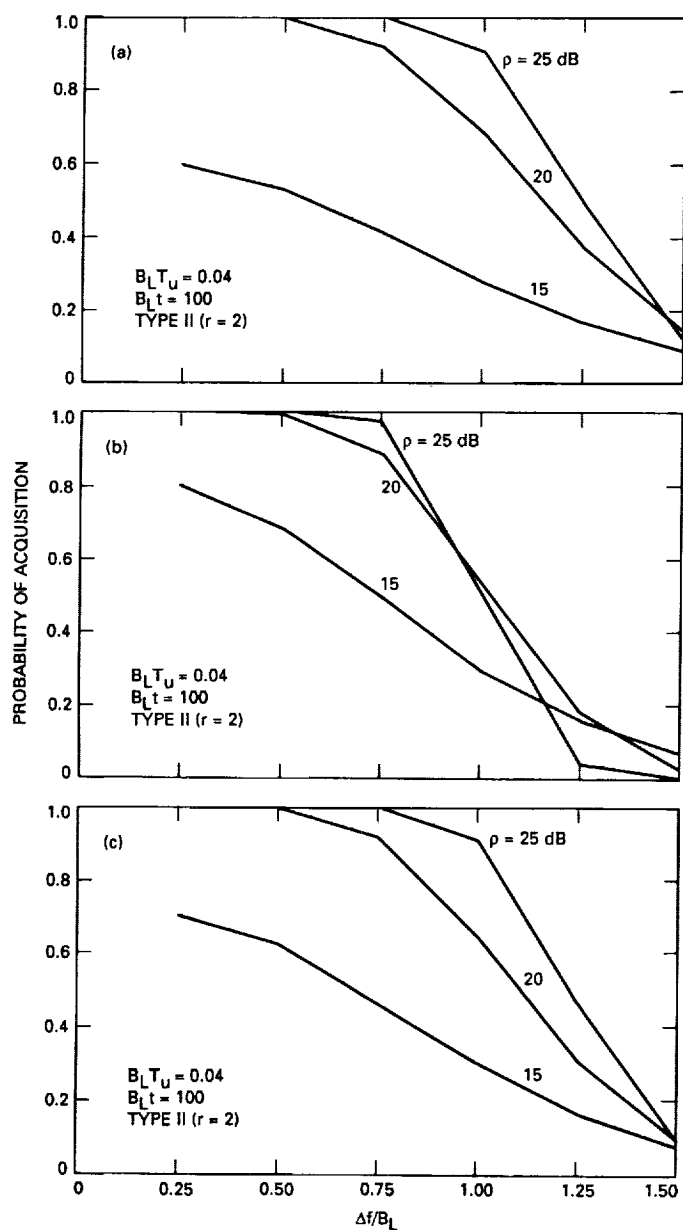


Fig. 8. Probability of acquisition versus $\Delta f/B_L$ for type-II loops: (a) Costas cross-over loop, (b) generalized Costas loop, and (c) MAP estimation loop.

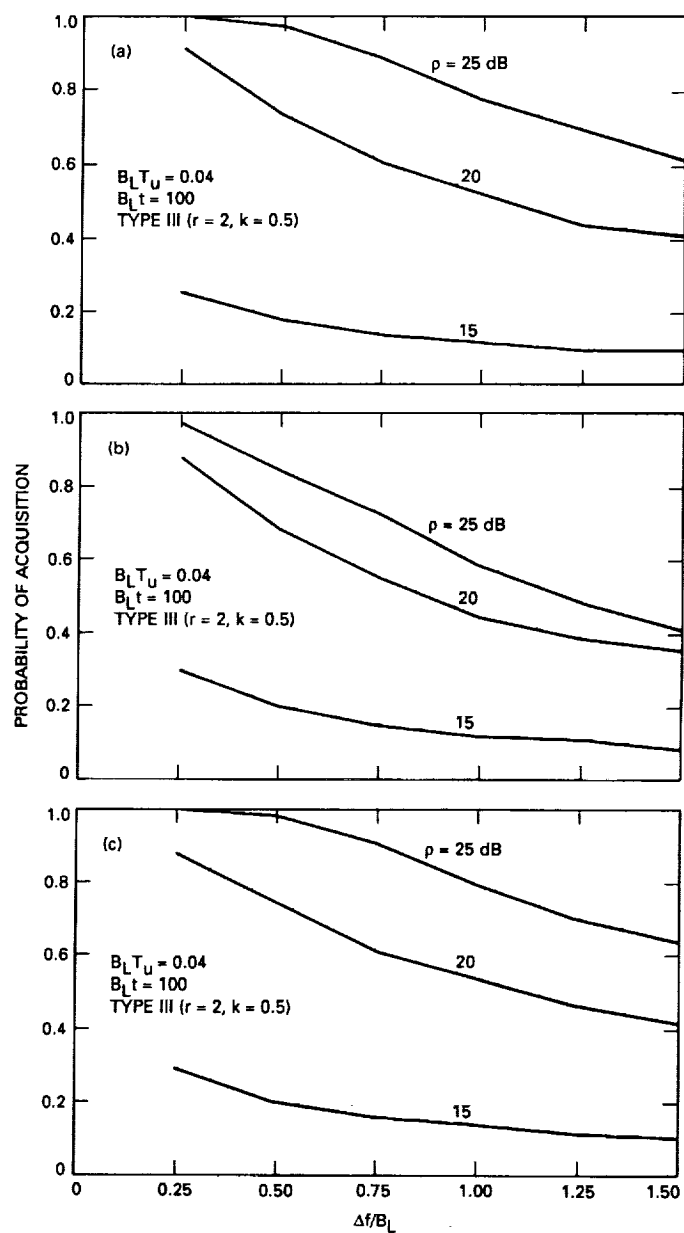


Fig. 9. Probability of acquisition versus $\Delta f/B_L$ for type-III loops: (a) Costas cross-over loop, (b) generalized Costas loop, and (c) MAP estimation loop.

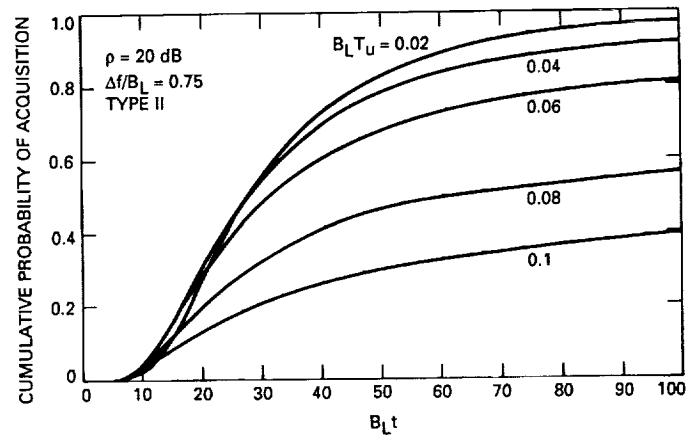


Fig. 10. Cumulative probability of acquisition versus $B_L t$ for the Costas cross-over loop for varying $B_L T_u$.

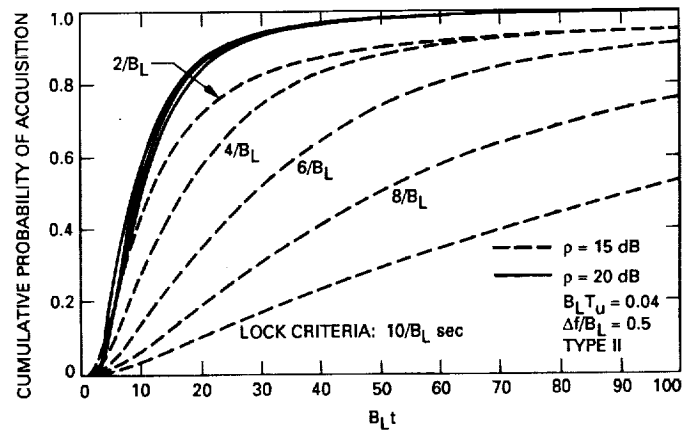


Fig. 11. Cumulative probability of acquisition versus $B_L t$ for the Costas cross-over loop for different lock criteria.

7/3-32
272329
148.

TDA Progress Report 42-100

N90-21893

February 15, 1990

Digital Doppler Extraction Demonstration With the Advanced Receiver

S. Hinedi, R. Bevan, and H. Del Castillo
Communications Systems Research Section

P. Kinman and D. Chong
Telecommunications Systems Section

R. Labelle
Radio Frequency and Microwave Subsystems Section

This article describes a digital Doppler extraction demonstration with the Advanced Receiver II (ARX II) tracking Pioneer 10 and Voyager 2. The measured results are compared with those of the Block IV receiver that was operating in parallel with the ARX II. It is shown that the ARX II outperforms the Block IV receiver in terms of Allan variance of the Doppler residuals, the amount of which depends on the scenario of interest.

I. Introduction

The Advanced Receiver (ARX) [1] was modified and upgraded to accommodate higher telemetry rates and to enable digital Doppler extraction. The previous receiver (ARX I) underwent several telemetry demonstrations [2-4] at Goldstone, but was unable to perform Doppler extraction because the sampling clock was not fixed, but was an integer multiple of the symbol rate. This is referred to as "synchronous sampling" (synchronous with respect to the symbol epoch). In the new receiver (ARX II) [5], the sampling clock is fixed and is derived from the Frequency and Timing Subsystem (FTS). Hence, sampling is performed asynchronously with the symbol period and may result in a noninteger number of samples per symbol. However, a fixed time base is present in the system and digital Doppler extraction can be performed.

The architecture of the ARX II is discussed in [5] along with the various signal-processing schemes employed to perform the functions of carrier and subcarrier tracking and symbol synchronization, among others. In this article, only the phase-locked loop is described, as it was used to track the residual component of the received signal from the two spacecraft.

Both analog and digital loop closures can be mechanized with the ARX II and both are depicted in Fig. 1. A software command can switch between both methods of loop closure without losing the carrier lock.

In the analog loop closure, the incoming signal is at 70 MHz with about 17-MHz intermediate frequency (IF) bandwidth. It is mixed with an 80-MHz locked reference

to produce the 10-MHz IF, which is sampled at 40 MHz using an 8-bit analog-to-digital converter (A/D). The down-conversion to baseband in-phase and quadrature (I and Q) samples is then accomplished by digital mixing using look-up tables. The baseband signal has about an 8-MHz baseband bandwidth and thus a decimation by two is performed to process the signal at the sufficient 20-MHz processing rate (effective sampling rate). The Q samples are accumulated to reduce the rate to the more appropriate loop update rate (50 Hz–1 kHz) to enable a software implementation of the loop filter $F(z)$. The filter output is used to adjust the frequency of a numerically controlled oscillator (NCO), which nominally operates at 4 MHz. The output of the NCO drives a digital-to-analog converter (DAC), producing a 4-MHz analog sine wave, which is mixed with an 84-MHz fixed reference to produce the 80-MHz locked signal that closes the loop.

The operation of the digital loop closure is similar to that of the analog loop closure, except that the 80-MHz signal is a fixed reference and an NCO (running at 10 MHz) is used to close the loop. In this case, the A/D operates on a signal that is not centered at exactly 10 MHz, but rather shifted according to the instantaneous Doppler.

The ARX II accepts an input signal in the 200–400-MHz range from the IF Distribution Assembly output and downconverts it to the 70-MHz range using an external synthesizer (typically a Marconi), as shown in Fig. 2 for the digital loop closure. A total-power automatic gain control (AGC) circuit operates on the signal to ensure that the resulting 10-MHz IF signal lies within the dynamic range of the A/D.

The residual component of the received signal at the output of the AGC can be modeled as

$$r(t) = \sqrt{2P_c} \sin(\omega_i t + \theta_c) + n(t) \quad (1)$$

where P_c denotes the carrier power, ω_i the IF frequency (70-MHz) in rad/sec, and θ_c the carrier phase in radians. The data modulation is ignored here as only the residual carrier component is of interest. The narrowband noise $n(t)$ can be written as

$$n(t) = \sqrt{2}n_c(t) \cos(\omega_i t + \theta_c) - \sqrt{2}n_s(t) \sin(\omega_i t + \theta_c) \quad (2)$$

where $n_c(t)$ and $n_s(t)$ are statistically independent, stationary, band-limited, white-Gaussian-noise processes with one-sided spectral density N_0 watts/Hz and one-sided bandwidth W Hz (roughly 17 MHz in this case). The 70-MHz IF signal is downconverted to 10 MHz and sampled

at 40 MHz. Following I and Q mixing, the samples are accumulated to produce

$$I_{cr}(n) = \sqrt{P_c} \cos \phi_c(n) + n_I(n) \quad (3)$$

$$Q_{cr}(n) = \sqrt{P_c} \sin \phi_c(n) + n_Q(n) \quad (4)$$

where $n_I(n)$ and $n_Q(n)$ are independent Gaussian random variables with variances $N_0/2T_u$, where T_u denotes the loop update period and $\phi_c(n)$ the carrier phase error. The Q_{cr} samples update the loop filter, which is given by

$$F(z) = G_1 + \frac{G_2}{1-z^{-1}} + \frac{G_3}{(1-z^{-1})^2} \quad (5)$$

where

$$G_1 = rd/T_u \quad (6)$$

$$G_2 = rd^2/T_u \quad (7)$$

$$G_3 = krd^3/T_u \quad (8)$$

and

$$d = 4B_L T_u (r - k) / (r - k + 1) \quad (9)$$

where B_L is the design loop bandwidth in hertz, r is typically 2 or 4 and is equal to 4ξ (where ξ is the damping ratio), and k is a type-III loop gain parameter ($k = 0$ for a type-II loop) with typical values ranging from 1/4 to 1/2.

Digital Doppler extraction consists of recording the phase estimate of the incoming waveform time tagged to station time using a Time Code Translator (TCT). The resolution of Doppler data depends heavily on the number of bits used in the NCO to represent phase and frequency.

In the digital loop closure, the NCO accumulates fractional phase using 32 bits, but only the most significant 12 bits are used in the sin/cos look-up table. This results in a phase resolution of 0.24 mcycle (2^{-12}). The phase is recorded, external to the NCO, using 48 bits for the integer part and 12 bits for the fractional part. The NCO, in this case, operates at a nominal frequency of 10 MHz and is clocked by a 40-MHz clock. This results in 9.3-mHz ($4 \times 10^7/2^{32}$) frequency resolution.

In the analog loop closure case, the fractional phase is recorded using 14 bits, two of which are accumulated external to the NCO. The resulting phase resolution is 0.06 mcycle, better than that of the digital loop closure. The NCO,

in this case, is actually operating at a nominal frequency of 16 MHz with a 20-MHz clock, but it effectively produces a 4-MHz signal because of the additional bits. The resulting frequency resolution is 1.16 mHz ($2 \times 10^7 / 2^{34}$), better than the digital loop case.

However, the improved resolution for the analog loop case is at the expense of limiting the maximum nominal frequency to one-fourth of the clock rate (the practical limit of an NCO is roughly 40 percent of the clock rate). That in turn limits the dynamic excursion of the frequency to ± 1 MHz in the analog loop closure, whereas it is about ± 5 MHz in the digital closure. Time tagging is performed at ± 1 μ sec accuracy with a ± 1 nsec stability. The logging rate can be chosen by software from 1, 2, 5, 10, or 20 Hz.

II. Demonstration Description and Results

The demonstration consists of operating the ARX II in parallel with the Block IV receiver; both track the same spacecraft. Both receivers record data for an identical period of time. The data are later analyzed to quantify and compare the respective performances (see Fig. 3). In all the results to follow, the ARX II was operating with a 500-Hz loop update rate.

A. Allan Variance

The X-band downlink of Voyager 2 (channel 14, 8415.000000 MHz) and the S-band downlink of Pioneer 10 (channel 7, 2292.407407 MHz) were tracked to determine the relative effectiveness of the ARX II and the Block IV receiver in Doppler measurements.

Doppler residuals from four data arcs, as defined in Table 1, are reported here. The first arc is from the Voyager-2 X-band downlink tracked on Greenwich Day of Year (DOY) 195, the second from the Pioneer-10 downlink tracked on DOY 195, and the last two from the Pioneer-10 downlink tracked on DOY 201. During all of these tracks, the spacecraft operated in a three-way Doppler mode, thus eliminating problems associated with the spacecrafts' on-board oscillators. For each of these arcs, the downlink was simultaneously tracked by the ARX II and a Block IV receiver. The downconverted carrier phase as measured by the ARX II was recorded once per second and the Doppler count as measured by the Block IV Doppler extractor was recorded once per second.

Use was made of two time scales during the analysis: a 24-hour clock for Greenwich Mean Time and the number of seconds past the January 1, 1950 epoch. This "J 50" time was needed to do a linear interpolation on the uplink

frequencies as reported in the Program Frequency Records (Table 2).

The Doppler observable is calculated for the ARX II as

$$f_{DO} = 48Gf_E + f_{NCO} - f_1 - f_2 - f_3 \quad (10)$$

where f_{DO} is the Doppler observable, G the transponder ratio (880/221 for the Voyager track and 240/221 for the Pioneer tracks), f_E the exciter frequency, f_{NCO} the difference in phase as measured one second apart by the ARX II, f_1 the frequency of the first local oscillator (8.1 GHz for the Voyager track, 2.0 GHz for the Pioneer tracks), f_2 the frequency of the second local oscillator, and finally, f_3 , which is 80 MHz for digital closure and 74 MHz for analog closure of the ARX II loop.

The exciter operates at approximately 44 MHz, whereas the Program Frequency Records report the exciter frequency divided by two (approximately 22 MHz). The exciter frequency is found by interpolating the frequency ramps of Table 2 and then multiplying by two. The second local oscillator is a Marconi synthesizer; its setting for each of the tracks is given in Table 3. The final term in Eq. (10) depends on whether the ARX II had digital or analog loop closure; this information is given in Table 4.

Prior to DOY 201, there was a hardware clocking problem that caused the cycle count recorded by the ARX II to be low by exactly one cycle on some (randomly occurring) samples and these were processed by a cycle correction algorithm that was used to correct the data. The DOY 201 data did not need to be processed by this algorithm because the hardware problem was fixed by that date.

The Doppler residual is calculated from the Doppler observable as

$$f_{\Delta} = f_{DO} - f_M - f_T \quad (11)$$

where f_{Δ} is the Doppler residual, f_M the modeled (based on trajectory) Doppler observable, and f_T the tropospheric correction.

Equation (11) applies to both the ARX II Doppler residual and to the Block IV Doppler residual. The Block IV Doppler observable, the modeled Doppler observable, and the tropospheric correction are all available on the REGRES file provided by navigation personnel.

Doppler residuals, spaced one second apart, were calculated for both receivers for each of the four data arcs defined in Table 1. The Doppler residual record for part of the Voyager 195 data arc is shown in Fig. 4. In that figure, a bias was added to the Block IV residuals to displace them from the ARX II residuals so that the two traces could be discerned. In fact, the means of those two traces are in excellent agreement. The mean values of the Doppler residuals for the ARX II and for the Block IV receiver are given in Table 5. It is noted that the two receivers agree on the magnitude of the Doppler effect to less than 1 mHz.

The square-root Allan variance is used here to characterize the jitter in the Doppler residual measurements. It is defined as

$$\sigma_y(\tau) = \frac{1}{\sqrt{2\nu}} \sqrt{\langle [f_\Delta(t+\tau) - f_\Delta(t)]^2 \rangle} \quad (12)$$

where $\sigma_y(\tau)$ is the square-root Allan variance, τ the integration time, ν the downlink frequency, and $\langle \cdot \rangle$ denotes the time average over the entire data arc. The square-root Allan variances computed for a one-second integration time are contrasted in Table 6. Also given in Table 6 are modeled values of that quantity based on the assumption that downlink thermal noise dominates the frequency instability for such a short integration time. The modeled values came from [6], i.e.,

$$\sigma_y(\tau) = \frac{1}{2\pi\tau\nu} \sqrt{\frac{3B_L}{P_c/N_0}} \quad (13)$$

where P_c/N_0 denotes the carrier power-to-noise spectral density ratio (from Table 4) and B_L the one-sided loop bandwidth.

The modeled and measured variances for short integration times as presented in Table 6 agree closely for the Pioneer tracks, indicating that the Doppler measurement error really is dominated by downlink thermal noise. For the Voyager 195 track, the measured variance is somewhat higher than the modeled variance; this could be due to the carrier power not being constant during the pass.

The square-root Allan variance of the Doppler residual is plotted as a function of integration time in Figs. 5 and 6 for Voyager and Pioneer, respectively. These plots show a thermal noise dominance with reciprocal τ dependence for the smaller integration times, and the advantage of the ARX II's smaller loop bandwidth is made plain. For larger integration times, other error sources dominate and the bandwidth advantage of the ARX II is lost. With the

Pioneer data, the two traces actually begin to coalesce, indicating a common error source for the two receivers. This common error source, which dominates for larger integration times, is almost certainly solar plasma. The Sun-Earth-probe angle was 38.2 deg for Pioneer on DOY 195. The Pioneer 201 data arcs are not long enough to provide a statistically significant Allan variance for the longer integration times.

B. Receiver Tracking Performance

The ARX II acquires the carrier using an automated procedure, which consists of performing several software fast Fourier transforms (FFTs) around the frequency predict. The size and the rate of the FFTs are controlled by software, and the spectra are noncoherently averaged to reduce the noise. The largest peak is then selected and the loop enabled. The acquisition time depends on the speed of the FFT computation. A future upgrade is planned that will use a dedicated FFT board to perform the computations. This will dramatically reduce the acquisition time to less than a second.

For performance evaluation, the ARX II records several parameters. These include: carrier power-to-noise, spectral, density-level estimates, $\widehat{P_c/N_0}$; averaged and normalized I and Q samples; steady-state phase-error estimates $\widehat{\phi_{ss}}$, etc.

The averaged I sample can be used in real time to monitor the loop gain and, hence, the loop bandwidth. With proper normalization, the averaged I samples should be unity, assuming that the amplitude of the incoming signal is "close" to its predict and, hence, that the operating loop bandwidth is its designed value. That averaging should be performed when the steady-state error due to dynamics is negligible; otherwise, no information about loop gain can be deduced. A steady-state phase-error estimate can be obtained in real time by computing

$$\widehat{\phi_{ss}}(j) = \tan^{-1} \left(\frac{\sum_{n=j}^{n=j+L+1} Q_{cr}(n)}{\sum_{n=j}^{n=j+L+1} I_{cr}(n)} \right) \quad (14)$$

where L controls the estimation period. In the ARX II, $\widehat{\phi_{ss}}(j)$ can be sent to the monitor's screen at different rates ranging from once every second to once every minute. Another key parameter is the incoming P_c/N_0 , which is also estimated by the receiver according to

$$\widehat{\frac{P_c}{N_0}}(j) = \frac{\left(\sum_{n=j}^{n=j+L+1} I_{cr}(n) \right)^2}{2T_u \sum_{n=j}^{n=j+L+1} Q_{cr}^2(n)} \quad (15)$$

Figure 7 depicts the various parameters for a Pioneer track using analog loop closure with a type-III loop. In this case, the estimation period was 10 sec and the designed loop bandwidth about 0.6 Hz. It is clear that the averaged I is close to unity and the averaged Q close to zero, which indicates that the operating bandwidth is 0.6 Hz, as designed. The estimated P_c/N_0 is roughly 12.5 dB-Hz (compared with a 13 dB-Hz predict) and the steady-state phase error is approximately zero.

In Fig. 8, Pioneer results are also shown for the digital loop closure case with a type-II loop. The averaged Q is no longer zero as expected, due to the 7-deg steady-state phase error (since type-II loops cannot track Doppler rates with zero phase error). The resulting P_c/N_0 estimates are also degraded by about 0.6 dB to yield roughly 11.9 dB-Hz. Finally, Fig. 9 depicts the results for a Voyager 2 track using a type-III loop with digital closure. Note that

the normalized I is not unity, which implies that the operating loop bandwidth is actually 2.5 Hz (1.2 times the designed bandwidth of 2 Hz). The steady-state error is zero as expected and the estimated P_c/N_0 is about 31.3 dB-Hz (compared with a 32-dB-Hz predict). The estimation period for the Voyager data is one second due to the relatively "strong" signal.

III. Conclusion

This article described digital Doppler extraction with the ARX II and the results were compared to the Block IV receiver Doppler extractor. It was shown that digital Doppler extraction did not result in any degradation relative to Block IV, and in most cases performed better with the ARX II. From the standpoint of Doppler extraction performance, ARX II is an excellent candidate for future Block V receiver implementation.

Acknowledgments

We would like to thank Joseph Statman, Ernest Stone, and Ted Peng for their encouragement, and Robert Johnson and Miguel Marina for their help as test coordinators. We would like to thank the personnel of DSS 14 and their representatives in the Modified Configuration Control Board for allowing us to be out there at a time when they were busy preparing for the Voyager 2 Neptune encounter. We thank the Pioneer 10 and Voyager 2 Projects for allowing us to tap into their signals. We thank Gene Goltz and Gil Roldan of the Navigation Systems Section for preparing Block IV REGRES files, and Stanley Mak of the TDA Mission Support and DSN Operations Section for preparing predicts. Many others contributed ideas and background information that led to the successful conclusion of these tests; these people are mainly from the following organizations: the Navigation Systems Section, the DSN Data Systems Section, the TDA Engineering Section, and the TDA Mission Support and DSN Operations Section.

References

- [1] D. Brown and W. Hurd, "DSN Advanced Receiver: Breadboard Description and Test Results," *TDA Progress Report 42-89*, Jet Propulsion Laboratory, Pasadena, California, pp. 48-66, May 15, 1987.
- [2] D. Brown, W. Hurd, V. Vilnrotter, and J. Wiggins, "Advanced Receiver Tracking of Voyager 2 Near Solar Conjunction," *TDA Progress Report 42-93*, Jet Propulsion Laboratory, Pasadena, California, pp. 75-82, May 15, 1988.
- [3] W. Hurd, D. Brown, V. Vilnrotter, and J. Wiggins, "Telemetry SNR Improvement Using the DSN Advanced Receiver With Results for Pioneer 10," *TDA Progress Report 42-93*, Jet Propulsion Laboratory, Pasadena, California, pp. 64-74, May 15, 1988.
- [4] V. Vilnrotter, W. Hurd, and D. Brown, "Optimized Tracking of RF Carriers With Phase Noise, Including Pioneer 10 Results," *TDA Progress Report 42-91*, Jet Propulsion Laboratory, Pasadena, California, pp. 141-157, November 15, 1987.
- [5] S. Hinedi, "Functional Description of the Advanced Receiver," *TDA Progress Report 42-100*, Jet Propulsion Laboratory, Pasadena, California, February 15, 1990.
- [6] J. A. Barnes, et al., "Characterization of Frequency Stability," *IEEE Transactions on Instrumentation and Measurement*, vol. IM-20, no. 2, pp. 105-120, May 1971.

Table 1. Definition of tracks

Track	Start time		End time	
	hh:mm:ss	J 50 Time, sec	hh:mm:ss	J 50 Time, sec
Voyager 195	06:41:00	1247553660	07:52:00	1247557920
Pioneer 195	14:11:00	1247580660	14:52:00	1247583120
Pioneer 201A	12:18:00	1248092280	12:35:00	1248093300
Pioneer 201B	14:46:00	1248101160	15:07:00	1248102420

Table 2. Data from Program Frequency Records

Pass	Time, J 50, sec	Start frequency, Hz	Ramp rate, Hz/sec
V195	1247552633	0.2201533673236050D + 08	0.2295000000000000D - 02
V195	1247557733	0.2201534843686050D + 08	0.2235000000000000D - 02
P195	1247573336	0.2198478999999950D + 08	0.0000000000000000D + 00
P201	1248089640	0.2198462999999950D + 08	0.0000000000000000D + 00

Table 3. Second local oscillator frequency

Track	Second local oscillator (f_2), Hz
Voyager 195	244416020
Pioneer 195	222078470
Pioneer 201A	222094640
Pioneer 201B	222094640

Table 4. Receiver parameters

Track	P_c/N_0 , dB-Hz	Block IV One-sided loop bandwidth, Hz	ARX II One-sided loop bandwidth, Hz	ARX II Loop closure
Voyager 195	30	54	2.0	Digital
Pioneer 195	11	2.5	2.0	Digital
Pioneer 201A	11	2.5	1.0	Digital
Pioneer 201B	11	2.5	0.6	Analog

Table 5. Mean of Doppler residual

Track	Mean of Doppler residual, Hz	
	Block IV	ARX II
Voyager 195	-0.130	-0.130
Pioneer 195	-0.153	-0.153
Pioneer 201A	-0.148	-0.147
Pioneer 201B	-0.149	-0.149

Table 6. Square-root Allan variance of Doppler residuals

Track	1-second (square-root) Allan variance of Doppler residuals			
	Block IV		ARX II	
	Modeled	Measured	Modeled	Measured
Voyager 195	7.6×10^{-12}	1.1×10^{-11}	1.5×10^{-12}	2.3×10^{-12}
Pioneer 195	5.4×10^{-11}	5.1×10^{-11}	4.8×10^{-11}	4.7×10^{-11}
Pioneer 201A	5.4×10^{-11}	5.4×10^{-11}	3.4×10^{-11}	2.9×10^{-11}
Pioneer 201B	5.4×10^{-11}	5.4×10^{-11}	2.6×10^{-11}	1.9×10^{-11}

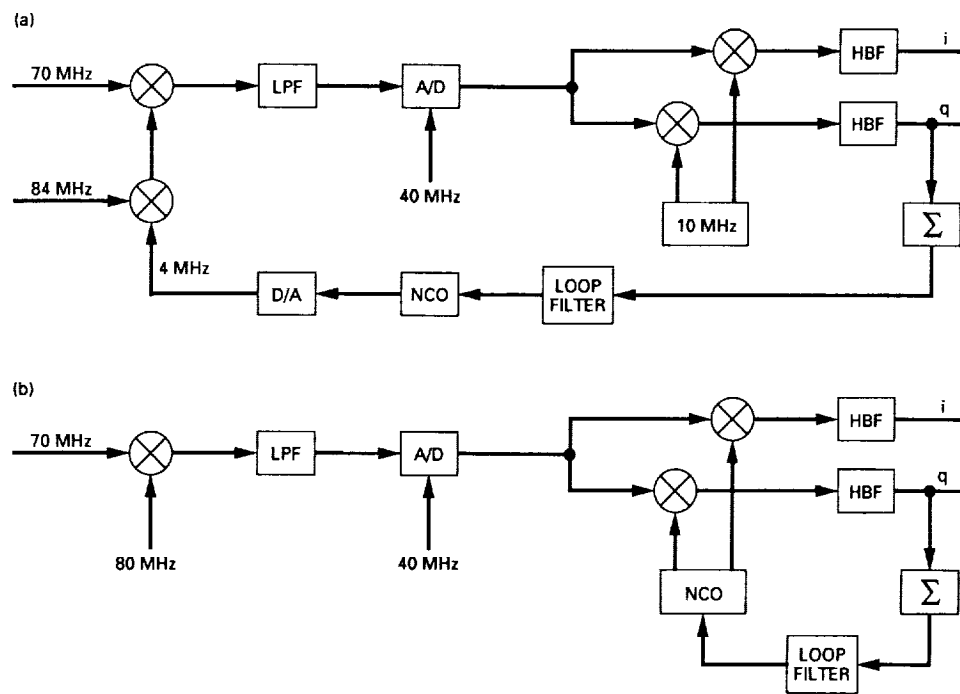


Fig. 1. Analog versus digital loop closure.

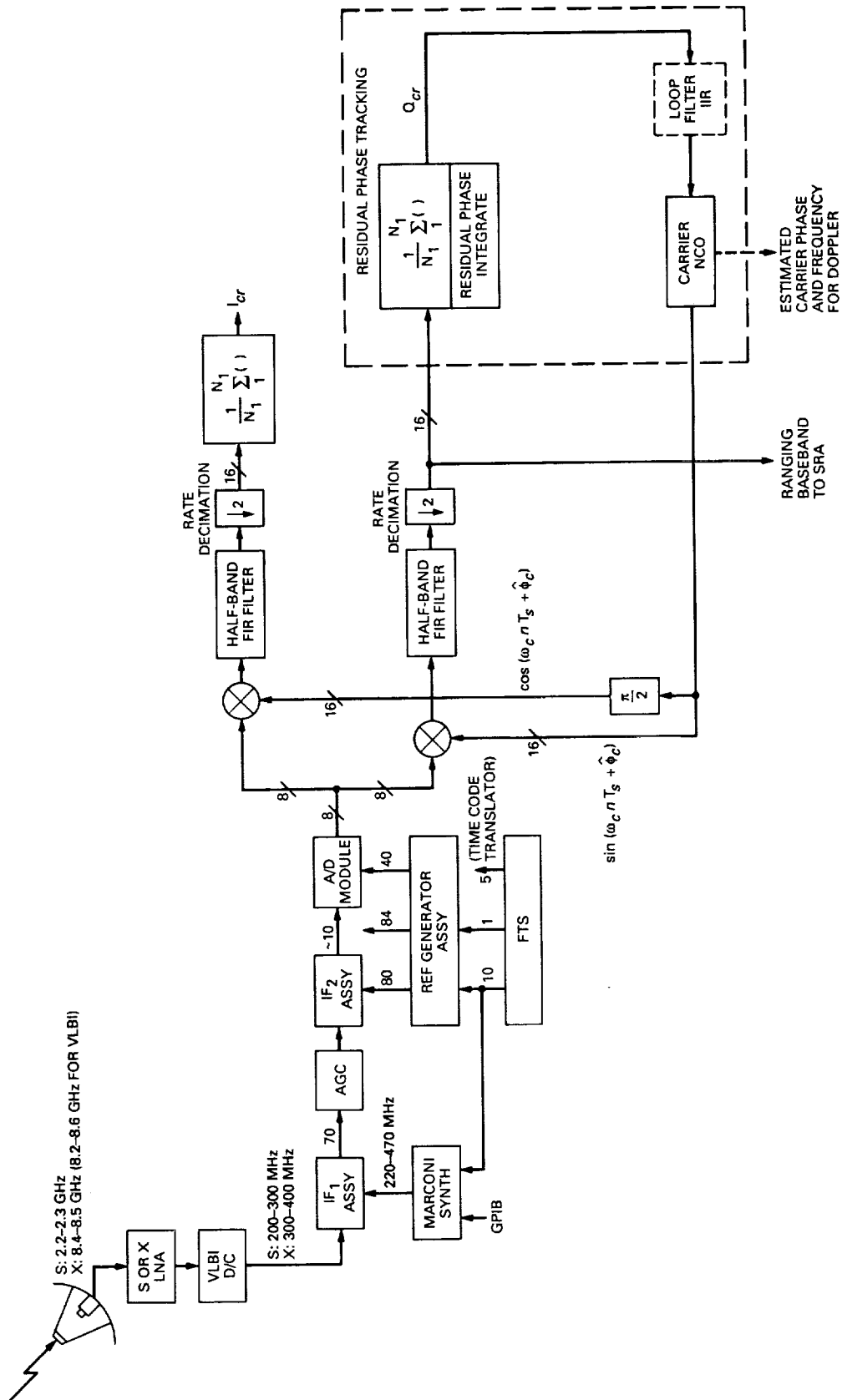


Fig. 2. ARX II phase-locked loop.

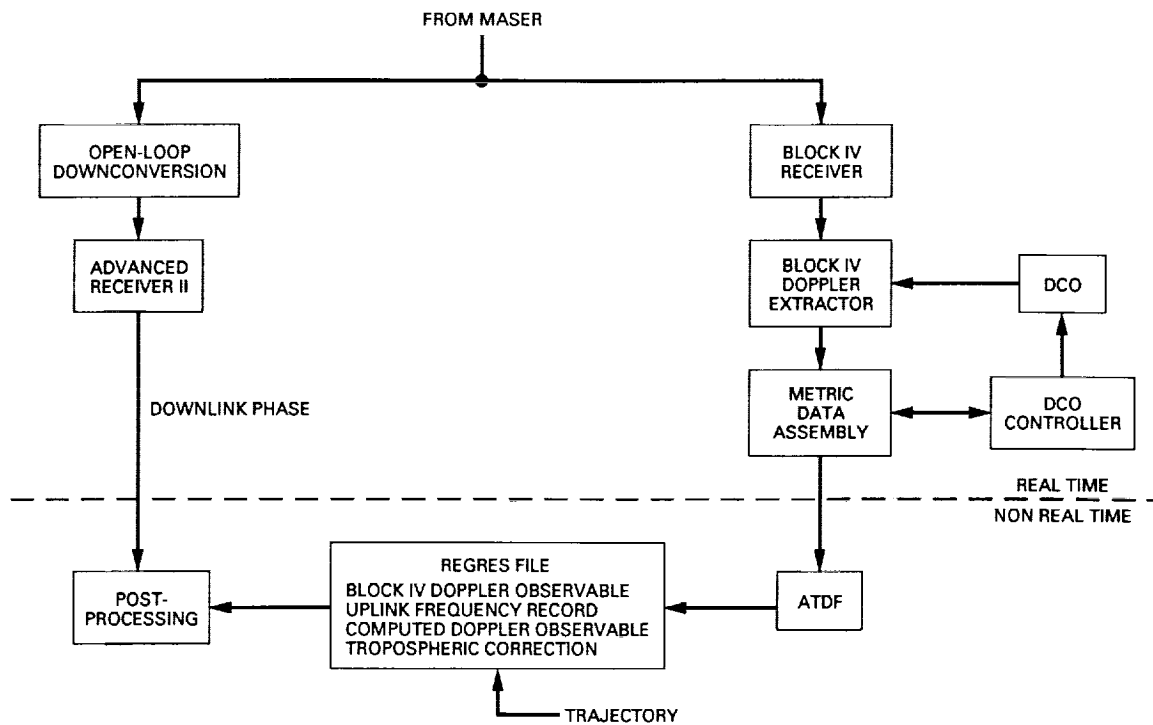


Fig. 3. Experiment description.

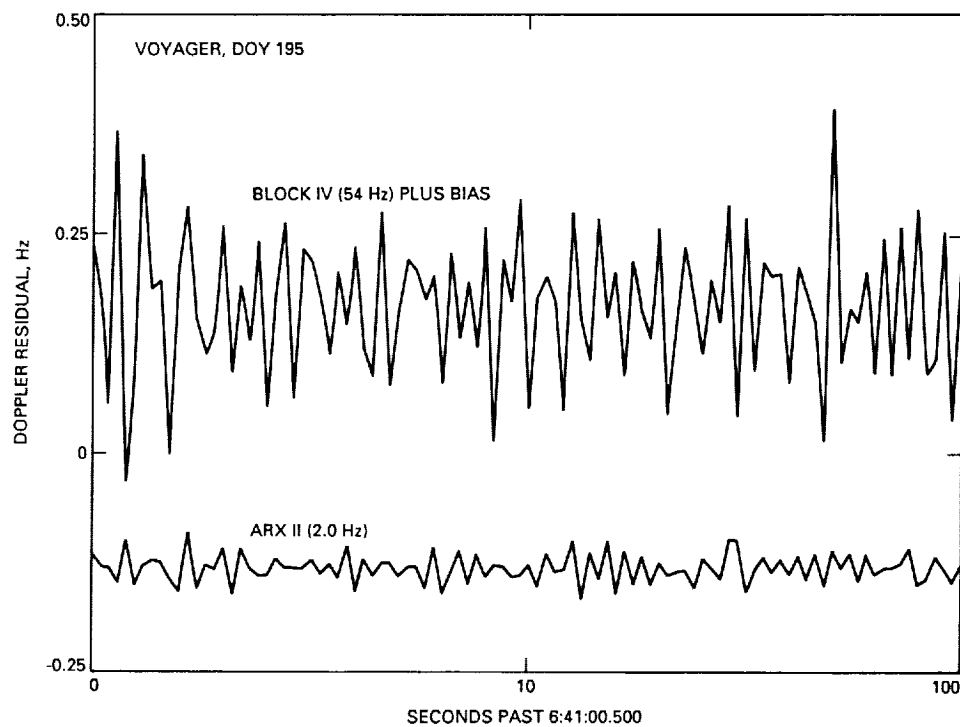


Fig. 4. Doppler residual for Voyager 2, DOY 195.

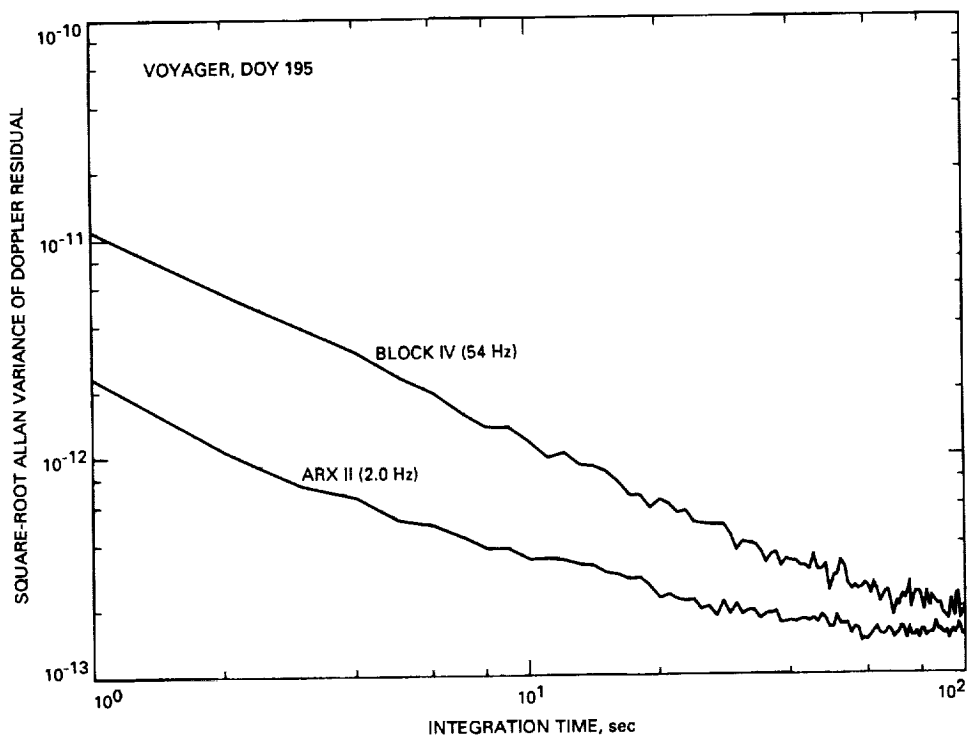


Fig. 5. Allan variance of Doppler residual for Voyager 2, DOY 195.

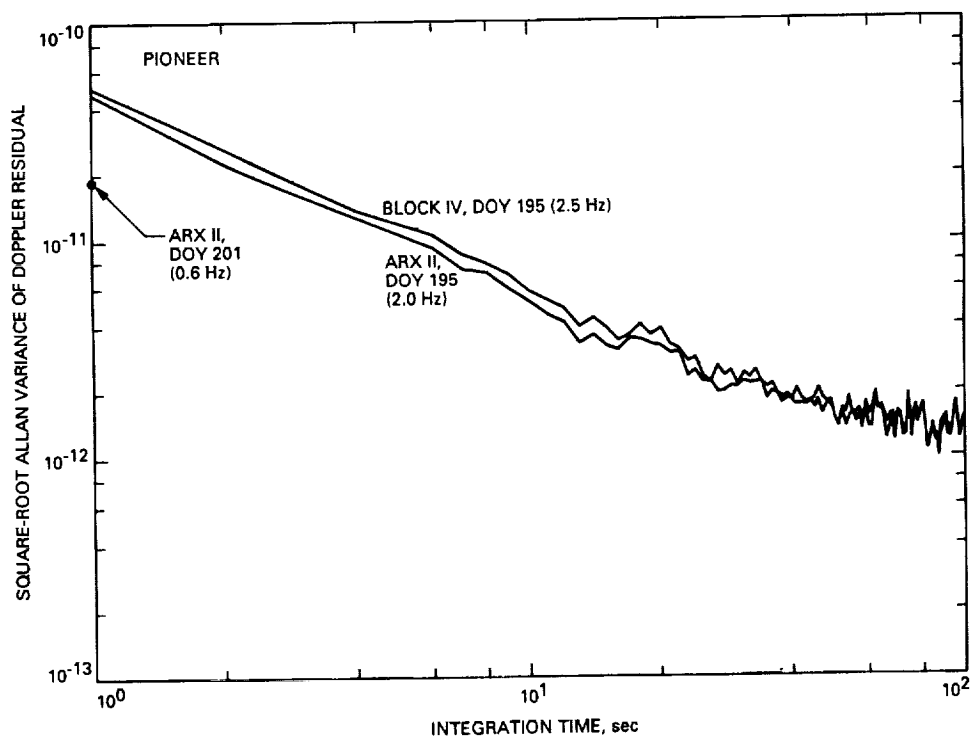


Fig. 6. Allan variance of Doppler residual for Pioneer 10, DOY 195.

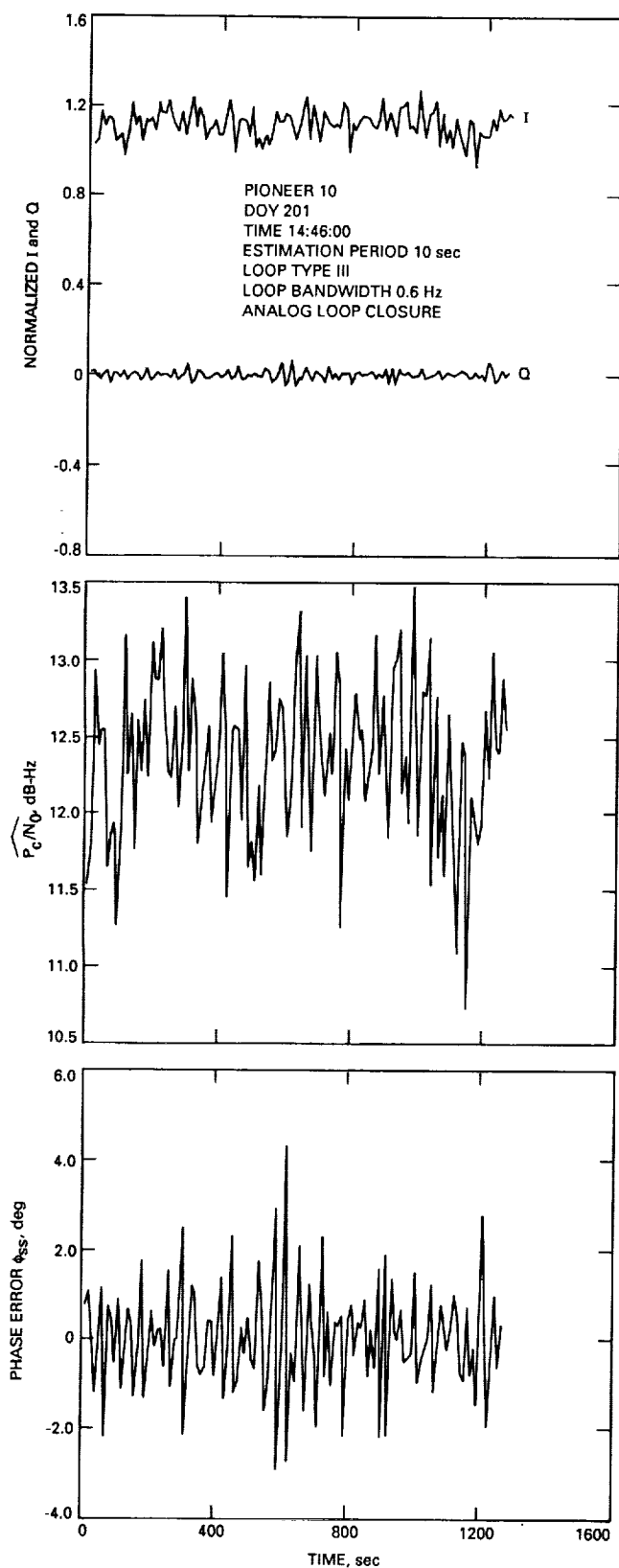


Fig. 7. ARX II data for Pioneer 10, DOY 201.

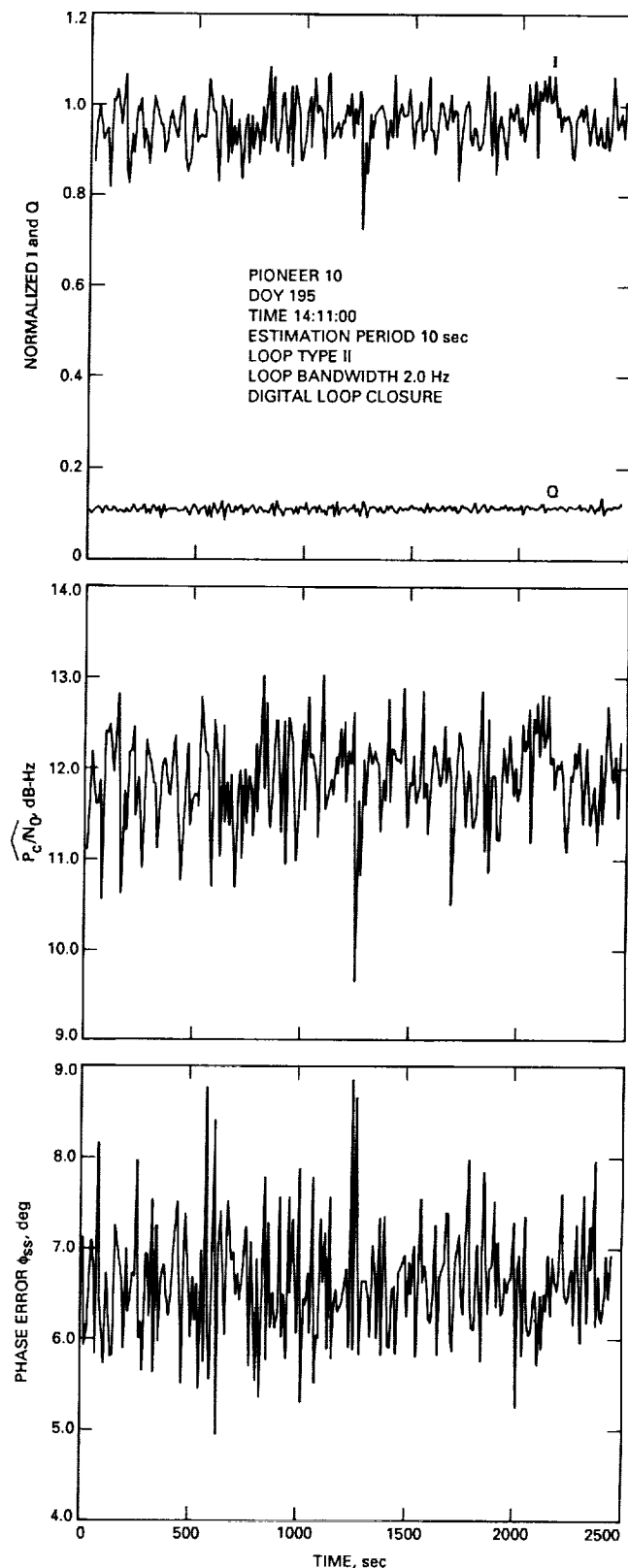


Fig. 8. ARX II data for Pioneer 10, DOY 195.

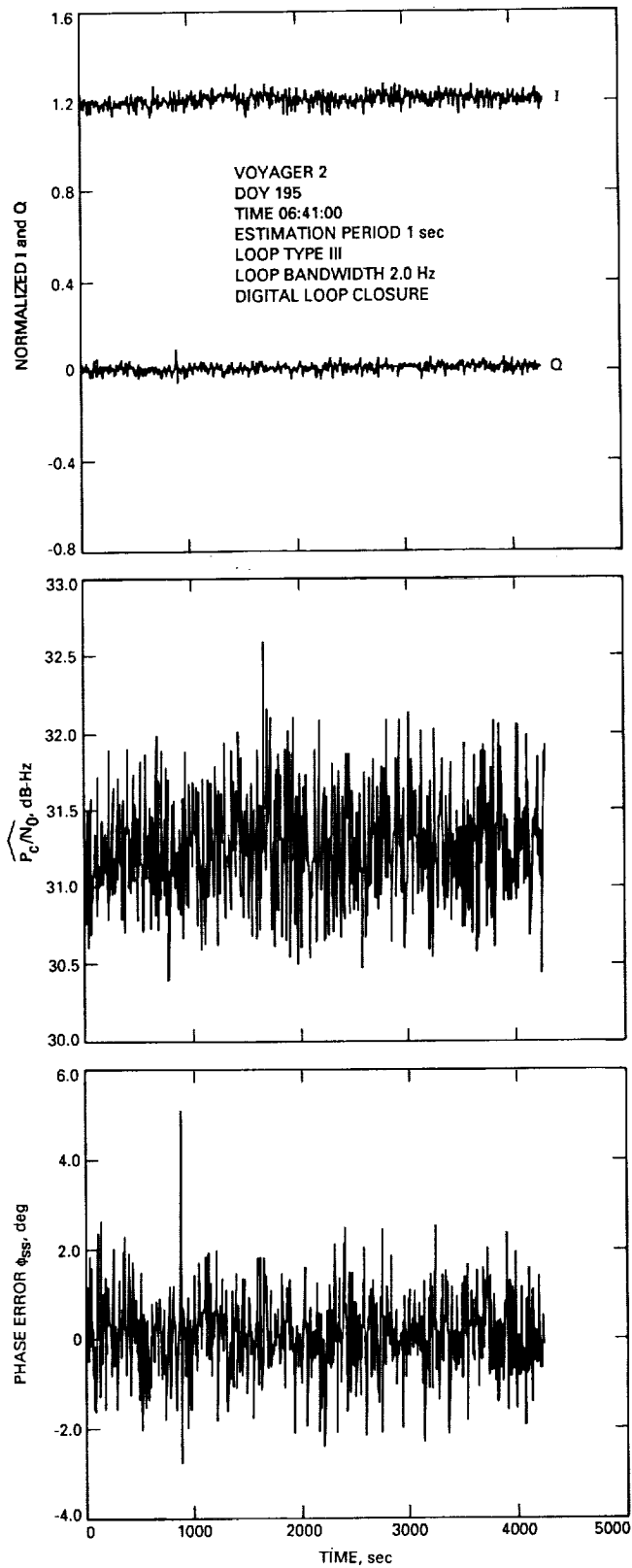


Fig. 9. ARX II data for Voyager 2, DOY 195.

74-32
272330
68.

N90-21894

TDA Progress Report 42-100

February 15, 1990

Exact Closed-Form Expressions for the Performance of the Split-Symbol Moments Estimator of Signal-to-Noise Ratio

S. Dolinar

Communications Systems Research Section

Previously, the performance of the split-symbol moments estimator (SSME) of signal-to-noise ratio (SNR) has been evaluated by means of approximate expressions for the estimator mean and variance. These are asymptotic formulas in the sense that they become accurate as the number of estimator samples gets large. In the present article, exact closed-form expressions are obtained for the same quantities. These expressions confirm the accuracy of the previously derived asymptotic results, and, unlike the asymptotic formulas, they are useful even when the number of samples is small. It is also shown that the conventional split-symbol estimator can be trivially scaled to form a signal-to-noise ratio estimator which is precisely unbiased (as long as the estimate is based on more than two split-symbols).

I. Introduction

The so-called split-symbol moments estimator (SSME) of the received symbol signal-to-noise ratio (SNR) was first suggested in [1]. In [2], approximate expressions were derived for the mean and variance of the SSME. These approximate expressions were later used as the basis for a more detailed performance analysis in [3], which included the effects of intersymbol interference as a result of filtering the received symbol stream.

The approximations in [2] result from expanding the exact equations for the mean and variance in a series involving the central moments of the SNR estimator and then ignoring all central moments higher than second or-

der. This type of approximation relies on the fact that successively higher central moments of the estimator go to zero faster than lower-order central moments as the number of samples gets large.

A natural question is: Under what conditions is it valid to throw away the higher-order terms? The results quoted for the mean and variance of the SSME (Eqs. 28 and 31 of [2]) include terms which go to zero as the number of sampled split-symbols, n , gets large. These terms are proportional to $1/n$. How large does n have to be before the $1/n$ terms are guaranteed to dominate the $1/n^2$ terms? Does the answer to this question depend on the magnitude of the true SNR? For example, the expression for the signal-to-noise ratio of the estimator (Eq. 32 of [2])

exhibits peculiar behavior when the true SNR is very small. In this case, the equation says that the signal-to-noise ratio of the estimator actually decreases with n until n gets to be about as large as $1/\text{SNR}$.

In this article, exact closed-form expressions are obtained for the mean and variance of the SSME when the symbol stream is unfiltered. These exact results confirm the accuracy of the previously known asymptotic expressions. Unlike the asymptotic formulas, the exact expressions are useful even when the number of sampled split-symbols, n , is small. Furthermore, the basic techniques developed for this derivation might also be useful in analyzing the more complicated case of filtered data.

II. Derivation of the Result

Using the notation of [3], the split-symbol moments estimator of SNR is denoted by SNR^* and expressed as

$$\text{SNR}^* = \frac{m_p}{2(\frac{1}{4}m_{ss} - m_p)} \quad (1)$$

where m_p and m_{ss} are average values over n symbols of the product and sum squared, respectively, of the outputs of half-symbol accumulators,

$$m_p = \frac{1}{n} \sum_{j=1}^n Y_{\alpha j} Y_{\beta j} \quad (2a)$$

$$m_{ss} = \frac{1}{n} \sum_{j=1}^n (Y_{\alpha j} + Y_{\beta j})^2 \quad (2b)$$

$$Y_{\alpha j} = \sum_{i=0}^{\frac{N_s}{2}-1} y_{ij} \quad (3a)$$

$$Y_{\beta j} = \sum_{i=\frac{N_s}{2}}^{N_s-1} y_{ij} \quad (3b)$$

$$\begin{aligned} y_{ij} &= \sqrt{S}d_j + n_{ij} \\ &= \frac{m}{N_s}d_j + n_{ij} \end{aligned} \quad (3c)$$

$i = 0, \dots, N_s - 1 \quad j = 1, \dots, n$

Here, N_s is the number of samples per symbol and is assumed to be an even integer, as in [3], and $m = N_s\sqrt{S}$. The datum y_{ij} is assumed to be an unfiltered sample of the j th

received symbol d_j ($d_j = \pm 1$). The noise samples n_{ij} are assumed to be independent with zero mean and equal variance σ^2/N_s . The parameters m and σ are the conditional mean and standard deviation (given d_j), respectively, of the whole-symbol accumulator outputs ($Y_{\alpha j} + Y_{\beta j}$). The symbol signal-to-noise ratio SNR is given in terms of m and σ by

$$\text{SNR} = \frac{m^2}{2\sigma^2} \quad (4)$$

Conditioned on the sequence of symbol values $\mathbf{d} = (d_1, \dots, d_n)$, the y_{ij} 's are independent Gaussian random variables, as are the vectors of half-symbol accumulator outputs $\mathbf{Y}_\alpha = (Y_{\alpha 1}, \dots, Y_{\alpha n})$ and $\mathbf{Y}_\beta = (Y_{\beta 1}, \dots, Y_{\beta n})$. The statistics of \mathbf{Y}_α and \mathbf{Y}_β are given by

$$E\{Y_{\alpha j}|\mathbf{d}\} = E\{Y_{\beta j}|\mathbf{d}\} = \frac{m}{2}d_j \quad (5a)$$

$$E\{Y_{\alpha j}^2|\mathbf{d}\} = E\{Y_{\beta j}^2|\mathbf{d}\} = \frac{m^2}{4} + \frac{\sigma^2}{2} \quad (5b)$$

$$E\{Y_{\alpha j}Y_{\beta j}|\mathbf{d}\} = \frac{m^2}{4} \quad (5c)$$

$$E\{Y_{\alpha j}Y_{\alpha j'}|\mathbf{d}\} = E\{Y_{\beta j}Y_{\beta j'}|\mathbf{d}\} = \frac{m^2}{4}d_jd_{j'} \quad j' \neq j \quad (5d)$$

$$E\{Y_{\alpha j}Y_{\beta j'}|\mathbf{d}\} = \frac{m^2}{4}d_jd_{j'} \quad j' \neq j \quad (5e)$$

We begin the derivation of our result by rewriting Eq. (1) for SNR^* in the form

$$2\text{SNR}^* + 1 = \frac{\sum_{j=1}^n u_j^2}{\sum_{j=1}^n v_j^2} = \frac{U}{V} \quad (6)$$

where

$$u_j = Y_{\alpha j} + Y_{\beta j} \quad (7a)$$

$$v_j = Y_{\alpha j} - Y_{\beta j} \quad (7b)$$

$$U = \sum_{j=1}^n u_j^2 \quad (8a)$$

$$V = \sum_{j=1}^n v_j^2 \quad (8b)$$

The usefulness of expressing SNR^* in the form of Eq. (6) results from the fact that U and V are conditionally inde-

pendent random variables. Given the data vector \mathbf{d} , the random vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ are jointly Gaussian with

$$E\{u_j|\mathbf{d}\} = md_j \quad E\{u_j^2|\mathbf{d}\} = m^2 + \sigma^2 \quad (9a)$$

$$E\{v_j|\mathbf{d}\} = 0 \quad E\{v_j^2|\mathbf{d}\} = \sigma^2 \quad (9b)$$

$$E\{u_j v_j|\mathbf{d}\} = 0 = E\{u_j|\mathbf{d}\} E\{v_j|\mathbf{d}\} \quad (9c)$$

$$E\{u_j v_{j'}|\mathbf{d}\} = E\{v_j v_{j'}|\mathbf{d}\} = 0 \quad j' \neq j \quad (9d)$$

$$E\{u_j u_{j'}|\mathbf{d}\} = m^2 d_j d_{j'} = E\{u_j|\mathbf{d}\} E\{u_{j'}|\mathbf{d}\} \quad j' \neq j \quad (9e)$$

Equations (9a-e) imply that \mathbf{u} and \mathbf{v} are uncorrelated given \mathbf{d} . Because \mathbf{u} and \mathbf{v} are jointly Gaussian and uncorrelated, conditioned on \mathbf{d} , they are therefore conditionally independent given \mathbf{d} . Likewise the random variables U and V are conditionally independent given \mathbf{d} , because they are functions of \mathbf{u} and \mathbf{v} , respectively [4]. Thus, all of the moments of $2 \text{ SNR}^* + 1$ can be factored as follows:

$$E\{(2 \text{ SNR}^* + 1)^k\} = E_{\mathbf{d}}[E\{U^k|\mathbf{d}\} E\{V^{-k}|\mathbf{d}\}] \quad (10)$$

To evaluate the mean and standard deviation of SNR^* exactly, it is only necessary to compute the conditional moments $E\{U|\mathbf{d}\}$, $E\{U^2|\mathbf{d}\}$, $E\{V^{-1}|\mathbf{d}\}$, and $E\{V^{-2}|\mathbf{d}\}$, and then calculate the expectation over \mathbf{d} . All four of these conditional moments can be evaluated in closed form. The U -moments are simply the (positive) moments of a noncentral chi-squared random variable.

$$E\{U|\mathbf{d}\} = nE\{u_j^2|\mathbf{d}\} = n(m^2 + \sigma^2) \quad (11)$$

$$\begin{aligned} E\{U^2|\mathbf{d}\} &= nE\{u_j^4|\mathbf{d}\} + n(n-1)[E\{u_j^2|\mathbf{d}\}]^2 \\ &= n[(m^4 + 6m^2\sigma^2 + 3\sigma^4) + n(n-1)(m^2 + \sigma^2)^2] \\ &= n^2(m^2 + \sigma^2)^2 + n(4m^2\sigma^2 + 2\sigma^4) \end{aligned} \quad (12)$$

In Eq. (12), the first equality follows from the independence of the individual u_j 's, and the second equality follows from applying the general formula for the fourth moment of a nonzero-mean Gaussian random variable [5].

The V -moments are the negative moments of a central chi-squared random variable with n degrees of freedom. In the Appendix, it is shown that

$$E\{V^{-1}|\mathbf{d}\} = \frac{1}{(n-2)\sigma^2} \quad (\text{for } n > 2) \quad (13)$$

$$E\{V^{-2}|\mathbf{d}\} = \frac{1}{(n-2)(n-4)\sigma^4} \quad (\text{for } n > 4) \quad (14)$$

These V -moments are infinite if n is smaller than the indicated limits.

Note that all of the expressions in Eqs. (11), (12), (13), and (14) are independent of \mathbf{d} . Plugging these results into Eq. (10) leads to the following expressions for the moments of the estimator SNR^* (assuming $n > 2$ or $n > 4$ as appropriate):

$$E\{2 \text{ SNR}^* + 1\} = \frac{n(m^2 + \sigma^2)}{(n-2)\sigma^2} = (2 \text{ SNR} + 1) \frac{n}{n-2} \quad (15)$$

$$\begin{aligned} E\{(2 \text{ SNR}^* + 1)^2\} &= \left[(m^2 + \sigma^2)^2 + \frac{1}{n} (4m^2\sigma^2 + 2\sigma^4) \right] \\ &\quad \times \frac{n^2}{(n-2)(n-4)\sigma^4} \\ &= \left[(2 \text{ SNR} + 1)^2 + \frac{1}{n} (8 \text{ SNR} + 2) \right] \\ &\quad \times \frac{n^2}{(n-2)(n-4)} \end{aligned} \quad (16)$$

The mean and variance of SNR^* are calculated from Eqs. (15) and (16) as

$$\begin{aligned} E\{\text{SNR}^*\} &= \frac{1}{2} [E\{2 \text{ SNR}^* + 1\} - 1] \\ &= \text{SNR} + \frac{1}{n-2} (2 \text{ SNR} + 1) \quad (\text{for } n > 2) \end{aligned} \quad (17a)$$

$$\begin{aligned} \text{var}\{\text{SNR}^*\} &= \frac{1}{4} \text{var}\{2 \text{ SNR}^* + 1\} \\ &= \frac{1}{4} [E\{(2 \text{ SNR}^* + 1)^2\} - (E\{2 \text{ SNR}^* + 1\})^2] \\ &= \left[\frac{n}{n-2} \right]^2 \frac{1}{n-4} \left[2 \text{ SNR}^2 + (4 \text{ SNR} + 1) \frac{n-1}{n} \right] \\ &\quad (\text{for } n > 4) \end{aligned} \quad (17b)$$

In comparison, in [2] the following approximate asymptotic results were obtained:

$$E\{SNR^*\} \approx SNR + \frac{1}{n}(2 SNR + 1) \quad (18a)$$

$$\text{var}\{SNR^*\} \approx \frac{1}{n}[2 SNR^2 + 4 SNR + 1] \quad (18b)$$

By comparing Eq. (17) with Eq. (18) it is seen that the asymptotic expressions of [2] are accurate if $\frac{n-1}{n}$, $\frac{n-2}{n}$, and $\frac{n-4}{n}$ can all be approximated by 1.

In addition to the mean and variance, higher-order moments of the SSME may also be evaluated by applying known formulas for the moments of noncentral chi-squared random variables (U -moments) and for the moments of the reciprocal of a central chi-squared random variable (V -moments), and then solving Eq. (10) algebraically for the corresponding moments of SNR^* .

Finally, it is seen from Eq. (15) that it is easy to convert SNR^* into an unbiased estimator SNR^{**} (for $n > 2$) by making the definition

$$2 SNR^{**} + 1 = (2 SNR^* + 1) \frac{n-2}{n} \quad (19a)$$

i.e.,

$$SNR^{**} = SNR^* - \frac{1}{n}(2 SNR^* + 1) \quad (19b)$$

Then, assuming $n > 2$ for Eq. (20a) and $n > 4$ for Eq. (20b),

$$E\{SNR^{**}\} = SNR \quad (\text{for } n > 2) \quad (20a)$$

$$\text{var}\{SNR^{**}\} = \frac{1}{n-4} \left[2 SNR^2 + (4 SNR + 1) \frac{n-1}{n} \right] \quad (\text{for } n > 4) \quad (20b)$$

This should be the preferred form for the split-symbol moments estimator because, in addition to being unbiased, SNR^{**} also has a slightly smaller variance than SNR^* .

III. Conclusion

The technique introduced in this article enables one to compute exact closed-form expressions for the performance of the split-symbol moments estimator, for the case of unfiltered, undegraded data. The essential trick required for this computation was to algebraically manipulate the estimator formula into the form of a quotient of conditionally independent random variables. The calculations in this article confirm the accuracy of previously derived asymptotic expressions for unfiltered data. In addition, the exact performance expressions, unlike the asymptotic formulas, are useful even when the number of sampled split-symbols, n , is small. The exact formulas show that as long as $n > 2$ the conventional split-symbol estimator can be trivially scaled to form a signal-to-noise ratio estimator which is precisely unbiased. Furthermore, the same techniques developed here may be applied to obtain simplified expressions (though not closed form) for the more complicated case when the data is filtered.

Appendix

The Moments of the Reciprocal of a Chi-Squared Random Variable

Conditioned on \mathbf{d} , the random variable V is a central chi-squared random variable with n degrees of freedom. According to [6], its conditional probability function is therefore given by

$$p(V|\mathbf{d}) = \frac{V^{(n/2)-1} e^{-V/2\sigma^2}}{(2\sigma^2)^{n/2} \Gamma(n/2)} \quad V \geq 0 \quad (\text{A-1})$$

where $\Gamma(\cdot)$ is the gamma function. Thus, the conditional (negative) moments of V can be calculated (for $k > 0$) as

$$\begin{aligned} E(V^{-k}|\mathbf{d}) &= \int_0^\infty \frac{V^{(n/2)-k-1} e^{-V/2\sigma^2}}{(2\sigma^2)^{n/2} \Gamma(n/2)} dV \\ &= (2\sigma^2)^{-k} \frac{\Gamma[(n/2) - k]}{\Gamma(n/2)} \end{aligned}$$

$$\begin{aligned} &\times \int_0^\infty \frac{V^{(n/2)-k-1} e^{-V/2\sigma^2}}{(2\sigma^2)^{(n/2)-k} \Gamma[(n/2) - k]} dV \\ &= (2\sigma^2)^{-k} \frac{\Gamma[(n/2) - k]}{\Gamma(n/2)} \quad \text{if } n > 2k \\ &= \frac{1}{(n-2)(n-4)\cdots(n-2k)\sigma^{2k}} \quad (\text{A-2}) \end{aligned}$$

The next-to-last equality in Eq. (A-2) follows from noting that, if $n > 2k$, the rescaled integral in the preceding line is simply the integral of the probability density function of a chi-squared random variable with $n - 2k$ degrees of freedom. The last equality in Eq. (A-2) follows from the gamma function recurrence formula, $\Gamma(z + 1) = z\Gamma(z)$. Finally, it is easy to see that the integral in Eq. (A-2), and hence the corresponding moment of V , is infinite if $n \leq 2k$, because in this case $V^{(n/2)-k-1}$ becomes infinite near $V = 0$ at least as fast as V^{-1} .

References

- [1] L. D. Howard, "Signal and Noise Measurements in the NASA Deep Space Network," *20th URSI Conference*, Washington, D.C., August 10–19, 1981.
- [2] M. K. Simon and A. Mileant, "SNR Estimation for the Baseband Assembly," *TDA Progress Report 42-85*, vol. January–March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 118–126, May 15, 1986.
- [3] B. Shah and S. Hinedi, "Performance of the Split-Symbol Moments SNR Estimator in the Presence of Inter-Symbol Interference," *TDA Progress Report 42-98*, vol. April–June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 157–173, August 15, 1989.
- [4] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, New York: McGraw-Hill Book Company, 1965.
- [5] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Applied Mathematics Series, December 1972.
- [6] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*, New York: John Wiley and Sons, Inc., 1968.

515-32

272 331

118 TDA Progress Report 42-100

A VLSI Decomposition of the deBruijn Graph

O. Collins

Johns Hopkins University, Baltimore, Maryland

S. Dolinar, R. McEliece, and F. Pollara

Communications Systems Research Section

A new Viterbi decoder for convolutional codes with constraint lengths up to 15, called the Big Viterbi Decoder, is under development for the Deep Space Network. It will be demonstrated by decoding data from the Galileo spacecraft, which has a rate $1/4$, constraint-length 15 convolutional encoder on board. In this article, the mathematical theory underlying the design of the very-large-scale-integrated (VLSI) chips that are being used to build this decoder is explained. The deBruijn graph B_n describes the topology of a fully parallel, rate $1/\nu$, constraint length $n+2$ Viterbi decoder, and it is shown that B_n can be built by appropriately "wiring together" (i.e., connecting together with extra edges) many isomorphic copies of a fixed graph called a " B_n building block." The efficiency of such a building block is defined as the fraction of the edges in B_n that are present in the copies of the building block. It is shown, among other things, that for any $\alpha < 1$, there exists a graph G which is a B_n building block of efficiency $> \alpha$ for all sufficiently large n . These results are illustrated by describing a special hierarchical family of deBruijn building blocks, which has led to the design of the gate-array chips being used in the Big Viterbi Decoder.

I. Introduction and Summary

The n th order deBruijn graph B_n is the state diagram for an n -stage binary shift register. It is a directed graph with 2^n vertices, each labeled with an n -bit binary string, and 2^{n+1} edges, each labeled with an $(n+1)$ -bit binary string. The vertex labels represent the contents of the shift register at a given point of time. The label on an edge connecting one vertex to another represents the contents of the shift register preceded by the bit that is being input to

the shift register as it changes from one state to the next. Figure 1 is a representation of B_3 .

The deBruijn graph B_n gives the topology for a fully parallel Viterbi decoder for any rate $1/\nu$ convolutional code with a constraint length of $n+2$ ([3], Chapter 7). In such a decoder, a "butterfly" (a pair of add-compare-select units) must be located at each vertex of the graph and all communication between butterflies takes place along the edges of

the graph. In fact, the Deep Space Network (DSN) is currently developing such a decoder, called the Big Viterbi Decoder (BVD), for constraint-length 15 convolutional codes, to be used for the Galileo mission using a rate 1/4 code. The BVD has 2^{13} butterflies connected according to the topology of B_{13} . It is constructed from 256 identical gate-array chips, each containing 32 butterflies. These chips are arranged on 16 identical printed-circuit boards, each containing 16 chips. Of the 2^{14} "wires" (butterfly interconnections) in the decoder, 56 percent are internal to the chips, another 17 percent are internal to the boards, and 27 percent are inter-board, or "backplane" connections. Furthermore, these chips and boards are universal, in the sense that any deBruijn graph B_n with $n \geq 5$ can be built from copies of these same chips, and any B_n with $n \geq 9$ can be built from copies of these same boards. In this article, the theoretical background that led to the design of these chips and boards is given. See [1] and [5] for further details. (The BVD will be discussed at the end of the article—see Example K.)

II. Preliminaries About Strings

In this section, some notation is introduced and a few elementary facts about binary strings, which will be needed throughout the rest of the article, are established.

Definitions. A binary string is a sequence of bits, i.e., 0s and 1s. The *length* of a binary string x , denoted by $|x|$, is the number of bits in x . The *empty string* ϵ is the string with no bits. Thus, $|\epsilon| = 0$. The set of all strings of length n is denoted by $\{0, 1\}^n$. If x and y are two strings, the *concatenation* of x and y , denoted by xy or $x * y$, is the string obtained by following the bits of x by the bits of y . If $x = a * b * c$, then a is called a *prefix*, b is a *substring*, and c is a *suffix* of x . If $b * c$ isn't empty, then a is called a *proper prefix* of x ; if either a or c is nonempty, b is called a *proper substring* of x ; and, if $a * b$ isn't empty, c is called a *proper suffix* of x . If x is a nonempty binary string, then the symbol x^L (the *left part* of x) denotes the string obtained by removing the rightmost bit of x ; similarly, x^R (the *right part* of x) denotes the string obtained by removing the leftmost bit of x .

If S and T are sets of binary strings, then S *covers* T if every string in T has a substring in S . Similarly, S *prefixes* T if every string in T has a prefix in S . If T consists of a single string t , S covers or prefixes t . Also, S is *irreducible* if no string in S is a substring of any other. Finally, the *cost* of a set of strings S is $\text{cost}(S) = \sum_{s \in S} 2^{-|s|}$, where $|s|$ denotes the length of the string s .

Example (A). If $x = 1011$, then $|x| = 4$, $x^L = 101$, and $x^R = 011$. The set $S = \{10, 111\}$ covers $\{010, 100, 101,$

$110, 111\}$, and $\{1, 0000\}$ covers $\{0, 1\}^n$ for all $n \geq 4$. Similarly, $\{1, 000\}$ prefixes $\{1, 00000\}$, and $\{0, 10, 110, 111\}$ prefixes $\{0, 1\}^n$ for all $n \geq 3$. Also, $\{1, 000\}$ is irreducible, but $\{1, 001\}$ is not. The cost of the set $\{10, 111\}$ is $3/8$, $\text{cost}(\{1, 000\}) = 5/8$, and $\text{cost}(\{0, 1\}^n) = 1$, for all $n \geq 1$.

Theorem (1). If S is an irreducible set of strings, then every string x covered by S can be factored uniquely in the form $x = \lambda s \rho$, where $s \in S$, λ and ρ are (possibly empty) strings, and $(\lambda s)^L$ has no substring from S . This factorization will be called the *S-factorization* of x .

Proof: Since x is covered by S , x will have one or more substrings from S . Among these S -substrings, there will be a unique *leftmost* one, since no string in S covers any other. Call this unique leftmost S -substring s . Then plainly, $x = \lambda s \rho$ is the desired unique factorization.

Example (B). As noted above, $S = \{1, 0000\}$ is irreducible and covers all strings of length 4. The S -factorization of 1010 is $\epsilon * 1 * 010$, the S -factorization of 0101 is $0 * 1 * 01$, and the S -factorization of 0000 is $\epsilon * 0000 * \epsilon$.

Lemma (1). If S is irreducible and covers $\{0, 1\}^n$, then every string x of length n or greater will have a unique S -factorization, and if $x = \lambda s \rho$ is this factorization, then $|\lambda s| \leq n$.

Proof: Every string of length n or greater will have a substring of length n . This substring will be covered by S and, hence, so will x . Now let x be a string of length $\geq n$, and let $x = \lambda s \rho$ be its S -factorization, as described in Theorem (1). By definition of the S -factorization, $(\lambda s)^L$ is not covered by S . However, if $|\lambda s| > n$, then $|(\lambda s)^L| \geq n$, which would imply that $(\lambda s)^L$ is covered by S , a contradiction.

III. deBruijn Graphs and Subgraphs

The deBruijn graph B_n , which is the state diagram for an n -stage shift register, can be described as follows. There are 2^n vertices, each labeled with an n -bit binary string x . There is a directed edge from the vertex with label x to exactly two other vertices, viz., those with labels $0x^L$ and $1x^L$. The edge from x to $0x^L$ is labeled $0x$ and the edge from x to $1x^L$ is labeled $1x$. Similarly, there are exactly two edges directed into x , from $x^R 0$ and $x^R 1$, which are labelled $x0$ and $x1$. This definition is summarized in Fig. 2(a). For example, in Fig. 1, from the vertex 101 there are edges leading to $0(101)^L = 010$ and to $1(101)^L = 110$. Equivalently, one can define B_n by saying that it has 2^{n+1} edges, each labeled with an $(n+1)$ -bit binary string, and that the edge labeled X connects the vertices

with n -bit labels X^R and X^L . This alternative definition is summarized in Fig. 2(b). For example, in Fig. 1, the edge labeled 1011 is a directed edge from $(1011)^R = 011$ to $(1011)^L = 101$.

Next, the notion of a *subgraph* of a deBruijn graph needs to be defined. In this definition, and later, the symbols $E(G)$ and $V(G)$ stand for the number of edges and vertices in the graph G . Thus, for example, $E(B_n) = 2^{n+1}$ and $V(B_n) = 2^n$.

Definition. If H and G are graphs, H is called a *G-subgraph*, written $H \subseteq G$, if H has the same vertex set as G and an edge set that is a subset of the edge set of G . The *density* of a G -subgraph H , denoted by $\text{den}(H : G)$, is defined as $\text{den}(H : G) = E(H)/E(G)$.

Example (C). A B_n -subgraph of density 0 consists of 2^n isolated vertices, and a B_n -subgraph of density 1 is B_n itself. Figure 3 shows a B_3 -subgraph of density $6/16$, consisting of the eight vertices of B_3 and the six edges labeled $\{0010, 0011, 0100, 0101, 0101, 0111\}$.

Our goal is to build a large deBruijn graph B_N by connecting together multiple copies of a smaller graph, called a "building block." If one thinks of the building blocks as very-large-scale-integrated (VLSI) chips, it is natural to want to minimize the number of edges needed to connect the building blocks together. This goal leads to the following definition.

Definition. A graph H is a *building block* for a graph G if there exists G -subgraph \bar{H} , which is the disjoint union of several copies of H . The *efficiency* of H as a building block for G , denoted by $\text{eff}(H : G)$, is defined to be $\text{den}(\bar{H} : G)$. In other words, $\text{eff}(H : G)$ represents the fraction of the edges of G that are accounted for by the edges in the building blocks.

Theorem (2). If H is a building block for G , then

$$\text{eff}(H : G) = \frac{V(G)E(H)}{V(H)E(G)}$$

Proof: The G -subgraph \bar{H} has the same number of vertices as G and is the disjoint union of several disjoint copies of H . Since G has $V(G)$ vertices, and H has $V(H)$ vertices, this means that \bar{H} is the union of exactly $V(G)/V(H)$ copies of H . Since each of these copies of H has $E(H)$ edges,

$$E(\bar{H}) = E(H)(V(G)/V(H))$$

Thus,

$$\begin{aligned} \text{eff}(H : G) &= \text{den}(\bar{H} : G) \\ &= E(\bar{H})/E(G) \\ &= (E(H)(V(G)/V(H)))/E(G) \\ &= (V(G)E(H)/V(H)E(G)) \end{aligned}$$

Example (D). Any B_n -subgraph H is a building block of efficiency $\text{den}(H : B_n)$ for B_n . A B_n -subgraph of density 0 is a building block of efficiency 0 for any B_N with $N \geq n$. A B_n -subgraph of density 1 (i.e., B_n itself) cannot be a building block for any larger deBruijn graph, since the disjoint union of 2^{N-n} copies of B_n is a disconnected graph with the same number of edges as B_N , which is connected. In Fig. 4 are two copies of the graph H (from Fig. 3), relabeled and wired together with 20 new edges to form a graph isomorphic to B_4 . Since B_4 has 32 edges, and the two copies of H together have 12 edges, it follows that H is a building block for B_4 of efficiency $12/32 = 37.5$ percent. In fact, the graph in Fig. 3 is a building block of efficiency $3/8$ for any B_N with $N \geq 3$, as will be seen in Example (F), below. The building block of Fig. 3 is an example of what is called a *universal deBruijn building block*.

Definition. A *universal deBruijn building block* of order n is a B_n -subgraph that is a building block for any deBruijn graph B_N with $N \geq n$.

The following theorem shows that it is easy to compute the efficiency of any universal deBruijn building block.

Theorem (3). Let H be a universal deBruijn building block of order n . Then, for all $N \geq n$, $\text{eff}(H : B_N) = \text{den}(H : B_n)$. This common value will be called the *efficiency* of H as a deBruijn building block.

Proof: By Theorem (3), the efficiency of H as a building block for B_N is

$$\begin{aligned} \text{eff}(H : B_N) &= \frac{V(B_N)}{E(B_N)} \cdot \frac{E(H)}{V(H)} \\ &= \frac{2^N}{2^{N+1}} \cdot \frac{2^{n+1} \text{den}(H : B_n)}{2^n} \\ &= \text{den}(H : B_n) \end{aligned}$$

In the next section, a general construction for universal deBruijn subgraphs is described (Theorem 4), and in

Section V (Theorem 9) it will be seen that there exist universal deBruijn subgraphs whose efficiency approaches 1 as n approaches infinity.

IV. A General Construction for Universal deBruijn Subgraphs

In this section, the main theorem (Theorem 4) is presented, which gives a general construction for universal deBruijn building blocks. The key to this construction is a way of constructing a B_n -subgraph from a set of strings of length $\leq n$.

Definition. If S is a set of strings, $B_n(S)$ is defined to be the B_n -subgraph obtained by deleting from B_n all edges whose labels have a prefix in S .

Lemma (2). If S is irreducible, $E(B_n(S)) = 2^{n+1} \times (1 - \text{cost}(S))$; equivalently, $\text{den}(B_n(S) : B_n) = 1 - \text{cost}(S)$.

Proof: The 2^{n+1} edges in B_n are labeled with the $(n+1)$ -bit strings, and there are 2^{n+1} of them. For each $s \in S$, there are exactly $2^{n+1-|s|}$ $(n+1)$ -bit strings with s as a prefix. Since no $(n+1)$ -bit string can have two prefixes in S (no string in S is a prefix of any other, since S is irreducible), the subgraph $B_n(S)$ will have exactly $2^{n+1} - \sum_{s \in S} 2^{n+1-|s|} = 2^{n+1}(1 - \text{cost}(S))$ edges.

The main theorem of this article is the following.

Theorem (4). If S is irreducible and covers $\{0,1\}^n$, then $B_n(S)$ is a universal deBruijn building block of order n with efficiency $1 - \text{cost}(S)$.

The proof of Theorem (4) is postponed until several examples have been given and a stronger but more technical result has been stated and proved (Theorem 5).

Example (E). The set $S = \{0,1\}$ is irreducible, has cost 1, and covers $\{0,1\}^n$ for any $n \geq 1$. The corresponding subgraph $B_n(S)$ is a B_n -subgraph of density zero, and is plainly a building block of efficiency zero for any deBruijn graph with $N \geq n$.

Example (F). The set $S = \{1,000\}$ is irreducible, has cost $5/8$, and covers $\{0,1\}^3$. In this case, $B_3(S)$ is identical to the B_3 -subgraph in Fig. 3. Thus, Theorem (4) implies that the graph $B_3(\{1,000\})$ is a universal deBruijn building block with efficiency $3/8$, as asserted in Example (D).

The next theorem concerns a family of relabeled copies of the graph $B_n(S)$. If A is any binary string, the graph $B_n(S, A)$ is constructed from $B_n(S)$ by inserting the string A into each vertex or edge label just after the first (leftmost) occurrence of a substring from S . In Fig. 5, this construction is illustrated for the graph $B_3(\{1,000\})$.

Theorem (5). For all $N \geq n$,

$$\bigcup_{|A|=N-n} B_n(S, A) = B_N(S)$$

Example (G). In Fig. 6, Theorem (5) is illustrated by showing the two graphs $B_3(S, 0)$ and $B_3(S, 1)$ for $S = \{1,000\}$. When taken together, these two graphs form the graph $B_4(S)$, which is a subgraph of B_4 . Thus, by adding the 20 edges whose labels have a prefix from S (16 with prefix 1 and 4 with prefix 000), B_4 is obtained, and indeed, this is how one arrives at Fig. 4.

Proof of Theorem (5): Call the graph

$$\bigcup_{|A|=N-n} B_n(S, A)$$

appearing in the statement of the theorem the *union graph*. To prove Theorem (5), one needs to show that the union graph is a B_N -subgraph, and that its edges are exactly those whose labels have no prefix from S . To do this, one must prove the following four assertions (A always denotes a string of length $N - n$):

- (1) Every N -bit string occurs as a vertex label in some $B_n(S, A)$.
- (2) Every edge in $B_n(S, A)$ is an edge of the deBruijn graph B_N .
- (3) No edge label in $B_n(S, A)$ has an S -prefix.
- (4) Every $(N+1)$ -bit string without an S -prefix appears as an edge label on some $B_n(S, A)$.

Taken together, (1) and (2) show that the union graph is a B_N -subgraph, and (3) and (4) show that the edge labels in the union graph are the $(N+1)$ -bit strings without an S -prefix.

Proof of (1): Let X be an arbitrary N -bit string, and let $X = \lambda s \rho$ be its S -factorization. By Lemma (1), $|\lambda s| \leq n$, so that $|\rho| \geq N - n$. If the leftmost $N - n$ bits of ρ are denoted by A , then $\rho = A \rho'$ and, hence, $X = \lambda s A \rho'$. It follows that X appears as the label of the vertex $\lambda s \rho'$ in $B_n(S, A)$. For example, if $n = 3$, $S = \{1,000\}$, and $N = 8$, the 8-bit string $X = 01011110$ has S -factorization

0 * 1 * 011110. The first five bits of 011110 are 01111, and so 01011110 appears as the label on vertex 010 in $B_3(S, 01111)$.

Proof of (2): If an $(n+1)$ -bit edge label x in $B_n(S)$ has S -factorization $x = \lambda s \rho$, then neither λ nor ρ is empty: λ isn't empty, because no edge label in $B_n(S)$ has an S -prefix, and ρ isn't empty, since $|x| = n+1$, and by Lemma (1), any S -factorization has $|\lambda s| \leq n$. Thus, in $B_n(S)$, an edge with S -factorization $\lambda s \rho$ connects the vertices with labels $(\lambda s \rho)^R = \lambda^R s \rho$ and $(\lambda s \rho)^L = \lambda s \rho^L$. Furthermore, this representation of the vertex labels must in fact be the S -factorization of them, since an earlier occurrence of a substring from S in either $\lambda^R s \rho$ or $\lambda s \rho^L$ would imply an earlier occurrence of an S -string in $\lambda s \rho$. This means that in the graph $B_n(S, A)$, the edge $\lambda s \rho$ connects the vertices $\lambda^R s A \rho = (\lambda s A \rho)^R$ and $\lambda s \rho^L = (\lambda s A \rho)^L$. In other words, an edge with label X in $B_n(S, A)$ connects X^R to X^L , and by the definition in Fig. 2(b), this is an edge in the deBruijn graph B_n . For example, in Fig. (5), the edge with label 001A0 connects $001A0^R = 01A0$ to $001A0^L = 001A$, and this is an edge in the deBruijn graph $B_{|A|+3}$, for any string A .

Proof of (3): Let X be an edge label in the graph $B_n(S, A)$. Then by definition, X has the form $X = \lambda s A \rho$, where $\lambda s \rho$ is the S -factorization of an edge label in $B_n(S)$ with λ nonempty. If X had an S prefix, say s' , then either s' would be a prefix of $\lambda s \rho$, or s would be a proper prefix of s' . But both of these alternatives are impossible: s' cannot be a prefix of $\lambda s \rho$, since $\lambda s \rho$, being an edge of $B_n(S)$, has no S -prefix; and s cannot be a proper substring of s' , since no string in S is a proper substring of any other. Thus, every edge in $B_n(S, A)$ is an edge in $B_N(S)$, as asserted.

Proof of (4): To prove that every $(N+1)$ -bit string with no S -prefix occurs as an edge label in $B_n(S, A)$ for some A , let X be such a string and let $X = \lambda s \rho$ be its S -factorization, in which necessarily λ is nonempty. If A denotes the leftmost $N-n$ bits of ρ , then as above, $X = \lambda s A \rho'$. The string $\lambda s \rho'$ cannot have a prefix in S , for if s' were such a prefix, then either s' would be a prefix of X , or else s would be a proper substring of s' (since λ is nonempty), and both of these alternatives are impossible. Thus, $\lambda s \rho'$ is the label on an edge of $B_n(S)$, and so $X = \lambda s A \rho'$ appears as the label corresponding to that edge in the graph $B_n(S, A)$. For example, let $S = \{1, 000\}$, $n = 3$, and $N = 8$, and consider the nine-bit string $X = 001011100$, which has no S -prefix. The S -factorization of X is $X = 00 * 1 * 011100$. The first five bits of 011100 are 01110, and so X appears as the label on the edge 001 in the graph $B_3(S, 01110)$.

This completes the proof of Theorem (5).

This section concludes with the proof of Theorem (4).

Proof of Theorem (4): Theorem (5) explicitly shows that the union of 2^{N-n} copies of $B_n(S)$ forms a subgraph (namely, $B_N(S)$) of the big deBruijn graph B_N , and so B_N can be constructed simply by adding the edges missing from B_N to this union. Thus, B_n is a universal deBruijn building block. According to Theorem (3), the efficiency of a universal deBruijn building block is the same as its density and, by Lemma (2), the density of $B_n(S)$ is $1 - \text{cost}(S)$.

V. Construction of Low-Cost Covers for $\{0,1\}^n$

In Theorem (4), it was shown how to construct universal deBruijn building blocks from covering sets for $\{0,1\}^n$ of small cost, but only a few examples were cited. In this section, several general constructions for low-cost covers for $\{0,1\}^n$ will be given, thereby automatically producing [via Theorem (4)] efficient universal deBruijn building blocks.

To produce a cover for $\{0,1\}^n$, begin with an arbitrary irreducible set S of strings of length $\leq n$, which will be called a *precover* for $\{0,1\}^n$. In general, S will fail to cover a certain subset of $\{0,1\}^n$, which will be called $\text{Omit}_n(S)$. Denote the number of strings in $\text{Omit}_n(S)$ by $\text{omit}_n(S)$. Plainly, if $\text{Omit}_n(S)$ is adjoined to S , the resulting set, denoted by $C_n(S)$, will still be irreducible, will cover $\{0,1\}^n$, and its cost will be $\text{cost}(S) + 2^{-n} \text{omit}_n(S)$. This simple but useful construction is summarized in the following theorem.

Theorem (6). For any irreducible set S of strings of length $\leq n$, the set $C_n(S)$ is an irreducible cover of $\{0,1\}^n$ of cost $(\text{cost}(S) + 2^{-n} \text{omit}_n(S))$.

Example (H). If $S = \{1\}$, then $\text{Omit}_n(S) = \{00 \dots 0\}$, and $\text{omit}_n(S) = 1$ for all $n \geq 1$. Thus, $C_n(S) = \{1, 00 \dots 0\}$ is a cover for $\{0,1\}^n$ of cost $1/2 + 2^{-n}$, for all $n \geq 1$.

Example (I). If $S = \{10\}$, then for $n \geq 2$, $\text{Omit}_n(S)$ consists of the $n+1$ strings of the form $0^k * 1^{n-k}$ for $0 \leq k \leq n$. Thus, $\text{omit}_n(S) = n+1$, and $C_n(S)$ is a cover for $\{0,1\}^n$ of cost $1/4 + (n+1)/2^n$, for all $n \geq 2$.

Example (J). If $S = \{100, 1101\}$, then for $n \geq 4$, it can be shown that $\text{omit}_n(S) = 1 + \binom{n}{1} + \binom{n}{2}$ and, thus, $C_n(S)$ is a cover for $\{0,1\}^n$ of cost $3/16 + (1 + \binom{n}{1} + \binom{n}{2})/2^n$ for all $n \geq 4$.

If Theorems (4) and (6) are combined, it is found that if S is an irreducible set of strings of length $\leq n$, then $B_n(C_n(S))$ is a universal deBruijn building block of order n and efficiency $1 - \text{cost}(S) - 2^{-n} \text{omit}_n(S)$. For simplicity, $B_n(C_n(S))$ is denoted by $\bar{B}_n(S)$, and this fact is given as a theorem.

Theorem (7). If S is an irreducible set of strings of length $\leq n$, then the graph $\bar{B}_n(S)$ is a universal deBruijn building block of order n and efficiency $1 - \text{cost}(S) - 2^{-n} \text{omit}_n(S)$.

The following theorem is a partial generalization of Examples (H) and (I).

Theorem (8). Fix $m \geq 1$. If $S_m = \{10^{m-1}\}$, then as $n \rightarrow \infty$, $\text{omit}_n(S_m) = O(\alpha_m^n)$, where α_m is the largest positive root of the equation $z^m - 2z^{m-1} + 1 = 0$, which is strictly less than 2. Thus, for all $n \geq m$, $C_n(S)$ is a cover for $\{0, 1\}^n$ of cost $2^{-m} + O(\alpha_m/2)^n$, which approaches 2^{-m} as $n \rightarrow \infty$.

Proof: According to [2], if m is fixed, the generating function $f_m(z) = \sum_{n \geq 0} \text{omit}_n(S_m) z^n$ is given in closed form by $f_m(z) = 1/(1 - 2z + z^m)$. It follows then from the general theory of rational generating functions (see [4], Theorem 4.1.1), that $\text{omit}_n = O(\beta^n)$, where β is the reciprocal of the smallest positive root of the equation $1 - 2z + z^m = 0$, which is also the largest positive root of the "reciprocal" polynomial $P_m(z) = z^m - 2z^{m-1} + 1$. The largest root of $P_m(z)$ is strictly less than 2, since $P_m(1) = 0$, $P_m(2) = 1$, and $P'_m(z) > 0$ for $z > 2$.

Corollary. If c_n denotes the minimum cost for a cover for $\{0, 1\}^n$, then $\lim_{n \rightarrow \infty} c_n = 0$.

Proof: Theorem (8) implies that for any $m \geq 1$, $\lim_{n \rightarrow \infty} c_n \leq 2^{-m}$.

Remarks. We conjecture, but cannot prove, that $c_n = \Theta(1/n)$. However, McEliece and Swanson, in a forthcoming paper, will show that $c_n = \Omega(1/n)$ and $c_n = O(\log n/n)$. [The latter result is based on a more careful analysis of the type given in Theorem (8).]

In view of the connection between covers for $\{0, 1\}^n$ and universal deBruijn building blocks, the Corollary of Theorem (8) implies the following.

Theorem (9). There exist universal deBruijn building blocks whose efficiency is arbitrarily close to 1.

Although Theorem (8) gives an infinite family of reasonably cheap covers for $\{0, 1\}^n$, it does not produce the cheapest possible covers for all values of n . Indeed, Table 1 gives the cheapest known covers of $\{0, 1\}^n$, for $1 \leq n \leq 10$ and, therefore, also the most efficient universal deBruijn building blocks known, for orders up to 10. In every case, only the *precover* S is given, it being understood that the actual cover is the larger set $C_n(S)$. Notice that for $n \leq 7$, the "10...0" construction of Theorem (8) gives the best cover known, while for $8 \leq n \leq 10$ (and presumably for all larger values of n , too) the best cover is considerably more complicated. For $1 \leq n \leq 5$, it is believed that the values in Table 1 are the best possible. For larger values of n , however, improvements may be possible. For $n \geq 8$, the covers described in the table are based on the general "{100, 1101}" cover described in Example (J). For example, for $n = 8$, $\text{omit}(\{100, 1101\}) = 37$, so that $C_8(\{100, 1101\})$ is a cover for $\{0, 1\}^8$ of cost $1/8 + 1/16 + 37/256 = 85/256$. However, by trial and error, it is found that of the 37 strings in $\text{Omit}(\{100, 1101\})$, all but six are covered by $\{010101, 010111, 011111, 0000001, 0000101, 0000111\}$. Thus, if $\{100, 1101\} \cup \{010101, 010111, 011111, 0000001, 0000101, 0000111\}$ is used as a precover, Theorem (6) guarantees a cover of cost $1/8 + 1/16 + 1/64 + 1/64 + 1/64 + 1/128 + 1/128 + 1/128 + 6/256 = 72/256$, as shown in Table 1. (Using $\{10\}$ as a precover for precover for $n = 8$ results in a cover of cost $73/256$.)

VI. Hierarchical Building Blocks

It has been seen that the universal deBruijn building blocks described in Theorem (4) can be used to build deBruijn graphs of any size. Surprisingly, however, they can also be used as building blocks for larger universal deBruijn building blocks! This is useful in practice when many chips must be put on several boards, and the boards are then wired together to make the deBruijn graph. The main theorem here is the following.

Theorem (10). Suppose $k \leq n$. If S is irreducible and covers $\{0, 1\}^k$, T is irreducible and covers $\{0, 1\}^n$, and if S prefixes T , then $B_k(S)$ is a building block for $B_n(T)$. Furthermore,

$$\text{eff}(B_k(S) : B_n(T)) = \frac{1 - \text{cost}(S)}{1 - \text{cost}(T)}$$

Proof: Theorem (5) says that $B_n(S)$ is equal to

$$\bigcup_{A: |A|=n-k} B_k(S, A)$$

and so $B_k(S)$ is a building block for $B_n(S)$. But since every string in T has a prefix in S , then $B_n(S) \subseteq B_n(T)$, so that $B_k(S)$ is also a building block for $B_n(T)$. To calculate $\text{eff}(B_k(S) : B_n(T))$, Theorems (2) and (4) are used:

$$\begin{aligned} \text{eff}(B_k(S) : B_n(T)) &= \frac{V(B_n(T))E(B_k(S))}{V(B_k(S))E(B_n(T))} \\ &= \frac{2^n 2^{k+1} (1 - \text{cost}(S))}{2^k 2^{n+1} (1 - \text{cost}(T))} \\ &= \frac{1 - \text{cost}(S)}{1 - \text{cost}(T)} \end{aligned}$$

Lemma (3). If S is any set of strings, and if $k \leq n$, then $C_k(S)$ prefixes $C_n(S)$.

Proof: By definition, $C_k(S) = S \cup \text{Omit}_k(S)$ and $C_n(S) = S \cup \text{Omit}_n(S)$. A string $s \in \text{Omit}_n(S)$ is a string of length n with no substring from S . The k -bit prefix of s is a string of length k , which also has no substring from S , and so this prefix is in $\text{Omit}_n(S)$. Thus, every string in $C_n(S)$ is either in S or has a prefix in $\text{Omit}_k(S)$.

Theorem (11). If $k \leq n$, then $\bar{B}_k(S)$ is a building block for $\bar{B}_n(S)$, and

$$\text{eff}(\bar{B}_k(S) : \bar{B}_n(S)) = \frac{1 - \text{cost}(S) - 2^{-k} \text{omit}_k(S)}{1 - \text{cost}(S) - 2^{-n} \text{omit}_n(S)}$$

Proof: The proof follows from Theorem (10) and Lemma (3).

Example (K). Returning to the Big Viterbi Decoder mentioned briefly in Section I: the BVD requires the construction of the deBruijn graph B_{13} using 256 one-chip realizations of the graph $\bar{B}_5(\{10\})$, which, by Theorem (7) and Example (I), is a universal deBruijn building block of efficiency $18/32$ (see Fig. 7), and so it contains exactly $64 \cdot \frac{18}{32} = 36$ edges.

According to Theorem (11), $\bar{B}_5(\{10\})$ is a building block for $\bar{B}_9(\{10\})$ and, in the BVD, 16 of the $\bar{B}_5(\{10\})$ -chips are wired together on a printed-circuit board to make a $\bar{B}_9(\{10\})$ board. Now $\bar{B}_9(\{10\})$ is a universal deBruijn building block of efficiency $374/512$, and so it contains exactly $1024 \cdot \frac{374}{512} = 748$ edges. However, $16 \cdot 36 = 576$ of these edges are internal to the component chips, so that each $\bar{B}_9(\{10\})$ board actually has only $748 - 576 = 172$ printed wires. Finally, since $\bar{B}_9(\{10\})$ has efficiency $374/512$ as a deBruijn building block, in order to build B_{13} , there will be $2^{14} \cdot (1 - \frac{374}{512}) = 4416$ backplane wires, i.e., wires external to the board. In summary:

Unit type	Number of units	Wires/unit	Total wires
chip	256	36	9216
board	16	172	2752
backplane	1	4416	4416
			16384

References

- [1] O. Collins, F. Pollara, S. Dolinar, and J. Statman, "Wiring Viterbi Decoders (Splitting deBruijn Graphs)," *TDA Progress Report 42-96*, Jet Propulsion Laboratory, Pasadena, California, pp. 93-103, February 15, 1989.
- [2] L. J. Guibas and A. M. Odlyzko, "String Overlaps, Pattern Matching, and Nontransitive Games," *J. Comb. Theory A*, vol. 30, pp. 183-208, 1981.
- [3] R. J. McEliece, *The Theory of Information and Coding*, Reading, Massachusetts: Addison-Wesley, 1977.
- [4] R. Stanley, *Enumerative Combinatorics, Vol. 1*, Monterey, California: Wadsworth and Brooks-Cole, 1986.
- [5] J. Statman, G. Zimmerman, F. Pollara, and O. Collins, "A Long Constraint VLSI Viterbi Decoder for the DSN," *TDA Progress Report 42-95*, Jet Propulsion Laboratory, Pasadena, California, pp. 134-142, November 15, 1988.

Table 1. Cheapest covers of $\{0,1\}^n$, for $1 \leq n \leq 10$

n	Cost $\cdot 2^n$	Efficiency	S (precover)
1	2	0.000	{1}
2	3	0.250	{1}
3	5	0.375	{1}
4	9	0.438	{1} or {10}
5	14	0.563	{10}
6	23	0.641	{10}
7	40	0.688	{10}
8	72	0.719	{100, 1101, 010101, 010111, 011111, 0000001, 0000101, 0000111}
9	127	0.752	{100, 1101, 0000001, 0101011, 0101111, 0111111, 00001011, 00001111, 01010101}
10	229	0.776	{100, 1101, 01010101, 01010111, 01111111, 000000001, 000000101, 000000111, 000010101, 000010111, 000011111}

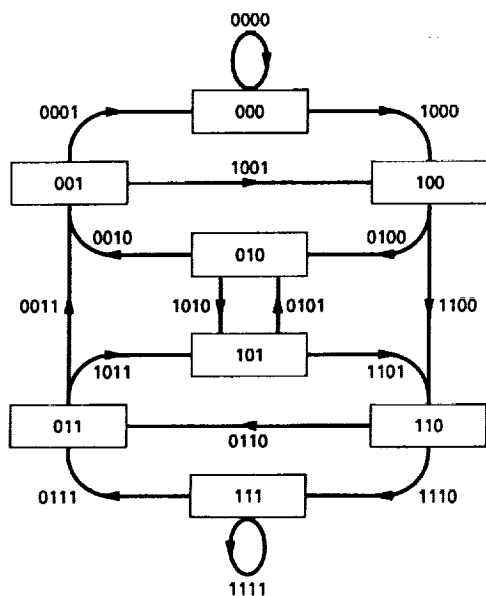


Fig. 1. The deBruijn graph B_3 .

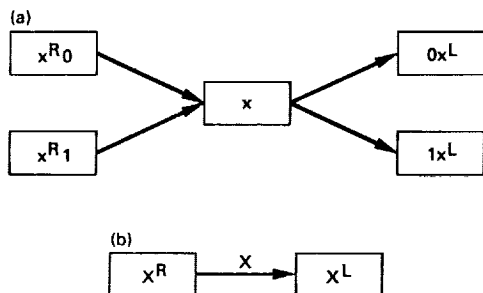


Fig. 2. Two equivalent definitions of the de-Bruijn graph B_n : (a) four vertices connected to the vertex labelled x (x is an n -bit string), and (b) vertices at the left and right ends of the edge labelled X (X is an $(n+1)$ -bit string).

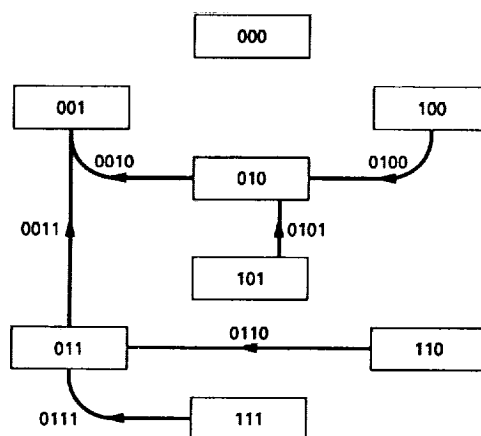


Fig. 3. A B_3 subgraph with density $6/16$. In the notation of Section IV, this is the graph $B_3(\{1,000\})$. In the notation of Section V, it is the graph $\bar{B}_3(\{1\})$. It is a universal deBruijn subgraph of efficiency $3/8$.

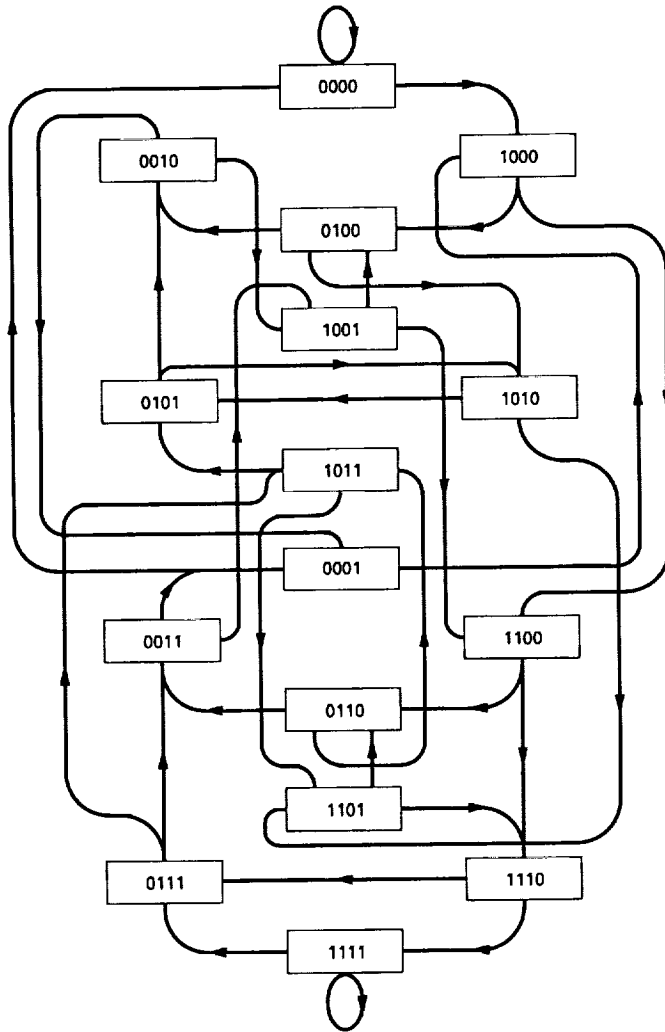


Fig. 4. Two copies of the B_3 subgraph H from Fig. 3, relabelled and wired together to make B_4 (edge labels omitted).

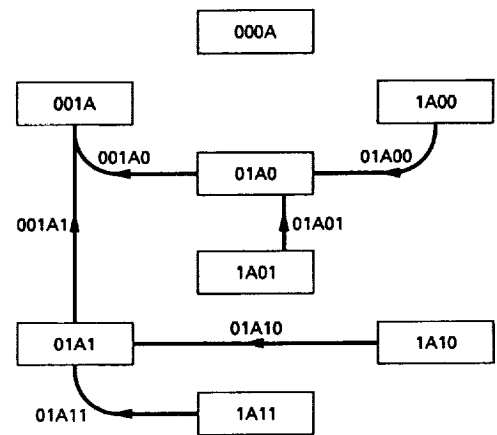


Fig. 5. The graph $B_3(S, A)$, obtained by inserting the symbol A immediately after the first occurrence of a substring from $S = \{1, 000\}$ in the corresponding label in $B_3(S)$ (Fig. 3).

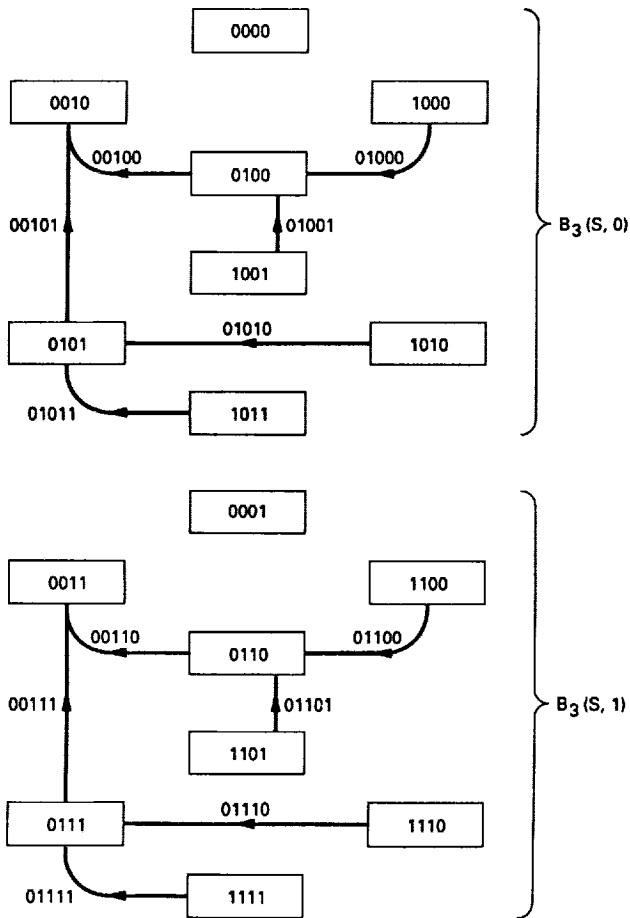


Fig. 6. Illustrating Theorem 4.6: $B_3(S,0) \cup B_3(S,1) = B_4(S)$ (compare with Fig. 4).

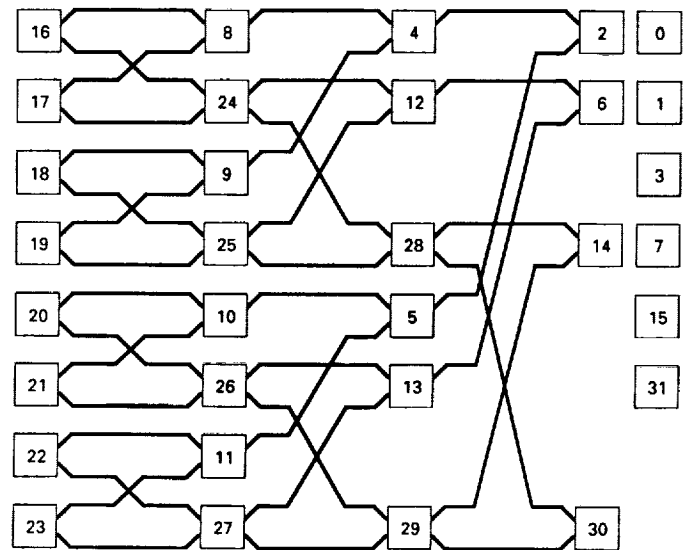


Fig. 7. One possible layout of the $B_5(\{10\})$ chip used to build the BVD. All edges are directed from left to right. The vertex labels shown are the decimal equivalents of the actual five-bit binary labels, and the edge labels have been omitted.

Validity of the Two-Level Model for Viterbi Decoder Gap-Cycle Performance

S. Dolinar and S. Arnold
Communications Systems Research Section

A two-level model has previously been proposed for approximating the performance of a Viterbi decoder which encounters data received with periodically varying signal-to-noise ratio. Such cyclically gapped data is obtained from the Very Large Array (VLA), either operating as a stand-alone system or arrayed with Goldstone. This approximate model predicts that the decoder error rate will vary periodically between two discrete levels with the same period as the gap cycle. It further predicts that the length of the gapped portion of the decoder error cycle for a constraint length K decoder will be about $K - 1$ bits shorter than the actual duration of the gap.

In this article the two-level model for Viterbi decoder performance with gapped data is subjected to detailed validation tests. Curves showing the cyclical behavior of the decoder error burst statistics are compared with the simple square-wave cycles predicted by the model.

The validity of the model depends on a parameter often considered irrelevant in the analysis of Viterbi decoder performance, the overall scaling of the received signal or the decoder's branch-metrics. This article examines three scaling alternatives: optimum branch-metric scaling and constant branch-metric scaling combined with either constant noise-level scaling or constant signal-level scaling.

The simulated decoder error cycle curves roughly verify the accuracy of the two-level model for both the case of optimum branch-metric scaling and the case of constant branch-metric scaling combined with constant noise-level scaling. However, the model is not accurate for the case of constant branch-metric scaling combined with constant signal-level scaling.

I. Introduction

Voyager's concatenated code performance using gapped data as obtained from the Very Large Array (VLA) arrayed with Goldstone or the VLA operating as a stand-

alone antenna was previously analyzed in [1]. This analysis rested on an assumed ad hoc approximate model for the errors produced by a Viterbi decoder when the signal-to-noise ratio (SNR) is subjected to variations between two discrete levels according to a fixed duty cycle. The model

of [1] predicted that the decoder error rate would also vary periodically between two discrete levels with the same period as the gap cycle. It further predicted that the duration of the gapped portion of the decoder error cycle for a constraint length K decoder would be about $K - 1$ bits shorter than the actual VLA data gap.

The two-level model in [1] was generalized from an argument presented in [2] applicable to an ideal case in which the SNR inside the gap is zero ($-\infty$ dB), while the SNR outside the gap is infinite ($+\infty$ dB). The arguments of [2] predicted that the two-level model would be precisely accurate in this extreme case.

Simulated performance curves for gapped data were presented in [2] and in a follow-up article [3]. The curves in [2] were compared with predictions of the generalized two-level model in the appendix of [1]. Results showed fairly good agreement between simulations and model predictions, but some small discrepancies were also noted.

The performance curves of [2] and [3] and the validity checks in the appendix of [1] were limited in that they only considered overall average Viterbi decoder error rates. This article looks at the decoder's error characteristics in finer detail by examining the cyclical variations of error rate over a gap cycle. This level of detail is necessary for a full direct validation of the model, rather than just an incomplete indirect confirmation of the model based on average error rates.

In the course of the new investigation, it was discovered that a parameter usually considered irrelevant in the analysis of Viterbi decoder performance, the overall scaling of the received signal or the decoder's branch metrics, must be carefully accounted for in order to ensure maximum-likelihood performance. When the received-signal characteristics (e.g., SNR) change abruptly, so should the scaling of branch metrics. If the signal characteristics fluctuate, a Viterbi decoder designed with constant branch metric scaling is not a maximum-likelihood decoder unless the received signal is rescaled optimally. Since hardware and software versions of the Viterbi decoder often have constant branch-metric scaling built in, the effects of nonoptimum received signal scaling on these decoders must be evaluated.

Fortunately, the scaling issue seems moot for most of the VLA applications and analysis performed up to now. The optimum weights for combining the VLA and Goldstone received signals happen to provide the exact scaling required for optimum performance by a decoder with constant branch-metric scaling. The simulations of [2]

and [3] appear to have been based on near-optimum scaling. However, most of the Viterbi-decoder simulation software developed recently for nongapped scenarios is based on nonoptimum received-signal scaling, which can greatly widen the effective gap if used by a decoder with constant branch-metric scaling. Therefore, extreme care must be exercised in adapting existing software to gapped scenarios.

II. Method of Analysis

A. Decoder Error Cycles

When the SNR varies over a gap cycle, the errors made by the decoder can be expected to vary cyclically with the same period. In order to determine the variation of decoder performance over a gap cycle, it is necessary to obtain some estimate of the decoder's instantaneous error rate at every point within the cycle, rather than to simply average its error rate over the entire simulation run. Knowledge of the decoder's error-cycle characteristics is especially important when concatenated coding is used and the convolutionally decoded output must be further decoded according to an outer code.

As a measure of decoder performance, we chose to calculate the *burst inclusion probability* (BIP) rather than the bit error rate (BER) or the symbol error rate (SER). A given bit is included within an error burst if it is not part of a string of $K - 1$ or more consecutive correctly decoded bits, where K is the constraint length of the convolutional code. This definition is equivalent to the definition of an error burst given in [4]. The burst inclusion probability at any particular point in the gap cycle is the probability that a bit at that location relative to the gap will be included within a decoder error burst. The BIP can be used to give reasonable estimates of either the BER or the SER for various symbol sizes, and the cyclical variation of the BIP accurately reflects the cyclical variation of the decoder's instantaneous BER or SER.

If the SNR is constant over a gap cycle, the gap is effectively missing and the BIP is also a constant over the entire cycle. However, when the SNR switches between two levels, SNR_0 inside the gap and SNR_1 outside the gap, the BIP will also oscillate. The model of [1] predicts the magnitudes of the peaks and troughs of this oscillation and the width of the peaks. It also predicts instantaneous jumps from peak to trough and vice versa. The simulations reported here investigate the accuracy of these predictions. How big are the actual peaks and troughs? How wide are the actual peaks? What are the actual rise times and fall times?

The two-level model of [1] predicts simple square-wave error cycles. The BIP should oscillate between two levels, BIP_0 and BIP_1 , for “gapped” and “ungapped” bits, respectively. A discrete jump from level BIP_1 to level BIP_0 should occur at the beginning of every gap, and a discrete jump from level BIP_0 to level BIP_1 should occur $K - 1$ bits earlier than the end of every gap. Furthermore, the model predicts that these two discrete levels, BIP_0 and BIP_1 , are the same as the constant BIP levels that would prevail at constant SNR levels SNR_0 and SNR_1 , respectively. The model also predicts the same square-wave cycles for the decoder’s BER and SER, but these predictions are only tested indirectly in this article by measuring the burst inclusion probability.

B. Description of Simulations Performed

Simulations were run for a Viterbi decoder operating on data encoded by the NASA-standard (7,1/2) code. The simulated decoder generated its own random noise samples, based its decisions on unquantized metrics, and used a decoding truncation window of 64 bits.

The decoder was confronted with a periodically varying SNR. The SNR varied between two levels in synchronism with a gap cycle of fixed period T . For a data rate R , the gap cycle measured in decoded bits is of length RT . Within every gap cycle is a gap of length G , or RG bits, during which the SNR is at its lower level, $SNR = SNR_0$. For the remainder of each cycle, the SNR is at its higher level, $SNR = SNR_1$.

In the simulations reported here, the full gap cycle was taken to be $RT = 1123$ bits, and the gapped portion of the cycle was $RG = 35$ bits. This corresponds to the VLA gap cycle parameters used in [3] for the highest Voyager-Neptune data rate of $R = 21.6$ kbps.

Each simulation decoded a total of 800,000 bits, just over 712 full gap cycles. Separate simulations were run for many different combinations of SNR_0 and SNR_1 .

III. Viterbi Decoder Scaling Considerations

Viterbi decoders are supposed to operate as maximum-likelihood decoders. However, real-world Viterbi decoders, such as the DSN’s maximum-likelihood convolutional decoder (MCD), may be designed on the assumption of constant SNR or slowly varying SNR. Similarly, software simulations of Viterbi decoder operations usually implicitly assume nonvarying SNR. Such decoders and simulations of decoders are significantly suboptimum if

the received-signal characteristics change abruptly within a time period comparable to a few constraint lengths.

There are two types of scaling considerations that become important when the SNR is time-varying. How is the received signal scaled? How are the decoder’s branch metrics scaled? With nonvarying SNR, the overall scaling of the received signal and the branch metrics is irrelevant. They may all be multiplied or divided by any convenient scale factor and the decoder will produce the same results. When the SNR is time-varying, the scale factor does matter. To be optimum, the scaling of the branch metrics and/or the received signal must change as the signal and noise characteristics vary with time.

A. Optimum Branch-Metric Scaling

The model for the received symbol r_k is

$$r_k = m_k d_k + \sigma_k n_k \quad (1)$$

where $d_k = \pm 1$ is the k th encoded symbol, n_k is zero-mean, unit-variance Gaussian noise, and m_k , σ_k are the signal level and noise standard deviation, respectively, for the k th received symbol. The conditional probability density function for a set of received symbols $\{r_k\}$ computed along a path corresponding to encoded symbols $\{c_{ik}\}$ for the i th codeword ($c_{ik} = \pm 1$) is

$$p_i(\{r_k\}) = \prod_k \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-(r_k - m_k c_{ik})^2 / 2\sigma_k^2} \quad (2)$$

A maximum-likelihood metric¹ M_i can be obtained from Eq. (2) by taking logarithms and ignoring some terms which are independent of i .

$$M_i = \sum_k \frac{1}{2} [|r_k| - c_{ik} r_k] \alpha_k \quad (3)$$

where α_k is the optimum branch-metric scaling factor,

$$\alpha_k = \frac{m_k}{\sigma_k^2} \quad (4)$$

When the signal characteristics are constant with time, the scale factor α_k is also constant and may be disregarded, because an overall scaling of all metrics does not

¹ The metric in Eq. (3) is called the “sign-magnitude metric” because $\frac{1}{2} [|r_k| - c_{ik} r_k]$ equals $|r_k|$ if c_{ik} and r_k have opposite signs, or 0 if they have the same sign. This metric is obtained by subtracting the usual correlation metric from $\sum_k |r_k| \alpha_k$ and dividing by 2.

matter. Both hardware decoders and simulated decoders often have this assumption built into their computation of branch metrics, i.e., they assume $\alpha_k = 1$ (or some other suitable constant).

B. Received-Signal Scaling

If the signal characteristics vary with time, scaling may also be performed on the received signal itself. Some sort of automatic gain control may be functioning to keep the received signal within the dynamic range of the receiver. When baseband combining is performed, additional weights are applied to each separate component of the received signal. If these scale factors change quickly enough during a gap cycle, the resulting variation in α_k must be accommodated by the Viterbi decoder, or else the decoder is not maximum likelihood.

If the decoder cannot adjust the optimum branch-metric weights α_k dynamically, optimum performance may still be achieved if the received signal is scaled in such a way that $\alpha_k = 1$ for all k . This optimum scaling of the received signal is

$$r_k = \rho_k d_k + \sqrt{\rho_k} n_k \quad (\text{optimum scaling}) \quad (5)$$

where ρ_k is twice the instantaneous channel symbol SNR,

$$\rho_k = \frac{m_k^2}{\sigma_k^2} = \frac{2}{N} \left(\frac{E_b}{N_0} \right)_k \quad (6)$$

In Eq. (6), $1/N$ is the rate of the convolutional code, and $(E_b/N_0)_k$ is the bit-energy-to-noise ratio for the k th symbol.

Two other nonoptimum received-signal scalings have usually been assumed in software simulations of the Viterbi decoder. Sometimes the signal level is assumed to be a constant (± 1), and sometimes the standard deviation of the noise is assumed to be constant. These two standard scaling assumptions can be written compactly in terms of the parameter ρ_k as

$$r_k = d_k + \frac{n_k}{\sqrt{\rho_k}} \quad (\text{constant signal-level scaling}) \quad (7)$$

or

$$r_k = \sqrt{\rho_k} d_k + n_k \quad (\text{constant noise-level scaling}) \quad (8)$$

C. Scaling Used in Existing Software Simulations and DSN Hardware

Historically, most versions of the Communications Systems Research Section's Viterbi decoder software have

assumed constant branch-metric scaling. The DSN's hardware MCD for the NASA-standard (7,1/2) code also makes use of constant branch-metric scaling. The Big Viterbi Decoder (BVD) currently under development has programmable branch-metric scaling.²

Most recent simulation software has incorporated constant signal-level scaling of the received symbols. Some older simulations, apparently³ including the VLA gapped data simulations reported in [2] and [3], were based on constant noise-level scaling.

The DSN's current hardware uses an automatic gain control circuit (AGC), but the response time is far too slow⁴ to cause any fluctuation during a VLA gap. In a VLA-Goldstone arrayed system, optimum combining weights must be applied in order to fully utilize the SNR of each array component. It can be shown from Eq. (20) of [5] that the scaling produced by applying the optimum weights for baseband combining is equivalent to the optimum scaling for a decoder using constant branch-metric scaling, provided that the signal characteristics at the reference antenna are unchanging. This should be true if Goldstone is the reference antenna. In a VLA stand-alone system, the scaling will be optimum if the received data is ignored during the gap, i.e., if all the branch metrics during the gap are set to zero.

D. Effects of Received-Signal Scaling on Decoders With Constant Branch-Metric Scaling

If the decoder is constrained to have constant branch-metric scaling, the decoder's error characteristics are very much influenced by the received-signal scaling. Some appreciation of these differences may be gained by reanalyzing the ideal gap scenario upon which the model of [1] is based, as explicated in [2]. If the SNR outside the gap is infinite and the SNR inside the gap is zero, the decoder should have perfect state information at all places outside the gap, including the two bit times positioned at the brink of the gap. Since the state information at the end of the gap includes $K - 1$ bits that were transmitted during the gap, knowledge of the state at the end of the gap allows the decoder to correctly decode $K - 1$ gapped bits in addition to all the bits outside the gap.

² Oliver Collins, personal communication.

³ The scaling assumed in [2] and [3] was not documented, and this tentative conclusion can only be inferred by deciphering the original computer code, some of which was written in assembly language.

⁴ Joseph Statman, personal communication.

However, this is not true if nonoptimum branch-metric scaling is used. For example, consider the case of *constant branch-metric scaling combined with constant signal-level scaling*. At the leading edge of the gap, the accumulated metrics will have all reached some distribution of steady-state values corresponding to finite signal and zero noise according to Eq. (7). At the very next bit time, if the SNR drops nearly to zero, the noise component of Eq. (7) will completely overwhelm the data component, and the branch metrics will all be random and huge. Large random differences in the branch metrics will wipe out any steady-state differences in accumulated state metrics that were built up over the entire ungapped period. Large random branch metrics will continue to be added throughout the gap, resulting in a state-metric distribution at the end of the gap which is completely random and very large. Whichever state metric is smallest at the end of the gap is likely to be smallest by a large margin over the second smallest metric. After the gap is finished, the noise component of Eq. (7) returns to zero, the branch metrics revert to their usual magnitudes and all point toward the correct state. However, the metric differences arising from adding these smaller branch metrics following the gap are unlikely to be large enough to overcome the huge accumulated metric difference in favor of the random state with the smallest metric at the end of the gap. Thus, the decoded state path will almost assuredly pass through this random state. The best-metric state path traced backward from this random state will find its way to another random state at the start of the gap.

Thus, with constant signal-level scaling, the decoded state path passes through a random state at the beginning of the gap and another random state at the end of the gap. The decoded bits are totally random from the beginning of the gap to the end. Also, there are additional decoded bit errors before the beginning of the gap along the best-metric path from the correct state to the first random state, and after the end of the gap along the best-metric path from the last random state back to the correct state. Thus, the gap is effectively widened by some number of bits on both ends.

Now consider the case of *constant branch-metric scaling combined with constant noise-level scaling of the received signal*. In this case, all the accumulated metrics at the beginning of the gap will have reached the same relative steady-state distribution as before, but all differences in accumulated metrics will be magnified by a scale factor $\sqrt{\rho}$ related to the high SNR outside the gap. The branch metrics during the gap will be just as random as before, but their magnitudes will be small relative to the accumulated metric differences at the beginning of the gap. At the end

of the gap, a random state will have the best metric, but the metrics of all states will be small relative to the metric differences that are obtained along the path portions outside the gap. After the end of the gap, the decoder will eventually find the correct state and will trace the correct state path all the way back to the trailing edge of the gap. The branch-metric information accumulated after the gap outweighs the random state-metric differences existing at the end of the gap, and therefore the correct state path may be traced all the way to the gap edge.

Constant noise-level scaling thus corresponds to the ideal case considered originally, where the decoder knows the correct state both at the beginning and at the end of the gap. The random path followed between these two states during the gap corresponds to a string of random decoded bits which is $K-1$ bits shorter than the actual length of the gap. A similar argument reaches the same conclusion for the case of optimum scaling, or, equivalently, for arbitrary received-signal scaling combined with optimum branch-metric scaling.

In summary, both optimum scaling and constant noise-level scaling produce decoded error cycles that precisely match the predicted model in the extreme case of zero SNR inside the gap and infinite SNR outside the gap. In contrast, constant signal-level scaling does not fit the model even in this ideal case. The simulation results in the next section generalize these conclusions to nonideal cases having nonzero, finite SNR levels.

IV. Decoder Error-Cycle Curves

The burst inclusion probability (BIP) was evaluated as a function of location within a gap cycle for the three types of received-signal scaling defined above, assuming that the branch-metric computations used a constant scale factor $\alpha_k = 1$. As noted earlier, the case of "optimum" scaling of the received signal is equivalent to arbitrary signal scaling combined with optimum branch-metric scaling, so this case shows the performance of a true maximum-likelihood decoder. The BIP was evaluated for various combinations of SNR_0 and SNR_1 . These curves were plotted against a background showing the effective gap width and the constant error levels predicted by the two-level model in [1].

Figures 1(a), (b), and (c) show the error-cycle curves for a series of cases with $SNR_0 = 0$ (the VLA stand-alone case considered in [1]). The SNR_1 values for the five curves plotted were 1 dB, 2 dB, 3 dB, 4 dB, and $+\infty$ dB. Figure 1(a) plots these five curves for optimum branch-metric

scaling, Fig. 1(b) plots the same curves assuming constant noise-level scaling, and Fig. 1(c) plots them for constant signal-level scaling.

Figures 2(a), (b), and (c) and 3(a), (b), and (c) show the error-cycle curves for a series of cases with SNR_0 3 dB lower than SNR_1 (the VLA-Goldstone equal array case considered in [1]). The SNR_1 values for the four curves plotted in Figs. 2(a), (b), and (c) are 1 dB, 2 dB, 3 dB, and 4 dB. The SNR_1 values for the three curves plotted in Figs. 3(a), (b), and (c) are 1.8 dB, 2 dB, and 2.2 dB.

Figures 4(a), (b), and (c) show the error-cycle curves for a series of cases with SNR_0 2 dB lower than SNR_1 . The SNR_1 values for the four curves plotted in Figs. 4(a), (b), and (c) are 1 dB, 2 dB, 3 dB, and 4 dB.

In each figure, dashed horizontal lines show constant BIP levels corresponding to constant SNR levels indicated on the right vertical axis. These are the constant BIP levels predicted by the two-level model for the decoder error-cycle curves. Dashed vertical lines on each figure show the predicted effective beginning and end of the gap. The effective beginning equals the beginning of the actual gap, and the effective end is $K - 1$ bits earlier than the actual end. If the two-level model is correct, the simulated decoder error-cycle curves should be framed exactly by a square wave constructed from the appropriate combination of dashed horizontal and vertical lines.

It is seen in Figs. 1(a) and (b) that the two-level model of [1] is precisely correct⁵ (as expected) for the extreme case of zero SNR inside the gap and infinite SNR outside the gap. When the SNR outside the gap is not infinite, the burst inclusion probability experiences a vertical drop-off precisely at the points predicted by the model, but this sharp drop-off slackens before reaching the predicted error rate outside the gaps. Thus, the effective gap length is approximately $RG - K + 1$ bits as predicted by the model, but it is widened slightly due to the "wings" that appear in the curves of Figs. 1(a) and (b). The model of [1] *slightly underestimates* the decoder's errors by squaring off these curves and eliminating these wings. This finding is consistent with the slight underestimate of decoder errors by the two-level model shown in Fig. A-1 of [1] for the VLA stand-alone case.

In contrast, the curves in Fig. 1(c) show that the two-level model of [1] *significantly underestimates* the effective width of the gaps, as predicted by the arguments at the end of the previous section. Not only is the region of maximum error rate longer than the effective gap length $RG - K + 1$ by about 4 to 5 bits on each side, but the drop-off from the maximum error rate is not as sharp and the "wings" are much wider.

Figures 2(a), (b), 3(a), (b), and 4(a), (b) show less definitive results for the cases of nonzero SNR inside the gap. There are clearly wings that extend the effective length of the gaps but at greatly reduced error levels, as in Figs. 1(a) and (b). However, the burst inclusion probability inside the gap is also rounded off (as compared to the model's prediction) and does not rise vertically to the predicted steady-state level. Figure 5 is a replot of one of the curves of Fig. 2(a) on a linear scale. On this figure it can be clearly seen that the model *slightly overestimates* the decoder's total errors. The area under the curve saved by not quickly rising to the steady-state level inside the gap is greater than the area lost by not quickly falling to the steady-state level outside the gap. This finding offers an explanation for the curves in Fig. A-2 of [1], where it was seen that the two-level model predicted slightly worse performance than the simulation for the case of VLA-Goldstone equal array (3-dB gaps).

Figures 2(c), 3(c), and 4(c) show the same types of tradeoffs. The error-cycle curves do not fall as sharply as predicted to their steady-state levels outside the gap, nor do they rise as sharply as predicted to their steady-state levels inside the gap. However, the error-cycle peaks are clearly several bits wider than the corresponding ones in Figs. 2(a), (b), 3(a), (b), or 4(a), (b).

V. Conclusions

In this study the two-level model for Viterbi decoder performance with gapped data was subjected to more-detailed validation tests than were performed in [1]. Detailed curves showing the cyclical behavior of the decoder error burst statistics were compared with the simple square-wave cycles predicted by the model.

The simulated decoder error-cycle curves roughly verify the accuracy of the two-level model for both the case of optimum branch-metric scaling and the case of constant branch-metric scaling combined with constant noise-level scaling of the received signal. Minor discrepancies between the model and the simulation include a broadening of the effective gap by several bits on both sides, but at a greatly reduced error level, whenever the SNR outside the gap is

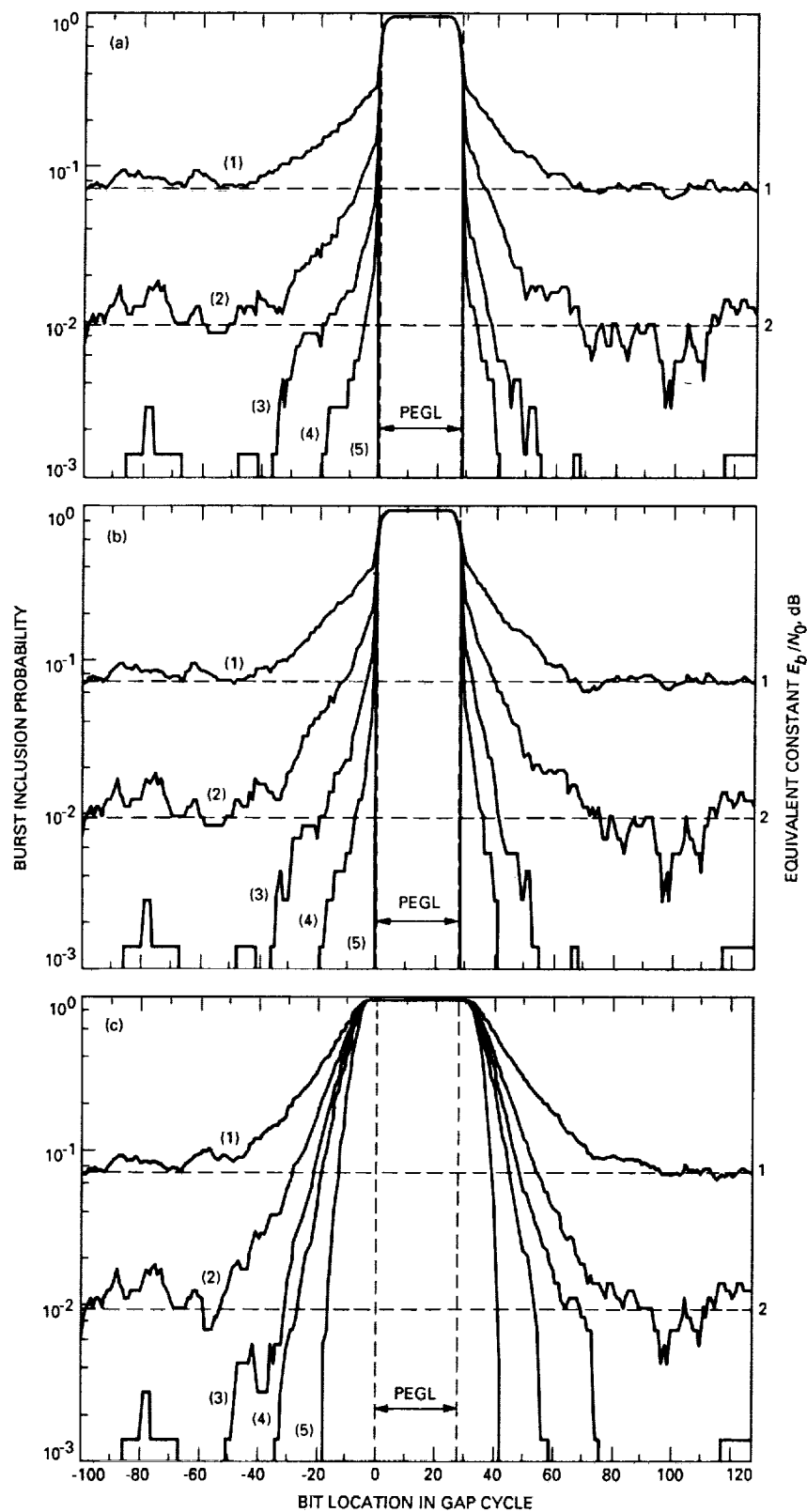
⁵ The slight rounding of the error-cycle curves near the edges of the effective gap window occurs because, in our definition, bursts must start and end with actual error bits. Some random bits at the beginning or end of the effective gap which happen to be decoded correctly are therefore not classified as being included within a burst, even though they are random.

not infinite, and a lowering of the gap error rate, especially near the edges, whenever the SNR inside the gap is not zero. These compensating effects lead to a slight net underestimate of the overall error rate by the model in the case of a VLA stand-alone system, and a slight net overestimate of the error rate in the case of the VLA arrayed equally with Goldstone. Both of these effects are small and do not critically diminish the validity of the model.

The model is not accurate for the case of constant branch-metric scaling combined with constant signal-level scaling. The effective gap in this case is significantly wider than that predicted by the model, and the error rate falls much less rapidly than predicted outside the gap. Therefore, it is important to avoid this combination of received-signal scaling and branch-metric scaling both in hardware systems and in software simulations.

References

- [1] S. J. Dolinar, "VLA Telemetry Performance With Concatenated Coding for Voyager at Neptune," *TDA Progress Report 42-95*, vol. July-September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 112-133, November 15, 1988.
- [2] L. J. Deutsch, "The Performance of VLA as a Telemetry Receiver for Voyager Planetary Encounters," *TDA Progress Report 42-71*, vol. July-September 1982, Jet Propulsion Laboratory, Pasadena, California, pp. 27-39, November 15, 1982.
- [3] L. J. Deutsch, "An Update on the Use of the VLA for Telemetry Reception," *TDA Progress Report 42-72*, vol. October-December 1982, Jet Propulsion Laboratory, Pasadena, California, pp. 51-60, February 15, 1983.
- [4] R. L. Miller, L. J. Deutsch, and S. A. Butman, *On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes*, JPL Publication 81-9, Jet Propulsion Laboratory, Pasadena, California, September 1, 1981.
- [5] D. Divsalar, D. Hansen, and J. H. Yuen, "The Effect of Noisy Carrier Reference on Telemetry with Baseband Arraying," *TDA Progress Report 42-63*, vol. March and April 1981, Jet Propulsion Laboratory, Pasadena, California, pp. 128-135, June 15, 1981.

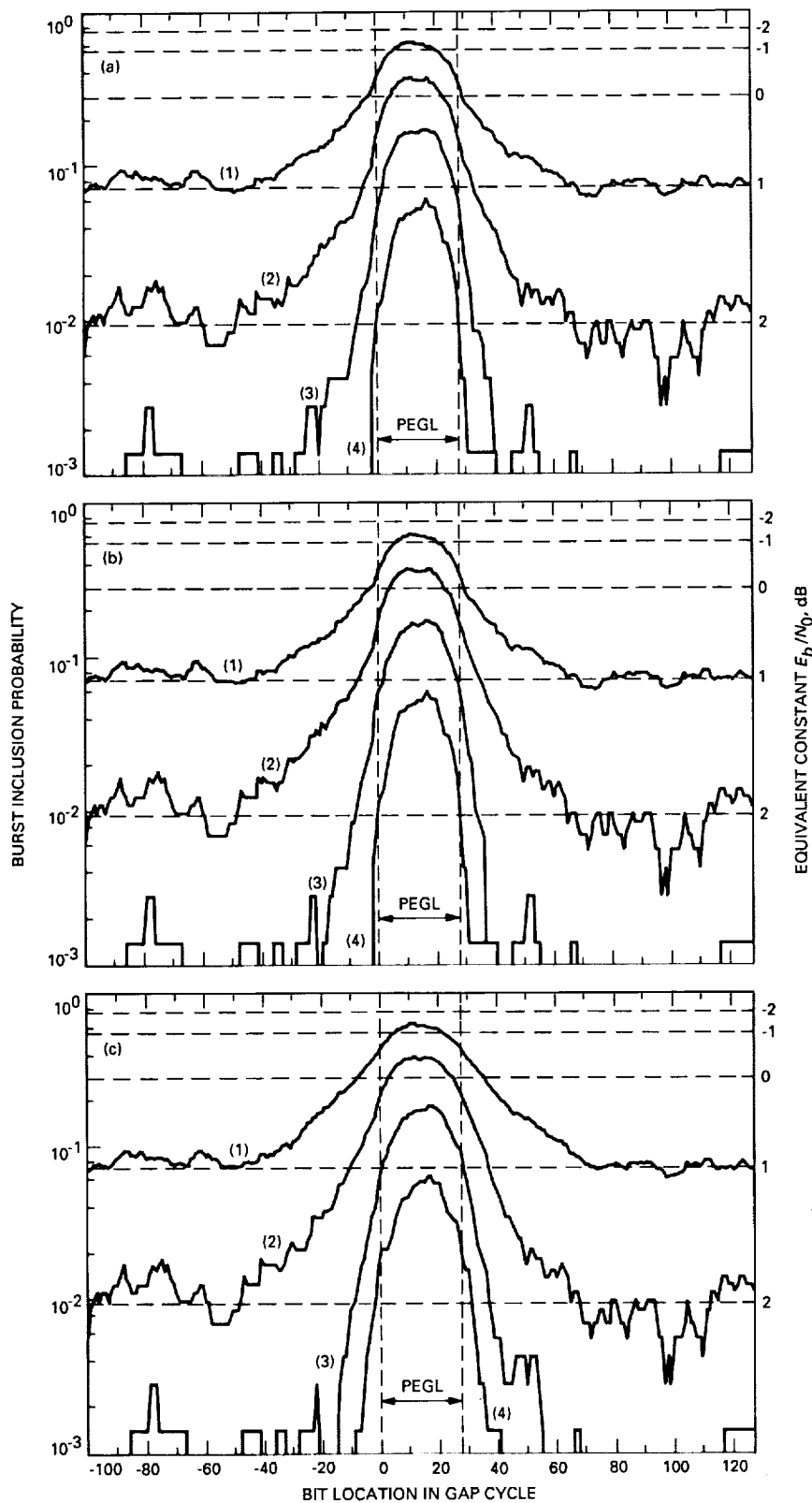


CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	$-\infty$
(2)	2	$-\infty$
(3)	3	$-\infty$
(4)	4	$-\infty$
(5)	∞	$-\infty$

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	$-\infty$
(2)	2	$-\infty$
(3)	3	$-\infty$
(4)	4	$-\infty$
(5)	∞	$-\infty$

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	$-\infty$
(2)	2	$-\infty$
(3)	3	$-\infty$
(4)	4	$-\infty$
(5)	∞	$-\infty$

Fig. 1. Decoder error-cycle curves for VLA stand-alone system using (a) optimum branch-metric scaling, (b) constant noise-level scaling, and (c) constant signal-level scaling. (PEGL = predicted effective gap length.)



CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-2
(2)	2	-1
(3)	3	0
(4)	4	1

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-2
(2)	2	-1
(3)	3	0
(4)	4	1

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-2
(2)	2	-1
(3)	3	0
(4)	4	1

Fig. 2. Decoder error-cycle curves for VLA arrayed equally with Goldstone (3-dB gaps) using (a) optimum branch-metric scaling, (b) constant noise-level scaling, and (c) constant signal-level scaling. (PEGL = predicted effective gap length.)

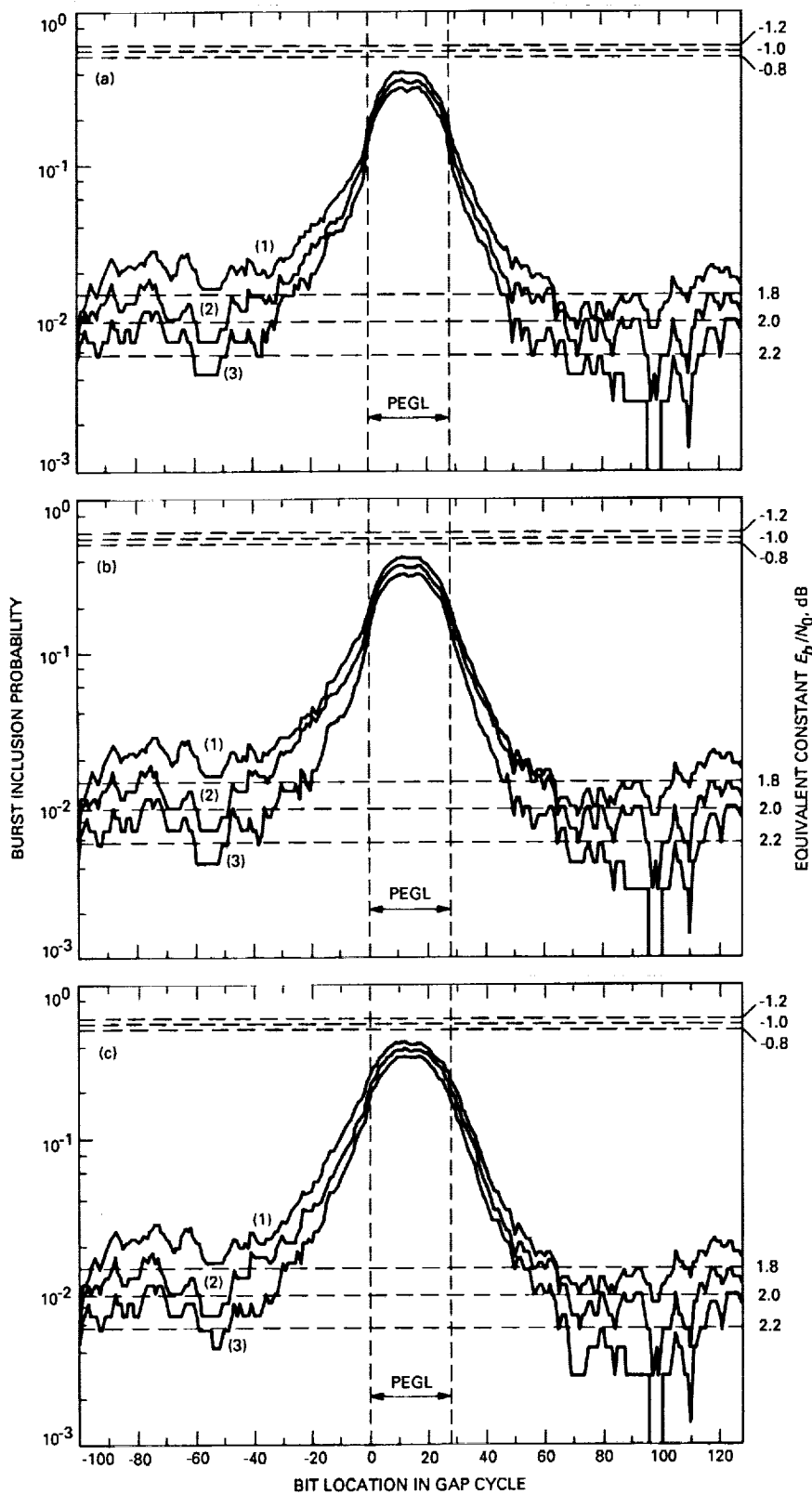
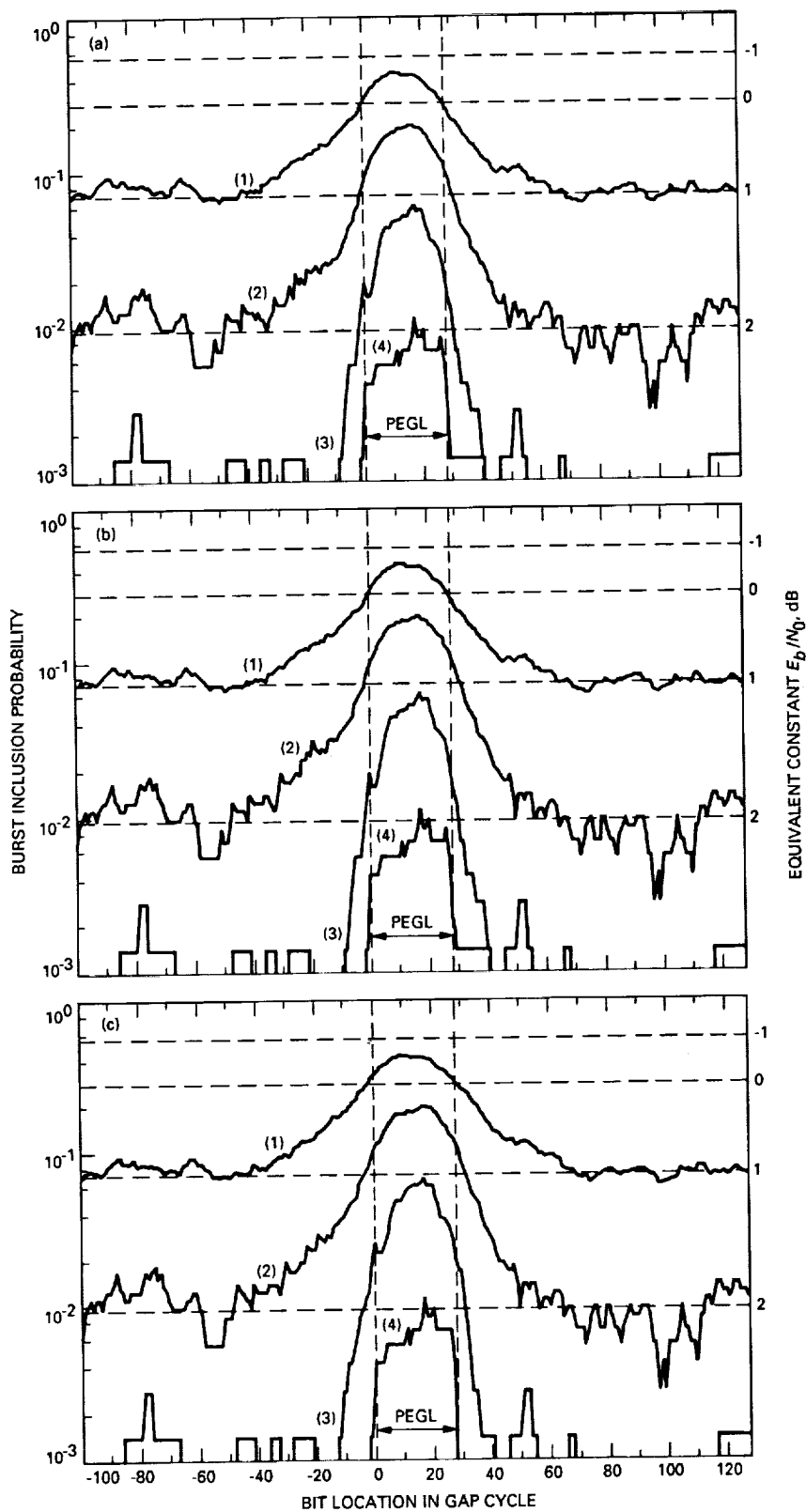


Fig. 3. Decoder error-cycle curves (outside gaps 1.8–2.2 dB) for VLA arrayed equally with Goldstone (3-dB gaps) using (a) optimum branch-metric scaling, (b) constant noise-level scaling, and (c) constant signal-level scaling. (PEGL = predicted effective gap length.)

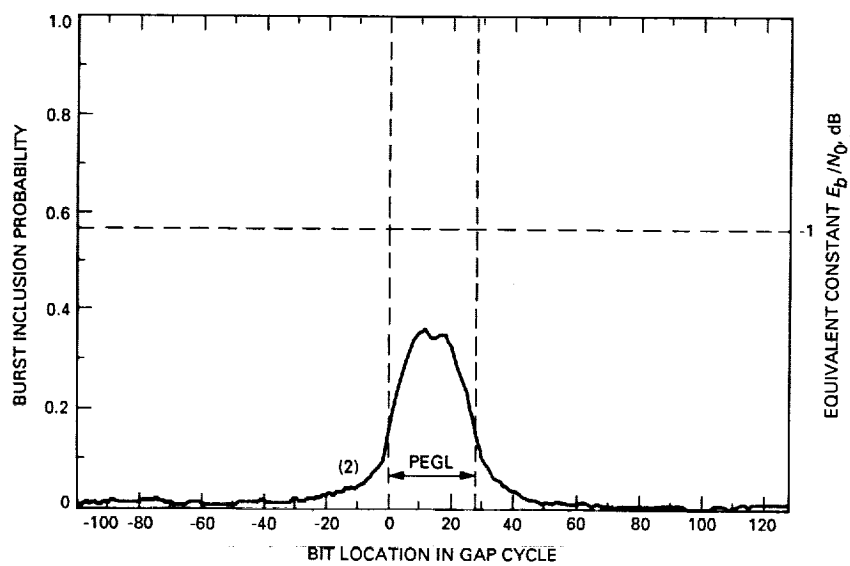


CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-1
(2)	2	0
(3)	3	1
(4)	4	2

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-1
(2)	2	0
(3)	3	1
(4)	4	2

CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(1)	1	-1
(2)	2	0
(3)	3	1
(4)	4	2

Fig. 4. Decoder error-cycle curves for VLA arrayed unequally with Goldstone (2-dB gaps) using (a) optimum branch-metric scaling, (b) constant noise-level scaling, and (c) constant signal-level scaling. (PEGL = predicted effective gap length.)



CURVE NO.	E_b/N_0 , dB	
	OUTSIDE GAP	INSIDE GAP
(2)	2	-1

Fig. 5. Decoder error-cycle curves for VLA arrayed equally with Goldstone (3-dB gaps) using optimum branch-metric scaling. (PEGL = predicted effective gap length.)

5/7-32
232333
118
N90-21897

Finding the Complete Path and Weight Enumerators of Convolutional Codes

I. Onyszchuk¹

Communications Systems Research Section

A method for obtaining the complete path enumerator $T(D, L, I)$ of a convolutional code is described. A system of algebraic equations is solved, using a new algorithm for computing determinants, to obtain $T(D, L, I)$ for the (7,1/2) NASA standard code. Generating functions, derived from $T(D, L, I)$, are used to upper bound Viterbi decoder error rates. This technique is currently feasible for constraint length $K < 10$ codes.

A practical, fast algorithm is presented for computing the leading nonzero coefficients of the generating functions used to bound the performance of constraint length $K < 20$ codes. Code profiles with about 50 nonzero coefficients are obtained with this algorithm for the experimental $K = 15$, rate 1/4, code in the Galileo mission and for the proposed $K = 15$, rate 1/6, 2-dB code.

I. Introduction

Convolutional codes such as the (7,1/2) NASA standard have been used for satellite and deep-space communications during the past 20 years. In 1971, Viterbi [1] defined generating functions for upper bounding error probabilities of convolutional codes on memoryless channels. In practice, for codes with more than eight states, these functions are still unknown, so error bounds have been evaluated using numerical matrix multiplications [2,8], which require extensive computations for each channel noise level.

As an alternative, algorithms have also been developed to calculate the first few coefficients (the distance and bit error profiles) of these enigmatic generating functions [4,5]. Unfortunately, lists of these numbers are bulky (see Tables 2 and 3), and the minimum number of terms required to approximate the decoder error bounds depends upon the code rate and channel noise level.

A code's complete path enumerator $T(D, L, I)$ contains the number of paths having identical triples: weight, length, number of input 1s. The least-magnitude pole of the weight enumerator $T(D) = T(D, 1, 1)$ determines the point at which the union bounds [1] diverge, while additional poles and residues yield the dominant terms in the

¹ Also a student in the Electrical Engineering Department, California Institute of Technology, Pasadena, California.

partial fraction expansion of $T(D)$. Concise, yet very accurate, analytic approximations to several generating functions' coefficients may be obtained from these dominant terms. Furthermore, the poles and residues of $T(D)$ may eventually help unlock the structure of convolutional codes.

The generating functions $T(D)$ and $T(D, L, I)$ for the (7,1/2) NASA standard code (among many others) were obtained with a simple determinant algorithm described in this article. In addition, a new algorithm is presented for computing the initial coefficients of $T(D)$, $\partial T(D, L, I)/\partial L$ and $\partial T(D, L, I)/\partial I$ at $L=I=1$. These numbers are used to upper bound a Viterbi decoder's event, node, bit, and symbol error rates. The techniques explained here may be adapted, with modifications for code nonlinearity and Euclidean instead of Hamming distances, to find weight enumerators of trellis codes [3,5].

II. The Complete Path Enumerator

A binary, rate k/n , feed-forward convolutional encoder is defined by kn binary generator polynomials $g_{ij}(x)$, each representing the transfer function from the i th input to the j th output (x is a delay operator). Let \underline{g}_{ij} be the vector whose r th component is the coefficient of x^r in $g_{ij}(x)$. The encoder's memory is $m = \sum_{i=1}^k \max_j [\deg g_{ij}(x)]$, which for rate $1/n$ codes is $K-1$. If the i th encoder memory cell contains s_i then the encoder is in state $s = \sum_{i=1}^m s_i 2^{i-1}$. The encoder's state diagram is a directed graph whose edges (corresponding to branches in the associated trellis diagram) have labels $a_{i,j} = D^d L I^b$ if there is an edge from state j into state i while $a_{i,j} = 0$ otherwise. During a transition from state j into state i , b is the number of ones input to the encoder and the number of ones output by the encoder is $d = \sum_{h=1}^n \underline{g}_{1h} \cdot [2j + (i \bmod 2)]_2$, where \cdot is a modulo-2 inner product and $[k]_2$ denotes the binary representation of the integer k .

Let X_s be the trivariate generating function of all *simple* paths: those from state 0 into state s via nonzero states. X_0 counts all simple paths into state 0, called *fundamental* paths. Note that X_s is indexed by the destination state (s) while the source state (0) is constant. Now define \underline{A} as the $2^m - 1$ by $2^m - 1$ *adjacency matrix* of the encoder graph with state 0 and its edges removed. The entry in row i and column j of \underline{A} is $a_{i,j}$, which is nonzero only if there is a directed edge connecting state j to state i (they are adjacent). Now the following set of linear equations is constructed:

$$\underline{A} [X_1, X_2, \dots, X_{2^m-1}]^T = [a_{1,0}, a_{2,0}, \dots, a_{2^m-1,0}]^T$$

By Cramer's Rule,

$$X_i = \frac{\det(\underline{A}_i)}{\det(\underline{A})}$$

for $i > 0$, where \underline{A}_i is \underline{A} with all column i entries $a_{r,i}$ replaced by $a_{r,0}$ for all rows r . A code's *complete path enumerator* (by weight, length, and number of input 1s)

$$T(D, L, I) = X_0 = \sum_{j=1}^{2^k} a_{0,z_j} X_{z_j}$$

where $\{z_j\}$ are the 2^k states having edges into state 0 (so $a_{0,z_j} \neq 0$).

As an example with $k=1$ for simplicity, the $m=2$ encoder in Fig. 1 is in state $s = 2s_1 + s_0$. The corresponding state diagram in Fig. 2 leads to the equations $\underline{A} [X_1, X_2, X_3]^T = [a_{1,0}, 0, 0]^T$:

$$\begin{bmatrix} 1 & -LI & 0 \\ -DL & 1 & -DL \\ -DLI & 0 & 1-DLI \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} = \begin{bmatrix} D^2 LI \\ 0 \\ 0 \end{bmatrix}$$

Now $T(D, L, I) = X_0 = D^2 L X_2 = D^5 L^3 I / (1 - DLI - DL^2 I)$, so the code's weight enumerator is

$$T(D) = \frac{D^5}{1-2D} = \sum_{d=5}^{\infty} 2^{d-5} D^d = \sum_{d=d_{\text{free}}}^{\infty} p(d) D^d$$

where $p(d) = 2^{d-5}$ is the number of weight d fundamental paths, and $d_{\text{free}} = 5$ is the code's free distance.

III. Reducing an Adjacency Matrix

When $k=1$, there are only $3(2^m-1)-2$ nonzero out of $(2^m-1)^2$ entries in \underline{A} , and they are located in a special pattern. Since \underline{A} is sparse if $m > 3$, reduction is useful before computing the determinant. The following example, different from the one in the previous section, illustrates the reduction procedure. The code with generator polynomials $g_{11}(D) = 1 + D + D^3$ and $g_{12}(D) = 1 + D + D^2 + D^3$

has an adjacency matrix $\underline{\mathbf{A}}$ (with $L = I = 1$ to simplify entries)

$$\begin{bmatrix} 1 & 0 & 0 & (D^2)-1 & 0 & 0 & 0 \\ -D^2 & 1 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 0 & -D^2 & 0 & 0 \\ 0 & -D & 0 & (0)1 & 0 & -D & 0 \\ 0 & -D & 0 & 0 & 1 & -D & 0 \\ 0 & 0 & -D & 0 & 0 & 1 & -D \\ 0 & 0 & -D & 0 & 0 & 0 & 1-D \end{bmatrix}$$

The round brackets (parentheses) above indicate values in $\underline{\mathbf{A}}_{2^m-1}$ that are different from those in $\underline{\mathbf{A}}$. Since $k = 1$, the determinants of only these two matrices are needed for $T(D, L, I)$, and this notation will lead to their simultaneous computation.

As in Gaussian elimination, $-a_{r, \lfloor r/2 \rfloor}$ times row $\lfloor r/2 \rfloor$ is added to rows $r = 2$ to 2^m-1 so that $\underline{\mathbf{A}}$ becomes zeroed below its main diagonal for columns 1 to $2^{m-1}-1$:

$$\begin{bmatrix} 1 & 0 & 0 & (D^2) - 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & (D^4) - D^2 & -1 & 0 & 0 \\ 0 & 0 & 1 & (D^2) - 1 & -D^2 & 0 & 0 \\ 0 & 0 & 0 & (D^5)1 - D^3 & -D & -D & 0 \\ 0 & 0 & 0 & (D^5) - D^3 & 1-D & -D & 0 \\ 0 & 0 & 0 & (D^3) - D & -D^3 & 1 & -D \\ 0 & 0 & 0 & (D^3) - D & -D^3 & 0 & 1-D \end{bmatrix}$$

Therefore, $\det(\underline{\mathbf{A}})$ equals the determinant of the resulting lower right 2^{m-1} by 2^{m-1} submatrix. To further reduce $\underline{\mathbf{A}}$, row $\lfloor r/2 \rfloor$ times $-a_{r, 2^{m-1} + \lfloor r/2 \rfloor}$ is added to each row $r = 2^{m-3}$ to 2^{m-1} (4 to 5 here), so that columns $2^{m-1} + 2^{m-2}$ to 2^{m-2} (both 6 here) are zeroed above the main diagonal. Now $\det(\underline{\mathbf{A}}) = \det(\underline{\tilde{\mathbf{A}}})$, where $\underline{\tilde{\mathbf{A}}}_{2^m-1}$ and $\underline{\tilde{\mathbf{A}}}$ are the new lower right 2^{m-1} by 2^{m-1} submatrices:

$$\begin{bmatrix} (D^4 + D^5)1 - D^2 - D^3 & -D - D^4 & 0 & -D^2 \\ (D^4 + D^5) - D^2 - D^3 & 1 - D - D^4 & 0 & -D^2 \\ (D^3) - D & -D^3 & 1 & -D \\ (D^3) - D & -D^3 & 0 & 1 - D \end{bmatrix}$$

To zero column 2^{m-1} of $\underline{\tilde{\mathbf{A}}}$ (above the diagonal entry $\tilde{a}_{2^{m-1}, 2^{m-1}} = a_{2^{m-1}, 2^{m-1}} \neq 0$), $-\tilde{a}_{r, 2^{m-1}}/\tilde{a}_{2^{m-1}, 2^{m-1}}$ times row 2^{m-1} is added to each row $r = 1$ to $2^{m-1}-1$. Define

$(\underline{\mathbf{B}}_{2^m-1})$ and $\underline{\mathbf{B}}$ as the resulting upper left 2^{m-2} by 2^{m-2} submatrices,

$$\frac{\begin{bmatrix} (D^4 + D^5 - D^6)1 - D - D^2 - D^3 + D^4 & -D + D^2 - D^4 \\ (D^4 + D^5 - D^6) & -D^2 - D^3 + D^4 & 1 - 2D + D^2 - D^4 \end{bmatrix}}{1 - D}$$

This reduction method simultaneously produces two dense 2^{m-2} by 2^{m-2} matrices $(\underline{\mathbf{B}}_{2^m-1})$, $\underline{\mathbf{B}}$ with the same determinants as the corresponding original sparse 2^m-1 by 2^m-1 matrices $(\underline{\mathbf{A}}_{2^m-1})$, $\underline{\mathbf{A}}$.

IV. A Determinant Algorithm

The following algorithm yields the determinant of any $N \times N$ matrix $\underline{\mathbf{B}}$ having entries from a Euclidean domain, such as the set of all polynomials with integer coefficients. A sequence of matrices $\underline{\mathbf{B}}^{(N)}$, $\underline{\mathbf{B}}^{(N-1)}$, $\underline{\mathbf{B}}^{(N-2)}$, ..., $\underline{\mathbf{B}}^{(1)}$ is computed with each matrix having the same determinant (up to sign) as $\underline{\mathbf{B}}$. Starting with $\underline{\mathbf{B}}^{(N)} = \underline{\mathbf{B}}$, step j in the algorithm produces the numerator $b_{i,k}^{(j-1)}$ of each entry in row i and column k of $\underline{\mathbf{B}}^{(j-1)}$:

for $j = N$ to 2 (step)

for $i = j-1$ to 1 (row index)

for $k = j$ to 1 (column index)

$$b_{i,k}^{(j-1)} = \frac{b_{i,k}^{(j)} b_{j,j}^{(j)} - b_{i,j}^{(j)} b_{j,k}^{(j)}}{b_{j+1,j+1}^{(j+1)}}$$

Naturally, if $b_{j+1,j+1}^{(j+1)} = 0$ prior to step j , then any column k such that $b_{j+1,k}^{(j+1)} \neq 0$ must first be interchanged with column $j+1$. This operation negates the determinant, so a counter t is incremented to record the event. If no such column k exists, row $j+1$ is zero, so the algorithm is stopped and $\det(\underline{\mathbf{B}}) = \det(\underline{\mathbf{B}}^{(j)}) = 0$. Also note that $b_{N+1,N+1}^{(N+1)} = 1$ initially.

The following example with $N = 4$ illustrates the above reduction procedure. When $j = 4$, after $-b_{34}^{(4)}/b_{44}^{(4)}$, $-b_{24}^{(4)}/b_{44}^{(4)}$, and $-b_{14}^{(4)}/b_{44}^{(4)}$ times row 4 are added to rows 3, 2, and 1, respectively (corresponding to $j = N = 4$ in the

algorithm), in order to zero all entries in column 4 above $b_{44}^{(4)}$,

$$\underline{\mathbf{B}}^{(3)} = \begin{bmatrix} \frac{b_{11}^{(3)}}{b_{44}^{(4)}} & \frac{b_{12}^{(3)}}{b_{44}^{(4)}} & \frac{b_{13}^{(3)}}{b_{44}^{(4)}} & 0 \\ \frac{b_{21}^{(3)}}{b_{44}^{(4)}} & \frac{b_{22}^{(3)}}{b_{44}^{(4)}} & \frac{b_{23}^{(3)}}{b_{44}^{(4)}} & 0 \\ \frac{b_{31}^{(3)}}{b_{44}^{(4)}} & \frac{b_{32}^{(3)}}{b_{44}^{(4)}} & \frac{b_{33}^{(3)}}{b_{44}^{(4)}} & 0 \\ b_{41}^{(4)} & b_{42}^{(4)} & b_{43}^{(4)} & b_{44}^{(4)} \end{bmatrix}$$

Then, after the $j = 3$ step during which $-b_{23}^{(3)}/b_{33}^{(3)}$ and $-b_{13}^{(3)}/b_{33}^{(3)}$ times row 3 are added to rows 2 and 1,

$$\underline{\mathbf{B}}^{(2)} = \begin{bmatrix} \frac{b_{11}^{(2)}}{b_{33}^{(3)}} & \frac{b_{12}^{(2)}}{b_{33}^{(3)}} & 0 & 0 \\ \frac{b_{21}^{(2)}}{b_{33}^{(3)}} & \frac{b_{22}^{(2)}}{b_{33}^{(3)}} & 0 & 0 \\ \frac{b_{31}^{(3)}}{b_{44}^{(4)}} & \frac{b_{32}^{(3)}}{b_{44}^{(4)}} & \frac{b_{33}^{(3)}}{b_{44}^{(4)}} & 0 \\ b_{41}^{(4)} & b_{42}^{(4)} & b_{43}^{(4)} & b_{44}^{(4)} \end{bmatrix}$$

Since only elementary row operations on $\underline{\mathbf{B}}$ have been performed, $\det(\underline{\mathbf{B}}) = \det(\underline{\mathbf{B}}^{(3)}) = \det(\underline{\mathbf{B}}^{(2)})$. The final step ($j = 2$) produces a $\underline{\mathbf{B}}^{(1)}$ matrix identical to $\underline{\mathbf{B}}^{(2)}$ except with $b_{11}^{(1)}/b_{22}^{(2)}$ 0 0 0 in the first row. Therefore, $\det(\underline{\mathbf{B}}) = \det(\underline{\mathbf{B}}^{(1)}) = b_{11}^{(1)}$. This result generalizes to any nonsingular $N \times N$ matrix $\underline{\mathbf{B}}$.

Lemma. Entry $b_{1,1}^{(1)} = \det(\underline{\mathbf{B}}^{(1)}) = (-1)^t \det(\underline{\mathbf{B}})$, where t is the total number of column interchanges performed by the algorithm in computing $b_{1,1}^{(1)}$.

Proof. For each value of j from N to 2, when $k = j$, the algorithm makes $b_{i,j}^{(j-1)} = 0$ for all $1 \leq i \leq j-1$ so that all column j entries above the main diagonal become 0. Since $\underline{\mathbf{B}}^{(1)}$ is zero above its main diagonal, $\det(\underline{\mathbf{B}})$ is the product of $(-1)^t$ and the diagonal entries in $\underline{\mathbf{B}}^{(1)}$.

$$\det(\underline{\mathbf{B}}) = (-1)^t \det(\underline{\mathbf{B}}^{(1)})$$

$$= (-1)^t \prod_{j=1}^N \frac{b_{j,j}^{(j)}}{b_{j+1,j+1}^{(j+1)}}$$

$$= (-1)^t b_{1,1}^{(1)} \quad \bullet$$

Expanding $b_{i,k}^{(j)} b_{j,j}^{(j)} - b_{i,j}^{(j)} b_{j,k}^{(j)}$ using the equation in the algorithm shows that $b_{j+1,j+1}^{(j+1)}$ divides this expression so there are never any remainders. This is expected because the algorithm implements a recursive factorization of the determinant written as a sum of products of matrix entries. The algorithm differs from standard Gaussian elimination because the diagonal entries in the reduced matrices are not made equal to 1. Also, after calculations with a particular value of j are completed, all entries' denominators are previous pivot numerators.

V. Path Distance, Length, and Bit Error Approximations

The complete path enumerator $T(D, L, I)$ for the $m = 6$, rate $1/2$, NASA standard code was obtained by using the preceding algorithm to simultaneously compute determinants of the two 16 by 16 reduced matrices $\underline{\mathbf{B}}_{2m-1}$ and $\underline{\mathbf{B}}$. The 76 poles of $T(D)$ for this code are plotted on the complex plane (Fig. 3) along with the unit circle for reference. Using only the six least-magnitude poles (indicated by the large points in Fig. 3), an approximation to the partial fraction expansion of $T(D)$ is

$$\begin{aligned} T(D) &\approx D^{10} \left[\frac{r_1}{1 - \alpha_d D} + \frac{r_1}{1 + \alpha_d D} + \frac{r_2}{1 - \alpha_b D} \right. \\ &\quad \left. + \frac{r_2}{1 + \alpha_b D} + \frac{r_2^*}{1 - \alpha_b^* D} + \frac{r_2^*}{1 + \alpha_b^* D} \right] \\ &= D^{10} \sum_{k=0}^{\infty} \left[r_1 (\alpha_d)^k + r_1 (-\alpha_d)^k + r_2 (\alpha_b)^k \right. \\ &\quad \left. + r_2 (-\alpha_b)^k + r_2^* (\alpha_b^*)^k + r_2^* (-\alpha_b^*)^k \right] D^k \\ &= \sum_{j=0}^{\infty} \left[2r_1 (\alpha_d)^{2k} + 4\text{Re}\{r_2 (\alpha_b)^{2k}\} \right] D^{2k+10} \end{aligned}$$

where $\alpha_d = 2.3876225$ is the reciprocal of the least-magnitude pole of $T(D)$, $\alpha_b = 1.657193e^{-0.983418j}$ is the reciprocal of the pole with next smallest magnitude, * means complex conjugate, and $j = \sqrt{-1}$. The residues are

$$r_1 = \frac{-P(\alpha_d^{-1})}{\alpha_d^{-1} Q'(\alpha_d^{-1})} \quad \text{and} \quad r_2 = \frac{-P(\alpha_b^{-1})}{\alpha_b^{-1} Q'(\alpha_b^{-1})}$$

where $D^{10}P(D)/Q(D) = T(D)$ and $Q'(D)$ is the derivative of $Q(D)$. The other poles in Fig. 3 may be ignored because their magnitudes are greater than 0.8, and the corresponding residues have magnitudes less than 0.07.

Define $p(d)$, $i(d)$, and $\ell(d)$ as the coefficients of D^d in $T(D)$, $\partial T(D, L, I)/\partial L$, and $\partial T(D, L, I)/\partial I$, respectively, at $L = I = 1$ (these generating functions are shown in the Appendix). The number of fundamental paths having weight $2k + 10$ is

$$p(2k+10) \approx 6.82(2.3876225)^{2k} \\ + 4.25(1.657193)^{2k} \cos(0.310 - 1.967k)$$

Similarly, the terms corresponding to the six least-magnitude poles of $T(D)$ in the partial fraction expansions of $\partial T(D, L, I)/\partial L$ and $\partial T(D, L, I)/\partial I$ at $L = I = 1$ were used to obtain the approximations

$$\ell(2k+10) \approx (77.725 + 22.625k)(2.3876225)^{2k} \\ + 39.3(1.657193)^{2k} \cos(0.485 - 1.967k) \\ + (2k+1)7.2676(1.657193)^{2k} \cos(0.383 - 1.967k) \\ i(2k+10) \approx (24.474 + 12.018k)(2.3876225)^{2k} \\ + 9.942(1.657193)^{2k} \cos(0.575 - 1.967k) \\ + (2k+1)2.8723(1.657193)^{2k} \cos(0.366 - 1.967k)$$

which have a relative error < 0.0001 for $k > 4$.

A rate $1/n$ Viterbi decoder's bit error rate (BER) on a binary-input, output-symmetric [2], discrete memoryless channel is bounded by

$$\text{BER} \leq \sum_{d=d_{\text{free}}}^{\infty} i(d)P_d$$

where P_d is the probability that the decoder outputs a fundamental path having distance d from the one transmitted. The probability that a b -bit symbol is decoded incorrectly is bounded by [9]

$$\text{SER}_b \leq \sum_{d=d_{\text{free}}}^{\infty} [(b-1-m)p(d) + \ell(d)] P_d$$

For the additive white Gaussian noise (AWGN) channel with bit signal-to-noise ratio E_b/N_0 ,

$$P_d = Q(\sqrt{2dRE_b/N_0})$$

where $Q(x)$ is the Gaussian integral function [2, p. 62]. On a binary symmetric channel with crossover probability p ,

$$P_{2i} = P_{2i-1} = \sum_{e=i}^{2i-1} \binom{2i-1}{e} p^e (1-p)^{2i-1-e} \leq \binom{2i-1}{i} p^i$$

[6]. For decoders using integer metrics, as for example on a binary-input, output-quantized AWGN channel, P_d can also be computed exactly [2, p. 291].

VI. Algorithms for Profiles of Convolutional Codes

Finding the complete path enumerator of codes with memory greater than 8 currently seems infeasible. In these cases, an algorithm for distance profiles [4,5] may be used to calculate the first few nonzero coefficients of the generating functions used for error bounds. However, these methods are fairly complex and some require extra computation to ensure that the output is correct. In this section, Viterbi's algorithm, with survivors replaced by vectors of integers that count paths, lengths, or bits, is applied on a noiseless channel to compute $p(d)$, $\ell(d)$, and $b(d)$ values. Rate $1/n$ codes are treated first to simplify the discussion. Define $\text{out}_0[s]$ and $\text{out}_1[s]$ as the number of ones that the encoder outputs going from state $s_0 = \lfloor s/2 \rfloor$ and $s_1 = s_0 + 2^{m-1}$ into state s . Analogous to a state metric, the entry in row $s > 0$ and column 0 of a matrix W , referred to as $W[s][0]$, will contain the *least weight* of any simple path with length $\leq T$ trellis branches. For $t = 1$ to coeffs (a parameter described later), $W[s][t]$ will be the *number* of simple paths of weight $W[s][0] + t - 1$ and length $\leq T$ branches into state $s > 0$. For state 0, $W[0][t]$ is always kept at 0, except $W[0][1] = 1$. The entries in a second matrix B count either the total number of ones input to the encoder (when the variable $\text{len} = 0$) or the total length in trellis branches (when $\text{len} = 1$), of all simple paths having length $\leq T$ (again $B[0][t] = 0$ always). These matrices are obtained for successive values of T starting with 1 by extending, one branch length at a time (an algorithm 'step'), the code trellis starting from state 0 only. Thus the longest pathlength (T) explored by the algorithm equals the number of 'steps' executed. The algorithm terminates after step T^* when W has reached values that will never change, which also forces B to remain constant.

Then since $W[s][0]$ is the least weight of any simple path into state $s > 0$, $d_{\text{free}} = W[2^{m-1}][0] + \text{out}_1[0]$. Also,

$$W[2^{m-1}][d - d_{\text{free}} + 1] = p(d)$$

$$B[2^{m-1}][d - d_{\text{free}} + 1] = b(d) \quad (\ell(d) \text{ if } \text{len} = 1)$$

for $d = d_{\text{free}}$ to $d_{\text{free}} + \text{coeffs} - 1$.

Two versions of the basic algorithm above are presented in C language format in Algorithm 1 and Table 1. In Algorithm 1, matrices P and A store *previous* W and B entries corresponding to simple paths of length $\leq T$, which are used to compute new W and B matrices for length $\leq T+1$ simple paths. When change remains 0 after step T^* , W (and thus B) will never change because $W[s][t] = P[s][t]$ for all s and t . $P[s][0]$ is initialized to 999 for $s > 0$, $P[0][1] = 1$, and all other array values are initialized to 0. If any second array index $t + \text{offset}$ is ≤ 0 in the $W[s][t]$ and $B[s][t]$ instructions, the array referenced is simply ignored.

The algorithm requires storage for $2^{m+1}(\text{coeffs} + 2)$ integers and the amount of work per step is proportional to this number. The number of steps executed, T^* , equals the length, in trellis branches, of the longest fundamental path(s) having weight $d_{\text{free}} + \text{coeffs} - 1$. The parameter coeffs should be set equal to $\lceil 10n/k \rceil$ because using this many nonzero terms in the union bounds gives results with three significant digits of precision when the bounds are tight enough to be useful.

Setting $\text{coeffs} = 0$ and ignoring offset0 , offset1 , A , and B yields a simple and fast algorithm for finding d_{free} . About 2^{m+1} bytes of storage and 2.1 CPU seconds (on a computer executing 12 million instructions per second) were required to obtain d_{free} for the $m = 14$ Galileo and "2-dB" codes.

For rate k/n codes with $k > 1$, a state s is partitioned into s_1, s_2, \dots, s_k where s_i corresponds to the contents of the i th encoder shift register. The output weight as the encoder enters state s is $\text{out}_j[s]$ and the input is

$j \in [0 \dots 2^k - 1]$. New $W[s][t]$ and $B[s][t]$ values are computed using at most 2^k entries from each of the P and A matrices.

For codes with $2^k \ll 2^m$ (such as rate $1/n$ with $m > 3$), Algorithm 1 may be improved by looping through groups of 2^k states called butterflies (see Fig. 4) instead of individual states and by computing $W[s][t]$ and $B[s][t]$ in place [7]. This reduces the storage memory required by almost one-half because the "double-buffering" matrices P and A are eliminated. If the entire vector $W[2^{m-1}]$ remains the same for m consecutive steps, it will never change because there is a trellis path of length $\leq m$ branches between any two states. The algorithm shown in Table 1, which incorporates these improvements, produced the first $10n$ nonzero coeffs in the profiles of two $K = 15$ codes: the rate $1/4$ Galileo code (Table 2) and the proposed "2-dB", rate $1/6$ code (Table 3). These profiles took only a few minutes of CPU to generate and required storage of 7 and 8.5 Mbytes, respectively, when 4 bytes were used for each integer. These memory requirements could be further reduced by storing each integer in the smallest number of bytes needed (1 for $W[s][0]$ to $W[s][15]$ in the $K = 15$ codes above).

Algorithm 1. A simple profile algorithm

```

do {
  for (s = 1 to  $2^m - 1$ ) {
    s0 =  $\lfloor s/2 \rfloor$ ; s1 =  $s0 + 2^{m-1}$ ;
    bit =  $\text{len} + (1 - \text{len}) * (s \bmod 2)$ ;
     $W[s][0] = \min(P[s0][0] + \text{out}_0[s], P[s1][0] + \text{out}_0[s]);$ 
     $\text{offset0} = W[s][0] - P[s0][0] - \text{out}_0[s];$ 
     $\text{offset1} = W[s][0] - P[s1][0] - \text{out}_0[s];$ 
    for (t = 1 to coeffs) {
       $W[s][t] = P[s0][t + \text{offset0}] + P[s1][t + \text{offset1}];$ 
       $B[s][t] = A[s0][t + \text{offset0}] +$ 
         $A[s1][t + \text{offset1}] + \text{bit} * W[s][t];$ 
    }
    change = 0;
    for (s = 1 to  $2^m - 1$ )
      for (t = 0 to coeffs)
        if ( $P[s][t] \neq W[s][t]$ ) { change = 1;
           $P[s][t] = W[s][t]; B[s][t] = A[s][t];$  }
    } while (change  $\neq$  0);

```

References

- [1] A. J. Viterbi, "Convolutional Codes and their Performance in Communication Systems," *IEEE Trans. Comm.*, COM-19, pp. 751-772, October 1971.
- [2] A.J. Viterbi and J.K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.
- [3] E. Zehavi and J.K. Wolf, "On the Performance Evaluation of Trellis Codes," *IEEE Trans. Info. Theory*, vol. IT-33, pp. 196-202, March 1987.
- [4] L. Chevillat and R. Johannesson, "A Fast Algorithm for Finding the Distance Spectrum of Convolutional Codes," *IEEE Trans. Info. Theory*, to appear.
- [5] M. Rouanne and D.J. Costello, Jr., "An Algorithm for Computing the Distance Spectrum of Trellis Codes," *IEEE Journal on Selected Areas of Comm.*, vol. SAC-21, August 1989.
- [6] L. Van de Meeburg, "A Tightened Upper Bound on the Error Probability of Binary Convolutional Codes with Viterbi Decoding," *IEEE Trans. Info. Theory*, vol. IT-20, pp. 389-391, May 1974.
- [7] C. M. Rader, "Memory Management in a Viterbi Decoder," *IEEE Trans. Comm.*, COM-29, pp. 1399-1401, September 1981.
- [8] P. J. Lee, "A Very Efficient Transfer Function Bounding Technique on Bit Error Rate for Viterbi Decoded, Rate $1/N$ Convolutional Codes," *TDA Progress Report 42-79*, vol. July-September, pp. 114-123, November 15, 1984.
- [9] R. J. McEliece and I. M. Onyszchuk, "A Symbol Error Upper Bound for Convolutional Codes," *Proc. 27th Allerton Conf.*, Monticello, Illinois, September 26-29, 1989.

Table 1. A profile algorithm for rate $1/n$ codes

```

k=0; stop = 0;
do {
  for (t = 0 to coeffs)
    { temp[t] = W[2m-1-k][t]; B[0][t] = W[0][t] = 0; }
  W[0][1] = 1;
  for (s = 0 to 2m-1 - 1) {
    s0 = s >> k;    (cyclically)
    s1 = s0 + 2m-1-k
    tw0[0] = min (W[s0][0] + out0[2s], W[s1][0] + out1[2s]);
    tw1[0] = min (W[s0][0] + out1[2s+1], W[s1][0] + out0[2s+1]);
    offset00 = tw0[0] - W[s0][0] - out0[2s];
    offset01 = tw0[0] - W[s1][0] - out1[2s];
    offset10 = tw1[0] - W[s0][0] - out1[2s+1];
    offset11 = tw1[0] - W[s1][0] - out0[2s+1];
    for (t = 1 to coeffs) {
      tw0[t] = W[s0][t + offset00] + W[s1][t + offset10];
      tw1[t] = W[s0][t + offset10] + W[s1][t + offset11];
      tb0[t] = B[s0][t + offset00] + B[s1][t + offset10] + len * tw0[t];
      tb1[t] = B[s0][t + offset10] + B[s1][t + offset11] + tw1[t];
    }
    for (t = 0 to coeffs) {
      W[s0][t] = tw0[t]; W[s1][t] = tw1[t];
      B[s0][t] = tb0[t]; B[s1][t] = tb1[t]; }
    k = k + 1; if (k = m) k = 0; change = 1;
    if (W[2m-1-k][coeffs] > 0) change = 0;
    for (t = 0 to coeffs)
      if (W[2m-1-k][t] ≠ temp[t]) change = 1;
    if (change = 1) stop = 0;
    if (change = 0 and stop < m - 1) {change = 1; stop ++; }
  } while (change ≠ 0);

```

Table 2. Galileo code profiles

Distance d	Fundamental paths $p(d)$	Bit errors $i(d)$	Total lengths $\ell(d)$
35	2	6	7
36	1	2	5
37	4	16	22
38	2	8	12
39	3	11	17
40	5	20	28
41	6	24	46
42	17	76	122
43	24	126	214
44	29	180	285
45	39	255	438
46	66	416	721
47	94	628	1071
48	121	850	1478
49	175	1313	2260
50	277	2086	3643
51	415	3361	5855
52	639	5304	9388
53	934	8010	14161
54	1273	11452	20271
55	1906	17550	31381
56	2878	27332	49172
57	4054	39750	71705
58	5978	60788	109808
59	8864	92738	167861
60	12966	139556	253134
61	18984	210112	383008
62	27949	317798	581467
63	41092	479512	878975
64	60126	720858	1323152
65	87799	1080933	1987235
66	128712	1622990	2992979
67	189880	2451782	4530508
68	278589	3682496	6817868
69	408780	5534126	10261968
70	598271	8283100	15386816
71	875283	12380669	23050515
72	1286052	18596544	34662286
73	1888299	27885609	52045238
74	2768375	41727376	78013493
75	4057688	62421220	116865844
76	5953416	93419654	175122289
77	8732134	139709066	262220198
78	12809968	208928290	392628663
79	18786484	312181796	587384902
80	27548175	466271448	878292728
81	40412499	696477455	1313354906
82	59269748	1039725314	1962719710

Table 3. Rate 1/6 "2-dB" code profiles

Distance d	Fundamental paths $p(d)$	Bit errors $i(d)$	Total lengths $\ell(d)$
56,57	1,5	2,15	3,19
58,60	1,3	2,12	2,14
61,62	5,12	25,56	35,84
63,64	11,5	43,24	67,40
65,66	8,11	44,62	68,95
67,68	8,11	48,62	76,98
69	27	167	267
70	30	162	277
71	36	216	363
72	54	366	573
73	74	464	785
74	89	610	998
75	94	670	1104
76	126	912	1524
77	163	1209	2022
78	226	1676	2814
79	290	2236	3785
80	369	2920	4993
81	493	4051	6846
82	574	4780	8168
83	767	6571	11236
84	979	8562	14687
85	1182	10474	18250
86	1574	14282	24860
87	1996	18516	32193
88	2618	24594	43183
89	3407	32955	57577
90	4238	41914	73499
91	5353	53757	94399
92	7006	71430	126401
93	8932	92712	164631
94	11418	120946	214330
95	14401	155175	275986
96	18467	202902	361135
97	24039	268439	479664
98	30325	344146	616671
99	38662	446878	800288
100	49690	583672	1048171
101	63930	762130	1371587
102	81742	990268	1785532
103	103839	1278325	2308219
104	133335	1666564	3012971
105	170357	2159215	3912282
106	217467	2801764	5081111
107	278512	3640320	6613934
108	356223	4721974	8592622
109	456347	6135943	11180051
110	583546	7956498	14517787
111	746528	10327464	18861798
112	954389	13376948	24468075
113	1220261	17333391	31748900
114	1562164	22493842	41242505
115	1997088	29126250	53477032

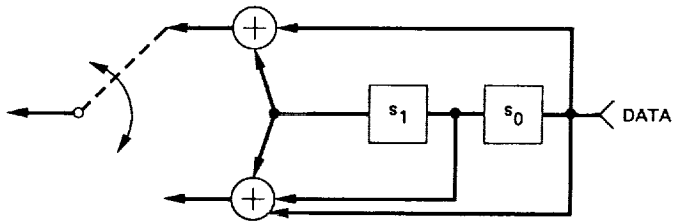


Fig. 1. A rate 1/2, 4-state encoder.

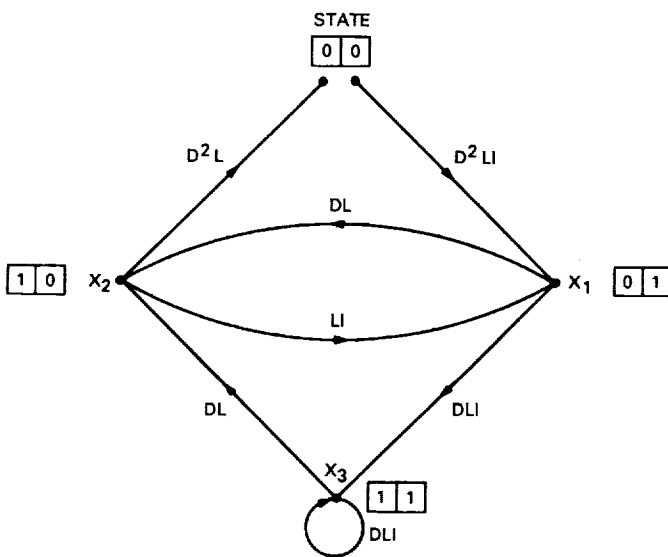


Fig. 2. State diagram of the encoder in Fig. 1.

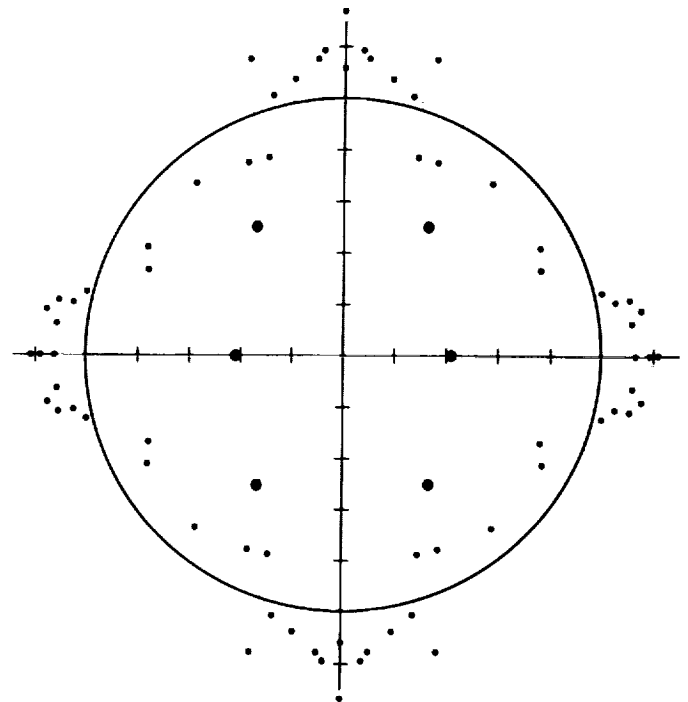


Fig. 3. Poles of $T(D)$ for the (7,1/2) NASA code.

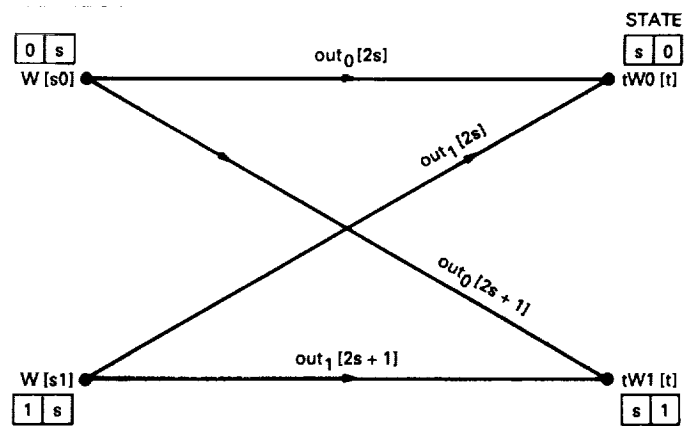


Fig. 4. A rate 1/n code butterfly (number s).

Appendix

Generating Functions for the (7,1/2) NASA Code

Using the determinant algorithm in section IV, the complete path enumerator $T(D, L, I)$ for the $m = 6$, rate $1/2$ NASA code was found to contain 1529 numerator and 2799 denominator trivariate terms. The code's weight enumerator, $T(D) = T(D, 1, 1)$, is

$$11D^{10} - 6D^{12} - 25D^{14} + D^{16} + 93D^{18} - 15D^{20} - 176D^{22} - 76D^{24} + 243D^{26} + 417D^{28} - 228D^{30} - 1156D^{32} - 49D^{34} + 2795D^{36} + 611D^{38} - 5841D^{40} - 1094D^{42} + 9575D^{44} + 1097D^{46} - 11900D^{48} - 678D^{50} + 11218D^{52} + 235D^{54} - 8068D^{56} - 18D^{58} + 4429D^{60} - 20D^{62} - 1838D^{64} + 8D^{66} + 562D^{68} - D^{70} - 120D^{72} + 16D^{76} - D^{80}$$

$$1 - 4D^2 - 6D^4 - 30D^6 + 40D^8 + 85D^{10} - 81D^{12} - 345D^{14} + 262D^{16} + 844D^{18} - 403D^{20} - 1601D^{22} + 267D^{24} + 2509D^{26} + 389D^{28} - 3064D^{30} - 2751D^{32} + 2807D^{34} + 8344D^{36} - 1960D^{38} - 16133D^{40} + 1184D^{42} + 21746D^{44} - 782D^{46} - 21403D^{48} + 561D^{50} + 15763D^{52} - 331D^{54} - 8766D^{56} + 131D^{58} + 3662D^{60} - 30D^{62} - 1123D^{64} + 3D^{66} + 240D^{68} - 32D^{72} + 2D^{76}$$

$$= 11D^{10} + 38D^{12} + 193D^{14} + 1331D^{16} + 7275D^{18} + 40406D^{20} + \dots$$

The other two generating functions used to compute error bounds, $\partial T(D, L, I)/\partial L$ and $\partial T(D, L, I)/\partial I$ at $L = I = 1$, both have denominators equal to the square of $T(D)$'s denominator above. Their numerators are, respectively,

$$121D^{10} - 387D^{12} - 706D^{14} + 1460D^{16} + 3970D^{18} - 6157D^{20} - 11643D^{22} + 8725D^{24} + 28677D^{26} + 12195D^{28} + 88D^{30} - 170654D^{32} - 306124D^{34} + 817895D^{36} + 1637616D^{38} - 2879440D^{40} - 6106837D^{42} + 8568521D^{44} + 18636083D^{46} - 22431469D^{48} - 48921504D^{50} + 52678351D^{52} + 113105887D^{54} - 112260733D^{56} - 232580537D^{58} + 217337170D^{60} + 426400859D^{62} - 379787502D^{64} - 696667758D^{66} + 592954735D^{68} + 1013294336D^{70} - 815739185D^{72} - 1311124721D^{74} + 968225450D^{76} + 1509511967D^{78} - 955561827D^{80} - 1548537967D^{82} + 721812022D^{84} + 1419185285D^{86} - 302418615D^{88} - 1166004400D^{90} - 173248817D^{92} + 861869027D^{94} + 545751792D^{96} - 574515412D^{98} - 713158178D^{100} + 345328990D^{102} + 676371119D^{104} - 186475599D^{106} - 515274530D^{108} + 89761092D^{110} + 326707300D^{112} - 38067910D^{114} - 174942675D^{116} + 14010022D^{118} + 79516060D^{120} - 4391424D^{122} - 30654965D^{124} + 1145504D^{126} + 9969013D^{128} - 241287D^{130} - 2706667D^{132} + 39344D^{134} + 603670D^{136} - 4651D^{138} - 107908D^{140} + 354D^{142} + 14883D^{144} - 13D^{146} - 1488D^{148} + 96D^{152} - 3D^{156}$$

and

$$36D^{10} - 77D^{12} - 140D^{14} + 813D^{16} + 269D^{18} - 4414D^{20} + 321D^{22} + 14884D^{24} - 5273D^{26} - 40509D^{28} + 39344D^{30} + 83884D^{32} - 177469D^{34} - 111029D^{36} + 608702D^{38} - 29527D^{40} - 1820723D^{42} + 817086D^{44} + 4951082D^{46} - 3436675D^{48} - 12279246D^{50} + 10300306D^{52} + 27735007D^{54} - 25648025D^{56} - 56773811D^{58} + 55659125D^{60} + 104376199D^{62} - 106695512D^{64} - 170819460D^{66} + 180836818D^{68} + 247565043D^{70} - 270555690D^{72} - 317381295D^{74} + 356994415D^{76} + 360595622D^{78} - 415401723D^{80} - 364292177D^{82} + 426295756D^{84} + 328382391D^{86} - 385686727D^{88} - 264812337D^{90} + 307287819D^{92} + 191225378D^{94} - 215144035D^{96} - 123515898D^{98} + 131946573D^{100} + 71124860D^{102} - 70570661D^{104} - 36310569D^{106} + 32722089D^{108} + 16308558D^{110} - 13052172D^{112} - 6380604D^{114} + 4433332D^{116} + 2147565D^{118} - 1265046D^{120} - 612040D^{122} + 297721D^{124} + 144665D^{126} - 56305D^{128} - 27569D^{130} + 8232D^{132} + 4066D^{134} - 874D^{136} - 435D^{138} + 60D^{140} + 30D^{142} - 2D^{144} - D^{146}$$

An Adaptive Vector Quantization Scheme

K.-M. Cheung

Communications Systems Research Section

Vector quantization is known to be an effective compression scheme to achieve a low bit rate so as to minimize communication channel bandwidth and also to reduce digital memory storage while maintaining the necessary fidelity of the data. However, the large amount of computations required in vector quantizers has been a handicap in using vector quantization for low-rate source coding. In this article an adaptive vector quantization algorithm is introduced that is inherently suitable for simple hardware implementation because it has a simple architecture. It allows fast encoding and decoding because it requires only addition and subtraction operations.

I. Introduction

Vector quantization can be used as a source coding technique for speech and images by mapping a sequence of continuous or discrete vectors into a digital sequence suitable for communication over a digital channel or storage in a digital medium. The goal is to reduce the volume of the data for transmission over a digital channel and also for archiving to a digital storage medium, while preserving the required fidelity of the data. This is particularly important to many present and future NASA missions, as the volume of speech and image data in the foreseeable future would become prohibitively large for many communication links or storage devices. It was shown in [1] that a well-designed vector quantization scheme can provide a high compression ratio with good reconstructed quality in the mean-square error sense.

Unlike scalar quantization, where the actual coding of continuous or discrete samples into discrete quantities is done on single samples, the input data of a vector quantization encoder are multi-dimensional blocks of data (input vectors). Most traditional vector quantization schemes [1,2] work as follows: A codebook, which consists of codeword vectors of the same dimension as the input data vectors, is first generated by training a subset of the source data by the LBG (Linde-Buzo-Gray) algorithm [1]. The codebook is then transmitted through the channel. An input vector is encoded by first comparing it with codewords in the codebook. The codeword closest to the input vector is chosen as the quantization vector to represent the input vector. The index of this chosen codeword is then transmitted through the channel. Compression is achieved since fewer bits are used to represent the codeword index than the quantized input data. Notice that the generation of

the codebook and the choice of the best codeword for each input vector involve computing the distortion between the input vector and each codeword in the codebook; these processes are usually computationally intensive, especially when the codebook size is large. In this article, we introduce an adaptive vector quantization scheme that does not require codebook generation and is relatively computationally simple.

Bentley et al. [3] and Elias [4] describe a simple heuristic move-to-front protocol to do lossless high rate data compression on text data. This protocol maintains a sequential list of vectors so that the more frequently accessed vectors are near the front. In this article, we modify this protocol and apply it to the design of an adaptive vector quantization scheme to do lossy low rate data compression on speech or image data. This vector quantization scheme exploits the locality of reference by assuming the same vectors are used more frequently over short intervals and then fall into a long period of disuse. It uses a self-organizing list, which will be explained in Section II, as an auxiliary data structure such that the vectors that are more likely to be transmitted are near the front of the list. This dynamic data structure has two advantages: the data are presorted to allow the encoder to efficiently and quickly match the input vectors with the codewords in the list. It also provides inherent adaptability to local statistics by keeping the more recently used codewords in the front of the list while shifting the long unused codewords out of the list. This scheme also has implementation advantages. It has a simple architecture for hardware implementation, and it only requires addition and subtraction operations to allow fast encoding and decoding. This scheme is particularly suitable for real-time compression of audio and video data (both intraframe coding and interframe coding) that will be a large part of communications in future NASA missions. This scheme also has potential uses in commercial areas like high-definition television and audio/video communications using telephone lines. The results reported here are, however, preliminary.

Section II outlines the basic algorithm and some variations of it. Section III describes the performance analysis method and the experimental results, and Section IV gives the concluding remarks.

II. Basic Algorithm

Unlike the traditional vector quantization schemes, this scheme does not require the computationally intensive process of generating a fixed codebook in the first place. The encoder and decoder keep an identical codeword table of the same size. The table is constructed on the fly

by exploiting the local statistics of the data as the vector quantization encoder is reading in the input vectors. The codewords in the table, and the positions of codewords in the table, are constantly updated according to a self-organizing protocol such that the more recently used codewords are near the front of the table. Codewords that have not been used for a long time would eventually reach the end of the list and be discarded. The following is a more quantitative discussion of this algorithm.

Suppose the original p -dimensional digital data file size is $L_1 \times L_2 \times \dots \times L_p$ pixels and each pixel's intensity is represented by k bits. Let each vector represent a block of $l = s_1 \times s_2 \times \dots \times s_p$ pixels, $s_i | L_i$, for $i = 1, \dots, p$. Let the codeword table have size N , $N = 2^m - 1$, and let the all ones' m -tuple (i.e., m ones'), denoted by a , be reserved to indicate the transmission of an uncoded vector while the rest of the m -tuples are used to represent the indices of codewords in the table. Let $\mathbf{r}(t)$ be the input vector and $\mathbf{c}_i(t)$ be the i th element in the codeword table at time t . Both $\mathbf{r}(t)$ and $\mathbf{c}_i(t)$ have the same dimensionality $s_1 s_2 \dots s_p$. Let $D(\mathbf{r}(t), \mathbf{c}_i(t))$ be the distortion function associated with the vector quantization coder. The encoder computes $D(\mathbf{r}(t), \mathbf{c}_i(t))$ for $i = 0, \dots, N - 2$. Let s be the index of the codeword corresponding to the minimum distortion with $\mathbf{r}(t)$. That is,

$$s = \text{index achieving } \min_{i=0, \dots, N-2} D(\mathbf{r}(t), \mathbf{c}_i(t))$$

The minimum distortion $D(\mathbf{r}(t), \mathbf{c}_s(t))$ is then compared to a preset threshold T . If $D(\mathbf{r}(t), \mathbf{c}_s(t)) \leq T$, the index s is sent. The codeword table is then updated as follows:

$$\text{tmp} \leftarrow \mathbf{c}_s(t) \quad (1)$$

$$\mathbf{c}_i(t+1) \leftarrow \mathbf{c}_{i-1}(t) \quad i = s, \dots, 1 \quad (2)$$

$$\mathbf{c}_0(t+1) \leftarrow \text{tmp} \quad (3)$$

If $D(\mathbf{r}(t), \mathbf{c}_s(t)) > T$, the reserved all ones' m -tuple a is sent, which is followed by the transmission of the uncoded input vector $\mathbf{r}(t)$. The codeword table is then updated as follows:

$$\mathbf{c}_i(t+1) \leftarrow \mathbf{c}_{i-1}(t) \quad i = N-2, \dots, 1 \quad (4)$$

$$\mathbf{c}_0(t+1) \leftarrow \mathbf{r}(t) \quad (5)$$

The decoder on the other end maintains the codeword table, which is identical to the encoder's. When a legitimate index s of the codeword table is received, the decoder outputs the codeword corresponding to s and updates the ta-

ble as in Eqs. (1), (2), and (3). When the reserved m -tuple a and the uncoded input vector $r(t)$ are received, the decoder outputs $r(t)$ and updates the tables using Eqs. (4) and (5). The encoding and the decoding algorithms are summarized in Fig. 1 and Fig. 2, respectively.

The speed of the vector quantization encoder can be greatly enhanced by performing a partial search for the codeword of minimum distortion rather than a full search. One efficient and simple way is to compute $D(r(t), c_i(t))$ sequentially along the self-organized list of codewords and to pick the first codeword $c_s(t)$ in the list with $D(r(t), c_s(t)) \leq T$. Another partial search algorithm is to search the first n codewords in the list, where $n < N - 1$, and pick the minimum if it is less than T . In most cases these partial search schemes can give the codeword of minimum distortion. This is because the codewords in the table, and the positions of codewords in the table, are constantly updated according to a self-organizing protocol such that the more recently used codewords, thus the most likely codewords with minimum distortion, are near the front of the table. The performance degradation from using partial search algorithms instead of full search is in most cases unnoticeable. The speed of the encoder can also be increased by choosing a simpler distortion measure like the sum of absolute difference between pixels in $r(t)$ and $c_i(t)$.

This scheme has the following advantages. It is inherently suitable for simple hardware implementation. The encoder and decoder maintain their codeword tables, which can be implemented using a set of registers, by performing only simple shifting operations. The computation of distortion between input vector and codeword vector at the encoder side requires only simple addition and subtraction operations. It allows fast encoding and decoding, and unlike many traditional vector quantization schemes that take two or more passes over the data, this scheme requires only one pass.

This scheme has one disadvantage for communications systems requiring rigid data formatting. Unlike most conventional vector quantization encoding schemes that keep the compression rate of the input data files constant and let the distortion drift, this scheme bounds the distortion and lets the compression rate vary according to the degree of activity in the data. Thus the sizes of the compressed data files cannot be a priori specified, and this would create difficulties in specifying the data format of the overall communication system. One way to circumvent this problem is to iteratively re-encode an input data file using different distortion thresholds until the compressed file size matches a certain preset data format. Another solution is

to restrict the number of uncoded vectors per data file to be sent such that the compressed file size remains constant for each data file. The first solution increases the overall computation time and the second solution degrades the overall distortion performance.

III. Performance Analysis and Experimental Results

The compression ratio of a vector quantization scheme on a data file is defined to be the ratio of the original data file size to the compressed data file size. A distortion measure D is an assignment of a cost $D(r(t), c_i(t))$ of reproducing any input vector $r(t)$ by a code vector $c_i(t)$. Given such a distortion measure, the performance of a vector quantization scheme can be measured in terms of the compression ratio versus the average distortion $E[D(r(t), c_i(t))]$ between the input and the final reproduction.

A good distortion measure should be computationally tractable to allow easy analysis and subjectively meaningful so that large or small measured distortions closely correspond to bad and good subjective quality. In this article, we choose to use the conventional squared error distortion measure, which is defined as follows:

$$D(r(t), c_i(t)) = \|r(t) - c_i(t)\|^2 = \sum_{j=0}^{l-1} (r^j(t) - c_i^j(t))^2 \quad (6)$$

It is also common practice to measure the performance of a system by the signal-to-noise ratio (or signal-to-quantization-noise ratio) in dB. That is,

SNR =

$$10 \log_{10} \frac{2^{2k}}{\frac{1}{L_1 L_2} \sum_{i=0}^{L_1-1} \sum_{j=0}^{L_2-1} (x_{orig}(i, j) - x_{decoded}(i, j))^2} \quad (7)$$

where $x(i, j)$ is the intensity value of the (i, j) -th pixel.

Both planetary and nonplanetary images (8 bits per pixel) with various degrees of activity were used in this data compression experiment, and the compression ratio, the root-mean-square-error per pixel, and SNR per-

formance are tabulated in Table 1. The original picture and a decoded picture (compression ratio 8.325 and SNR 34.68 dB) of "Mercury" are given in Fig. 3 and Fig. 4, respectively. The original picture and a decoded picture (compression ratio 4.163 and SNR 31.57 dB) of "USC411" are given in Fig. 5 and Fig. 6, respectively.

An inherent disadvantage of low-rate vector quantization image coding is the annoying presence of blockiness in the decoded image. Because of the mismatch between the pixel values of two spatially adjacent codewords, discontinuity is developed at the block boundaries and a staircase effect shows at the edges. This degradation is usually the most noticeable and most annoying quantization error. Simple image enhancement techniques such as prefiltering and postfiltering can be applied to the decoded image to

eliminate the obvious blockiness in the decoded version of the picture.

IV. Concluding Remarks

In this article, a locally adaptive vector quantization scheme, which has a simple architecture to facilitate hardware implementation, was introduced. This scheme adapts efficiently to local statistics by using a simple heuristic method for self-organizing sequential search introduced in [3] and [4]. This scheme is fast, simple and robust, and has potential use for real-time compression of audio and video data in future NASA missions and in commercial areas such as high-definition television, facsimile, and audio/video communications using telephone lines. Performance in noise is a big, open issue.

References

- [1] R. M. Gray, "Vector Quantization," *IEEE ASSP Magazine*, April 1984, vol. 1, pp. 4-29.
- [2] N. Jayant and P. Noll, *Digital Coding of Waveforms-Principles and Applications to Speech and Video*, Prentice-Hall, 1984.
- [3] J. Bentley, D. Sleator, R. Tarjan, and V. Wei, "A Locally Adaptive Data Compression Scheme," *Communications of the ACM*, April 1986, vol. 29, pp. 320-330.
- [4] P. Elias, "Interval and Recency Rank Source Coding: Two On-Line Adaptive Variable-Length Schemes," *IEEE Transactions on Information Theory*, January 1987, vol. 33, pp. 3-10.

Table 1. Compression ratio versus distortion performances

File name	File size	C.R.	RMSE/pixel	SNR, dB
Mercury	256 × 256	8.32	4.72	34.68
		5.49	3.66	36.90
USC411	256 × 256	4.16	6.75	31.57
		2.58	4.60	34.85
Saturn6	800 × 800	18.75	7.40	30.75
		12.54	3.70	36.80
		9.29	3.00	38.63
Lena	512 × 512	15.94	20.88	21.77
		6.91	14.12	25.16
		3.91	10.08	28.10
		2.58	7.42	30.75

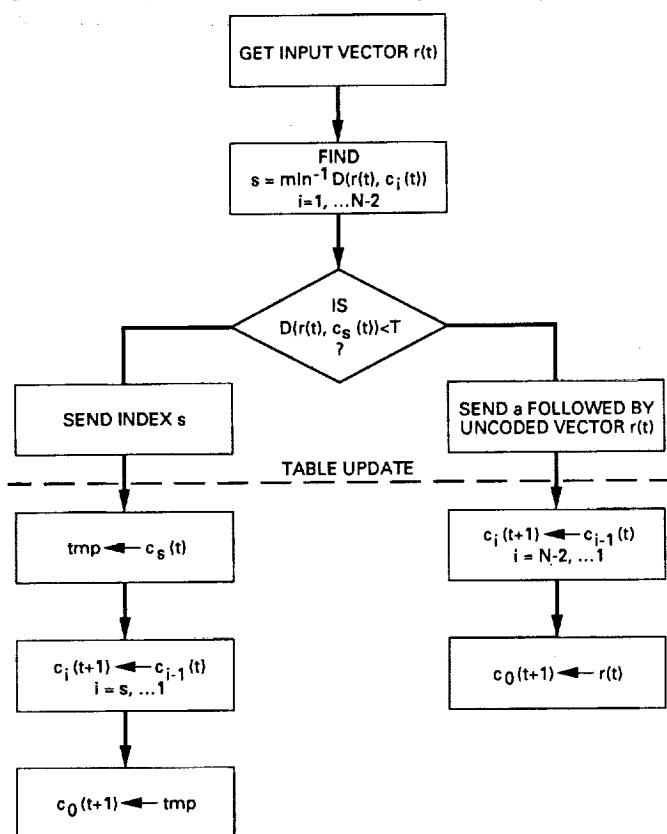


Fig. 1. Encoding algorithm.

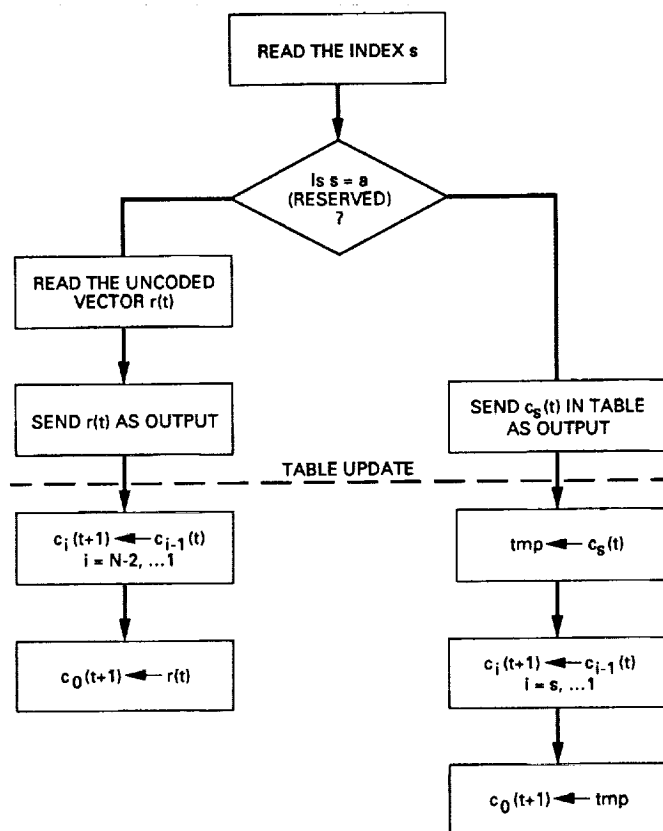


Fig. 2. Decoding algorithm.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

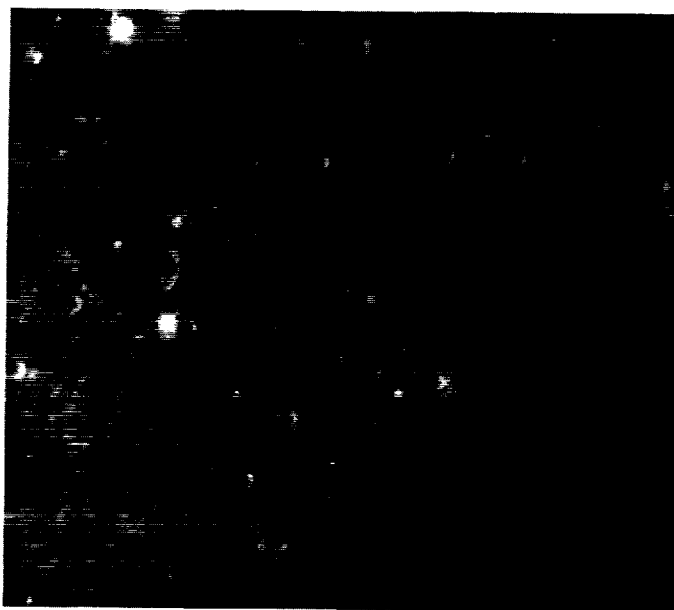


Fig. 3. Original Image of "Mercury".



Fig. 5. Original Image of "USC411".

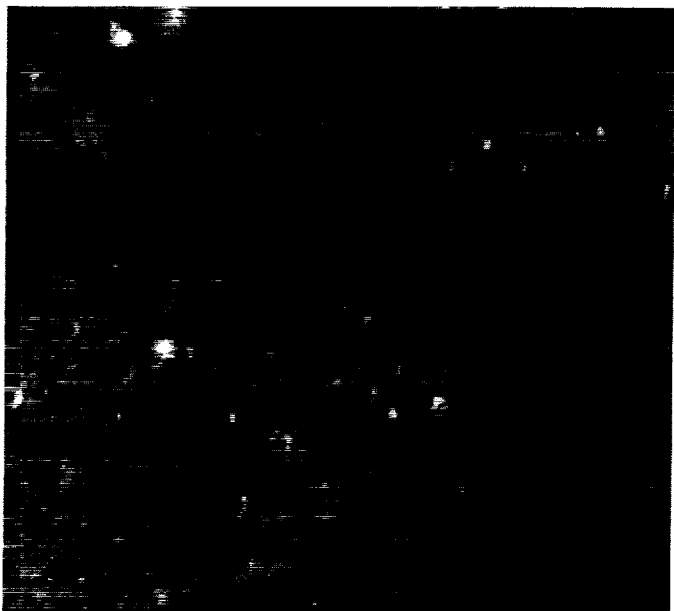


Fig. 4. Reconstructed Image of "Mercury".



Fig. 6. Reconstructed image of "USC411".

59-32
272335
138
N90-21899

A VLSI Design for a Systolic Viterbi Decoder

T. K. Truong and E. Satorius
Communications Systems Research Section

M. T. Shih and I. S. Reed
Department of Electrical Engineering, University of Southern California

A systolic Viterbi decoder for convolutional codes is developed. This decoder uses the trace-back method to reduce the amount of data needed to be stored in registers. It is shown that this new algorithm requires a smaller chip size and achieves a faster decoding time than other existing methods.

I. Introduction

Convolutional coding with Viterbi decoding [1] is a powerful method for forward error correction. As a consequence there is a growing need for implementing the Viterbi decoder in VLSI in deep-space communication [2].

There are two classes of algorithms established for realizing the Viterbi decoder. These are the register-exchange and trace-back methods [3]. Unfortunately, each class has drawbacks. Both require a substantial amount of storage space and hardware for even a moderate speed while the latter, for long constraint lengths, also needs a long decoding time.

In this article, the Viterbi decoding algorithm is first presented for comparison with other methods. For this example, a modified trace-back algorithm is created and shown to require a minimal number of storage devices and a short decoding time. Next, a systolic architecture is developed for this modified trace-back decoding algorithm. It is demonstrated for this new trace-back architecture that the tracking, updating, and storing of the hypothesized in-

formation sequences for the Viterbi decoder can be accomplished simultaneously during a single clock cycle. Finally, a suitable VLSI implementation is suggested for this new systolic architecture.

II. Example of a Viterbi Decoder for a (3,1/2) Convolutional Code

Let K and R be the constraint length and the rate, respectively, of what is denoted a (K, R) convolutional code (CC). In this section, the nature of the Viterbi decoding algorithm is demonstrated by using a (3,1/2) convolutional code.

Consider the example of a (3,1/2) CC with the generator polynomial $G(x) = (x^2 + x + 1, x^2 + 1)$. The encoder for this code is shown in Fig. 1 where one observes that the encoder is composed of a two-stage shift register $A = (A_1, A_0)$ with three modulo-2 adders and a multiplexer for converting a parallel to a serial output, where A_i for $(i = 1, 2)$ denotes a one-bit register. Let the information sequence be $\underline{u} = (u_1, u_2, u_3, u_4, \dots) = (0, 0, 1, 0, \dots)$. Af-

ter encoding the information sequence $\underline{u} = (0, 0, 1, 0, \dots)$, the output code word is given by $\underline{v} = (v_1, v_2, v_3, \dots) = (00, 00, 11, 10, \dots)$.

The trellis diagram for a particular $(3, 1/2)$ CC is shown in Fig. 2. In each column of the trellis there are $2^K = 2^2 = 4$ states of the shift register. These are the four possible states of the shift register of the encoder. Each information bit causes the shift register to change state. This is represented by a branch from the present node to the next node. Each branch in the j th column is labeled with the single output code-word frame v_j . The upper branch leaving a node at time unit $j - 1$ represents the "zero" input bit $u_j = 0$, while the lower branch represents the "one" input bit $u_j = 1$. The code word that corresponds to the information sequence $\underline{u} = (0, 0, 1, 0, 1, 0, 0, \dots)$ is shown in Fig. 2 as a heavy-line path.

Assume that a code word $\underline{v} = (v_1, v_2, v_3, \dots)$ is transmitted over a binary symmetric channel (BSC) and that the received sequence is $\underline{r} = (r_1, r_2, r_3, \dots)$. The branch metric from state S_l at time unit $j - 1$ to S_k at time unit j is defined by

$$d_{j-1,j}(l, k) = \|r_j - v_j\| \quad (1)$$

where $d_{j,j-1}(l, k)$ denotes Hamming path distance from state S_l at time unit $j - 1$ to S_k at time unit j . Here also $\|r_j - v_j\|$ is the Hamming weight of the difference of the two binary vectors r_j and v_j . The partial path metric for a state S_k up until the end of the first j branches of a path is denoted by $M_j(k)$. If there is a state S_l at time $j - 1$ which can change to state S_k at time j along the given path, then the partial path metric is expressible by the difference equation

$$M_j(k) = M_{j-1}(l) + d_{j,j-1}(l, k) \quad (2)$$

where $M_0(k) = 0$ for $k = 0$. The smallest metric for all paths terminating at state S_k at time j is given by

$$P_j(k) = \min M_j(k) \quad (3)$$

where the minimum is over all possible partial-path metrics that end at state S_k at time j . $P_j(k)$ is the weight of what is called the survivor path to state S_k at time unit j .

The Viterbi decoder for a convolutional code finds the survivor path, the path of minimum metric, which reaches a given state at time j . This survivor path is dependent on the input information bits so that the decoded bits are read off easily along this path.

The Viterbi algorithm is outlined as follows:

- (1) Beginning at time unit $j = 1$, compute the partial metric for the single path entering each state. Store the path (survivor) and its metric for each state.
- (2) Increase j by 1. Compute the partial metric for all the paths entering a state by adding the branch metric entering that state to the metric of the connecting survivor at the preceding time unit. For each state, store the path with the smallest metric (the survivor), together with the metric, and eliminate all other paths.
- (3) If $j < L_c$, repeat step (2). Otherwise, stop.

In the above, L_c is the length of the code word. The final survivor with the smallest metric can be used to decode the information bit along this path.

As an example of the above Viterbi algorithm, assume that a code word of this $(3, 1/2)$ CC is $\underline{v} = (v_1, v_2, v_3, \dots) = (00, 00, 11, 10, 00, 10, 11, 00, 11, 01, 01, 00, 10, 11, 11, 10, 11, 11, 01, 01)$ and this code word is transmitted over a BSC. The received sequence is $\underline{r} = (r_1, r_2, r_3, \dots) = (00, 01, 11, 10, 10, 10, 11, 00, 11, 11, 01, 00, 10, 10, 11, 10, 11, 11, 01, 01)$. The Viterbi decoder for this convolutional code is illustrated in Fig. 3. At the beginning of the trellis diagram, one observes from Fig. 3 that only a single path enters each state. For example, states S_0 and S_2 follow from state S_0 during the first time unit. By step (2), the partial path metric for state S_0 at time unit 1 is $M_1(0) = d_{0,1}(0, 0) = \|r_1 - v_1\| = 0$. Thus the survivor path has metric $P_1(0) = M_1(0) = 0$. Similarly, one obtains the partial metric for state S_2 as $P_1(2) = M_1(2) = 2$.

At the second time unit, only single paths enter states S_0, S_2, S_1 , and S_3 . Thus the weights of survivors at $j = 2$ are found from $P_2(0) = M_2(0) = M_1(0) + d_{1,2}(0, 0) = 1$, $P_2(2) = M_2(2) = M_1(0) + d_{1,2}(0, 2) = 1$, $P_2(1) = M_2(1) = M_1(2) + d_{1,2}(2, 1) = 4$, $P_2(3) = M_2(3) = M_1(2) + d_{1,2}(2, 3) = 2$.

At time unit 3, there are two branches entering each state. One of these two paths entering state S_0 is obtained by the algorithm using the following results: First $M_3(0) = M_2(0) + d_3(0, 0) = 3$ or $M_3(0) = M_2(1) + d_3(1, 0) = 4$; hence, $P_3(0) = \min \{3, 4\} = 3$ for state S_0 ; similarly, one yields $P_3(2) = 1$ for state S_2 , $P_3(1) = 2$ for state S_1 , and $P_3(2) = 2$ for state S_3 .

For the remaining frame times, the same procedure yields the survivor path segments from each partial path metric for each state. In Fig. 3, the metrics of the survivors

are denoted at each node. Assume that input information bits stop at time unit 20 and that one chooses the smallest metric among the four nodes. Thus, at time unit 20, the smallest metric among survivors is $P_{20}(1) = 4$. This survivor path that reaches state S_1 at time unit 20 is heavy lined in Fig. 3. For this path, the decoded information bits are (0, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1, 0).

Define W_i to be a window of length L at time unit i , where the times of the start and end of the window are i and $i + L - 1$, respectively. It is shown in [4] that if the length of the information bit stream is large, the decoding-window length L is usually several times the constraint length. At time unit L , the decoder chooses the surviving path that reaches the state with the smallest survivor by the first decoding window W_1 . The first branch in W_1 can be decoded as the first information bit. This decoding window then shifts one time unit to be the next decoding window W_2 . This new decoding window can be used to decode the second decoded information bit. The following section discusses two different methods for realizing the Viterbi decoder for a decoding window of length $L = 5K$.

III. Methods for Realizing the Viterbi Decoder

There are two methods for approximately realizing the Viterbi decoder [3]. The first method, called the register-exchange method, calls for all paths to be stored for each of the 2^K states. At each time unit, a new branch is processed by comparing the partial path metrics. Then certain registers are interchanged corresponding to the paths that survived the comparison, and a new information bit is added at one end of each register. After $5K$ branches have been processed, the first bit of register, corresponding to the smallest survivor, is shifted out as the first decoded bit. The register exchange algorithm is illustrated in [5]. This algorithm for long constraint-length code requires a substantial amount of storage space and hardware for even moderate decoding speeds.

The second method, called the trace-back method, does not store the actual information sequence but instead stores the results of each comparison. After $5K$ branches have been processed, the trellis connections are recalled in reverse order. That path traced back through the trellis diagram is used to decode the first bit. The following constitutes the trace-back algorithm.

To trace back a survivor path through the trellis diagram, one needs to store the trellis connections for each state at each time unit. For example, the survivor paths from time unit 3 to time unit 4 in Fig. 3 are shown in detail

in Fig. 4(a). At time unit 3, the state transitions to state S_0, S_2, S_1 , and S_3 at time unit 4 are the states S_1, S_1, S_2 , and S_3 , respectively. In Fig. 4(a), if one needs to trace the path from S_2 at time unit 4 back to S_1 at time unit 3, then one needs the information that a state transition from S_1 at time unit 3 to state S_2 occurred at time unit 4. Note that the last bit of state vector S_1 , i.e., the last bit of 01, shifts out of the shift register A prior to the time that state S_1 changes to S_2 . Thus this last bit of state vector S_1 needs to be stored in order to trace state S_2 back to S_1 . Let $y_j(k)$ be this one-bit information needed to trace state S_k at time unit j back to the state of the survivor at time unit $j - 1$. By this definition, $y_4(2) = 1$. As shown in Fig. 4(a), the values $y_4(0), y_4(2), y_4(1)$, and $y_4(3)$ are obtained as 1, 1, 0, and 1, respectively.

To trace a state vector back to its previous state vector survivors, one notes first by the example in Fig. 4(a) that the current state S_2 at time unit 4 is the result of shifting a 1 into register A . This bit is the first bit of state vector 10 of S_2 and should be deleted when traced back. Also the $y_4(2)$ should be linked to the last bit of this state vector. Thus one can determine the previous state vector to be 01 by deleting the first bit of state vector 10 of S_2 and concatenating with the $y_4(2)$, which is bit "1". Every previous state of states S_0, S_2, S_1 , and S_3 can be easily obtained by using the same method. These previous states are S_1, S_1, S_2 , and S_3 , respectively; they are shown in Fig. 4(b).

As mentioned in Section II, one can use each decoding window W_i to decode the i th decoded bit. Note that the decoding window length $L = 5K = 10$ in our example. To decode the first decoded bit, first one uses the first decoding window W_1 to construct the survivors by the Viterbi algorithm in Section II. Also one obtains the $y_i(k)$ for $1 \leq i \leq 10$. Let \underline{y}_i be the column vector with elements $y_i(k)$ where $k = 0, 2, 1, 3$, i.e., $\underline{y}_i = [y_i(0), y_i(2), y_i(1), y_i(3)]^T$. One obtains the \underline{y}_i for $1 \leq i \leq 10$ to be $\underline{y}_1 = [0, 0, X, X]^T$, $\underline{y}_2 = [0, 0, 0, 0]^T$, $\underline{y}_3 = [0, 0, 0, 0]^T$, $\underline{y}_4 = [1, 1, 0, 1]^T$, $\underline{y}_5 = [1, 1, 0, 1]^T$, $\underline{y}_6 = [0, 0, 0, 0]^T$, $\underline{y}_7 = [1, 0, 1, 1]^T$, $\underline{y}_8 = [0, 1, 0, 0]^T$, $\underline{y}_9 = [0, 0, 0, 0]^T$, and $\underline{y}_{10} = [1, 0, 0, 0]^T$. In the above, the $\underline{y}_1 = [0, 0, X, X]^T$ means that there are not any state changes to state S_1 or S_3 at time unit 1. Among these survivors, one chooses the state vector with the smallest metric of the survivors at time unit 10 to be $X = 11$ (state S_3). These \underline{y}_i s and state vector $X = 11$ are shown in Fig. 5(a). In Fig. 5(a), the elements in each column \underline{y}_i represent the associated $y_i k$ for $k = 0, 2, 1, 3$. One then recursively traces the survivor from state vector 11 at time unit 10 back to time unit 1 using the method shown in Fig. 4(b). First, at time unit 9, the state vector is obtained as $DMSB(11) * y_{10}(3) = 1 * 0 = 10$ where

$DMSB(11)$ denotes the state vector 11 without the first bit and the operation $*$ denotes the concatenation operation. Then at time unit 8, the state vector is obtained as $DMSB(10) * y_9(2) = 0 * 0 = 00$. Also, the state vector at time unit 7, 6, 5, 4, 3, 2, 1 can be obtained sequentially as 00, 01, 10, 01, 01, 10, 00. Finally, one obtains the first decoded bit as $Z = MSB(00) = 0$ where $MSB(00)$ denotes the first bit of 00 since the input information bit is the first bit of this state vector at time unit 1.

To decode the second decoded bit, the second decoding window W_2 is needed. Thus one needs the survivor at time unit 11. The state vector with the smallest metric at time 11 is $X = 01$ and $y_{11} = [1, 1, 1, 0]^T$, which are shown in Fig. 5(b). Using the same method above, one can trace state vector 01 at time unit 11 back to time unit 2, shown sequentially, and one obtains the state vector $X = 00$ at time unit 2. Thus the second decoded bit is $Z = MSB(00) = 0$.

The same procedure is used again to decode the following decoded bit. This procedure stops when the time unit is equal to $L_c + 5K + 1$ where L_c is the length of the code word.

The trace-back algorithm is summarized as follows. Note that because the start state is S_0 at time unit 0, one assigns metric $M_0(k) = \infty$ for $k \neq 0$.

- (1) Initially let $M_0(0) = 0$ and $M_0(k) = \infty$ for $k \neq 0$.
- (2) For each $k = 0, 1, 2, 3$, find an $l \in \{0, 1, 2, 3\}$ such that $M_{j-1}(l) + d_j(l, k)$ is minimum. Then

$$M_j(k) = M_{j-1}(l) + d_j(l, k)$$

$$y_j(k) = LSB(l)$$

- (3) If $j < 5K$, go to (2); otherwise, find an $m \in \{0, 1, 2, 3\}$ such that $M_j(m)$ is minimum. Then

$$X = m$$

Also, for $i = j, j-1, \dots, j-5K+1$,

$$X = DMSB(X) * y_i(X)$$

Then

$$Z = MSB(X)$$

- (4) If $j = L_c + 5K$, stop; otherwise, if $j \geq L_c$, $j \leftarrow j+1$ and go to (3); otherwise, $j \leftarrow j+1$ and go to (2).

$MSB(k)$ denotes the first bit of the binary representation of k . $LSB(l)$ denotes the last bit of binary representation of l . $DMSB(X)$ denotes the sequence of bits of state vector X without the first bit. L_c is the length of the code word. Step (2) is used to compute the partial path metric and store the information for choosing each survivor. Step (3) is used to trace back the survivor to find the decoded bit. First, the state vector of the state with the smallest metric of the survivor is assigned to be X when $j \geq 5K$. Then one can use this state vector to trace back the survivor by the data y_i . Once the trace-back procedure is finished, the decoded bit is denoted by Z . Note that in step (2), one needs $(5K \times 2^K)$ -bits storage space to store $y_j(k)$. Also in step (3), the trace-back procedure takes about $5K$ cycles for each decoded bit. This method requires a long decoding time.

IV. The Systolic Viterbi Decoder

The advantage of this systolic Viterbi decoder is that the trace-back operation is accomplished by processing a systolic array of registers in a pipeline fashion instead of waiting for the whole trace-back procedure. As a consequence, this systolic structure reduces the decoding time. However, if one traces back the survivor from time j to $j-1$, simultaneously the survivor is selected forward from time j to $j+1$. Thus one needs extra storage space to store the information needed to choose the survivor path at time unit j . As a result, one needs about twice the amount of storage space to store $y_j(k)$ as that needed in the original trace-back algorithm. This is expressed in detail in the following description of the operation of the systolic Viterbi decoder.

The systolic structure of the trace-back algorithm is illustrated in Fig. 6. As shown in Fig. 6, this systolic-array structure consists of a selection unit and $10K - 1 = 19$ path units. The selection unit processes step (2) of the trace-back algorithm, i.e., at time unit t , it computes the metrics of survivors of all nodes, selects the state vector with the smallest metric, and stores the information $y_t(k)$ for selecting each survivor. Note that this selection unit operates recursively. The inputs of the selection unit are the received sequence r_t and the metrics of the previous survivors $P_{t-1}(l)$. The outputs are the metrics of the survivors $P_t(l)$, the information for selecting survivors, and the state vector m of the state with the smallest metric.

Step (3) of the trace-back algorithm is modified and implemented by $10K - 1 = 19$ path units in a pipeline manner. Each path unit consists of 4-bit registers Y_i and 2-bit registers X_i for odd i or only a 4-bit register Y_i for

even i . The register Y_1 is used to store the elements of column vector \underline{y}_j at time unit j , and the register Y_i for $1 < i \leq 19$ is used to store the data shifted from register Y_{i-1} . The register X_1 is used to store the state vector m of the state with the smallest metric, and the register X_i for $i = 3, 5, \dots, 19$ is used to store the data shifted from register X_{i-2} ; the one-bit register Z is used to store the decoded bit. The Register Transfer Language (RTL) developed in [6] is used to describe in detail the operation of this systolic structure as shown in Fig. 6. At time unit t , the contents in each register Y_i for $i = 1, 2, \dots, 18$ is transferred to the following register Y_{i+1} and the $y_t(k)$ s for $k = 0, 1, 2, 3$, which are the output of selection units, are stored in the register Y_1 . If $t \geq 5K$, the contents in register X_i without the first bit for odd $i \leq 17$ is concatenated with the corresponding element in register Y_i , where the address of this element is the contents of X_i , to generate a new state vector. This new state vector is then transferred to the following register X_{i+2} . Also, the state vector m of the state with the smallest metric is transferred to the register X_1 , and the first bit of the contents of register X_{19} is stored in register Z as the decoded bit.

Using the same example as that in Section III, one obtains the contents in each register at different time units as shown in Fig. 7. Note that at time unit $10K = 20$, one obtains the first decoded information bit stored in register Z and then one after each time unit. One obtains the decoded information bits sequentially. Each odd path unit stores 2^k bits of data in Y_i and k bits in X_i , while each even path unit stores 2^K bits of data in Y_i .

Note also that it is possible to further reduce the number of registers. As shown in Fig. 7 at time unit 19, the first bit of X_{19} is the last bit of X_{17} at time unit 18. Thus the registers Y_{18} , Y_{19} , and X_{19} are not needed in this $(3, 1/2)$ Viterbi decoder, and register Z is directly connected to register Y_{17} to store the first bit of the contents of Y_{17} . Also, the first decoded bit is decoded at time unit 19 instead of time unit 20. If a structure is similar to that used in a $(3, 1/2)$ convolutional code, the general systolic Viterbi decoder can be obtained.

It has been shown [3, 7] that one can use any state vector to trace back if the window length is equal to $5K$. This is due to the fact that all survivor paths most likely will merge within a window length of $5K$. Then the operation of selecting the state vector with the smallest metric of the survivor in the selection unit is eliminated. As a result, this tracking, updating, and storing of the hypothesized information sequence can be accomplished simultaneously during a single clock cycle. This systolic-array structure minimizes the interconnections between components. Hence, this new architecture is suitable for a VLSI implementation.

V. Conclusion

The realization of the systolic Viterbi decoder for convolutional codes is demonstrated. This fully parallel architecture of decoding requires a minimum amount of storage space and decoding time. This makes it possible to readily implement a Viterbi decoder with VLSI circuits.

References

- [1] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*, New York: McGraw-Hill, 1979.
- [2] R. L. Miller, L. J. Deutsch, and S. A. Butman, *On the Error Statistics of Viterbi Decoding and the Performance of Concatenated Codes*, JPL Publication 81-9, Jet Propulsion Laboratory, Pasadena, California, September 1, 1981.
- [3] G. C. Clark, Jr. and J. B. Cain, *Error-Correction Coding for Digital Communications*, New York: Plenum Press, 1981.
- [4] R. E. Blahut, *Theory and Practice of Error Control Codes*, Massachusetts: Addison-Wesley Publishing Company, May 1984.
- [5] R. J. McEliece, *The Theory of Information and Coding*, Massachusetts: Addison-Wesley Publishing Company, 1977.
- [6] T. C. Bartee, I. L. Lebow, and I. S. Reed, *Theory and Design of Digital Machines*, New York: McGraw-Hill Book Company, 1962.
- [7] S. Lin and D. J. Costello, Jr., *Error Control Coding*, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1983.

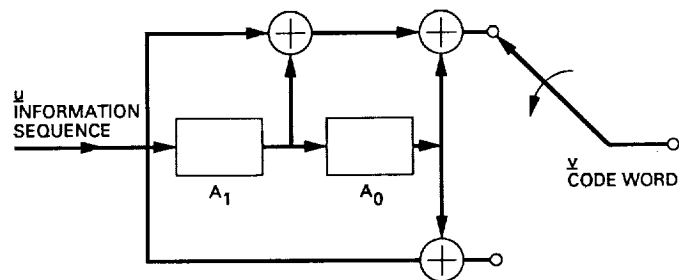


Fig. 1. A (3,1/2) convolutional code.

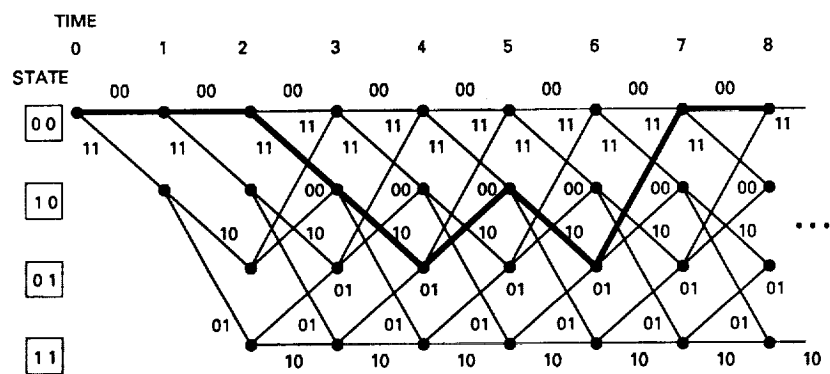


Fig. 2. Trellis diagram for a (3,1/2) convolutional code.

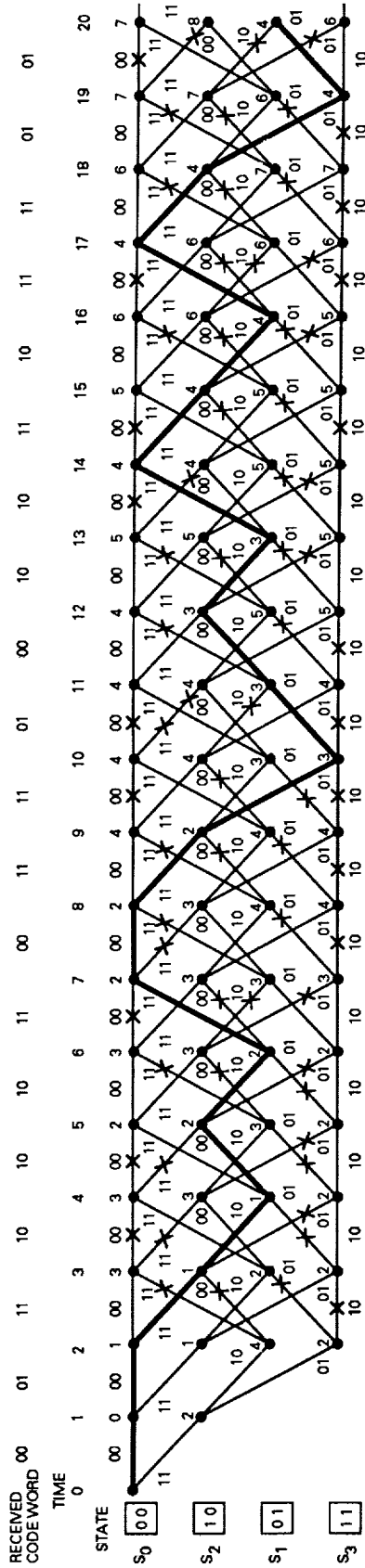
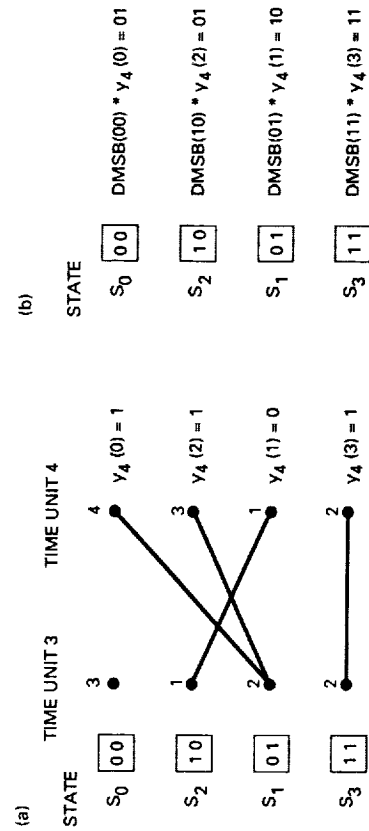


Fig. 3. Viterbi algorithm for a binary symmetric channel.

Fig. 4. Survivor paths from time unit 3 to time unit 4 and the previous states of S_0 , S_2 , S_1 , and S_3 : (a) the assignment of $y_4(k)$ in the trace-back algorithm; (b) the traced-back state vector for each state.

(a)	y_{-1}	y_{-2}	y_{-3}	y_{-4}	y_{-5}	y_{-6}	y_{-7}	y_{-8}	y_{-9}	y_{-10}	
STATE											
00	0	0	0	1	1	0	1	0	0	1	
10	0	0	0	1	1	0	0	1	0	0	
01	X	0	0	0	0	0	1	0	0	0	
11	X	0	0	1	1	0	1	0	0	0	X = 11

(b)	y_{-2}	y_{-3}	y_{-4}	y_{-5}	y_{-6}	y_{-7}	y_{-8}	y_{-9}	y_{-10}	y_{-11}	
STATE											
00	0	0	1	1	0	1	0	0	1	1	
10	0	0	1	1	0	0	1	0	0	1	
01	0	0	0	0	0	1	0	0	0	1	
11	0	0	1	1	0	1	0	0	0	0	X = 01

Fig. 5. The column vector y_i , the state vector X with the smallest metric: (a) the data of y_i in decoding window W_1 for decoding the first bit; (b) the data of y_i in decoding window W_2 for decoding the second bit.

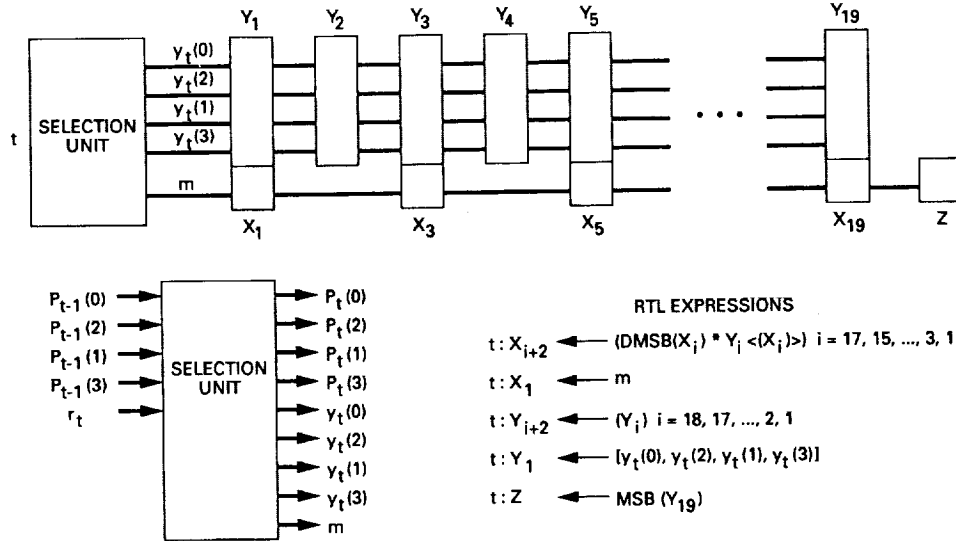


Fig. 6. The structure of the systolic Viterbi decoder.

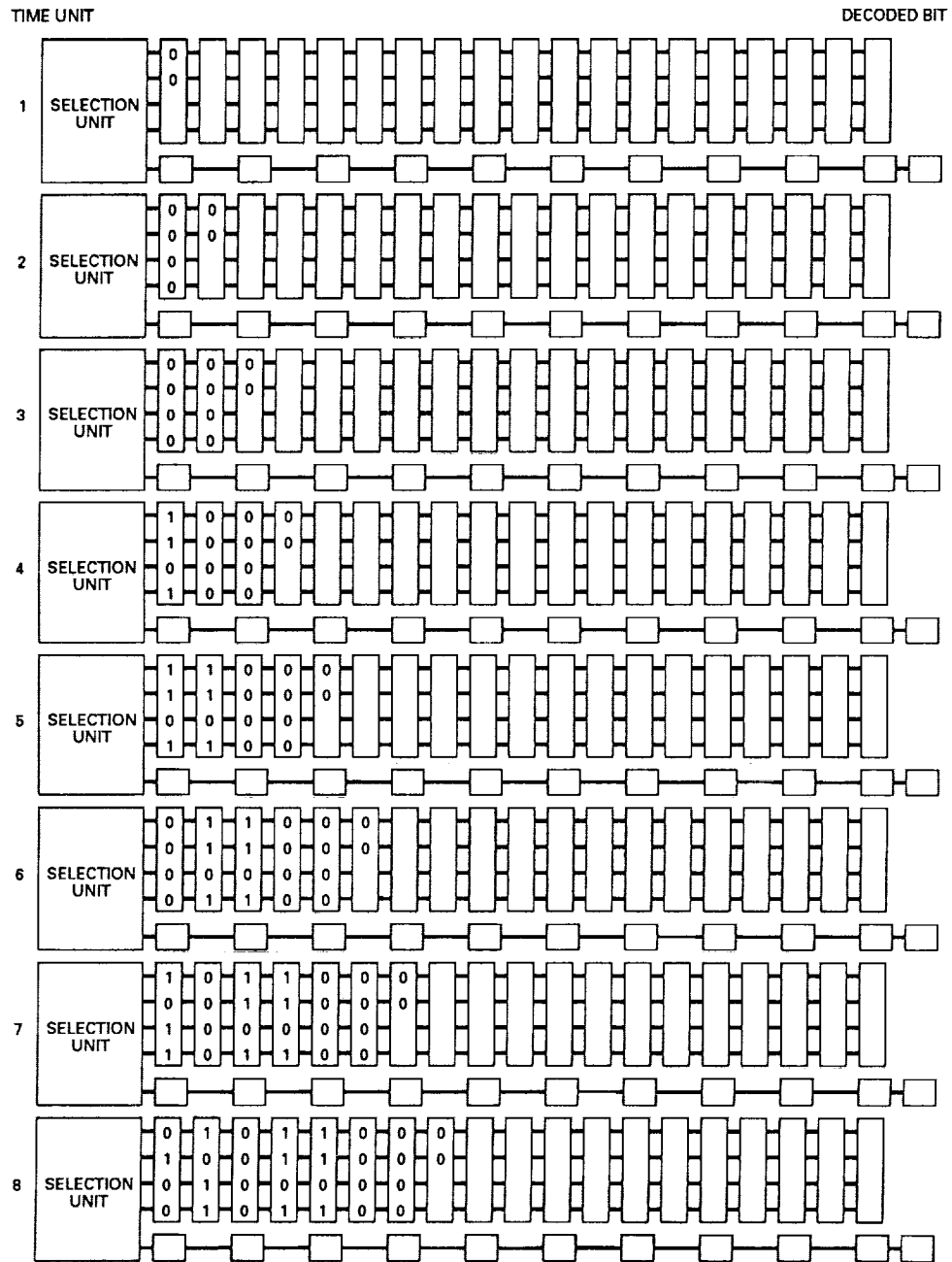


Fig. 7. The systolic Viterbi decoder for a (3,1/2) convolutional code.

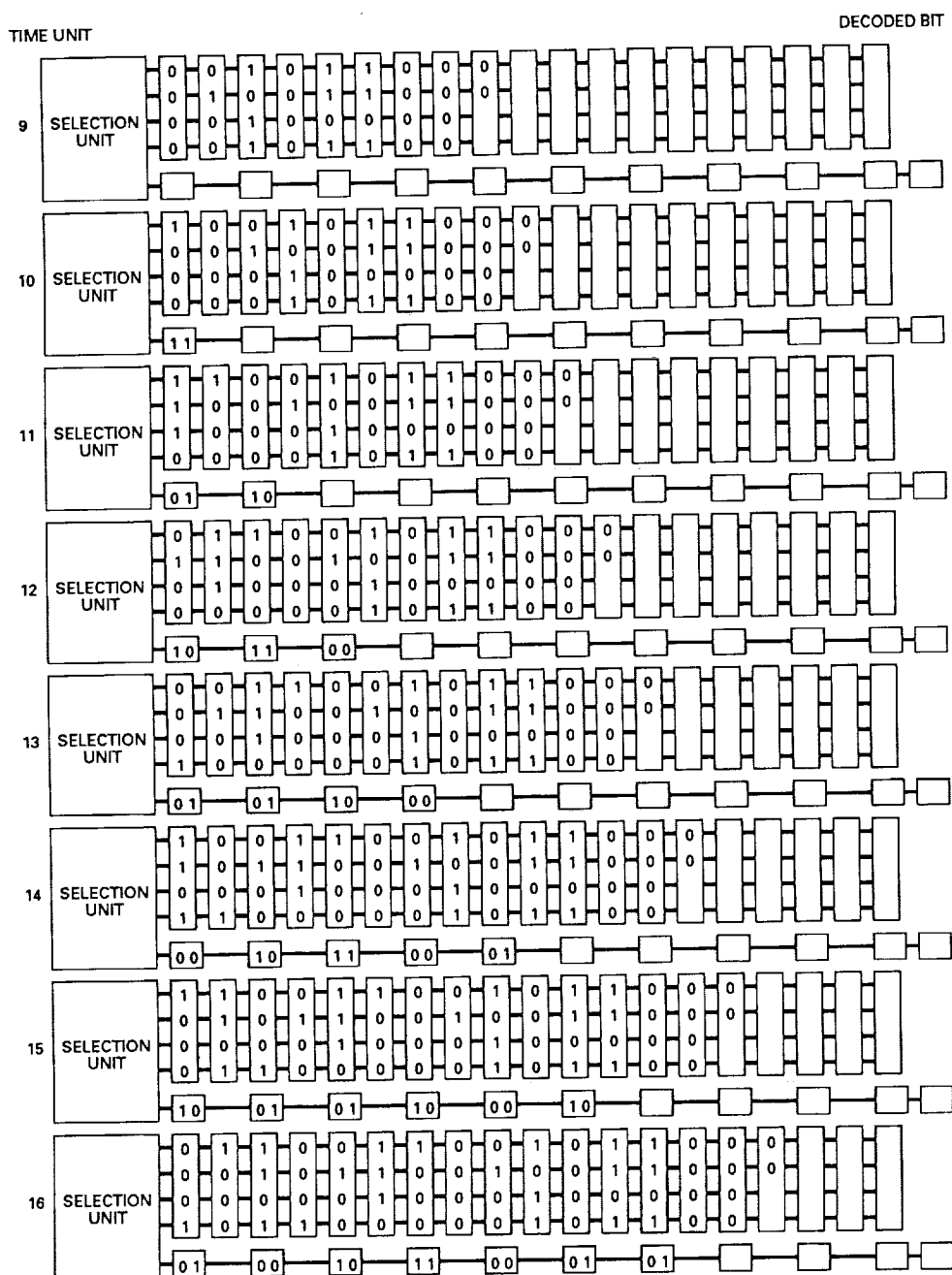


Fig. 7. (contd)

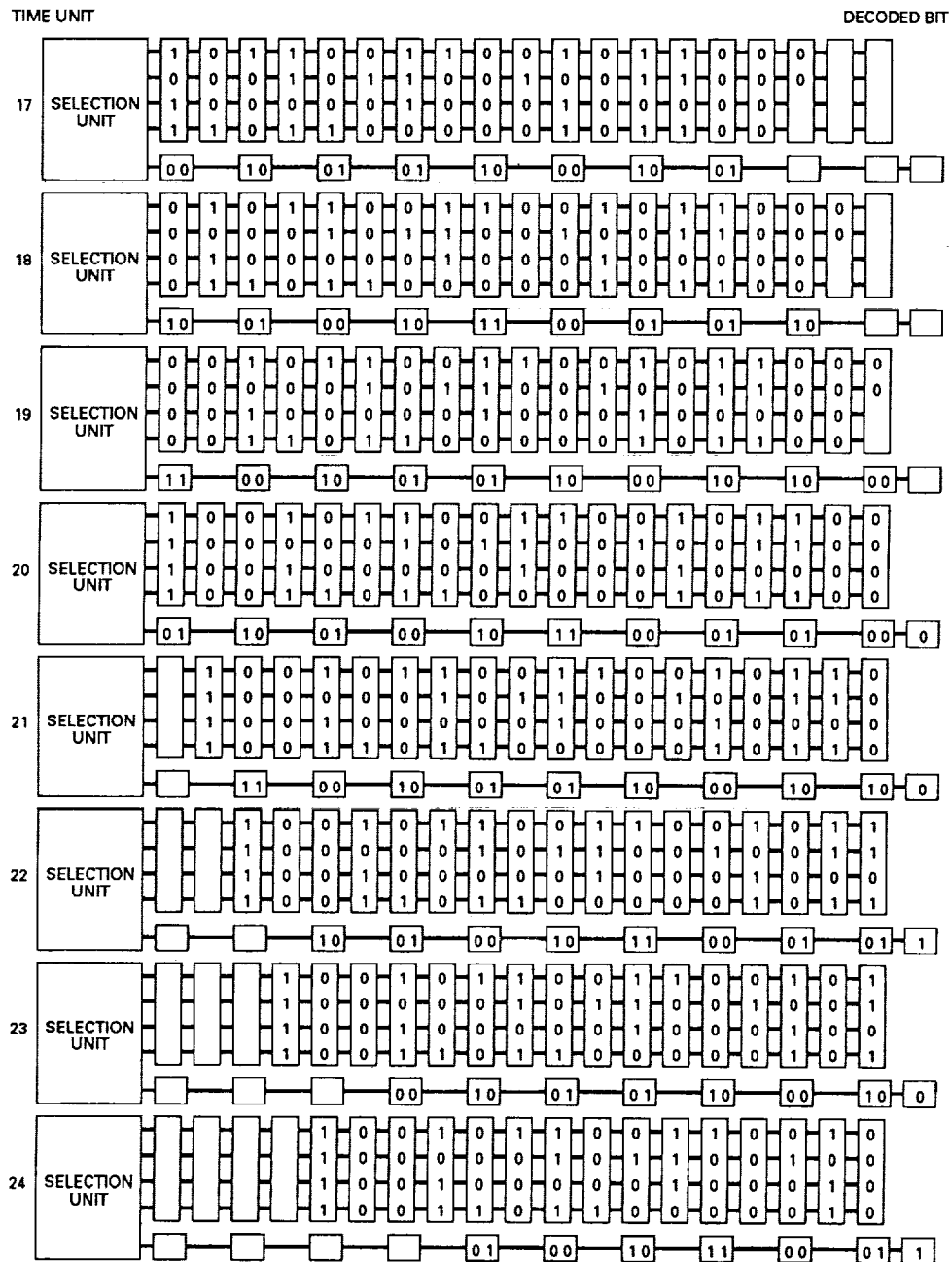


Fig. 7. (contd)

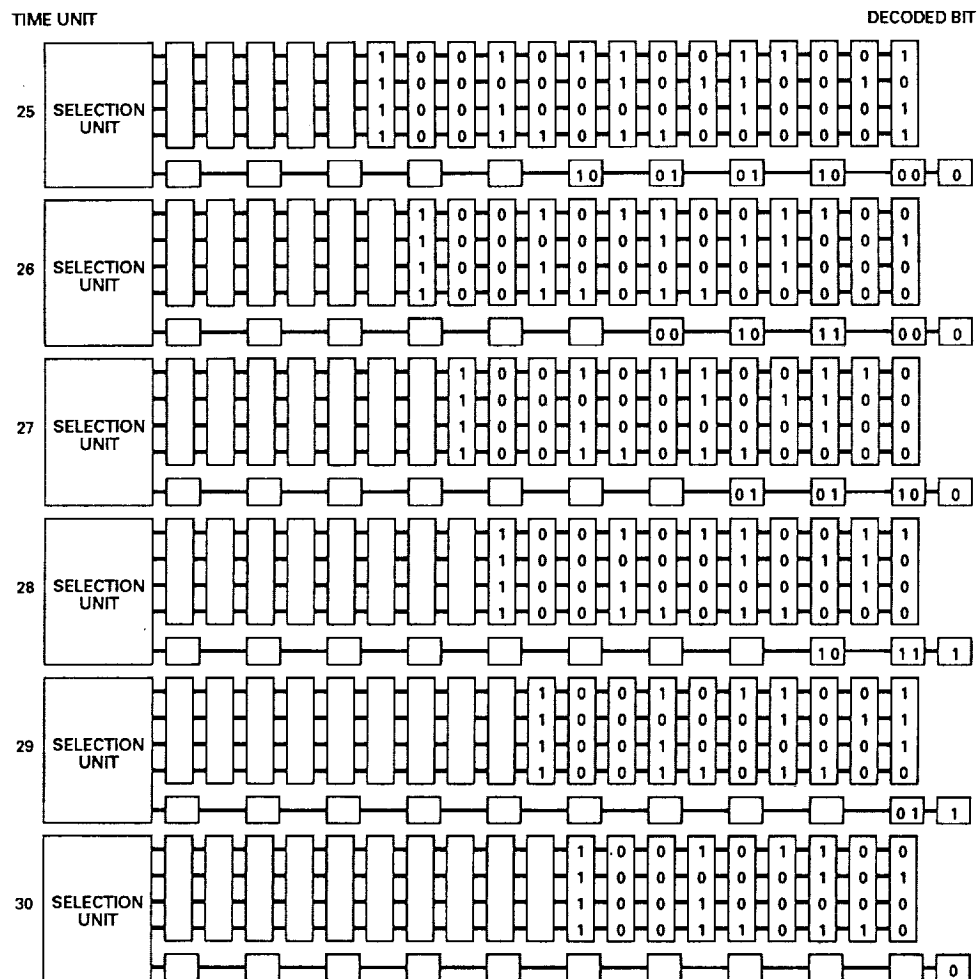


Fig. 7. (contd)

Radar RFI at Goldstone DSS 12 and DSS 16

S. D. Slobin and T. K. Peng
Telecommunications Systems Section

Radio frequency interference (RFI) from the DSS 14 Goldstone Solar System Radar (GSSR) was investigated at DSS 12 and DSS 16 with the goal of assisting in the choice of the location of future DSN antennas. Total power measurements at both locations were made at the S-band carrier frequency of 2320 MHz. X-band measurements at the carrier frequency of 8495 MHz could not be made. Exciter-chain output spectrum and klystron output spectrum measurements were made at S- and X-bands using a probable worst-case modulation of the radar signal (short pseudorandom number [PN] code length and short pulse length). Based on these measurements it is estimated that RFI levels in the DSN receiving bands at both sites (above 10-deg elevation) would be below -192 dBm for a 1-Hz bandwidth.

I. Introduction

During January through May 1989, a series of tests was carried out at Goldstone to determine whether S- and X-band solar system radar experiments using the DSS 14 70-m antenna would create radio frequency interference (RFI) at new DSN receiving antennas which might be located at the present DSS 12 or DSS 16 sites. All tests were carried out in clear weather. The S-band radar frequency is 2320 MHz, and the DSN downlink band extends from 2200 to 2300 MHz. The X-band radar frequency is 8495 MHz, and the DSN downlink band extends from 8400 to 8440 MHz.

Modulating the radar signal causes power to extend away from the carrier frequencies; depending on the pulse and code lengths of the modulation, this power may be a problem to spacecraft signal reception. The power level decreases with frequency away from the carrier; therefore received power is highest at the upper end of the S- and X-receiving bands. DSS 16 does not have an X-band receiver

and the DSS 12 X-band receiver would not tune up as high as 8495 MHz. The modulated radar signals were estimated to be too weak to be received by the DSN receivers, yet the RFI requirements were below even that low level of detectability.

A primary driver for the RFI measurements was the need to determine a low-noise location for the new 34-m beam-waveguide DSN receiving antenna at Goldstone (DSS 18). Colocation with the existing 26-m antenna (DSS 16) was under consideration.

The plan for doing these tests was as follows:

- (1) Make received-power measurements at the S- and X-band radar carrier frequencies at DSS 12 and DSS 16. These measurements would be made using all orientations of the transmitting and receiving antennas. X-band measurements were later found impossible to make with existing equipment.

- (2) Using the *total* received power levels, predict the worst-case received power per spectral line at the tops of the DSN receiving bands under the condition of worst-case modulation (short pulse length and short code length). Included in this modeling is an estimate of the radar-transmitter input and output filtering.
- (3) Measure the transmitter exciter-chain output spectrum (klystron-amplifier input spectrum) and klystron output spectrum at DSS 14 to verify estimates used in the RFI modeling.
- (4) Determine whether deduced power at the top of the receiving bands exceeds the DSN limit of -192 dBm per 1-Hz bandwidth. This limit is 6 dB below the noise power of a 20-K system (6 dB below -186 dBm).

The following tests were carried out:

- (1) January 27, 1989: S-band total-power RFI at DSS 16 (26-m, X-Y antenna)
- (2) February 27, 1989: S-band total-power RFI at DSS 12 (34-m, HA-Dec antenna)
- (3) March 31, 1989: S-band total-power RFI at DSS 12
- (4) April 9, 1989: S- and X-band radar modulation spectrum at klystron input (exciter-chain output) at DSS 14
- (5) May 17, 1989: S-band total-power RFI at DSS 16
- (6) May 22, 1989: S- and X-band radar modulation spectrum at klystron output at DSS 14

Throughout the test period, there were FAA and USAF restrictions placed on both transmitter power and transmitter pointing. At high power (400 kW, S-band), the DSS 14 antenna could not point lower than 15-deg elevation; below that, down to 10.6 deg, the power was limited to 105 kW. The X-band restrictions, had that frequency been used, would have been more severe due to the much higher X-band gain of the transmitting antenna. In addition, azimuth restrictions were placed on DSS 14 pointing. Sometimes the DSS 14 antenna could point toward the receiving antenna; sometimes that azimuth was totally off-limits. During one test (Test 5), there was one azimuth restriction for the first half and another during the second half, and only then was it possible to point directly toward DSS 16. The two sets of received-power data thus represented two different multipath/refraction conditions, and it was necessary to merge these data in an indirect way in order to create a consistent set of full-sky

data at the receiving antenna. Considering the azimuth, elevation, and power restrictions of the transmitting antenna, and the pointing differences and restrictions of the receiving antennas (DSS 16 is mechanically restricted from looking at the horizon due east or west; DSS 12 cannot look near the northern horizon toward DSS 14), only with great difficulty were the disparate data sets merged to create a consistent RFI model of the Goldstone environment.

II. RFI Field Test Results

A. Test 1 (DSS 16, S-Band, Total Power, January 27, 1989)

Figure 1 shows a received radar continuous wave (CW) signal as seen on a spectrum analyzer display during the first DSS 16 test. The view here is a 100-kHz portion of the IF band, centered at 520 MHz. Adjacent to the radar signal is a calibrated test signal, which is injected in front of the station's low-noise amplifier (S-band paramp). In this case, the test signal is offset 10 kHz below the radar signal and has an amplitude of -70 dBm. Thus, the received radar signal is determined to have a strength of -83 dBm. These are CW (single-line) signals and thus are of "infinitesimal" width. The line widths shown in the figure are convolutions of the individual lines and the 100-Hz resolution bandwidth of the spectrum analyzer. It should be noted that this was the highest-strength signal received during any of the tests and arose because both the DSS 14 antenna (transmitting 100 kW) and the DSS 16 receiving antenna were pointed at low elevation angles toward a group of hills located west of Goldstone Lake. These relative antenna orientations would not be used during normal radar and spacecraft tracking operations. The noise floor of the spectrum analyzer is shown to be -130 dBm (for the 100-Hz resolution bandwidth). With the narrowest resolution bandwidth available (10 Hz), radar signals as low as -140 dBm would be visible.

Figure 2 shows a series of received power measurements made at 2320 MHz on the DSS 16 antenna pointing at an elevation angle of about 4.5 deg, as low as the antenna could point along the horizon. At the DSS 16 azimuth of about 292 deg ($Y = -68$ deg), there exist a number of hills west of Goldstone Lake rising to a height of 1000 to 1500 ft above the elevation of DSS 16. With both antennas pointing toward those hills, received powers in the -83 -dBm to -90 -dBm range were observed (depending on the elevation angle of DSS 14). This type of isolated high RFI is referred to as a "hot spot." With DSS 14 pointing toward DSS 16, and DSS 16 pointing toward DSS 14 (azimuth = 351.4 deg, $Y = -8.6$ deg), powers in the range of -105 dBm to -120 dBm were observed,

with DSS 14 transmitting both 400 kW and 100 kW at low elevation angles. At other azimuths along the northern horizon (4.5-deg elevation), DSS 16 received powers as low as -127 dBm.

Based on far-field assumptions of antenna gain as a function of angle off beam peak, it was predicted that for the case of the two antennas pointing at each other (DSS 14, 400 kW, 15-deg elevation, and DSS 16, 4.5-deg elevation, 9.7 km apart), the received power at DSS 16 would be -14 dBm, from which one would infer a terrain/hill isolation effect of at least 91 dB. Even if the DSS 14 off-axis gain is 30 dB lower than modeled (due to incorrect far-field assumptions), the inferred hill isolation would still be 61 dB, a probable lower limit to the effect of the terrain.

B. Test 2 (DSS 12, S-Band, Total Power, February 27, 1989)

Figure 3 shows received power levels at the DSS 12 HA-Dec antenna. For the limited set of measurements, received powers lie in the -120 -dBm to -130 -dBm range, except for two points: (1) -105 dBm with DSS 12 pointing toward DSS 14 at the former's lowest elevation possible (32 deg for that hour angle/declination), and (2) -115 dBm with DSS 12 pointing at zero elevation angle near the southeastern horizon (azimuth = 150 deg). Although the 15.9-km separation should result in 4 dB more path loss than for the DSS 14/16 test, similar power levels were received. In any case, it appears that the power entering the DSS 12 receiving system is fairly independent of antenna pointing, as though the entire spillover for that antenna saw an enclosing "RFI blackbody" at a fairly uniform power level.

C. Test 3 (DSS 12, S-Band, Total Power, March 31, 1989)

A more comprehensive set of tests was made at DSS 12 approximately one month after Test 2. This time both test-signal comparison and AGC voltage were used to measure the 2320-MHz received radar signal. Figure 4 shows a polar plot of the AGC level (in dB) relative to the zenith power level of -127 dBm. The data taken cover virtually the entire sky from horizon to zenith, excluding the unreachable northern horizon. Except for the $+22$ dB point (-105 dBm), all data points lie in the range of -113 to -135 dBm. The highest point on this curve ($+22$ dB) is, as seen in Fig. 3, when DSS 12 points toward DSS 14 at about a 35-deg elevation angle. Other high points are at the horizon near azimuth 210 deg (14 dB above the zenith level) and low in the sky near azimuth 105 deg ($+11$ dB).

D. Test 5 (DSS 16, S-Band, Total Power, May 17, 1989)

During this test, azimuth restrictions on the DSS 14 transmitting antenna would not allow complete DSS 16 sky coverage for one consistent experiment condition. During the first two hours of the test, DSS 14 pointed at an azimuth of 195.1 deg, which was as far east as possible to the DSS 14/16 azimuth line of 171.2 deg. During the second two hours of the test, DSS 14 pointed toward DSS 16 at an azimuth of 171.2 deg. Throughout the test, DSS 14 transmitted 400 kW at an elevation of 15 deg. RFI was measured relative to a calibrated test signal as described above. Toward the middle of the test, an FET amplifier was installed, which allowed measurement of signals as low as -145 dBm. The only common set of data for the two DSS 14 azimuths was the DSS 16 scan at $Y = -30$ deg (westward). It was found that on the average the received power level *dropped* 6 dB when DSS 14 was pointed directly toward DSS 16, as opposed to the first half of the data taken when DSS 14 was pointed 24 deg westward (azimuth = 195.1 deg). This RFI decrease may be due to increased "line-of-sight" hill blockage, whereas the higher levels with nondirect pointing may be due to some multipath or diffraction effect.

Figure 5 shows the X-Y antenna coordinate system with received power levels indicated. All data shown without parentheses correspond to the line-of-sight (azimuth = 171.2 deg) condition. The $Y = 0$ points are the original data points (shown in parentheses, taken during the first two hours) decreased by the 6-dB loss determined above.

For the direct-pointing case (DSS 14 azimuth = 171.2 deg, toward DSS 16), received RFI power ranged from about -125 to -140 dBm, except for some hot spots near the northern horizon below 10-deg elevation, and a particular hot spot at an elevation of 26 deg and azimuth of 327 deg (see Fig. 5). For the nondirect pointing case (DSS 14 azimuth = 195.1 deg), this RFI range would be -119 to -134 dBm (6 dB higher), with correspondingly higher hot spots.

III. Radar Spectrum Tests

A. Theoretical Background

The purpose of these tests was to determine if there exists sufficient filtering or bandwidth limiting in the uplink exciter chain or the klystron amplifier to cause the modulated radar-spectrum amplitude at the top of the DSN bands to fall below the -192 -dBm CCIR (International Radio Consultative Committee) limitation. Al-

though the modulated spectrum falls off in amplitude away from the carrier frequency, this alone is not sufficient to bring the RFI level down to -192 dBm at 2300 MHz.

Figure 6 shows the theoretical spectrum of a modulated S-band radar signal. The curve shown represents the envelope of nearly three hundred lines between 2295 MHz and 2320 MHz (the carrier frequency). The modulation used for this representation is that which would give the highest sideband power level for realistic radar signals transmitted by the DSN antennas (PN code length $p = 31$, digit period $t = 0.375$ μ sec). It is seen that relative to the power at the carrier frequency of 2320 MHz, the power at 2300 MHz is down only 28 dB. It can be shown [1] that for any particular total power transmitted, the power of the peak line level of the modulated envelope is 15 dB below the total power level. The powers in all the individual lines add p to the total transmitted power. Thus by measuring the total RFI received power (CW, unmodulated) at the carrier frequency, one can equate that level to the power in all the lines of the *modulated* spectrum. As an example, if the total CW RFI power received is -120 dBm, the power in all the lines together of a *modulated* spectrum is -120 dBm. Since the envelope peak is 15 dB below the total, the modulated-spectrum absolute power level can thus be determined. In Fig. 6, for a total power of -120 dBm, the power level *per line* at 2320 MHz is thus -135 dBm, and the power *per line* at 2300 MHz is 28 dB below that, or -163 dBm. This power is 29 dB *above* the CCIR limit of -192 dBm. Obviously, additional filtering is required.

Figure 7 shows the theoretical spectrum of an X-band radar signal modulated identically to the S-band signal described above ($p = 31$, $t = 0.375$ μ sec).

In the above examples, the assumed total received RFI power is -120 dBm. This represents an average upper limit of the S-band measurements described above. Of course, several hot spots were detected, but the antenna orientations for those measurements are not considered representative of normal DSN spacecraft tracking/radar pointing configurations.

X-band RFI measurements could not be made due to equipment limitations. Theory would indicate, however, that for the same transmitted powers, X-band RFI would be less than that at S-band, due to lower terrain diffraction (hence multipath) effects at that shorter wavelength. To first order, there appear to be no obvious antenna-related RFI differences such as spillover or quadripod diffraction. In the analyses that follow, it is assumed that X-band received RFI total power is -120 dBm (as in the S-band case), which may represent a worst-case scenario.

B. Test 4 [DSS 14, Radar Spectrum at Exciter-Chain Output (Klystron Input), S- and X-Bands, April 9, 1989]

Figure 8 shows the transmitter exciter-chain output spectrum measured at the klystron-amplifier input for code length $p = 31$, code period $t = 0.375$ μ sec. Normalizing the spectrum to -120 dBm total received power (-135 -dBm envelope peak) shows that at 2300 MHz the single peak has a power level of about -188 dBm.

The normalization to -120 dBm is somewhat arbitrary, but represents an average upper limit to the RFI received total power during the DSS 12 and DSS 16 tests, as described above. Obviously several hot spots existed during those tests, but the average RFI environment that existed was typically in the -120 -dBm to -130 -dBm range. At -130 dBm, the 2300-MHz signal shown here would have an absolute level of -198 dBm, 6 dB *below* the CCIR limit. Also shown is the theoretical modulated spectrum envelope for $p = 31$, $t = 0.375$ μ sec. It is seen that at 2300 MHz, 25 dB of additional filtering has actually occurred, compared to the 29 dB needed.

Similar tests were made at X-band, and the modulated spectrum is shown in Fig. 9 along with the theoretical spectrum for the same code length and period. It is seen that significantly more filtering exists and the normalized power level at 8440 MHz (for -120 dBm total RFI power) is below -200 dBm, substantially below the CCIR -192 -dBm limit.

C. Test 6 (DSS 14, Radar Spectrum at Klystron Output, S- and X-Bands, May 22, 1989)

Although the S- and X-band klystron input spectra described above indicate that for the RFI power levels measured at DSS 12 and DSS 16 sufficient filtering exists to meet CCIR RFI limits, additional tests were carried out to determine if undesired harmonics in the klystron existed or if additional klystron filtering would reduce the RFI problem.

Figure 10 shows the S-band 395-kW klystron output with $p = 31$ and $t = 0.375$ μ sec. Comparison with the klystron input in Fig. 8 shows that no harmonics can be seen and that the output has slightly lower sidebands. Thus, it is evident that the klystron has an additional bandwidth-limiting effect. At 2300 MHz, the normalized power level (for -120 dBm total power) is about -202 dBm, a decrease of power deduced from the input spectrum of 14 dB (the input level being -188 dBm). Thus, for this case, even for RFI received-power levels

of -110 dBm (higher than virtually all the experimental field-test data, except the rare hot spots), the CCIR RFI limit of -192 dBm is met at 2300 MHz. At frequencies below this, the received power would decrease even more because of the steep rolloff of the measured modulated spectrum.

Figure 11 shows the X-band klystron *input* spectrum (measured on May 22, 1989, $p = 31$, $t = 0.375 \mu\text{sec}$) that is significantly different than that shown in Fig. 9. The difference in the two tests is that the synthesizers in the exciters differed, the first one producing more noise than the second. The klystron *output* spectrum taken the same day as the Fig. 11 *input*, using the same exciter input, is shown in Fig. 12. Comparison of Figs. 11 and 12 indicate virtually identical spectra, both with steeper rolloff than the earlier Fig. 9 spectrum. There appears to be no additional X-band klystron bandwidth-limiting effect, and it is assumed that an output spectrum corresponding to the Fig. 9 input will also be identical. In any case, since the Fig. 9 spectrum appears to satisfy the CCIR RFI limit by a substantial margin, use of the narrower-band modulation spectrum of the Fig. 12 case can only improve matters.

IV. Summary and Conclusions

A. Total Power Measurements

Figure 13 shows a chronological presentation of all 102 S-band RFI data points measured in the four field tests described in Section II. The 44 data points shown as "+" are for receiving-antenna elevations of less than 10 deg. Virtually all of the hot-spot points are at those low elevation angles. In particular, the high measurements above -100 dBm in the first 21 points were made at 4.5-deg elevation in a localized hot-spot direction (see Fig. 2). Only four data points with powers greater than -120 dBm are for elevations of 10 deg and higher. These four points represent only two distinct pointing conditions of the receiving antennas. Points 28 and 40 were obtained in Tests 2 and 3, respectively. The antenna pointing angles were practically the same for these two points. Points 66 and 67 were taken with identical pointing of the DSS 16 antenna during Test 5 (see Fig. 5). Thus, there are only two high-elevation RFI hot spots in this data set. The antenna pointing angles for the four points described above are shown in Table 1.

Seventy-five points have powers of -120 dBm or lower. Disregarding the points below 10 deg, there are only two points each for DSS 12 and DSS 16 that are above

-120 dBm; fifty-four points are -120 dBm or lower. This supports the use of -120 dBm in the theoretical modeling described in Section III. Other than the general -120-dBm upper limit for each site, it is difficult to characterize one or the other of the sites as having a "better" RFI environment.

The low-elevation hot-spot RFI may be a result of metallic objects reflecting hill-diffracted radar signals into the face of the receiving antennas. Because the lower edge of all the antennas is very near the ground at low elevation angles (5-20 ft in most cases), nearby objects such as microwave relay towers and their 6-8-ft diameter dishes, floodlight posts, air-conditioning ducts and blowers, collimation towers, and water tanks may all be within the near-field cylindrical beam of the receiving antenna. Multipath from hill, to metallic reflector, to quadripod, to feedhorn may be a cause of RFI. Additionally, significant "isotropic" RFI may be due to sidelobe and backlobe response to radar signals reflected from the myriad metallic objects located within miles of each antenna. In other words, although the antenna might be pointed at a high elevation angle, the backlobes are receiving RFI from objects on the ground, behind the antenna. It is suggested that future investigations of these effects be made and tests carried out to develop methods of reducing this problem.

X-band RFI measurements could not be made, but it is believed (as stated in Section III.A) that X-band received powers would be somewhat less than the corresponding S-band powers for the same antenna-pointing configurations because of decreased diffraction/multipath effects at the higher frequency.

B. Exciter-Chain Output and Klystron Output Spectrum Tests

Measurements of exciter-chain output spectrum and klystron output spectrum at both S- and X-bands indicate significant bandwidth filtering beyond the mathematical rolloff of modulated radar signals. Using a theoretical model of a worst-case spectrum ($p = 31$, $t = 0.375 \mu\text{sec}$) and a total received power of -120 dBm, it is shown that the additional filtering presently existing in the radar up-link at DSS 14 (in excess of 20 to 30 dB) appears sufficient to reduce the received power levels at the top of the DSN receiving bands (2300 MHz and 8440 MHz) to below the CCIR limit of -192 dBm for a 1-Hz bandwidth. However, if the radar bandwidth is increased in the future, this limit may be exceeded.

Acknowledgments

The authors wish to thank the following people for their assistance in carrying out the Goldstone RFI study reported here: John Sosnowski, Erik Holmgren, Chuck Goodson, Dennis Choate, and Larry Sturgis helped us in making the RFI measurements; Reg Cormier carried out the transmitter spectrum tests; Al Banji and Ray Jurgens advised us on numerous radar matters; and Mike Brown and Dee Yeaman arranged operations support and scheduling for all these tests.

Reference

- [1] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

Table 1. Antenna-pointing parameters for four points with powers greater than -120 dBm

Point	Station	Azimuth, deg	Elevation, deg	RFI Power, dBm
28	DSS 12	333	32	-105
40	DSS 12	331	35	-105
66	DSS 16	326	26	-115
67	DSS 16	326	26	-112

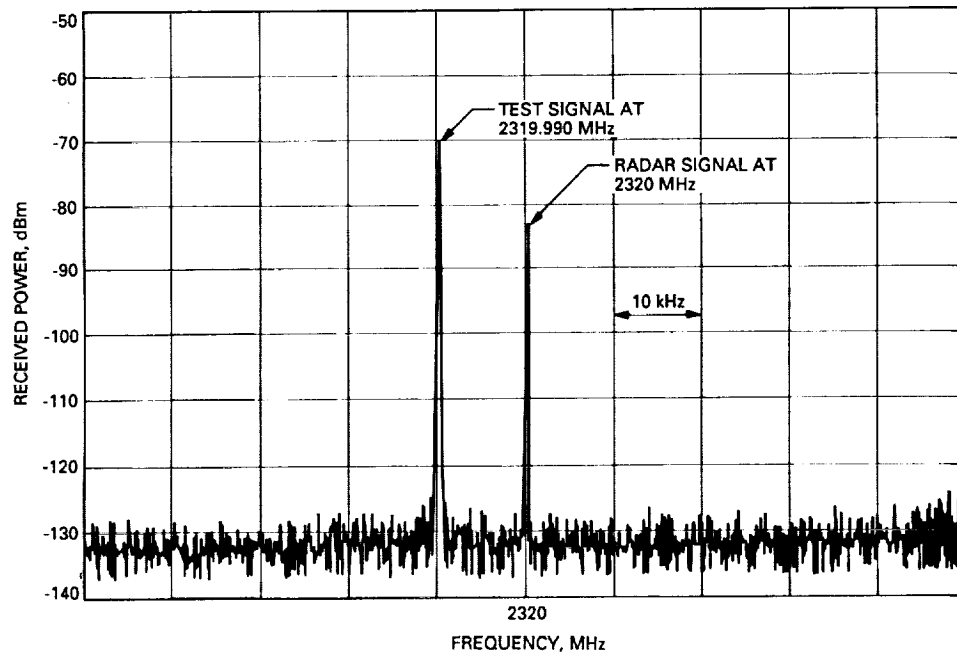


Fig. 1. Test 1: DSS 16, January 27, 1989, -70-dBm test signal and -83-dBm radar RFI at 2320 MHz.

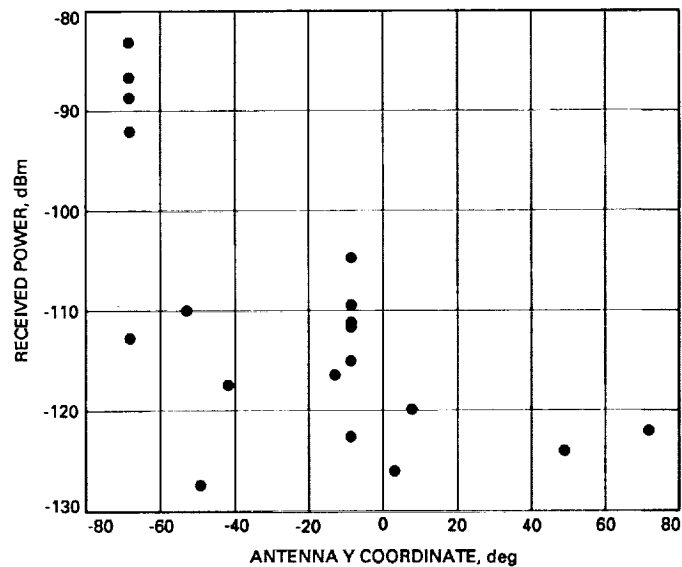


Fig. 2. Test 1: DSS 16, January 27, 1989, radar RFI versus DSS-16 pointing, elevation = 4.5 deg.

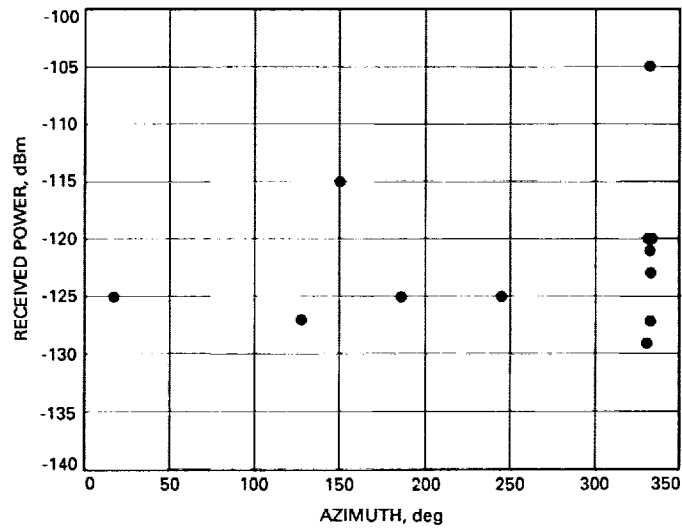


Fig. 3. Test 2: DSS 12, February 27, 1989, radar RFI versus DSS-12 pointing.

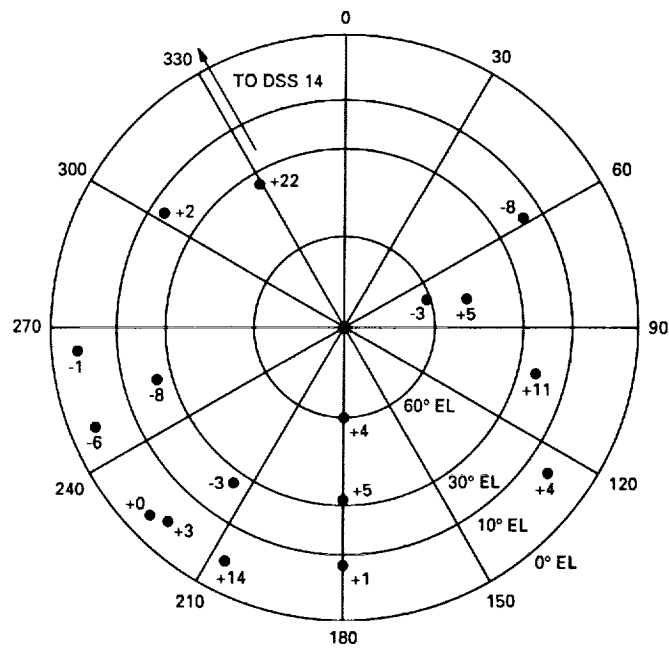


Fig. 4. Test 3: DSS 12, March 31, 1989, radar RFI versus DSS-12 azimuth-elevation pointing, referenced to zenith power level of -127 dBm.

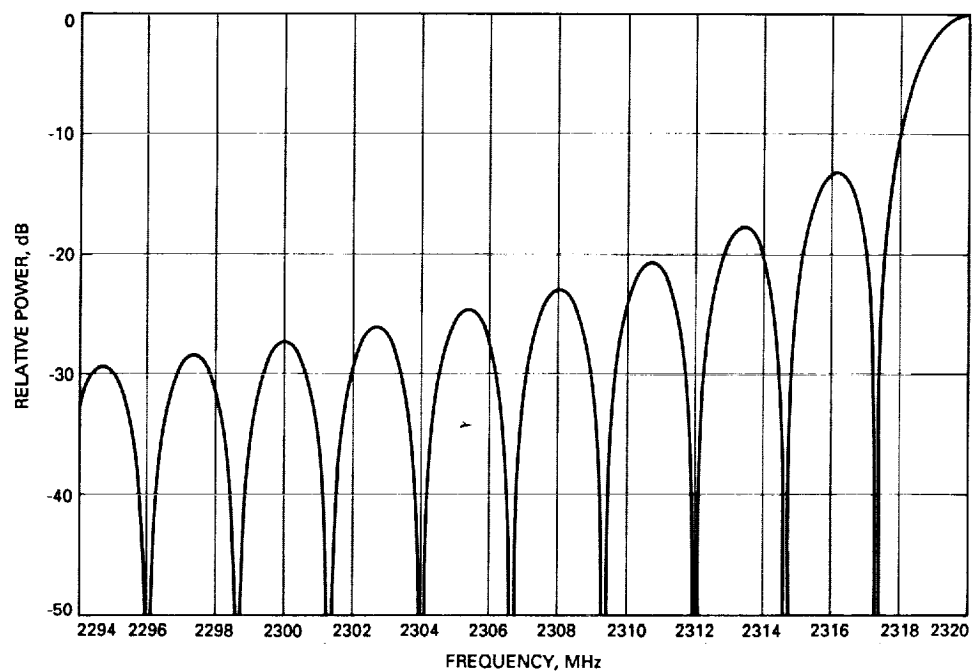


Fig. 6. Theoretical modulated S-band radar spectrum for $p = 31$, $t = 0.375 \mu\text{sec}$.

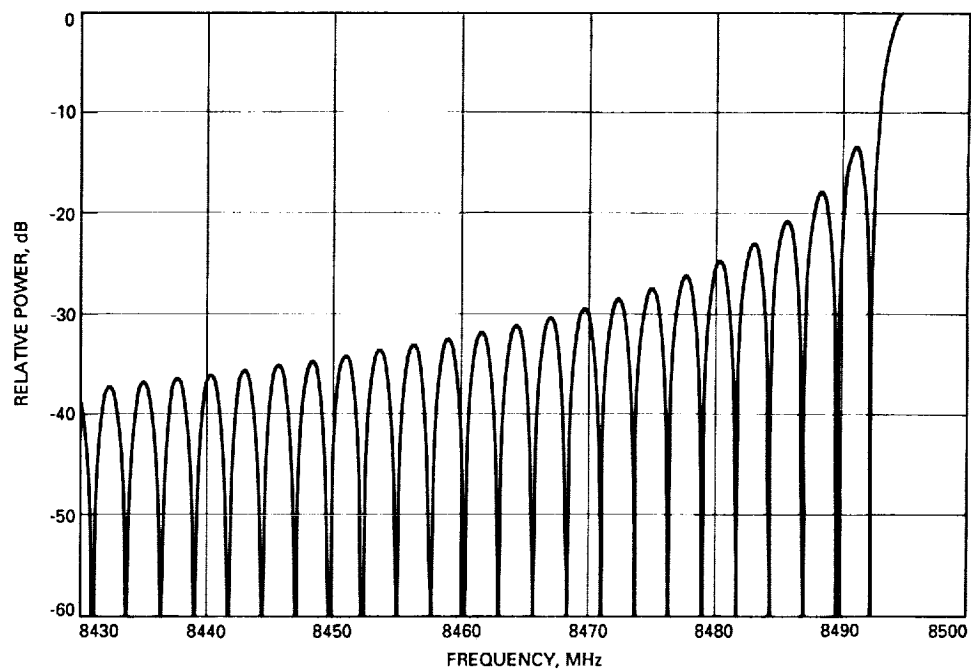


Fig. 7. Theoretical modulated X-band radar spectrum for $p = 31$, $t = 0.375 \mu\text{sec}$.

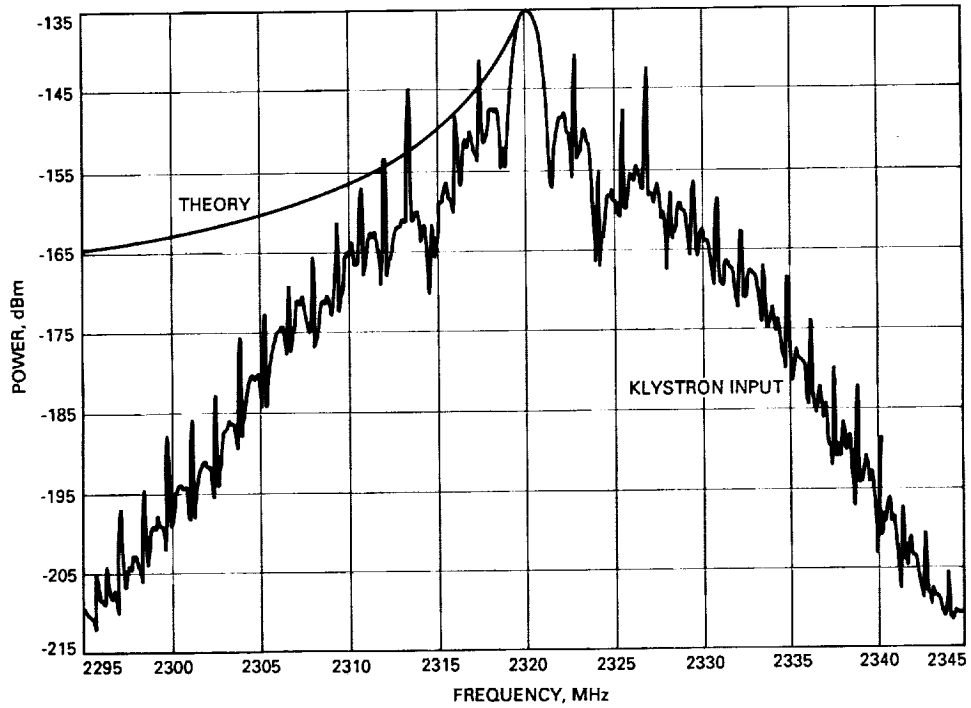


Fig. 8. Test 4: April 9, 1989, S-band exciter-chain output (klystron input) spectrum and theoretical modulated spectrum for $p = 31$, $t = 0.375 \mu\text{sec}$.

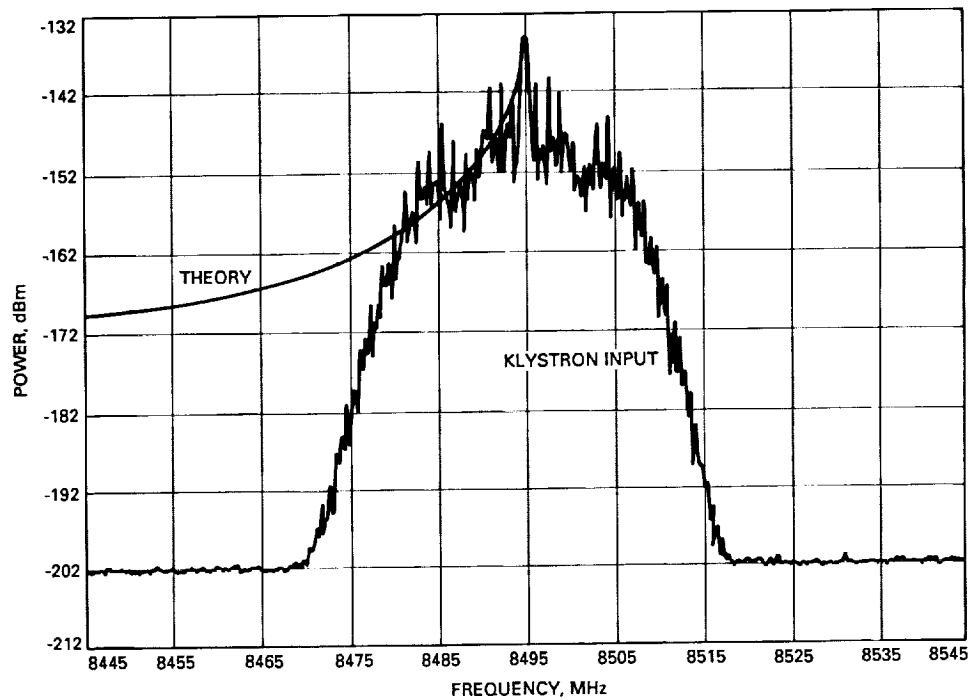


Fig. 9. Test 4: April 9, 1989, X-band exciter-chain output (klystron input) spectrum and theoretical modulated spectrum for $p = 31$, $t = 0.375 \mu\text{sec}$.

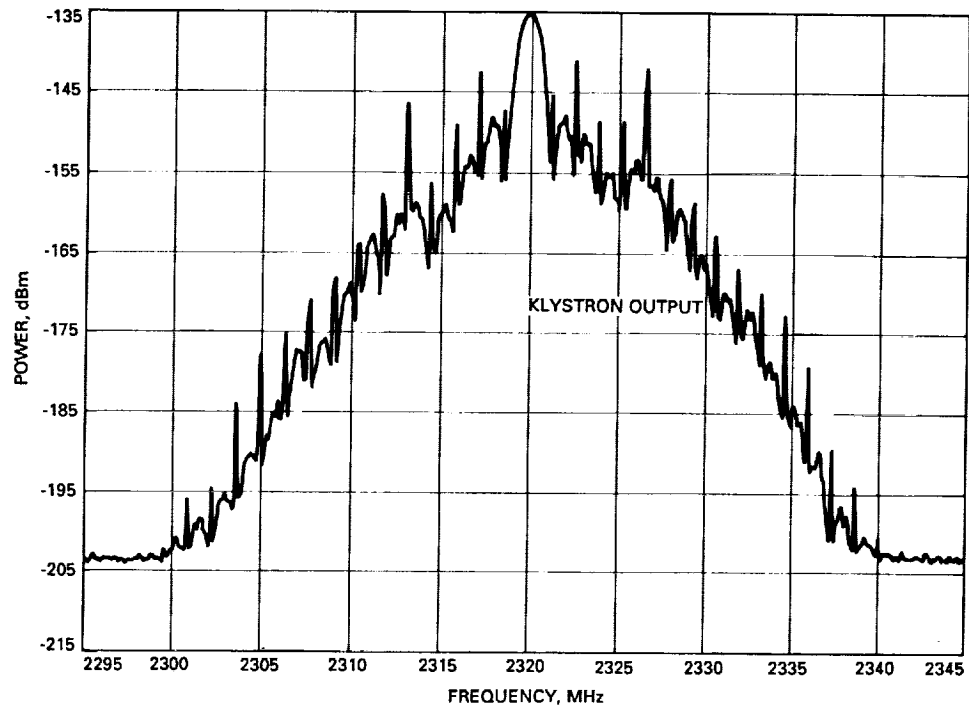


Fig. 10. Test 6: May 22, 1989, S-band klystron output spectrum, $p = 31$, $t = 0.375 \mu\text{sec}$.

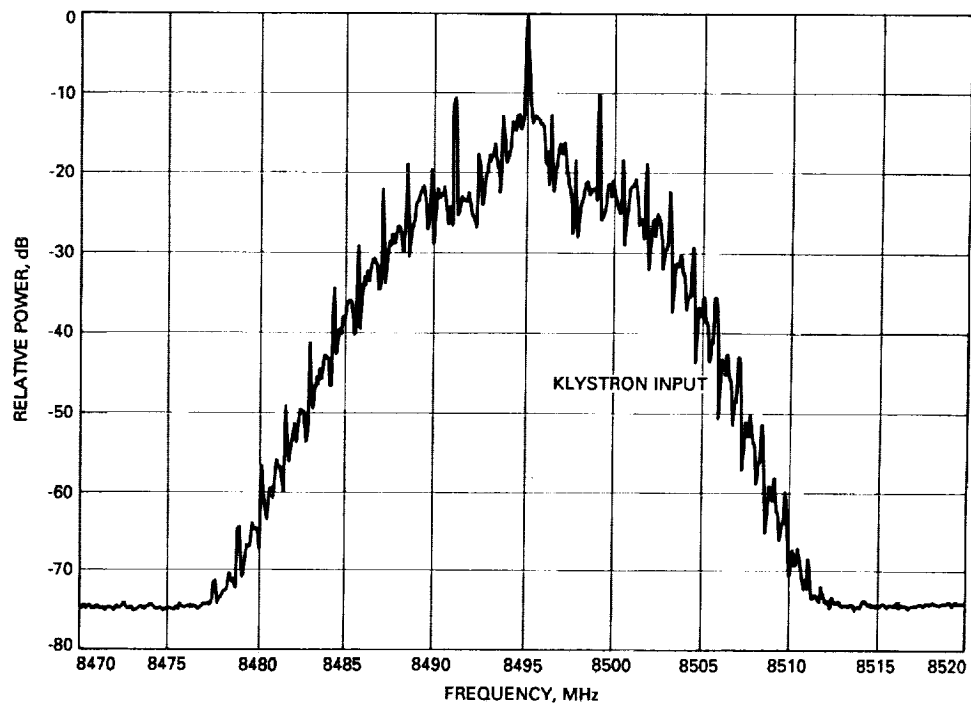


Fig. 11. Test 6: May 22, 1989, X-band klystron input spectrum, $p = 31$, $t = 0.375 \mu\text{sec}$ (compare with Fig. 10).

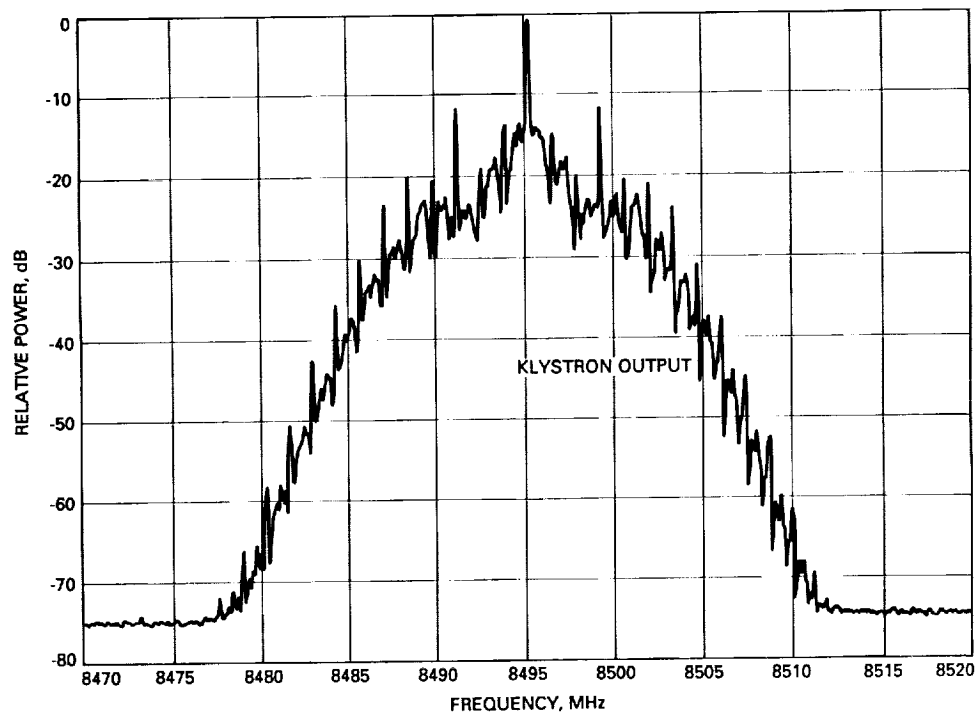


Fig. 12. Test 6: May 22, 1989, X-band klystron output spectrum, $p = 31$, $t = 0.375 \mu\text{sec}$.

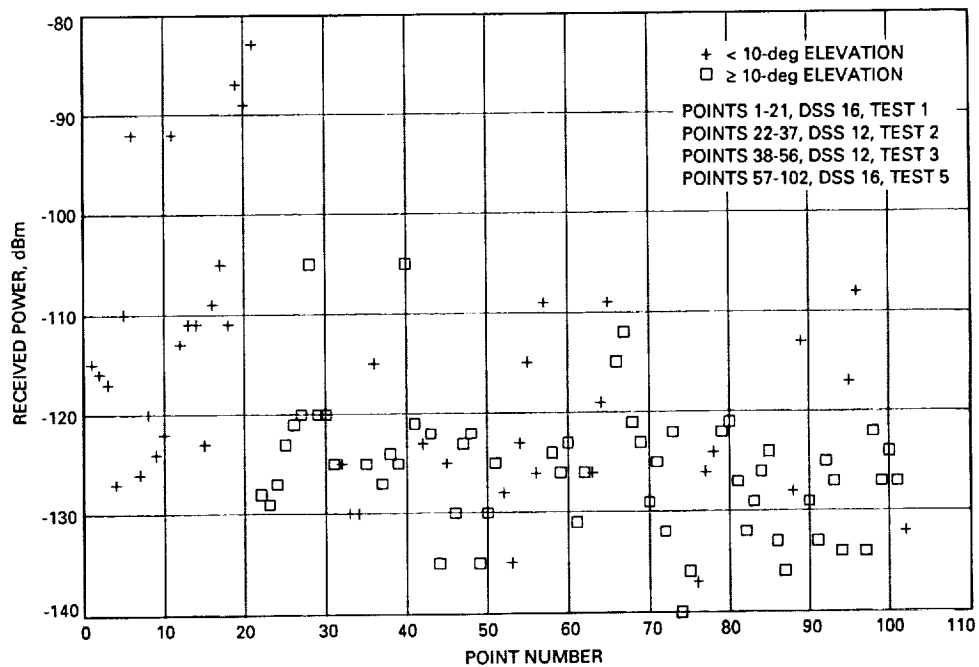


Fig. 13. S-band radar RFI data, all tests.

521-61
272 337

N90-21901

TDA Progress Report 42-100

February 15, 1990

48

Improving a Data-Acquisition Software System With Abstract Data Type Components

S. D. Howard

Communications Systems Research Section

Abstract data types and object-oriented design are active research areas in computer science and software engineering. Much of the interest is aimed at new software development. In this experiment, abstract data type packages developed for a discontinued software project were used to improve a real-time data-acquisition system under maintenance. The result saved effort and contributed to a significant improvement in the performance, maintainability, and reliability of the Goldstone Solar System Radar Data Acquisition System.

I. Introduction

Software components based on abstract data type design and initially developed for a new software system have been used to improve the performance and quality of a data-acquisition software system under maintenance. Although object-oriented design and abstract data types (ADTs) are active research areas in software engineering and computer science, much of the attention has been devoted to the employment of these concepts in the creation

of new software systems; incorporating ADTs into software systems under maintenance has been less widely explored.

ADT packages may be thought of as extensions to high-level programming languages. (A more complete discussion of this idea can be found in a paper by Liskov and Zilles [1].) Most widely used procedural languages provide facilities for declaring and manipulating basic data types. Integers, arrays, characters, and booleans are examples of basic data types available in high-level languages. Each

type has a set of operators that form part of the definition of the type. ADT packages provide additional data types and encapsulate the valid operations on those types. Type packages can be used to capture the domain-specific data abstractions that a general high-level language cannot provide. Integers, arrays, characters, and booleans are quite reusable in high-level languages. Similarly, ADT packages have potential for reusability in related applications.

The abstract data type components used in this experiment were developed for the Goldstone Solar System Radar Data Acquisition System. The architecture for the software system was based on an ADT component library from which specialized data-acquisition systems could be quickly assembled to support established science objectives and new experiments. The software was designed to replace prototype data-acquisition systems that had been used to collect radar returns for the planetary radar astronomy program since the inception of the High Speed Data Acquisition System. Reference 2 gives an overview of the High Speed Data Acquisition System.

The data-acquisition software ADT library was never completed. Work stopped in May 1988. To support observations of Mars during the fall 1988 opposition, effort was redirected toward the modification and enhancement of the Binary Phase Coded Ranging prototype system.

II. Assessment

The successful modification of the ranging prototype for the 1988 Mars opposition did not resolve all of the limitations of the prototype system. The performance of the ranging prototype had always fallen far below the estimate of possible performance. An interprocess communication mechanism based on polling, chosen for ease of implementation in the prototype system, was inadequate for operational use. Conversion to event-driven processing was a likely solution to the performance problem. However, the modification involved replacing parts of the real-time interprocess communication structure; a similar change to the ranging prototype had been attempted in the past without success.

This situation opened the opportunity to consider the inclusion of previously developed ADTs. Reliable components reduce the difficulty of modifications by minimizing errors introduced into the system. Two high-quality ADT-based packages managing event signals and interprocess message passing had been among the first deliveries to the ADT software library. After assessment of the modifications that would be necessary to the structure of the

prototype software to support event-driven real-time processing, it appeared that the new ADT packages could support the changes.

The benefits of using the new ADTs were clear. In addition to encapsulating data types and operators, ADT packages also encapsulate the important software quality attributes of maintainability, robustness, and reliability. The type packages that had been developed for the ADT software library were of high quality. Reliability was recognized in the initial work assessment to be a very important factor in the success of the enhancement to the ranging prototype. Errors in the interprocess communication mechanism of a real-time system can cause the system to transition to an unknown state and fail. The circumstances of these failures can be difficult to trace. The prototype software had not been designed for operational reliability. The maintenance history of the prototype indicated that the system was sensitive to change and prone to side-effect errors that were difficult to trace. Introducing unreliable new work into the interprocess communication mechanism had the potential for degrading the system to the point of unusability.

Successful use of the ADTs could also reduce the amount of time necessary to complete the ranging prototype system upgrade. One software engineer (the author) was available to complete the work. Without the implemented type packages, event signals and interprocess message passing would have to be designed, implemented, and tested for the prototype revision. It is estimated that four weeks of calendar time were required for the development of each ADT. That estimate includes work by two software engineers on specification, specification checking, test plan specification, ADT implementation, test implementation, and unit testing.

The most serious difficulty associated with using ADTs was that much of the ranging prototype had been implemented in FORTRAN and the ADTs were implemented in Pascal. A mechanism for invoking both of the Pascal type packages from FORTRAN had to be established. The greatest risk, of course, was that the incorporation of ADT components into programs already designed and implemented had not been attempted before. Modifications to the prototype that were done to support the 1988 Mars opposition involved replacing two processes in the prototype software with processes that had been rewritten and redesigned to include abstract data types. Designing programs to incorporate ADTs was a relatively well-understood task. Including ADTs in program structures that were not designed to accommodate them was not a familiar task; it was difficult to assess the scope of

changes that might be necessary to support the inclusion of ADTs. Underestimating the scope of modifications could result in the loss of approximately 3 to 4 work-months of engineering time.

III. Implementation

It was possible to gather more information about the feasibility of using the ADTs in the ranging prototype by investigating the parameter-passing mechanisms that would be necessary to call Pascal type packages from FORTRAN. The creation of two FORTRAN test shells quickly demonstrated that both of the candidate type packages could be cleanly invoked from FORTRAN.

The process in the real-time ranging prototype that managed the operator interface had been written in FORTRAN. Polling for interprocess messages in the operator program consumed excessive CPU time. Calls to the ADT procedures were added to the FORTRAN operator program; only minor modifications to the program structure were necessary to convert from polling to event-driven processing using the ADT-based operators.

Because the prototypes and the ADT packages used VAX/VMS operating system facilities, experimentation demonstrated that the new ADT interprocess communication structures were compatible with the older facilities in the prototypes. This was an unexpected advantage that allowed modification and testing of one process in the prototype at a time. Modification of the operator process was commenced in November 1988 and completed in January 1989. The demonstration of the event-driven operator program communicating successfully in the ranging system marked the beginning of the use of ADTs in software for which they had not been originally designed.

IV. Evaluation

The inclusion of the ADTs in the ranging prototype was remarkably free of difficulty. The availability of the components saved development and testing time. Considerable time could have been lost if changes and corrections to the ADTs had been necessary. No changes or corrections have been required. Instead, the time was available to spend on other tasks that needed attention in the prototype modification. Well-crafted and carefully checked ADT components improved the overall system maintainability and reliability.

Software design based on ADTs has interesting consequences for both software development and software main-

tenance. All of the software products that had been developed for the discontinued ADT software library are now potentially usable. Furthermore, this work established the concept of using previously developed ADT components in the maintenance of systems for which they had not been designed.

It is important to note that this approach is not a simple solution to software development problems. ADT-based software design, like many modern software engineering techniques (Fagan inspections, for instance), can cause "front-end loading" in the software development project; there is often a long preliminary phase while the type library is being constructed, during which results are not visible to those outside the software development team. Reliability and reusability are obtained only by careful initial work. Good design decisions in the definition (or specification) of ADT components are not achieved mechanically. Reliability is gained from careful craftsmanship, static program verification, and well-considered test strategies.

Care must be exercised in the generalization of these results. No attempt was made to control this experiment or remove bias; the primary objective was to successfully convert the prototype to event-driven processing. The prototypes and the ADT components were designed by the same software developer. It is difficult to determine how much this might bias the result. Further, this experience was limited to a small number of ADTs. Confidence in continuing to use ADTs in maintenance derives from an assessment that the circumstances under which this work took place were typical.

V. Conclusion

Successful use of ADT-based software components in the prototype system played a critical part in achieving a seven-fold improvement in ranging data throughput. This capacity was used as soon as it was available on radar interferometry observations of Mercury in the spring and fall of 1989. The ADTs remain unchanged since their delivery to the data-acquisition software ADT library in 1987; no errors have been detected in these components.

The satisfactory results of this work have led to further exploration. Recently, a new ADT-based software component was created specifically for adaptive maintenance in the ranging prototype. These experiences with ADT-based components have created another software maintenance option for the improvement of the Goldstone Solar System Radar Data Acquisition System.

Acknowledgment

The abstract data type-based software architecture of the data-acquisition system was the work of Dr. J. L. Robinett. Dr. Robinett and the author collaborated on the production of the two abstract data type packages that were described in this article.

References

- [1] B. H. Liskov and S. N. Zilles, "Programming with Abstract Data Types," *ACM SIGPLAN Notices*, vol. 9, no. 4, pp. 50-59, April 1974.
- [2] L. J. Deutsch, R. F. Jurgens, and S. S. Brokl, "Goldstone R/D High Speed Data Acquisition System," *TDA Progress Report 42-77*, vol. January-March 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 87-96, May 15, 1984.

522-39
272338

118

Structural Fatigue in the 34-Meter HA-Dec Antennas

R. A. Van Hek and B. P. Saldua
Ground Antenna and Facilities Engineering Section

Structural modifications to the 34-m hour-angle-declination antennas, coupled with the use of the antennas beyond their intended lifespans, have led to structural fatigue, as evidenced by damage to the declination drive gear and cracks on the structural members and gussets. An analysis and simulation were made of the main antenna structural members. The analysis showed that the total stress to the antenna structure substantially exceeds the maximum levels recommended by the American Institute of Steel Construction (AISC). Although each of the separate static conditions of stress is only 50 percent of the total stress and does not reach the AISC reduced yield limit, fatigue can and did occur, causing the material to crack in the weakest places.

I. Introduction

Three 26-m hour-angle-declination (HA-dec) antennas, designed for a lifespan of 20 years, were built in the early 1960s for the National Aeronautics and Space Administration's (NASA's) Deep Space Network. After 16 years of operation, the antennas were upgraded to support near-term and planned flight-project missions to outer planets. The upgrade increased the diameter of the antennas from 26 m to 34 m, improved surface dish tolerances, and added X-band receiving capability. The antenna modification was called the 26-Meter Antenna S-X Conversion Project [1].

The design required a structural weight increase of about 50 percent in both the HA and dec structures to achieve the desired improvements. In addition, counterweight was added to balance the rotating structures. The total added counterweight on the dec wheel amounted to 200 percent of the initial weight. The HA wheel also had a substantial weight increase. Because of these weight ad-

ditions to the original design and the tracking mode of the reflector, many stress-reversal conditions occurred in the structure. These stress reversals eventually caused "fatigue" in some areas. These fatigue conditions are discussed in this article.

II. Structural Modifications to the Original Design

Prior to the 34-m conversion, other modifications were made to the antennas that required increases in counterweight on both wheels. An S-band improvement in the mid-1960s consisted of the following changes:

- (1) The prime focus feed was changed to a Cassegrainian feed system.
- (2) A feedcone and subreflector were added to the reflector and apex.
- (3) A dewar hoist was added.

- (4) Surface panels were replaced with stiffer panels.
- (5) Counterweight was added to rebalance the structure.

The S-X Conversion Project, which was initiated in 1978, consisted of the following changes:

- (1) The outer ribs in the reflector were replaced with longer and stiffer ribs to increase the diameter to 34 m.
- (2) The inner structure was reinforced to increase stiffness.
- (3) The quadripod structure was replaced with a truss-type apex and legs to accommodate a more accurate subreflector.
- (4) The surface panels were replaced with S-X compatible panels.
- (5) The subreflector was replaced with a three-axis dual-function tilttable subreflector positioner.
- (6) The existing feedcone was replaced with a dual-frequency cone with dichroic plate and ellipsoid reflector.
- (7) Stiffening members were added to the hour-angle and declination wheel structures as well as to the pedestal frame.
- (8) Counterweight was added to both the hour-angle and the declination wheels.
- (9) The existing hydraulic drive was replaced with a new electric drive system, including drive skids.
- (10) Cable trays were added and new cables were provided to the feedcone and to the subreflector.

Before the S-X conversion modifications were made, an extensive optimization design was made of the structure with the aid of JPL Iterative Design of Antenna Structures (IDEAS) [2]. An extensive study was also performed on the bolted connections. All joints were checked for stiffness and acceptable stress levels. One major result of the study was the decision to stiffen all the bolted connections by welding all the structural joints, changing the original pin-jointed space frame to a rigid frame. Because of this structural joint stiffening, some members are now subject to end moments in addition to the axial forces; these stiffer joints reduce the structural deflections. Subsequently, the antenna movements cause stress reversals as well as reversal of moments in the connections.

The structural stiffness of the various subassemblies could not be maintained with the individual member stiffening due to the added weight. Additional accuracy was obtained by the conversion to a Cassegrainian system and by setting the panels at a specific antenna position.

Although extensive stiffening took place, the addition of all the counterweight caused large deflections in the declination wheel, although this did not affect the pointing accuracy.

After the S-X conversion was completed, regular antenna maintenance was increased, including the periodic inspections.

III. Antenna Structural Problems

The antenna structural problems all originated from the additional counterweights on the antenna, beyond the design-calculated amounts, that were required to balance the antenna structure. These additional weights resulted in increased forces acting on the antenna structure during operation and have led to the problems described below.

A. Gear-Mesh Separations

One of the most serious problems with the antenna is the excessive declination-axis bullgear-to-pinion separation that occurs as the antenna is moved from the east to the west horizon. The gear mesh separates when the reflector points to the horizon because the counterweight bends the rim and gear downward, causing a large deflection of the declination wheel. This reduced gear-mesh engagement has twice sheared off the pinion teeth at the DSS 12 antenna and three times at the DSS 42 and 61 antennas (Figs. 1 and 2). The gear-mesh separation increases during high winds, so the antennas are currently stowed before the wind reaches 45 mph.

B. Hour-Angle Bearing Deterioration

There are two bearings on the hour-angle shaft, the upper and the lower bearings. The lower bearing was designed with plenty of safety margin (of 2.5:1 in fact), while the upper bearing had a design safety margin of only 1.12:1. Because the antenna needed more counterweight than the design indicated, the actual safety margin was reduced to 0.98, and this margin decreased even further to 0.95 in winds of 30 mph. There has been excessive wear and denting in the bearing race for all three antennas. Fortunately, the slow speed of the antenna (approximately two revolutions per day) has made this bearing damage tolerable. However, analysis and observation indicates that the risk of bearing failure is increasing with use.

C. Structural Failures

Until March 1989, no major problems were found on the HA-dec antenna structures except for small weld cracks and occasional sheared and loose bolts, mostly in the secondary members and their connections. In April 1989, additional cracked welds were discovered in more significant areas (e.g., see Fig. 3), but action was postponed until after the planned annual inspection in July 1989. A close watch was kept to see if the cracks were getting worse, and this continued until the annual inspection.

On the day of the annual inspection, May 25, 1989, a major failure was discovered at DSS 12; it must have occurred between April and May 25. The failure occurred in a gusset plate near the declination bearing. There are four such plates, and a crack developed in three of the four plates. The gusset plates connect primary members from the wheel rim to the bearings; these members support the weight of the counterweight and the dec house when the antenna is looking east or west (Figs. 4 and 5). The longest crack was 4 in. long (Figs. 6 and 7), while the two other cracks were only 0.5 in. long. All the cracks looked recent. (Note that the cracks were very similar to the gusset plate crack that caused the NRAO 300-ft antenna to collapse.)

Immediate action was taken to repair the gussets and reinforce the structural connections (Fig. 8). All minor repairs were also made at this time.

IV. Analysis and Results

An analysis of the declination-wheel structural failures was made by the Ground Antenna and Facilities Engineering Section's Structural Group. The IDEAS finite-element model of the HA-dec antenna that was used in the original S-X conversion design was resurrected to find the maximum member forces for a 1.0-*g* loading in the X, Y, and Z axes. To determine whether the critical stress level was reached, member forces from the computer output were postprocessed in an HP 41C programmable calculator to obtain maximum and minimum forces in the structure for any combination of declination and hour angle. For an hour-angle/declination-angle axes-combination configuration, the angle between the gravity vector and the principal axes of the structural model becomes a function of the declination, hour, and latitude angles [3].

Specifications and loading conditions for A-7 steel with an ultimate yield strength of 33,000 psi were used (Fig. 10). Three loading conditions were applied to simu-

late three different positions of the antenna (Fig. 9). The loading conditions that were simulated are the following:

- (1) The reflector points at zenith and the declination wheel is vertical. Gravity loads in the Z direction.
- (2) The reflector points at the east or west horizon and the declination wheel is horizontal. Gravity loads in the X direction.
- (3) The reflector points at the south horizon and the declination wheel is vertical. Gravity loads in the Y direction.

The worst stresses occurred in loading condition (2) because a reversal of loads at the east or west position caused a reversal of stress in the members and their connections, which are the gusset plates. Thus the gusset plates are subject to stress reversals from (1) combined compressive stress plus bending stress due to eccentricity and (2) combined tensile stress plus bending due to eccentricity.

In addition, even if the critical stress levels were never reached, repeated loading and unloading might eventually result in failure. This phenomenon is known as "fatigue." In condition (2), where the loading is reversed from tension to compression and back again with an additional eccentricity in the connection causing a bending stress, fatigue is likely to occur. In fact the cracks in the plates occurred adjacent to the welding, where the parent material was weakened by weld undercuts.

The American Institute of Steel Construction [4] provides fatigue design parameters for most connections. The DSN HA-dec antenna structure falls under Category 1, which applies for a minimum of 20,000 cycles. This is equivalent to two applications per day for 25 years. (The DSN HA-dec antennas are now 27 years old.) The plate or gusset connection conforms to loading Condition 1 of Category E, where the allowable range of stress is 19,250 psi. The actual total stress from cyclic conditions, tension, compression, and bending amounts to 27,206 psi, substantially exceeding the range limit of 19,250 psi. Although each of the separate static conditions of stress is only 50 percent of the total stress and does not reach the reduced yield limit, fatigue can and did occur, causing the material to crack in the weakest places.

V. Conclusions

Estimates were made of the cost and feasibility of extending the lifespans of the HA-dec antennas for another

15 years and increasing their reliability. The result indicated great cost (on the order of \$5M) and a long downtime (6 to 8 months). Even after these improvements, the DSN would still be left with an antenna network limited in use to X/S-band only.

The TDA Office instead recommended that the three HA-dec antennas be replaced with 34-m az-el antennas similar to those in the existing DSN az-el antenna network. These antennas would be equipped with a center-fed beam-waveguide feed system with the capacity for multi-frequency transmitting plus improved performance. These plans have been approved and the antennas are currently

in the final design stages. Implementation will start in FY90, and the three antennas will be operational in mid 1993, 1994, and 1995. Until the new network is operational, increased maintenance and observation will be required to keep the 34-m HA-dec antennas operational.

It is planned to continue to inspect the antennas periodically and make repairs if required. The antennas will not be subjected to any further increase in loads because of additional microwave system modifications during their currently estimated future lifespan, which is planned to be complete when the 34-m beam-waveguide replacement antennas become operational.

References

- [1] *26-Meter Antenna S-X Conversion Project*, JPL Publication 82-26, Jet Propulsion Laboratory, Pasadena, California, March 15, 1982.
- [2] R. Levy, "Optimization of Antenna Structure Design," *Proceedings of the Eighth Conference on Electronics Computation (ASCE)*, Houston, Texas, pp. 114-129, February 1983.
- [3] M. S. Katow and R. Levy, "Computation of Gravity RMS for HA-dec Antennas," *TDA Progress Report 42-27*, vol. March-April 1975, Jet Propulsion Laboratory, Pasadena, California, pp. 139-147, July 15, 1975.
- [4] *Manual of Steel Construction*, Eighth Edition, American Institute of Steel Construction, Inc., Chicago, Illinois, pp. 5-86-5-91, 1980.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

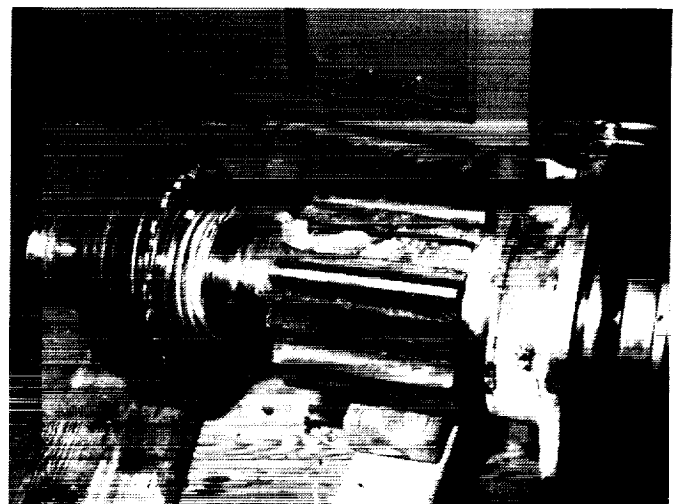
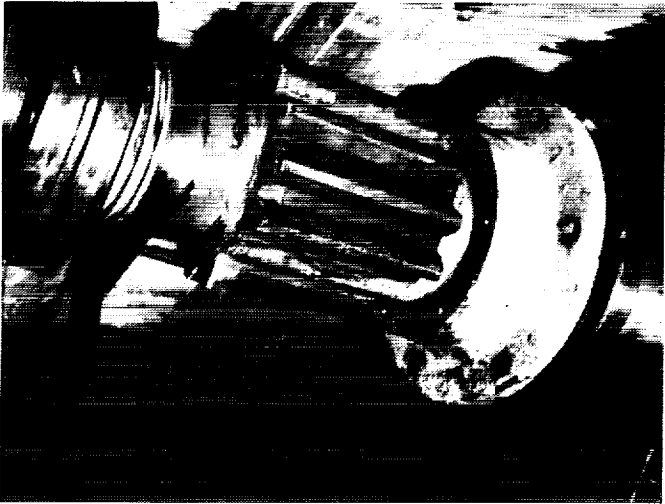


Fig. 1. Broken pinion teeth of declination wheel drives, DSSs 42 and 61.

ORIGINAL PAGE IS
OF POOR QUALITY

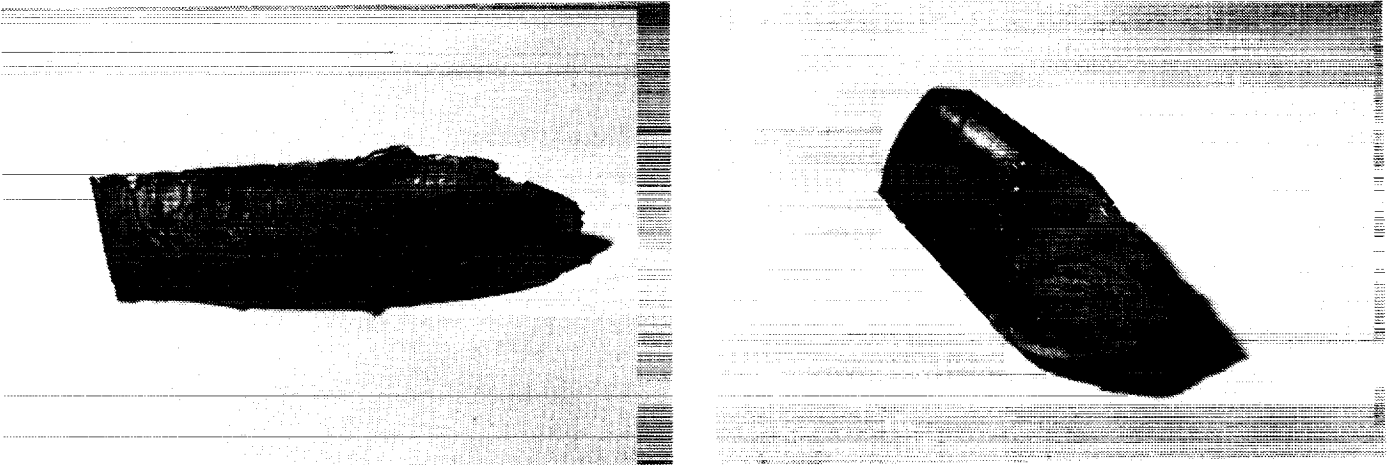


Fig. 2. Pieces of broken pinion teeth from the declination wheel drive, DSS 42.

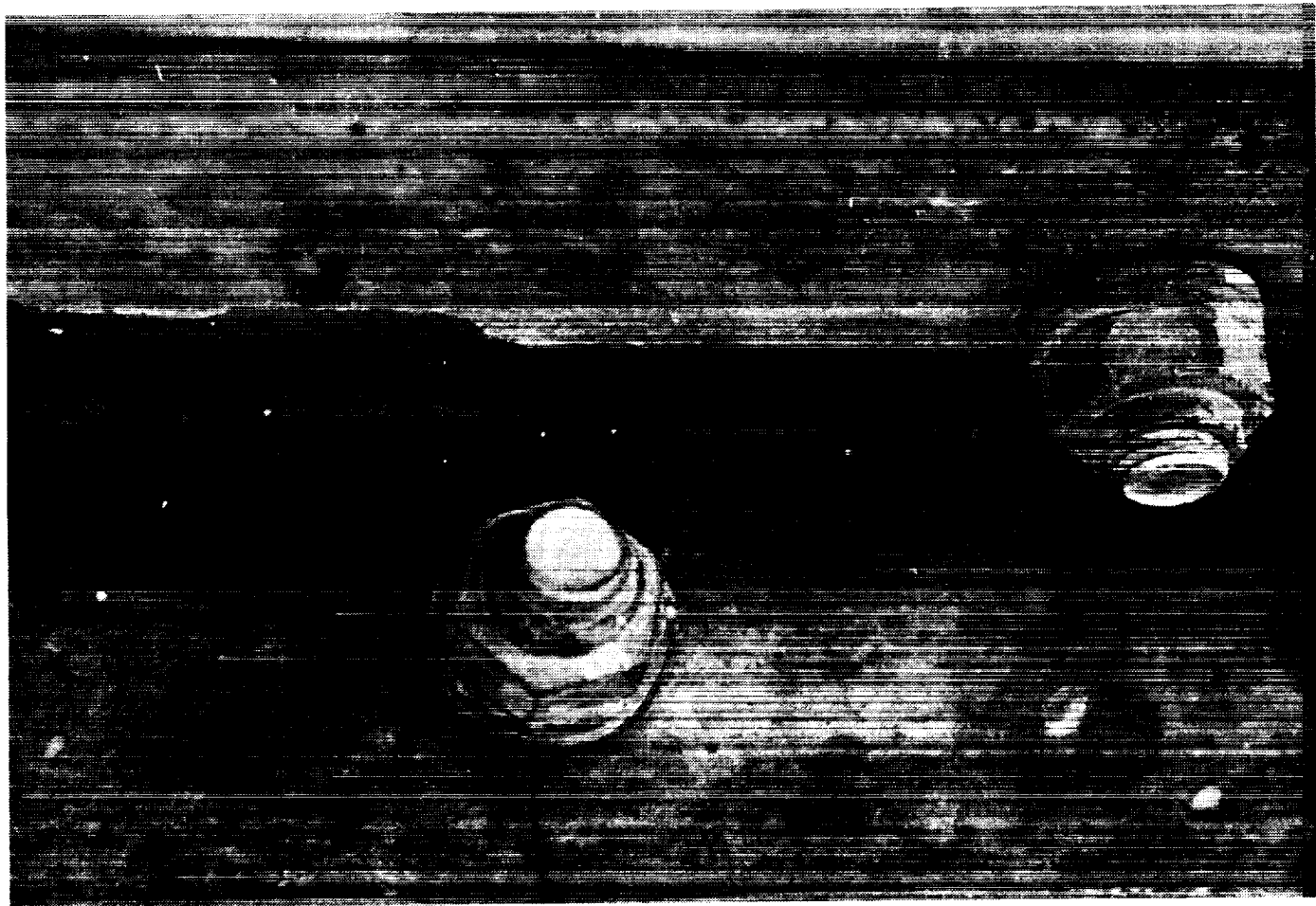


Fig. 3. Typical crack in the polar wheel structural member, DSS 12.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

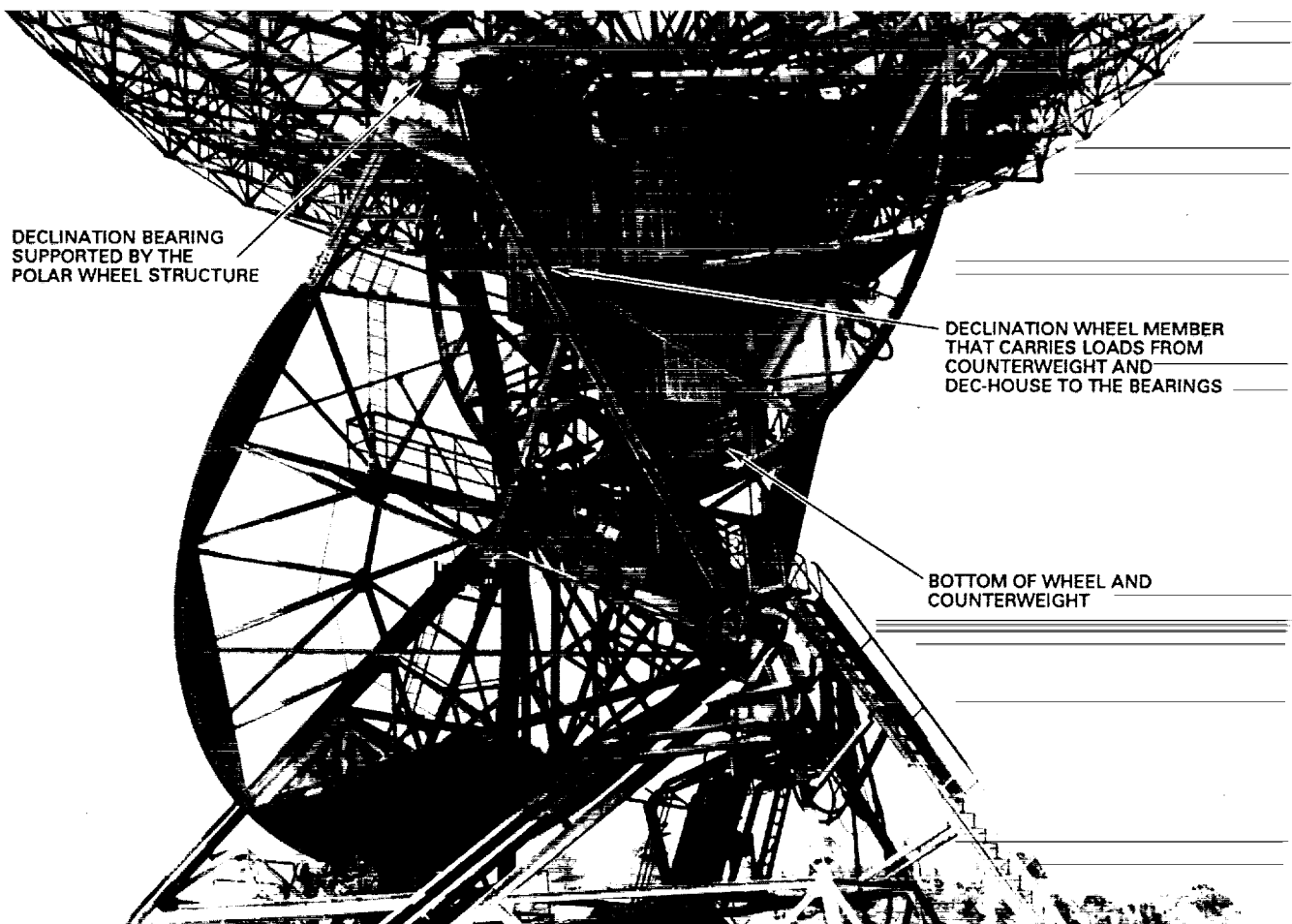


Fig. 4. The 34-m HA-dec structural support structure, DSS 12.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

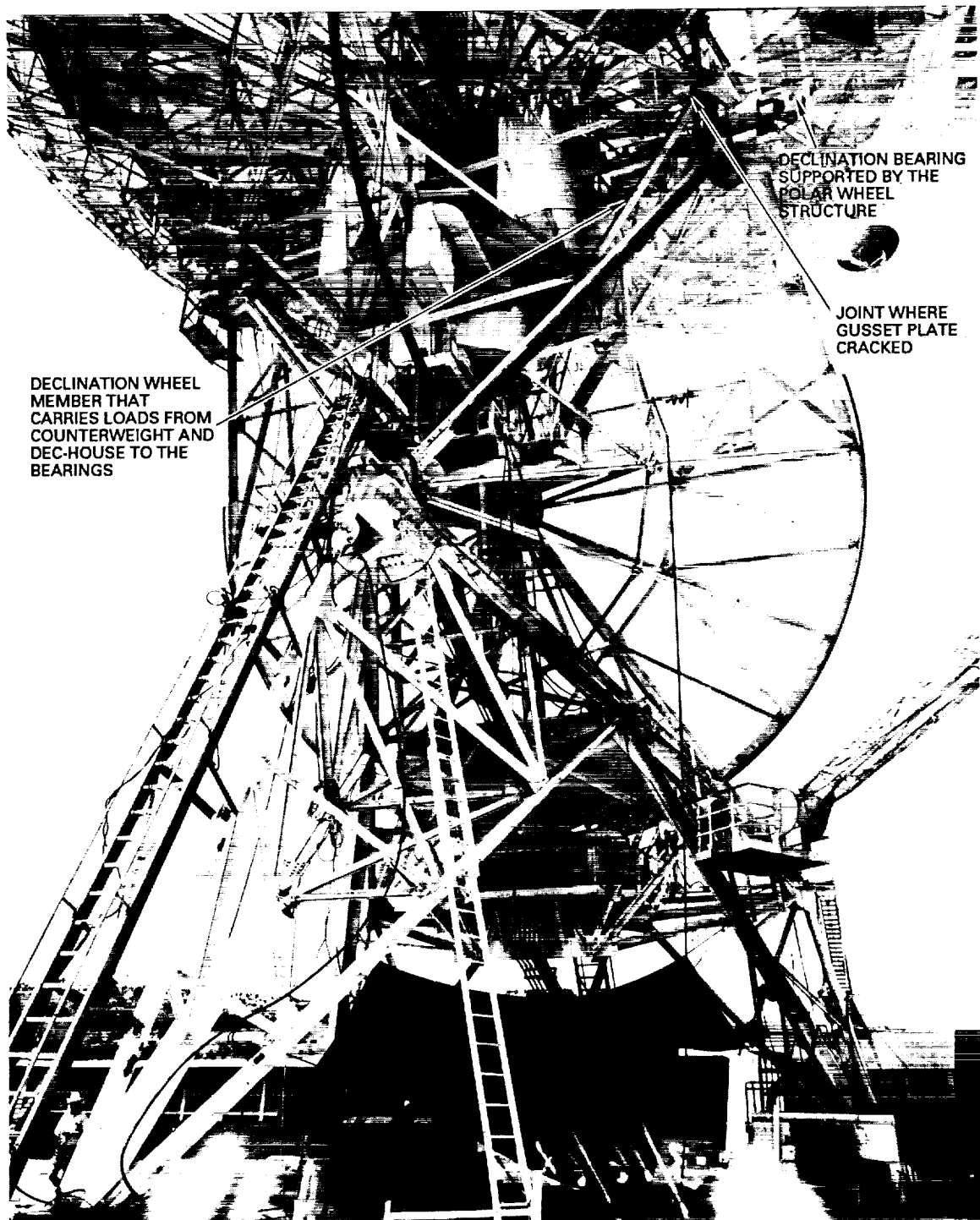


Fig. 5. Hour-angle and declination wheel support structure, DSS 12.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

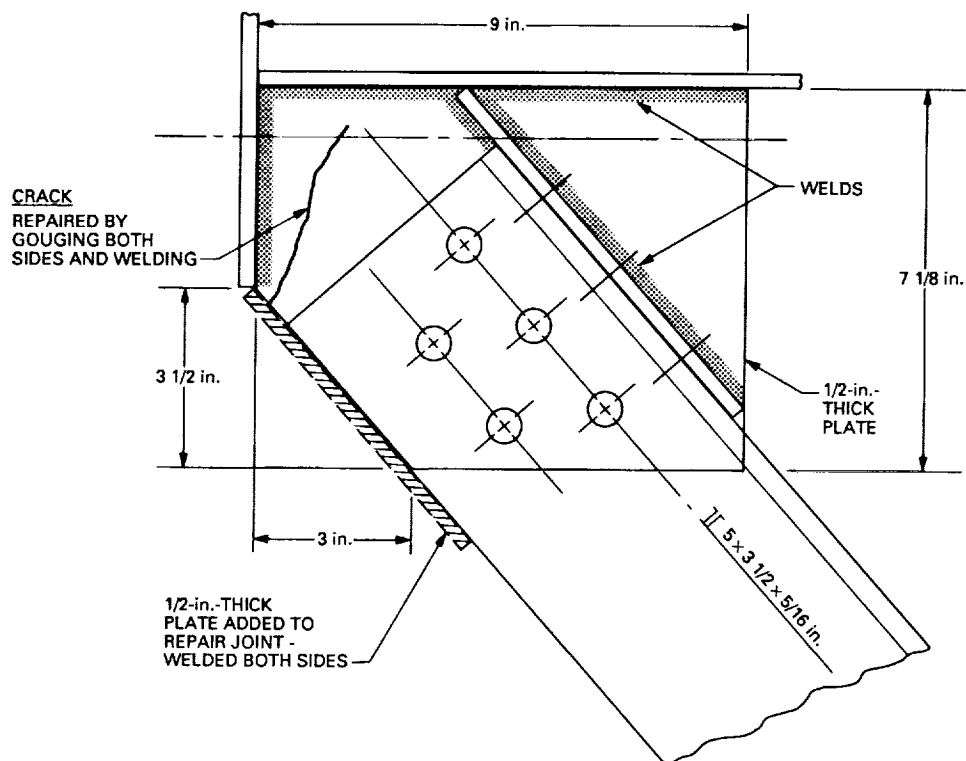


Fig. 6. Failed joint detail drawing, DSS 12.

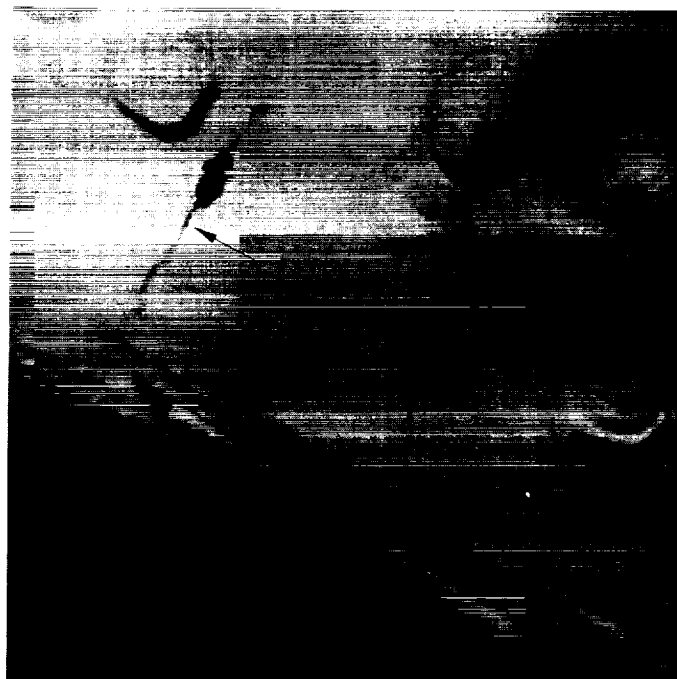


Fig. 7. Cracked gusset, inboard side, DSS 12.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH



Fig. 8. Repair of failed joint, DSS 12.

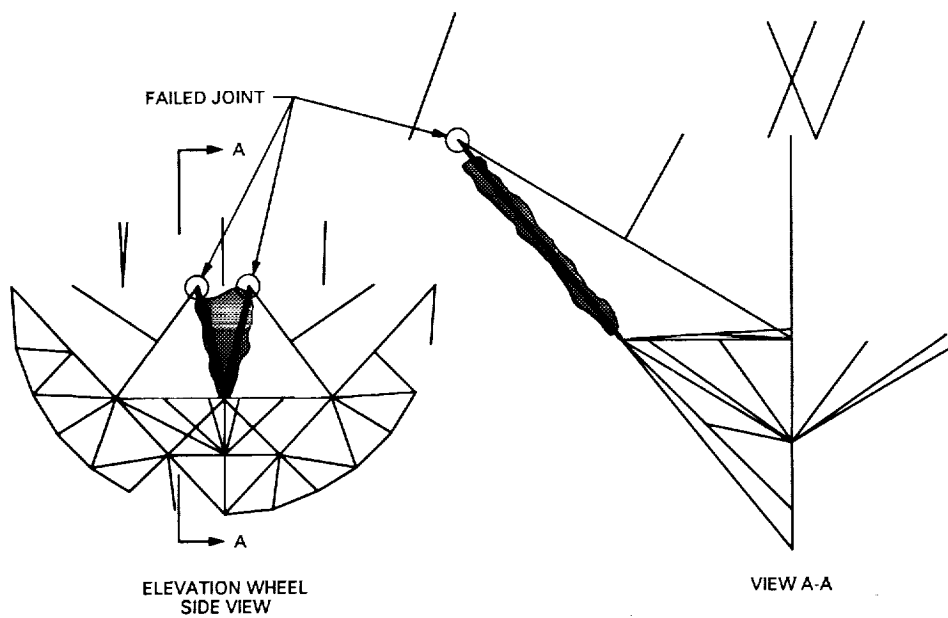


Fig. 9. Computer model of declination wheel.

SECTION B1 LOADING CONDITIONS; TYPE AND LOCATION OF MATERIAL

In the design of members and connections subject to repeated variation of live load stress, consideration shall be given to the number of stress cycles, the expected range of stress, and the type and location of member or detail.

Loading conditions shall be classified as in Table B1.

The type and location of material shall be categorized as in Table B2.

SECTION B2 ALLOWABLE STRESSES

The maximum stress shall not exceed the basic allowable stress provided in Sects. 1.5 and 1.6 of this Specification, and the maximum range of stress shall not exceed that given in Table B3.

Fillet welded connections	Base metal at intermittent fillet welds.	T or Rev.	E	
	Base metal at junction of axially loaded members with fillet welded end connections. Welds shall be disposed about the axis of the member so as to balance weld stresses.	T or Rev.	E	17,18,20
	Weld metal of continuous or intermittent longitudinal or transverse fillet welds.	S	F	5,17,18,21

^a "T" signifies range in tensile stress only; "Rev." signifies a range involving reversal of tensile or compressive stress; "S" signifies range in shear including shear stress reversal.
^b These examples are provided as guidelines and are not intended to exclude other reasonably similar situations.

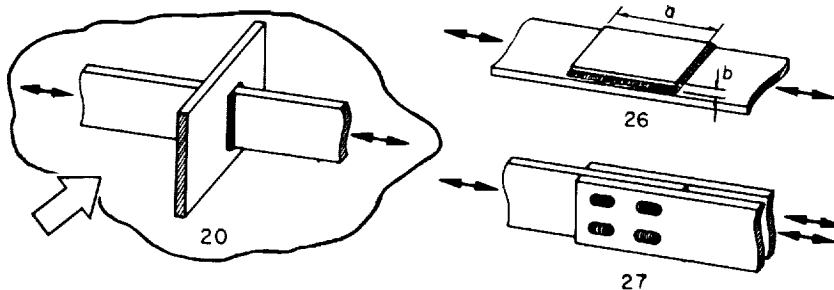


Fig. B1. Illustrative examples

TABLE B3
ALLOWABLE RANGE OF STRESS (F_{sr}), KSI

Category (From Table B2)	Loading Condition 1 F_{sr1}	Loading Condition 2 F_{sr2}	Loading Condition 3 F_{sr3}	Loading Condition 4 F_{sr4}
A	60	36	24	24
B	45	27.5	18	16
C	32	19	13	10 ^a
D	27	16	10	7
E	21	12.5	8	5
F	15	12	9	8

^a Flexural stress range of 12 ksi permitted at toe of stiffener welds on webs or flanges.

NOTE
THIS TABLE IS FOR A-36 STEEL AND SHOULD BE REDUCED FOR A-7 STEEL BY 9 PERCENT TO PRODUCE A VALUE OF 19.25.

Fig. 10. Reference chart excerpted from [4]. (Reprinted by permission of the American Institute of Steel Construction, copyright 1980.)

Mark IV-A DSCC (Magellan-Era) Telemetry System Description

D. L. Ross
TDA Engineering

This article provides an update to the description of the Deep Space Communications Complex portion of the Mark IV-A Telemetry System. This system, described in [1], is currently implemented at all signal processing centers. The upgrade of the telemetry system was undertaken primarily in support of the Voyager-Neptune Encounter and the Magellan mission. The Mark III Telemetry System is the predecessor of the Mark IV-A system and is described in [2].

I. Introduction

Implementation of the Deep Space Network (DSN) Mark IV-A (Magellan-era) Telemetry System was completed in early 1989 in time to support the Magellan launch and Voyager-Neptune Encounter. The DSCC portion of the DSN Telemetry System has been significantly changed in four subsystems as a result of the Magellan-era and Voyager-Neptune Encounter implementation.

A. Antenna Mechanical Subsystem

The 64-m subnet has undergone a major modification consisting of the enlargement of each of the three antennas to 70 m. This was done specifically to support the Voyager-Neptune Encounter and provided a gain increase of 2 dB at 8420 MHz (X-band) through enlargement and surface shape improvement.

B. Receiver-Exciter Subsystem

The Magellan Project required a telemetry system lockup within 60 seconds to maximize data return for an orbital mission that requires the downlink to be acquired

twice during each orbit of 3.15 h for three orbits per pass over a primary mission period of 250 days. To achieve this, the receiver was upgraded to provide automatic acquisition and tracking of the downlink carrier. This required the Deep Space Network to add fast Fourier transform (FFT) automatic signal acquisition hardware and firmware to all receivers. In addition, the Magellan orbit phase produces Doppler offsets which must be removed during tracking to maintain carrier lock and minimize signal-to-noise ratio (SNR) degradation. This required adding a programmable frequency synthesizer to the Block III receiver carrier loop and adding downlink tuning firmware to all receivers (Block III and Block IV).

C. Telemetry Subsystem

Several significant changes were made in the telemetry subsystem to meet Magellan requirements placed on the Deep Space Network.

The baseband assembly (BBA) was upgraded (1) to improve low data rate performance at data rates down to 10 b/s and (2) to improve acquisition time when the

subcarrier frequency is not precisely known. The existing acquisition bandwidth was 0.1 Hz, but the expected uncertainty approached 4 Hz. This was achieved by widening the acquisition bandwidth to 5 Hz or greater, depending on the data rate. A fast Fourier transform (FFT) was added to the BBA to meet the requirement of reducing acquisition time to 36 seconds or less for high data rates. Actual acquisition times have been far better than this, ranging from 12 to 20 seconds, depending on the accuracy of subcarrier predictions.

The maximum-likelihood convolutional decoder (MCD) was upgraded to provide processing of telemetry data at rates (from the previous 250 kb/s) up to 270 kb/s. This was necessary to support the maximum Magellan data rate of 268.8 kb/s. A serial data output (with its associated clock) was added to the MCD to provide an input to the frame synchronizer (described below). The MCD will now accept a node synchronization change command, which provides a quick switch over to the correct node in the event of an MCD false lock.

An upgraded MCD was added to Channel 2 of Telemetry Groups 3 and 4 to support concurrent processing of two coded streams of telemetry data in each group.

A frame synchronizer was added to both channels of groups 3 and 4 to enable frame synchronization of Magellan telemetry data. This is especially important for Magellan orbital data, since the data downlinked at 268.8 kb/s is recorded only at the Signal Processing Centers (SPCs) and then mailed back to JPL for analysis, which may take place as much as two weeks after recording. If the DSN is able to frame synchronize the high-rate data, then it is almost certain that good data are being recorded.

D. Ground Communications Facility (GCF) Data Records Subsystem

New equipment was added to the GCF subsystem to process and record Magellan telemetry data. The designation for this equipment is DSCC Data Records (DDR) assembly, which includes a high-density tape recorder that will record in excess of one hour of data at 268.8 kb/s on a single reel. The DDR formats the telemetry data into Standard Formatted Data Units (SFDUs), time-tags it with Earth-received time, stores it on magnetic tape, and transmits it to JPL in real time (except for 268.8 kb/s data which is stored only on magnetic tape and mailed to JPL).

Implementation of these changes, combined with previous capabilities, has prepared the network to support the following missions:

- (1) Galileo
- (2) Giotto Extended Mission
- (3) International Cometary Explorer (ICE)
- (4) Magellan
- (5) Pioneers 6 through 8
- (6) Pioneers 10 and 11
- (7) Pioneer Venus
- (8) Ulysses
- (9) Voyager Interstellar Mission

II. Key Characteristics

The key characteristics of the DSCC portion of the Mark IV-A Telemetry System are:

- (1) Auto acquisition and tracking of downlink carrier
- (2) Complete telemetry system lockup within one minute at high data rates
- (3) Process coded data at rates up to 270 kb/s
- (4) Baseband combining for up to six antennas
- (5) Four complete groups of telemetry equipment at each complex, each with the capacity to support any one of the above missions except for Magellan, which can only be supported by two of the four groups
- (6) Demodulation of Manchester coded (Bi ϕ -L) or NRZ-L or NRZ-M data
- (7) Symbol synchronization of demodulated data
- (8) Maximum-likelihood decoding of short-constraint-length convolutional codes and sequential decoding of long-constraint-length convolutional codes
- (9) Frame synchronization
- (10) Precise measurement of received signal level and system noise temperature
- (11) Centralized control by (and real-time reporting to) the monitor and control subsystem
- (12) Production of digital telemetry original data record (ODR) for each telemetry group with playback via local manual control; reduced playback rate for higher data rates as required

Automatic acquisition of the carrier, telemetry system lockup in one minute, and handling of increased data

rates have been implemented to meet Magellan requirements. New equipment providing frame-synchronization, formatting, and storage of high data rates in real time also meet Magellan Project requirements.

III. Project Data Rate and Coding Requirements

Requirements for new and existing deep space missions are listed in Table 1. The new missions to be supported during the Magellan era are Galileo, Giotto Extended Mission, Magellan, and Ulysses. The next section shows how deep-space mission requirements will be met.

IV. DSCC Conceptual Description

The DSCC block diagram in Fig. 1 provides a conceptual description of the portion of the Mark IV-A Telemetry System to be located at each DSCC. There will be one 70-m antenna and two 34-m antennas. The 70-m and 34-m standard antennas will be able to receive an S-band (2200–2300 MHz) plus an X-band (8400–8440 MHz) carrier simultaneously. Each of the 34-m high-efficiency antennas will receive one X-band (8400–8440 MHz) carrier or one S-band (2270–2300 MHz) carrier; however, the S-band (2270–2300 MHz) system temperature is higher than the 34-m standard antennas (see Table 2).

Table 2 gives the radio frequency reception characteristics for these antennas and indicates the distribution of masers, high-electron-mobility transistors (HEMTs), and field-effect transistors (FETs). The 70-m antennas will be equipped exclusively with masers whereas the 34-m antennas will also have HEMTs and FETs. At the 34-m standard antennas, the masers will provide deep-space support while the FETs support near-Earth missions with their broader (2200–2300 MHz) reception bandwidth requirements. New HEMTs have been installed at the 34-m high-efficiency antennas to support deep-space missions as a backup to the masers. New Block II-A X-band (8400–8440 MHz) masers have been provided for all 70-m antennas and 34-m high-efficiency antennas giving lower system temperatures than in the Mark III DSN. Also, the 34-m high-efficiency antennas provide an increase in X-band (8400–8440 MHz) gain over that of the existing 34-m standard antennas.

Also shown in Fig. 1 is a block diagram representation of the receiver-exciter and telemetry subsystems. Modified Block III and Block IV receivers will be used to receive and detect baseband signals. The Block III

and Block IV receivers have been modified to provide automatic acquisition and tracking of the downlink carrier leading to a 60-second overall lockup in conjunction with the telemetry subsystem. The telemetry subsystem is arranged to provide four telemetry groups, any of which can process data from any receiver and therefore from either a near-Earth mission or deep-space spacecraft. All groups will include one or two maximum-likelihood convolutional decoders (MCDs) and one telemetry processor assembly (TPA). The MCDs have been modified to increase the data processing rate to 270 kb/s and to provide a serial output to the frame synchronizers. Telemetry Groups 3 and 4 are equipped with a modified BBA, which will include the functions of baseband combining, subcarrier demodulation and symbol synchronization. Upgrades to the BBA included widening the acquisition bandwidth and reducing acquisition time.

Figure 2 is a functional block diagram of the BBA. Any combination of receiver outputs can be input to either subcarrier demodulator; or any single receiver output can be routed to either subcarrier demodulator, bypassing the combining function. The monitor and control function is performed from the TPA with no manual intervention required. The BBA is designed to accommodate NRZ-L, NRZ-M, or Bi ϕ -L symbol formats, subcarriers up to 2 MHz, and data rates from 4 s/s to 1 Ms/s with subcarrier, up to 4 Ms/s (NRZ) without subcarrier or up to 2 Ms/s for Bi ϕ -L. When the BBA is operated in the mode which combines the 70-m and two 34-m antenna basebands, a nominal system signal degradation of about 0.3 dB will result at the highest data rates. This includes an allowance for waveform, spectrum correlation and symbol timing losses and represents an improvement over the Mark III system. The soft symbol output from the BBA is passed to an MCD for decoding. The decoded bits from the MCD are then passed to one of two places:

- (1) a frame synchronizer where the data is synchronized to the transfer frame level and sent, frame-by-frame, to the DDR for formatting, or
- (2) the TPA where it is asynchronously formatted for transmission to JPL.

Telemetry Groups 1 and 2, as shown in Fig. 1, both include subcarrier demodulator assemblies (SDAs) and symbol-synchronizer assemblies (SSAs) as well as an MCD and a TPA. Except for the MCDs, which have been modified as indicated above, the hardware and configuration in Groups 1 and 2 remains unchanged for the Magellan era.

The 70-m and 34-m antennas can be arrayed by combining baseband signals and performing subcarrier demod-

ulation and symbol synchronization in the BBA in either Telemetry Group 3 or Telemetry Group 4. The combined signal is then decoded in the maximum-likelihood convolutional decoder. Subsequently it is either formatted for transmission to JPL in the telemetry processor assembly or sent to the frame synchronizer and then to the DSCC data records assembly (DDR) for formatting and transmission to JPL. When combining is not required, outputs from an antenna may also be routed to a subcarrier demodulator assembly (Groups 1 and 2) or to either BBA (Groups 3 and 4).

Any of the telemetry equipment groups can accept two data streams. In any group, one data stream is processed by channel 1 and one by channel 2. The performance parameters for channels 1 and 2 are listed in Tables 3 and 4, respectively. Comparing Table 1 with Table 3 and Table 4, it may be noted that Data Stream 1 in Table 1 is processed by channel 1, while Data Stream 2 is processed in channel 2. Also note that Groups 3 and

4 provide higher data rate capability (for example, for convolutionally coded data: 270 kb/s versus 135 kb/s) and higher subcarrier frequency capability (2 MHz versus 1 MHz). Note further that a second MCD has been added to Groups 3 and 4. This MCD was added in channel 2 during the Magellan-era upgrade and provides the capability of concurrently processing two convolutionally coded data streams (a Magellan requirement).

Future upgrades to the telemetry system include new receivers with wider bandwidths, new TPAs, MCD replacement, addition of a Reed-Solomon decoder and provision for high-rate recording. Implementation of this new equipment will enable processing of telemetry data at rates up to 2.2 Mb/s for a single stream or an aggregate of 3 Mb/s in each telemetry group. A fifth telemetry group will also be added, thus enabling the integration of the 26-m subnet into the SPCs for increased operational effectiveness. Future planned missions will require these additional improvements.

References

- [1] D. L. Ross, "Mark IV-A DSCC Telemetry System Description," *TDA Progress Report 42-83*, vol. July-September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 92-100, November 15, 1985.
- [2] E. C. Gatz, "DSN Telemetry System Mark III-77," *DSN Progress Report 42-49*, Jet Propulsion Laboratory, Pasadena, California, pp. 4-10, February 15, 1979.

Table 1. Single link requirements for deep-space missions

Mission	Data Stream 1	Data Stream 2
Galileo	Convolutionally coded; K=7, R=1/2; NRZ-L; up to 134.4 kb/s; Subcarrier: 360 kHz; Carrier: 8415/8420.4 MHz or Convolutionally coded; K=15, R=1/4; NRZ-L; 115.2 or 134.4 kb/s; Subcarrier: 720 kHz; Carrier: 8415/8420.4 MHz or Convolutionally coded; K=7, R=1/2; NRZ-L; up to 40 kb/s; Subcarrier: 22.5 kHz for data rates lower than 7.68 kb/s, 360 kHz for rates \geq 7.68 kb/s; Carrier: 2295/2296.5 MHz	Uncoded; NRZ-L; 40 b/s; Subcarrier: 22.5 kHz; Carrier: 2295/2296.5 MHz
Giotto Extended Mission	Convolutionally coded; NRZ-L; K=7, R=1/2; Subcarrier: 46.080 kHz for data rate of 360 b/s, 276.480 kHz for data rates of 5,760 b/s, 23,040 b/s and 46,080 b/s; Carrier: 2298.703704 or 8428.580248 MHz or Reed-Solomon* and convolutionally coded; NRZ-L; K=7, R=1/2; Subcarrier: 276.480 kHz for data rates of 5,760 b/s, 23,040 b/s, and 46,080; Carrier: 2298.703704 or 8428.580248 MHz	Uncoded; NRZ-L; Subcarrier: 46.080 kHz for data rate of 360 b/s; Carrier: 2298.70374 or 8428.580248 MHz or Reed-Solomon* coded; NRZ-L; Subcarrier: 276.480 kHz for data rates of 5,760 b/s, 23,040 b/s, and 46,080 b/s; Carrier: 2298.703704 or 8428.580248 MHz
ICE	Convolutionally coded; K=24, R=1/2; NRZ-L; 64 b/s; Subcarrier: 1024 Hz; Carrier: 2270.4 and 2217.5 MHz or Convolutionally coded; K=24, R=1/2; 512 to 2048 b/s; Bi ϕ -L; Carrier: 2270.4 and 2217.5 MHz	
Magellan	Convolutionally coded; NRZ-L; K=7, R=1/2; 115.2 or 268.8 kb/s; Subcarrier: 360 or 960 kHz for 115.2 kb/s data rate, 960 kHz for 268.8 kb/s; Carrier: 8425.8 MHz	Convolutionally coded; NRZ-L; K=7, R=1/2; 40 or 1200 b/s; Subcarrier: 22.5 kHz; Carrier: 2297.9 or 8425.8 MHz
Pioneers 6-8	Uncoded; NRZ-M; 8 to 512 b/s; Subcarrier: 512 Hz for 8 to 64 b/s, 2048 Hz for 128 b/s; Carrier: 2292.407407 MHz	
Pioneers 10/11	Uncoded; NRZ-L; 8 to 2048 b/s; Subcarrier: 32.768 kHz; Carrier: 2292.407407 MHz or Convolutionally coded; K=32, R=1/2; NRZ-L; 8 to 2048 b/s; Subcarrier: 32.768 kHz; Carrier: 2292.407407 MHz	
Pioneer Venus	Uncoded; NRZ-L; 8 to 4096 b/s; Subcarrier: 16 kHz; Carrier: 2293.888 or 2294.259 MHz or Convolutionally coded; K=32, R=1/2; NRZ-L; 8 to 2048 b/s; Subcarrier: 16 kHz; Carrier: 2293.888 or 2294.259 MHz	
Ulysses	Convolutionally coded; NRZ-L; K=7, R=1/2; 128 b/s to 8,192 b/s; Subcarrier: 65,536 Hz for rates up to 1,024 b/s, 131,072 Hz for rates of 2,048 b/s or greater; Carrier: 2293.148148 or 8408.209876 MHz	
Voyager	Convolutionally coded; K=7, R=1/2; NRZ-L; 10 to 115,000 b/s; Subcarrier: 360 kHz Carrier: 8420.432097 MHz	Uncoded; NRZ-L; 46.667 b/s; Subcarrier: 360 kHz; Carrier: 8420.432097 MHz or Uncoded; NRZ-L; 46.667 b/s; Subcarrier: 22.5 kHz; Carrier: 2296.481481 MHz

*Reed-Solomon data are not presently decoded at the DSN; only convolutionally encoded data are decoded at the DSN.

Table 2. Radio frequency reception characteristics

Parameter	Antenna		
	70 meter	34-meter standard	34-meter high efficiency
Frequency range, MHz			
S-band	2270-2300	2200-2300	2270-2300
X-band	8400-8440	8400-8440	8400-8440
Gain, dBi, non-diplexed mode			
S-band	63.34 ± 0.01	56.1 ^{+0.3} _{-0.7}	56.0 ± 0.2
X-band	74.17 ^{+0.2} _{-0.2}	66.2 ± 0.6	68.3 ± 0.2
System noise temp., K			
zenith, non-diplexed mode			
S-band with maser	18.5 ± 2	21.5 ± 2.5	
S-band with FET		110 ± 10	55 ± 5
S-band with HEMT			(DSS 45/65) 36 ± 3
X-band with maser	20.6 ± 2	21.5 ± 2.5	(DSS 15) 19.7 ± 2.0
X-band with HEMT			36 ± 6
The above values are representative only. For a complete set of specifications, see DSN Flight Project Interface Design Handbook, Document 810-5, Rev. D, Volumes 1 and 2.			

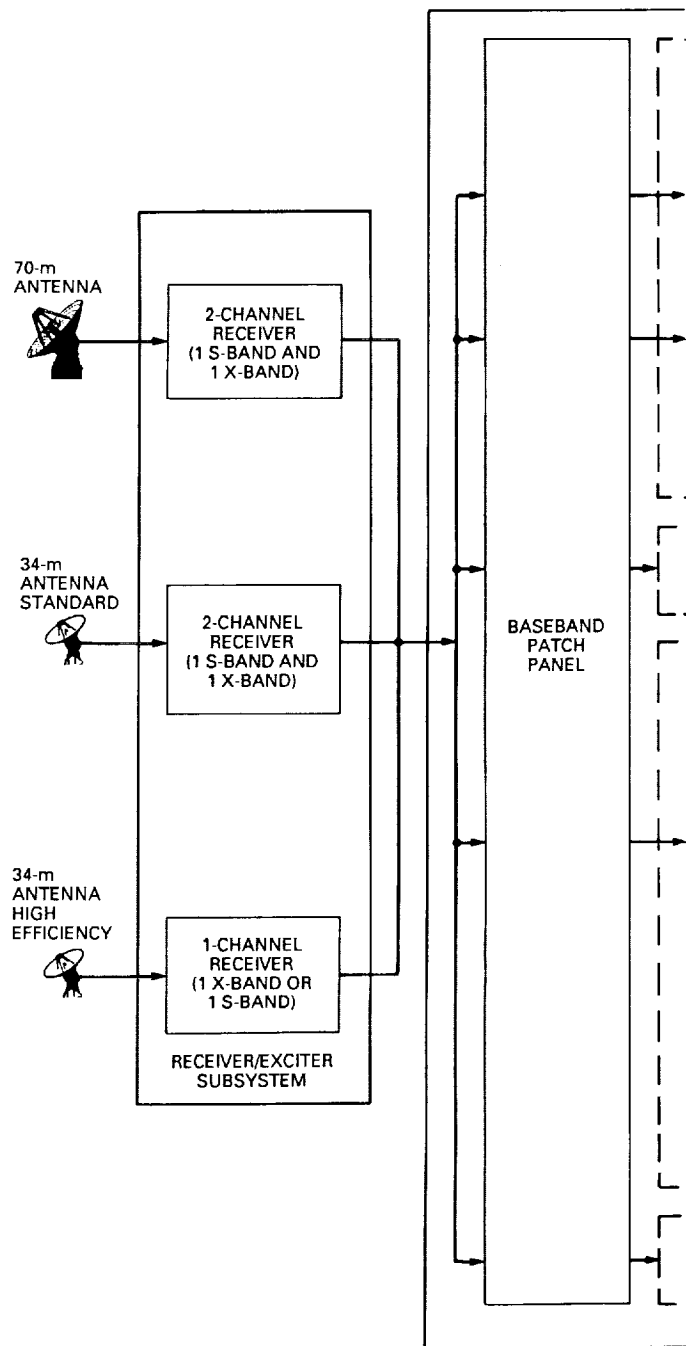
Table 3. DSCC telemetry s/s channel capabilities (telemetry groups 1 and 2)

Functions	Channel 1	Channel 2
Baseband Combining	N/A	N/A
Subcarrier Demodulation	100 Hz to 1 MHz, square wave or sine wave	100 Hz to 1 MHz square wave or sine wave
Symbol Synchronization	6 s/s to 268.8 ks/s	6 s/s to 268.8 ks/s
Data Format	NRZ-L, NRZ-M, Biφ-L	NRZ-L, NRZ-M, Biφ-L
Sequential Decoding	K=24 or 32; R=1/2; frame length selectable; 6 s/s to 20 ks/s	N/A
Maximum-Likelihood	K=7; R=1/2 or 1/3;	N/A
Convolutional Decoding	10 b/s to 134.4 kb/s	
Uncoded	6 b/s to 268.8 kb/s*	6 b/s to 268.8 kb/s*
*Record only with non-real-time playback above 250 kb/s.		

Table 4. DSCC telemetry s/s channel capabilities (telemetry groups 3 and 4)

Functions	Channel 1	Channel 2
Baseband Combining	Up to six basebands	N/A
Subcarrier Demodulation	10 kHz to 2 MHz, square wave or sine wave	10 kHz to 2 MHz square wave or sine wave
Symbol Synchronization	4 s/s to 4 Ms/s	4 s/s to 4 Ms/s
Data Format	NRZ-L, NRZ-M, Bi ϕ -L	NRZ-L, NRZ-M, Bi ϕ -L
Sequential Decoding	K=24 or 32; R=1/2; frame length selectable; 6 s/s to 20 ks/s	N/A
Maximum-Likelihood Convolutional Decoding	K=7; R=1/2 or 1/3; 10 b/s to 134.4 kb/s*	N/A
Frame Synchronization (Transfer Frame Level)	Up to 1.6 Mb/s	Up to 1.6 Mb/s
Uncoded	4 b/s to 500 kb/s*	4 b/s to 500 kb/s*
*Record only with non-real-time playback above 250 kb/s.		

FOLDOUT FRAME /



FOLDOUT FRAME 2 .

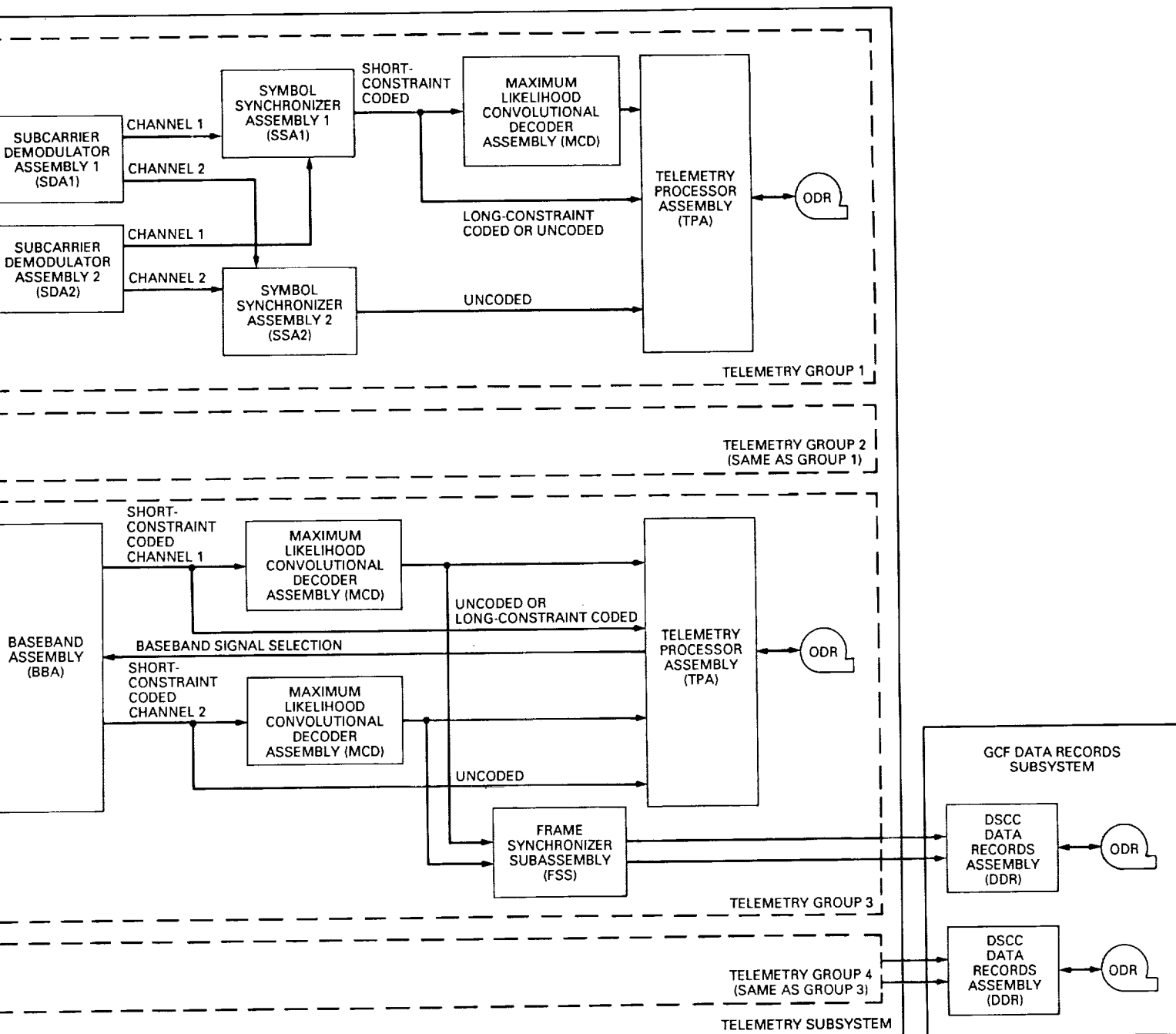


Fig. 1. Telemetry System: DSCC block diagram.

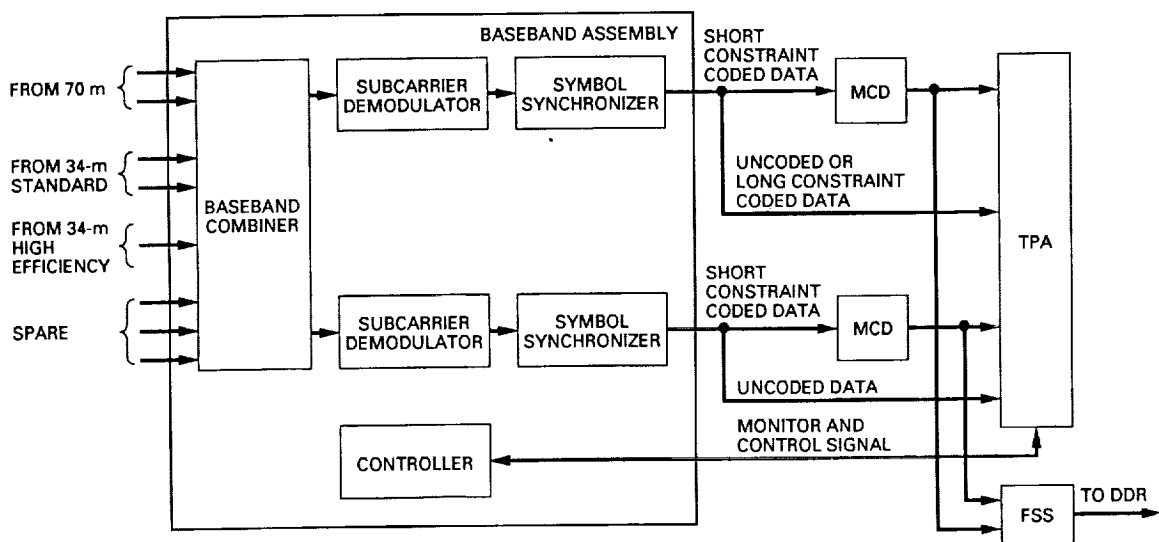


Fig. 2. Baseband assembly functional block diagram.

534-32
272340

A Higher Density VLBI Catalog for Navigating Magellan and Galileo

J. S. Ulvestad, O. J. Sovers, and C. S. Jacobs
Tracking Systems and Applications Section

The density of radio sources near the ecliptic in the astrometric JPL Very Long Baseline Interferometry (VLBI) catalog has been increased by over 50 percent since 1985. This density increase has been driven by the need for more sources for the VLBI navigation of the Magellan and Galileo spacecraft, but the sources also will be usable for Mars Observer and other future missions. Since the last catalog, including observations made through 1985, was published in 1988, a total of 21 radio sources has been added that fulfill the following criteria: (1) they lie within 10 deg of the ecliptic plane; (2) their correlated flux densities are above 0.2 Jy on at least one of the Deep Space Network intercontinental baselines at both 2.3 and 8.4 GHz; and (3) the source positions are known to better than 5 milliarcseconds (25 nanoradians). The density of such sources in the catalog has been increased from 15.6 per steradian to 25.2 per steradian. Ten more sources have been added that fulfill the last two criteria given above and lie between 10 deg and 20 deg from the ecliptic plane.

Analysis shows that there may be ~70 more sources with correlated flux densities above 0.2 Jy that are within ~20 deg of the ecliptic. However, VLBI navigation observations of the new and prospective sources with the 250-kHz bandwidth of the current operational system will require the use of two 70-m antennas in most cases. Including both old and new sources, if two 34-m antennas are used, there will be usable navigation sources within 10 deg of a spacecraft in only 30 percent of the ecliptic, and sources within 20 deg of a spacecraft over 70 percent of the ecliptic. If one 70-m antenna were used along with a 34-m telescope, or if a somewhat wider bandwidth system such as the Mark II VLBI system (effective bandwidth of 1.8 MHz) were used with two 34-m antennas, usable navigation sources would be within 10 deg of a spacecraft over about 65 percent of the ecliptic. Within 20 deg of a spacecraft, usable sources would exist over 98 percent of the ecliptic.

I. Introduction

The technique of Very Long Baseline Interferometry (VLBI) is used to produce astrometric catalogs of the positions of compact extragalactic radio sources (generically

referred to here as quasars). The catalogs are produced by making radio interferometric observations of compact radio sources and performing multiparameter fits, one output of which is a catalog of radio-source positions in a specified reference frame (e.g., [1,2]). Such a catalog can be

used for a variety of purposes, including monitoring of the Earth's orientation, measurement of plate tectonic motion, and navigation of interplanetary spacecraft. This article concentrates on the last use. Specifically, it addresses the problem of increasing the density of objects in the JPL VLBI catalog in regions of the sky that are important for planetary missions.

The general intent of the work described here has been to increase the density of sources in the catalog near the ecliptic, a great circle on the sky which approximates the path followed by all the planets in the solar system. In a technique called Δ VLBI that is used for spacecraft navigation, successive observations are made of the spacecraft and a quasar, with observables differenced to reduce errors. For navigation of a spacecraft using Δ VLBI, it is desirable to have at least one reference radio source near the spacecraft position at all times. Several different errors grow with increasing spacecraft/radio-source angular separation, so the smallest possible separations are desired. The maximum acceptable separation depends on the error budget and on the VLBI observable used. Errors grow rapidly with separation for interferometric delay measurements, but delay-rate measurement errors at the present time are dominated by wet troposphere fluctuations, which do not depend very strongly on spacecraft/quasar separation. A traditional goal has been to always have a reference source within 10 deg of the spacecraft being navigated.¹ A more accurate method of VLBI navigation is being developed [3]; it relies upon having several radio sources in approximately the same direction (to within about 30 deg), meaning that a high and moderately uniform VLBI source density is desirable along a spacecraft trajectory.

A spacecraft traveling to an outer planet may stay in a small part of the ecliptic for years. For example, the Voyager 2 spacecraft traversed less than 10 percent of the ecliptic on its 3.5-year trip from Uranus to Neptune. In such a case, the increased density of navigation reference sources is needed in only a fraction of the ecliptic. However, an inner planet mission, such as that of Magellan to Venus, may traverse the entire ecliptic in about a year. Then it is desirable to increase the density of sources all along the ecliptic. In the JPL VLBI catalog published in 1988 [1], the density of radio sources usable for navigation and within 10 deg of the ecliptic was 15.6 per steradian. Since the Magellan spacecraft will reach Venus in 1990, it was important to increase the density of the VLBI catalog before Magellan would go into orbit around Venus and

begin its radar mapping of the planet's surface. Thus, the immediate motivation for the work described here has been to prepare a new catalog for Magellan, although there are also benefits to Galileo, Mars Observer, and other upcoming missions.

This article describes the work done in recent years to increase the density of the JPL VLBI catalog in selected regions. Section II summarizes detection limits and looks briefly at some of the strategies employed for finding new candidate sources for the VLBI catalog. Section III describes VLBI experiments that have been done to check out these candidates as well as the analysis of source strength information. Section IV gives the end product of those experiments, with a catalog of positions and measured correlated flux densities for the sources. The distribution of sources along the ecliptic as a function of source strength is explored in some detail. Section V analyzes the prospects for finding more sources, and Section VI summarizes the main points of this article.

II. Hunting for More Ecliptic Plane Sources

A. Correlated Flux Density Detection Levels

If infinitely weak natural radio sources could be observed, the density of VLBI sources in the JPL catalog could be (almost) arbitrarily high and would be limited primarily by the time available for making VLBI observations and reducing the data. However, there are limitations imposed by nature and by the ground observing systems. Radio antennas have finite size as well as instrumental noise. They observe in the presence of other noise sources such as sky and ground radiation, and they have a limited observing bandwidth. Hence there is a calculable detection limit for the radio sources. Reference [4] contains some discussion of candidate identification and detection limits, and the reader is referred to that article for further information. The nominal detection limits for correlated flux density with the operational 250-kHz-bandwidth navigation VLBI system² (the "Block I" system) are ~ 0.2 Jy for two 70-m Deep Space Network (DSN) antennas, ~ 0.4 Jy for a baseline between a 70-m and a high-efficiency 34-m antenna, and ~ 0.8 Jy for two 34-m antennas.³ Such correlated flux densities are needed at both 2.3 GHz and 8.4 GHz observing frequencies. These limits were derived for specific signal-to-noise requirements; they assumed sys-

¹ J. B. Thomas, "An Error Analysis for Galileo Angular Position Measurements with the Block I Δ DOR System," JPL Engineering Memorandum 335-26 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 11, 1981.

² J. S. Border, "An Error Analysis for Magellan Differential Delay Rate Measurements," JPL Engineering Memorandum 335-96 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 23, 1987.

³ The unit of flux density, the jansky (Jy) is equal to 10^{-26} W·m⁻²·Hz⁻¹.

temperatures of 35 K (system temperature is a measure of the noise contributed by the telescope, receivers, and background), aperture efficiencies of ~ 60 percent, and a 250-kHz observing bandwidth. The 34-m antennas currently have higher system temperatures at 2.3 GHz than previously envisioned, so the correlated flux density limits at this frequency may be significantly (~ 20 – 30 percent) higher than the numbers given above for baselines involving 34-m antennas.

In considering the suitability of radio sources for the catalog, it is important to recognize that the relevant measure of source strength is the *correlated* flux density, which is always less than or equal to the *total* flux density. The correlated flux density is a measure of the strength of the part of the radio source that causes well-defined interference between the two telescopes making up an interferometer. For a point radio source, the correlated flux density is equal to the total flux density. However, for a more extended source, the correlated output of an interferometer is reduced because of the superposition of electromagnetic waves coming from more than one direction in the sky, which washes out the interference pattern. A priori, the correlated flux density can have any value between zero and the total flux density for a given radio source. There are no quantitative predictions of the expected distribution of the ratio between correlated and total flux densities, although it is known that certain classes of sources have, on average, a higher fraction of correlated flux density than other sources (see [4] for further discussion).

B. Lists of Candidate Sources

The best candidate sources for the JPL VLBI catalog have been selected and sieved over a number of years in the process of building that catalog, with particular emphasis on candidates near the ecliptic (e.g., [5]). In order to increase the catalog density in any region, more marginal candidates must be considered. In general, this means testing radio sources with lower total flux densities and/or with spectral flux distributions that are less likely to indicate a high ratio of correlated to total flux density. Because these are marginal candidates, the investment in antenna time and human time is large for each usable source that is found and observed enough to get an accurate position. Thus, the task of finding new sources is limited ultimately by the resources available as much as by the lack of candidate sources; at some point, the effort required to examine rather poor candidate sources becomes too great for the meager return expected.

Reference [4] outlines the basic method of finding new candidate radio sources. Because of the need to find sources for Magellan rather quickly, the step of searching

through single-telescope surveys was not done specifically for the work described in this article. Instead, emphasis was placed on objects that already had been observed interferometrically, had positions accurate to an arcsecond or better, and showed some promise of containing compact radio components of substantial strength. These objects came predominantly from two sources. First, there are lists of radio sources previously surveyed with intercontinental baselines in the DSN [4–7]. Many of these observations were made to find candidates for the VLBI catalog, but they were not followed up systematically, especially at the low end (0.2 – 0.5 Jy) of the correlated flux density scale. A second, smaller list of sources was selected from a survey of compact, low-frequency radio sources near the galactic center and anti-center [8]; these are the regions where the ecliptic and galactic planes intersect. Very Large Array (VLA) observations of these objects by one of the authors (J. S. Ulvestad) together with A. P. Rao and S. Ananthakrishnan of the Tata Institute of Fundamental Research provided a few more compact radio sources that were tried on intercontinental baselines.

III. Experiments and Data Analysis

The new candidate sources for the VLBI catalog have been observed in some of the regular DSN intercontinental VLBI experiments (Goldstone–Canberra or Goldstone–Madrid baseline) for catalog maintenance and enhancement, from mid-1986 through early 1989. The observations and analysis of such experiments are detailed in other papers describing the JPL VLBI catalog, such as [1]. All observations were dual-frequency, employing multiple frequency channels at S-band (2.3 GHz) and at X-band (8.4 GHz) in order to allow accurate delay measurements. The new data were processed using the JPL/Caltech Block II VLBI correlator described in [9]. This article reports results from the first extended usage of that correlator in the VLBI catalog effort. The bulk of the data was taken with the Mark II VLBI recording system [10], but two experiments are included which used the Mark III systems [11] that were installed recently in the DSN.

In the recent experiments, whose usable lengths ranged from ~ 8 to ~ 24 hours, a number of new candidate sources were observed, together with many of the objects already in the VLBI catalog. In this way, positions of the new sources were tied to the existing radio reference frame. For sources near the celestial equator, which includes most objects near the ecliptic, a baseline with a large north–south component is needed to get accurate source declinations. Thus, experiments on the DSN baseline between Goldstone and Canberra were quite important, since the

Goldstone-Madrid baseline gives poor north-south resolution.

The primary quantity of interest for the VLBI catalogs always has been the astrometric source positions. However, the correlated flux densities also are crucial for determining the suitability of radio sources for spacecraft navigation. Since the VLBI catalog experiments employ a larger bandwidth (effectively 1.8 MHz for Mark II, and up to 56 MHz for Mark III) than the 250 kHz of the Block I navigation system, these experiments can determine positions for sources that are not strong enough to be usable for navigation. Knowledge of the correlated flux densities and their variability with time and projected baseline gives information about which sources can be used for certain antenna combinations in order to get useful navigation data.

Correlated flux densities were estimated from the experiments between mid-1986 and late 1988. These were all Mark II VLBI experiments; there is not yet a working system to find correlated flux densities from Mark III catalog experiments. Initial values were determined by taking the measured correlation coefficients and calculating the correlated flux densities using the best available values for antenna gains (e.g., [12, 13]) and system temperatures. In most cases at X-band and some at S-band, system temperature measurements were read from strip charts. If no such data were available, zenith system temperatures were used along with a crude model for variation with elevation angle in order to get system temperatures for each observation. In such instances, an estimate of the total flux density of each source at both frequencies was used to determine the source's contribution to the system temperature.

Since there was no explicit amplitude calibration done for the VLBI experiments, it was necessary to make consistency checks for the set of experiments used in the correlated flux density analysis. The most straightforward check was based on the fact that most compact radio sources vary in strength, but the variations of different sources are completely independent. Therefore, in comparing any two experiments, half the sources should show increases in their correlated flux densities and half should show decreases.⁴

For each source in every experiment, an average correlated flux density was determined, usually from 2-4 measurements. Pairs of experiments using the same two DSN

complexes were compared by taking ratios of the average correlated flux densities at each frequency for each source common to both experiments. If the median value of the correlated-flux-density ratio at one of the frequencies was near 1.0, the experiment flux-density scales were considered consistent for that observing frequency; if it differed from 1.0 by more than 10-15 percent, the calibration of the two experiments was investigated further. In some cases, the choice of a different frequency channel (three were available at each band) resolved the discrepancy, implying that a particular experiment had an anomalous gain in one channel (e.g., due to mistuning of the maser receiving bandpass). In other instances, the "measured" system temperatures were lower than the zenith temperatures for some observations, implying that strip charts were miscalibrated or had changing calibration during an experiment. In these cases, the zenith system temperature and the crude atmospheric model were used to calculate the system temperature during a given observation, and these revised values sometimes brought the correlated flux density scales into agreement. Some experiments could never be brought into agreement and were discarded from the correlated flux density data set; for several experiments, the X-band amplitude data were discarded, but the S-band data were kept. Possible causes for these outlier experiments included poor antenna pointing (more important at X-band) and system gain variations.

The final flux-density data base included ten experiments on the Goldstone-Canberra baseline and ten on the Goldstone-Madrid baseline. Seven of those 20 experiments had little or no usable X-band flux-density data. A few individual observations were deleted from the experiments because of transient or short-term instrumental problems. Judging by the median correlated flux densities, final amplitude scales varied by ± 10 percent at S-band and as much as ± 15 percent at X-band. These are also the approximate errors on individual measurements, as estimated from the repeatability of measurements on sources observed at the same hour angle on two consecutive days. Estimated correlated flux density errors for individual source observations are found by adding ~ 0.03 Jy system noise in quadrature to 15 percent (at S-band) or 20 percent (at X-band) of the total correlated flux density. The 15 percent and 20 percent values, in turn, are roughly the scale errors and individual measurement errors added in quadrature. Errors in the *average* flux densities are somewhat fewer, but cannot be quantified simply for the general case. They depend on the details of observations for each source; for example, errors on different observations of a source in a single experiment can be highly correlated, because they reflect the same scale error in addition to (possibly) uncorrelated errors in system noise temperature and antenna pointing.

⁴ Source structure causes correlated flux densities to vary with time in a given experiment. However, in the absence of deliberate attempts to observe sources at times of minimum (or maximum) correlated flux density in a given experiment, structure should cause no overall trends for flux densities to be higher or lower in one experiment than in another.

IV. Results

Table 1 presents the astrometric data for a new VLBI catalog (the 1989-5 catalog) with a higher density of sources near the ecliptic. The source positions are given in the J2000 system, with the reduction to that system described more completely in [1].⁵ In this table, only those sources with formal declination errors less than five milliarcseconds (25 nanoradians) are included; such uncertainties are within the error budgets specified for Magellan and Galileo. There are 166 radio sources listed in this table. In contrast, the catalog published in [1] (1987-1) had only 105 radio sources with formal declination errors smaller than five milliarcseconds; the number of sources with accurate positions has grown by nearly 60 percent. Most of the new sources are near the ecliptic plane, as described later in this section.

Tables 2 and 3 contain the best estimates of the correlated flux densities of the radio sources listed in Table 1, from experiments between June 1986 and September 1988. The X-band data on the Madrid baseline come from experiments since May 1987, because the two 1986 experiments did not produce usable X-band amplitude data. No correlated flux densities are listed for objects for which there are no reliable amplitude data in the experiments between mid-1986 and late 1988. Since the source structures cause different correlated flux densities on different baselines, results are listed separately for the Goldstone-Canberra and Goldstone-Madrid baselines. At each frequency on each baseline, the tables give the average correlated flux densities, root-mean-square (RMS) variation of the correlated flux density measurements, maximum and minimum correlated flux density values, and the number of individual observations used to derive those numbers. Error estimates for individual observations were described in Section III, and are roughly 15 percent at S-band and 20 percent at X-band (plus 0.03 Jy added in quadrature at each frequency). Errors in the average flux densities are somewhat smaller, but are not quantified easily.

The catalog sources have been classified based on their suitability for VLBI navigation using different antenna pairs. Because of the higher system temperatures of the 34-m antennas at S-band, the correlated flux density cutoffs for different antenna pairs have been adjusted from those given in Section II. Category 1 sources typically have at least 0.8 Jy of correlated flux density at X-band and 1.0 Jy at S-band, so they should be usable for navigation with the

Block I system on baselines with two 34-m high-efficiency antennas. Sources in category 2 typically have correlated flux densities between 0.4 and 0.8 Jy at X-band, or 0.5 and 1.0 Jy at S-band, indicating that at least one 70-m antenna should be used for VLBI navigation. Category 3 sources have typical correlated flux densities in the 0.2–0.4 Jy (X-band) or 0.2–0.5 Jy (S-band) range, requiring 70-m pairs for navigation observations. Objects that are likely to be weaker than 0.2 Jy, and currently unusable for VLBI navigation, are relegated to category 4.

Table 4 gives the flux-density classifications for all the accurately positioned sources listed in Table 1. Classifications are given separately at S-band and X-band for most sources. The criteria for inclusion in a specific class are that the average correlated flux density, and at least 75 percent of the correlated flux densities from individual observations, are above the cutoff value for a given class. A "net" class for dual-frequency observations is also given; if the S-band and X-band classes differ, as they often do, this number is the larger of those given for the individual frequencies. Tabulating the separate values for the different frequencies as well as this dual-frequency classification facilitates selection of sources for single-frequency VLBI observations and also gives information on which band is likely to give the most trouble in dual-frequency data. Separate listings are given for each baseline (Goldstone-Canberra and Goldstone-Madrid), since radio source structure makes the correlated flux densities differ significantly between the two baselines. For some sources, the information in Table 4 is incomplete because there may not be correlated flux density values at both frequencies on both baselines in the experiments considered here. In those cases, older correlated flux density results (e.g., [5–7]) or values from the other baseline have been used to estimate a dual-frequency classification, which is included in Table 4 in parentheses.

It is crucial to recognize that many of the sources in the VLBI catalog are variable in strength on time scales on the order of a year, some by a factor of two or more. Therefore, 2- or 3-year-old correlated flux densities do not necessarily reflect current values. Some effort has been made to take variability and source structure into account by requiring that at least 75 percent of the observations have correlated flux densities above the cutoff for a certain class for inclusion in that class. Still, a source that is in the highest of the above-defined flux density categories based on its strength in 1988 may not be in that category when it is to be used for Magellan navigation in 1990 or 1991, so it may no longer be suitable for two 34-m antennas. It is hoped that more up-to-date correlated flux density values will be available when the Magellan orbital phase begins

⁵ Note that the full catalog accuracy for navigation can be achieved only via careful consideration of the reference frames in use for the radio source catalog, spacecraft navigation, and planetary ephemerides.

in August 1990, but there always will be enough of a time lag between amplitude measurements and navigation usage that sources can vary significantly.

Figures 1–3 show histories of correlated flux densities for several radio sources in the 1986–1988 period. In these figures, the correlated flux densities are plotted against interferometer hour angle (IHA). The IHA is the source hour angle with respect to the local meridian at the midpoint in longitude between the two observing stations. Figures 1(a) and 1(b) are plots of the quasar P 0420–01 at X-band on the Goldstone–Canberra and Goldstone–Madrid baselines, respectively. Note the decrease in correlated flux density by a factor of 3 or more between 1986 and 1988. Although it is more than 20 deg from the ecliptic, this source is a strong candidate for Δ VLBI navigation when Magellan is in the vicinity and two 34-m antennas are employed. P 0420–01 is listed in category 1, but there is no guarantee that it will be that strong in 1990 and 1991, as the past flux variability shows.

Figures 2(a) and 2(b) show the source P 1510–08 at both bands on the Goldstone–Madrid baseline. Its S-band correlated flux density increased by a factor of 3 or more between 1986 and 1988, which was sufficient to move it from class 2 to class 1. For part of the Magellan cruise VLBI campaign in early 1990, this is the best source visible on the Madrid baseline. If its flux density is the same as in 1986, it will not be detectable at S-band with a 34-m pair, and there would be *no* usable sources for the Goldstone–Madrid baseline when two 34-m antennas are used. Although on the average, as many sources increase in flux as decrease, practical limitations mean that *unmonitored* variability can only reduce the number of sources presumed to be available in a given flux density category. If five sources decrease in strength to drop out of class 1, and five quasars increase to move into that category, the increases do little good because there is no way of knowing which five sources can now be observed with smaller antennas.

Figures 3(a) and 3(b) show correlated flux density plots for a weaker radio source, 0536+145, at X-band on each of the two baselines. Note the different correlated flux densities on each baseline, which illustrate the reason that sources can have different flux density classifications on the two baselines. Inspecting Fig. 3(b), it is evident that a modest decrease in the flux density for 0536+145 can be enough to move it below the dividing line between class 2 and class 3 sources, so it is impossible to predict whether two 70-m antennas would be needed for this source at a given time.

Figures 4–6 display results of successive cuts on the sources in Table 1, based on their strength on the

Goldstone–Canberra baseline. These plots show source positions over the part of the celestial sphere between -40 and 40 deg declination, with 10-degree-radius circles drawn around each source position. In each plot, the ecliptic is shown as a sinusoidal curve.

There are 120 sources in categories 1, 2, and 3 on the Goldstone–Canberra baseline; i.e., 120 catalog sources are observable with two 70-m antennas using the Block I VLBI system. All 76 of them within 20 deg of the ecliptic, as well as some others in the plotted declination range, are shown in Fig. 4. Good sky coverage has been achieved, with significant “holes” along the ecliptic only at $3^{\text{h}}\text{--}5^{\text{h}}$ and $16^{\text{h}}\text{--}19^{\text{h}}$ right ascension, and a less prominent source deficit near 10^{h} right ascension. Considerable effort has been expended in the last several years to fill in the $3^{\text{h}}\text{--}5^{\text{h}}$ hole, but most of the new sources added to the catalog in that region are slightly weaker than 0.2 Jy most of the time. The $16^{\text{h}}\text{--}19^{\text{h}}$ region is where the ecliptic intersects the galactic plane near the galactic center. Confusion with galactic radio sources, and the consequent lack of source surveys in the region, make it difficult to find candidate VLBI sources.

Figure 5 is based on the 58 sources in categories 1 and 2, including 33 within 20 deg of the ecliptic. Note the generally sparser sky coverage. Figure 6 is based on the 18 sources in category 1 on the Canberra baseline, eight of which are within 10 deg of the ecliptic. This plot shows the inadvisability of navigating a spacecraft all the way around the ecliptic using only 34-m antenna pairs with the Block I system. For instance, the *only* usable sources within 15 deg of the ecliptic in the $0^{\text{h}}\text{--}12^{\text{h}}$ range of right ascension are P 0112–017 (a new source added since the last catalog was published) and OJ 287. In total, there will be a source within 10 deg of a spacecraft over only ~ 30 percent of the ecliptic.

Similar plots have been made for the Goldstone–Madrid baseline but are not included here. Their general appearance is similar to those for the Canberra baseline, except the hole near $16^{\text{h}}\text{--}19^{\text{h}}$ right ascension is worse because of the inability to observe sources below -26 deg declination. In general, the observations in this region also are constrained much more by the limited view periods for both spacecraft and quasar. For example, a source 15 deg east of the spacecraft at -22 deg declination may not be visible immediately following (and certainly not before) a VLBI observation of the spacecraft.

Figure 7 shows the same sources as does Fig. 6, but with the circles around the quasars expanded to 20 deg radius. Clearly, even if quasars as much as 20 deg from

the spacecraft are deemed acceptable, there are still significant holes in the source distribution for observations using 34-m pairs. About 70 percent of the ecliptic has a usable reference source within 20 deg. Increasing the allowable spacecraft/quasar separation to 25 deg would give coverage to approximately 90 percent of the ecliptic.

Since the strongest sources were in the catalog regardless of their ecliptic latitude, the density increase since the last published catalog has been confined almost entirely to the category 2 and 3 sources. The number of category 1, 2, and 3 sources (on at least one baseline) within 10 deg of the ecliptic has increased by 21 since the last published catalog, corresponding to a 60 percent density increase, from 15.6 to 25.2 per steradian. Another ten sources have been added between 10 deg and 20 deg of the ecliptic, and six more sources were added slightly farther to the north of the ecliptic in the direction near that of the galactic center. Thus, a total of 37 of the 43 new class 1, 2, and 3 sources added since the 1987-1 catalog facilitate navigation near the ecliptic plane. Figure 8 illustrates the change in the catalog; it is a version of Fig. 4 without circles surrounding each source position, and with the old and new sources indicated by different symbols.

V. Directions for Future Work

A. Source Lists and Surveys

In the effort to find yet more VLBI navigation sources, the radio source surveys in the astronomical literature have been studied systematically. This effort was similar to that described in [4], but some newer surveys and a much lower flux density limit were used. A list has been compiled that includes over 2,000 radio sources within ~ 20 deg of the ecliptic plane. The primary selection criterion is that these sources have total flux densities above ~ 0.25 Jy at an observing frequency of 5 GHz, although some sources have been selected from surveys at 2.7 GHz (e.g., [14]) or 1.4 GHz [15]. For the regions between -0.5 deg and 19.5 deg declination (except for the areas near the galactic plane), the MIT-Green Bank 5-GHz survey [16] has a completeness limit well below 0.25 Jy; the surveys in other areas generally are not complete to the desired selection limit. The summary list contains flux densities at several different frequencies when they are available, since spectral curvature is such an important indicator of source compactness. Table 5 shows a portion of this global source list. The literature also has been searched for all published interferometric observations of the radio sources in the large finding list, and lists have been compiled that summarize those observations. These databases are extremely useful in finding the best of the many candidate sources for possible VLBI observations in the future.

Depending on the degree of satisfaction with the current source density in the VLBI catalog, there are several options for finding more sources. The first is to use the above-described source lists to find the best candidates in some of the low-density regions of the catalog near the ecliptic plane, make moderately short baseline interferometric observations of those objects, and then try intercontinental VLBI on the sources that are still viable candidates. The area near 10^{h} right ascension is an obvious candidate for this procedure. Because of the confusion problem near the galactic center and the lack of prior surveys, the area near 18^{h} right ascension does not lend itself easily to such a screening procedure. Most of the known candidates in that area have been examined already. If a higher density of sources is desired, it may be necessary to perform a new source survey in the direction of the galactic center, probably using a short-baseline interferometer rather than a single telescope. After such a survey is made and the data analyzed, the standard post-survey screening procedures could be used. Doing a new source survey, even of this limited region, would require considerable resources, and must begin fairly soon in order to produce usable navigation radio sources by the time Galileo reaches Jupiter in 1995.

B. Expected Numbers of Sources That Could Be Added to the Catalog

Before any further efforts are made to find new VLBI catalog sources, it is useful to estimate the probabilities for finding new sources with adequate correlated flux densities for Δ VLBI navigation. In this connection, consider the distribution of sources currently in the JPL VLBI catalog. For brevity, only the Goldstone-Canberra baseline is examined here; the results are similar for the Goldstone-Madrid baseline. Since the most comprehensive source searches for the JPL VLBI catalog have been made near the ecliptic, assume for the sake of argument that *all* class 1 sources that are within 10 deg of the ecliptic are already in the VLBI catalog. Further, *assume* that the source density N as a function of correlated flux density S_c follows the relation

$$N(S_c > S_0) = N_0(S_c/S_0)^{-1.5}$$

where N_0 is the source density at correlated flux density S_0 . This relation is approximately true of source counts as a function of *total* flux density above 0.1 Jy, e.g., [17]; the assumption that it also holds for *correlated* flux density is arbitrary.

Given the numbers of sources in class 1 at S and X-bands, the above relation has been used to predict the numbers of class 2 and class 3 sources at each frequency.

Predictions for the combined (dual-frequency) class also have been made, although these are more uncertain because of the different flux density limits at the two frequencies. The predictions for class 2 sources match the observations quite well at X-band and for the dual-frequency category, while the S-band predictions are low by slightly more than one standard deviation. This provides some confidence that the above assumption about the number versus correlated flux density relation is not too far from reality. For dual-frequency observations, ~ 75 sources are predicted to be present in classes 1–3, but the VLBI catalog contains only 50. Hence there is a deficit of ~ 25 sources above 0.2-Jy correlated flux density; most of these would be category 3 sources that are not in the catalog.

Since the ecliptic is not a preferred region for the extragalactic radio sources used for VLBI navigation, the source density in each class should be similar at all ecliptic latitudes. Therefore there should be ~ 75 class 1, 2, and 3 sources between 10 deg and 20 deg from the ecliptic plane, whereas the current VLBI catalog has only 26. As many as 50 sources may be “missing” here, including 10 from categories 1 and 2 and close to 40 from category 3.

Under the assumptions given above, the result is that ~ 70 sources with correlated flux densities above 0.2 Jy at S-band and at X-band conceivably could be added to the JPL VLBI catalog at distances less than 20 deg from the ecliptic. If one were to assume that $N(S_c) \propto S_c^{-1.2}$, which is the shallowest possible relation allowed by the dual-frequency correlated flux density counts between categories 1 and 3 in the region within 10 deg of the ecliptic, only ~ 25 sources would be missing in the annulus between 10 deg and 20 deg from the ecliptic plane. This must be the *minimum* number of sources in classes 1–3 that are within 20 deg of the ecliptic and remain to be included in the VLBI catalog. The real source deficit probably is much closer to the original estimate of 70 objects.

The list of currently known candidate sources with over ~ 0.25 Jy total 5-GHz flux density and within ~ 20 deg of the ecliptic (see previous subsection) numbers $\sim 2,200$. This is about the number expected at 5 GHz within 20 deg of the ecliptic. (The fact that some sources are missing because deep surveys have not covered the entire region is balanced by the inclusion of sources selected at lower frequencies that would not make the flux-density cutoff at 5 GHz.) Of the $\sim 2,200$ sources mentioned above, $\sim 1,200$ have published interferometric (usually VLBI or VLA) observations. Finding a significant fraction of the ~ 70 missing sources would require interferometric observations of most of the remaining 1,000 objects as well as new source surveys in the regions where the ecliptic and galactic planes

intersect. Since most such sources will be observable only with two 70-m antennas in the current navigation VLBI system, such extensive searches would be useful only if (1) two 70-m antennas were to be used frequently for the navigation VLBI, or (2) a wider bandwidth VLBI system were used for navigation.

VI. Summary

A new JPL VLBI catalog has been presented here that has an increased density of radio sources near the ecliptic. The coordinate system of the catalog presented here has been based on an assumed right ascension for the reference source 3C 273 and will have a systematic rotation compared to planetary ephemerides. Therefore, the new astrometric catalog should not be used for navigation as presented here, because it is necessary to make certain that the catalog used for a specific mission is generated using the same parameters (e.g., DSN station locations and precession and nutation constants) that are used by the navigators for that mission.

Compared to the last published catalog, there are 31 new sources within 20 deg of the ecliptic plane that have position errors less than five milliarcseconds and correlated flux densities above 0.2 Jy on at least one of the two DSN intercontinental baselines. Twenty-one of those sources are within 10 deg of the ecliptic. The radio sources in the catalog have been divided into several categories depending on their correlated flux densities on the DSN baselines. Sources in class 1 can be used with the current navigation VLBI system on baselines with two 34-m antennas, a 34-m and a 70-m antenna, or two 70-m antennas. Class 2 sources require at least one 70-m antenna, while class 3 sources require 70-m antennas on both ends of the baseline. Class 4 sources are too weak for navigation using two 70-m antennas and a 250-kHz-bandwidth system. The categorization, which is based on correlated flux densities measured between 1986 and 1988, has been made to facilitate source selection and the determination of antenna requirements.

The returns for a given effort are diminishing as weaker candidate sources are considered. Finding new sources in the region of the ecliptic plane in the same direction as the galactic center may require a new source survey rather than depending upon surveys that have been published already. Counts of sources as a function of correlated flux density indicate that ~ 70 more sources above 0.2-Jy correlated flux density should exist within 20 deg of the ecliptic plane. Such sources, if added to the VLBI catalog, could be observed for VLBI navigation with two 70-m antennas. However, examining about 1,000 sources to find the new objects would be a considerable undertaking.

The JPL VLBI catalog always will have gaping holes for navigation when the current 250-kHz-bandwidth system is used with two 34-m antennas. At present, usable navigation sources will be within 10 deg of a spacecraft in only 30 percent of the ecliptic; only 70 percent of the ecliptic has usable sources within 20 deg. One option is to use one or two 70-m antennas all the time, but those telescopes are heavily oversubscribed. Another option is to increase the observing bandwidth. Using the 15-year-old Mark II VLBI system, which samples at 4 Mbits/sec (effective bandwidth of 1.8 MHz) instead of the 0.5-Mbits/sec rate of the Block I system, would increase the sensitivity by a factor of more than 2.5. All category 3 sources would

be reachable on a baseline between a 34-m and a 70-m antenna, and all Class 2 sources would be observable with two 34-m antennas. Using two 34-m antennas with this system, there would be a reference source within 10 deg of a spacecraft over 65 percent of the ecliptic, and a source within 20 deg over 98 percent of the ecliptic. For all class 3 sources to be reachable with two 34-m antennas, the bandwidth would have to be increased to ~ 6.25 MHz, with a sampling rate of at least 12.5 Mbits/sec. Then, more than 90 percent of the ecliptic would have a usable source within 10 deg. Using these wider bandwidths in a real-time system would require higher rates of data transmission from the VLBI stations to JPL.

Acknowledgments

The authors thank S. Lowe, J. Gotshalk, T. Vesperini, R. Dewey, and R. Branson for work done in scheduling and analyzing various VLBI experiments since 1986; and I. Chen for helping with the work to derive the correlated flux densities from our recent VLBI experiments. R. Linfield, R. Treuhaft, and K. Liewer made useful comments on a draft of this article.

References

- [1] O. J. Sovers, C. D. Edwards, C. S. Jacobs, G. E. Lanyi, K. M. Liewer, and R. N. Treuhaft, "Astrometric Results of 1978-1985 Deep Space Network Radio Interferometry: The JPL 1987-1 Extragalactic Source Catalog," *Astronomical Journal*, vol. 95, pp. 1647-1658, 1988.
- [2] C. Ma, T. A. Clark, J. W. Ryan, T. A. Herring, I. I. Shapiro, B. E. Corey, H. F. Hinteregger, A. E. E. Rogers, A. R. Whitney, C. A. Knight, G. L. Lundqvist, D. B. Shaffer, N. R. Vandenberg, J. C. Pigg, B. R. Schupler, and B. O. Rönnäng, "Radio Source Positions from VLBI," *Astronomical Journal*, vol. 92, pp. 1020-1029, 1986.
- [3] R. N. Treuhaft, "Deep Space Tracking in Local Reference Frames," *TDA Progress Report 42-94*, vol. April-June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 1-15, August 15, 1988.
- [4] J. S. Ulvestad and R. P. Linfield, "The Search for Reference Sources for Δ VLBI Navigation of the Galileo Spacecraft," *TDA Progress Report 42-84*, vol. October-December 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 152-163, February 15, 1986.
- [5] A. E. Wehrle, D. D. Morabito, and R. A. Preston, "Very Long Baseline Interferometry Observations of 257 Extragalactic Radio Sources in the Ecliptic Region," *Astronomical Journal*, vol. 89, pp. 336-341, 1984.

- [6] D. D. Morabito, A. E. Niell, R. A. Preston, R. P. Linfield, A. E. Wehrle, and J. Faulkner, "VLBI Observations of 416 Extragalactic Radio Sources," *Astronomical Journal*, vol. 91, pp. 1038–1050, 1986.
- [7] R. A. Preston, D. D. Morabito, J. G. Williams, J. Faulkner, D. L. Jauncey, and G. D. Nicolson, "A VLBI Survey at 2.29 GHz," *Astronomical Journal*, vol. 90, pp. 1599–1641, 1985.
- [8] A. P. Rao and S. Ananthakrishnan, "Interstellar Scattering in the Inner Parts of the Galaxy," *Nature*, vol. 312, pp. 707–711, 1984.
- [9] J. B. Thomas, *Interferometry Theory for the Block II Processor*, JPL Publication 87-29, Jet Propulsion Laboratory, Pasadena, California, October 15, 1987.
- [10] B. G. Clark "The NRAO Tape-Recorder Interferometer System," *Proc. IEEE*, vol. 61, pp. 1242–1248, 1973.
- [11] A. E. E. Rogers, R. J. Cappallo, H. F. Hinteregger, J. I. Levine, E. F. Nesman, J. C. Webber, A. R. Whitney, T. A. Clark, C. Ma, J. Ryan, B. E. Corey, C. C. Counselman, T. A. Herring, I. I. Shapiro, C. A. Knight, D. B. Shaffer, N. R. Vandenberg, R. Lacasse, R. Mauzy, B. Rayhrer, B. R. Schupler, and J. C. Pigg, "Very Long Baseline Radio Interferometry: The Mark III System for Geodesy, Astrometry, and Aperture Synthesis," *Science*, vol. 219, pp. 51–54, 1983.
- [12] S. D. Slobin, "DSS 14 64-Meter Antenna S- and X-Band Efficiency and System Noise Temperature Calibrations, September 1987," *TDA Progress Report 42-92*, vol. October–December 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 138–146, February 15, 1988.
- [13] S. D. Slobin and D. A. Bathker, "DSN 70-Meter Antenna X-Band Gain, Phase, and Pointing Performance, With Particular Application for Voyager 2 Neptune Encounter," *TDA Progress Report 42-95*, vol. July–September 1988, pp. 237–245, November 15, 1988.
- [14] J. G. Bolton, A. Savage, and A. E. Wright, "The Parkes 2700 MHz survey (fourteenth part): Catalogue for declinations -4° to -15° , right ascensions 10^h to 15^h ," *Australian Journal of Physics Astrophysical Supplement*, no. 46, p. 1, 1979.
- [15] J. J. Condon and J. J. Broderick, "A 1400 MHz Sky Survey. II. Confusion-limited Maps covering $19^h30^m < \alpha < 7^h30^m$, $-5^\circ < \delta < +82^\circ$," *Astronomical Journal*, vol. 91, pp. 1051–1057, 1986.
- [16] C. L. Bennett, C. R. Lawrence, B. F. Burke, J. N. Hewitt, and J. Mahoney, "The MIT-Green Bank (MG) 5 GHz Survey," *Astrophysical Journal Supplement*, vol. 61, pp. 1–104, 1986.
- [17] K. I. Kellermann and J. V. Wall, "Radio Source Counts and their Interpretation," in *International Astronomical Union Symposium No. 124, Observational Cosmology*, edited by A. Hewitt, G. Burbidge, and L. Fang, Dordrecht Reidel Publishing Company, pp. 545–564, 1987.

Table 1. Astrometric JPL VLBI catalog^a

Name	Right Ascension			Declination			R.A. error	Dec. error
	h	m	s	°	'	"	s	"
P 0008 - 264	0	11	1.24699355	-26	12	33.3773551	0.00010644	0.0015781
P 0019 + 058	0	22	32.44118862	6	8	4.2720531	0.00003601	0.0012603
P 0048 - 09	0	50	41.31734533	-9	29	5.2057755	0.00005632	0.0014180
P 0104 - 408	1	6	45.10803890	-40	34	19.9565151	0.00012592	0.0013105
P 0106 + 01	1	8	38.77104811	1	35	0.3203449	0.00003232	0.0007163
P 0111 + 021	1	13	43.14504680	2	22	17.3176826	0.00010071	0.0017404
P 0112 - 017	1	15	17.09989294	-1	27	4.5744796	0.00003303	0.0007690
P 0113 - 118	1	16	12.52201531	-11	36	15.4319487	0.00004358	0.0009582
P 0119 + 11	1	21	41.59500938	11	49	50.4154942	0.00003038	0.0007081
GC 0119 + 04	1	21	56.86158102	4	22	24.7399107	0.00005209	0.0013681
DA 55	1	36	58.59476987	47	51	29.1017977	0.00004802	0.0005116
0146 + 056	1	49	22.37082174	5	55	53.5720621	0.00003300	0.0007601
P 0201 + 113	2	3	46.65701680	11	34	45.4119221	0.00003083	0.0007475
P 0202 + 14	2	4	50.41391519	15	14	11.0450538	0.00003467	0.0006642
0212 + 735	2	17	30.81344376	73	49	32.6235037	0.00010937	0.0003974
GC 0221 + 06	2	24	28.42814525	6	59	23.3439920	0.00003330	0.0013038
DW 0224 + 67	2	28	50.05164546	67	21	3.0322165	0.00012631	0.0008321
P 0229 + 13	2	31	45.89403472	13	22	54.7180082	0.00003289	0.0010830
CTD 20	2	37	52.40566460	28	48	8.9922817	0.00003301	0.0005215
GC 0235 + 16	2	38	38.93012333	16	36	59.2759391	0.00003493	0.0006428
OD 166	2	42	29.17087987	11	1	0.7295586	0.00003987	0.0008338
OD 094.7	2	59	27.07664169	7	47	39.6424948	0.00004394	0.0019742
OE 400	3	3	35.24220862	47	16	16.2776761	0.00004975	0.0005190
0306 + 102	3	9	3.62348204	10	29	16.3426808	0.00003299	0.0026601
0309 + 411	3	13	1.96212709	41	20	1.1849592	0.00004334	0.0008132
3C 84	3	19	48.16013053	41	30	42.1061371	0.00008238	0.0011983
0326 + 277	3	29	57.66936164	27	56	15.4992592	0.00003858	0.0011890
P 0332 - 403	3	34	13.65454152	-40	8	25.3957294	0.00014044	0.0014000
NRAO 140	3	36	30.10766072	32	18	29.3432486	0.00004501	0.0006706
CTA 26	3	39	30.93773041	-1	46	35.7999706	0.00003202	0.0009034
0342 + 147	3	45	6.41651123	14	53	49.5600248	0.00003363	0.0010192
P 0402 - 362	4	3	53.74977936	-36	5	1.9097062	0.00010031	0.0012249
GC 0406 + 12	4	9	22.00870257	12	17	39.8489955	0.00003703	0.0010056
P 0420 - 01	4	23	15.80070198	-1	20	33.0637547	0.00002786	0.0007633
VRO 41.04.01	4	23	56.00981686	41	50	2.7176069	0.00021967	0.0032146
P 0425 + 048	4	27	47.57029577	4	57	8.3305196	0.00012503	0.0021880
3C 120	4	33	11.09551075	5	21	15.6227035	0.00006154	0.0012917
P 0434 - 188	4	37	1.48270317	-18	44	48.6120617	0.00006570	0.0012561
P 0438 - 43	4	40	17.17988751	-43	33	8.6003713	0.00014003	0.0013208
0440 + 345	4	43	31.63521824	34	41	6.6646666	0.00003922	0.0007564
P 0446 + 11	4	49	7.67106994	11	21	28.5984746	0.00002903	0.0018032
P 0451 - 28	4	53	14.64620104	-28	7	37.3188681	0.00034099	0.0041906
0500 + 019	5	3	21.19711453	2	3	4.6786075	0.00002983	0.0017863
0454 + 844	5	8	42.36483468	84	32	4.5443577	0.00036133	0.0004369
P 0506 + 101	5	9	27.45707126	10	11	44.6020917	0.00002725	0.0011530
P 0507 + 17	5	10	2.36914548	18	0	41.5833675	0.00002938	0.0006610
P 0528 + 134	5	30	56.41678287	13	31	55.1506072	0.00003060	0.0006461
P 0537 - 441	5	38	50.36122088	-44	5	8.9347405	0.00014966	0.0013899
0536 + 145	5	39	42.36600089	14	33	45.5626621	0.00002762	0.0006572
0544 + 273	5	47	34.14897656	27	21	56.8449776	0.00005573	0.0016400
DA 193	5	55	30.80561462	39	48	49.1662319	0.00003830	0.0004967
0556 + 238	5	59	32.03315606	23	53	53.9275968	0.00002972	0.0005826
0600 + 177	6	3	9.13027327	17	42	16.8116293	0.00002997	0.0006438
P 0607 - 15	6	9	40.94970020	-15	42	40.6743233	0.00027504	0.0040739
3C 166	6	45	24.09952596	21	21	51.2023276	0.00002969	0.0006424
0657 + 172	7	0	1.52557647	17	9	21.7010705	0.00004964	0.0010824
P 0722 + 145	7	25	16.80776214	14	25	13.7479152	0.00002813	0.0019852
DW 0723 - 00	7	25	50.63992029	-0	54	56.5440972	0.00004696	0.0010124

^a This catalog is a subset of the 1989-5 catalog, including all sources with formal declination errors of less than 5.0 milliarcseconds. Positions are given in J2000 coordinates, with the zero point of right ascension defined by the assumed right ascension of 3C 273.

Table 1 (contd)

Name	Right Ascension			Declination			R.A. error	Dec. error
	h	m	s	°	'	"	s	"
P 0727 - 11	7	30	19.11246024	-11	41	12.6005321	0.00003440	0.0009449
P 0735 + 17	7	38	7.39374592	17	42	18.9992145	0.00003255	0.0006931
P 0736 + 01	7	39	18.03391957	1	37	4.6186448	0.00010231	0.0038285
OI 363	7	41	10.70338957	31	12	0.2289329	0.00009196	0.0012888
DW 0742 + 10	7	45	33.05950366	10	11	12.6918275	0.00002903	0.0007084
GC 0743 + 25	7	46	25.87414169	25	49	2.1353629	0.00003262	0.0006390
B2 0745 + 24	7	48	36.10927896	24	0	24.1106879	0.00003958	0.0008046
P 0748 + 126	7	50	52.04575527	12	31	4.8275558	0.00003619	0.0007976
P 0754 + 100	7	57	6.64294522	9	56	34.8509802	0.00002652	0.0009451
OJ 425	8	18	15.99969677	42	22	45.4139981	0.00004286	0.0005609
P 0823 + 033	8	25	50.33837169	3	9	24.5196233	0.00002579	0.0007326
B2 0827 + 24	8	30	52.08616400	24	10	59.8208076	0.00003600	0.0005821
4C 71.07	8	41	24.36604803	70	53	42.1708564	0.00036869	0.0020193
OJ 287	8	54	48.87493186	20	6	30.6409927	0.00002475	0.0005294
OJ 499	9	3	3.99015987	46	51	4.1325722	0.00054797	0.0039037
P 0912 + 029	9	14	37.91348420	2	45	59.2452132	0.00005019	0.0009866
4C 39.25	9	27	3.01385886	39	2	20.8508520	0.00003726	0.0004605
1012 + 232	10	14	47.06547054	23	1	16.5697551	0.00004060	0.0006320
P 1034 - 293	10	37	16.07978339	-29	34	2.8134642	0.00008266	0.0011473
OL 064.5	10	41	17.16249455	6	10	16.9222264	0.00002980	0.0009321
3C 245	10	42	44.60589228	12	3	31.2591904	0.00035446	0.0047924
1044 + 719	10	48	27.62003187	71	43	35.9372599	0.00011409	0.0004905
P 1055 + 01	10	58	29.60522746	1	33	58.8229850	0.00002018	0.0007057
P 1104 - 445	11	7	8.69432359	-44	49	7.6190924	0.00018894	0.0015754
GC 1111 + 14	11	13	58.69511043	14	42	26.9511597	0.00003452	0.0006721
P 1123 + 26	11	25	53.71189910	26	10	19.9772741	0.00002800	0.0004920
P 1127 - 14	11	30	7.05268440	-14	49	27.3896707	0.00007075	0.0013349
GC 1128 + 38	11	30	53.28259214	38	15	18.5452991	0.00004249	0.0006856
P 1144 - 379	11	47	1.37079306	-38	12	11.0236144	0.00007819	0.0010893
P 1148 - 00	11	50	43.87085417	-0	23	54.2050281	0.00007624	0.0022801
P 1222 + 037	12	24	52.42189082	3	30	50.2925291	0.00003350	0.0009782
3C 273	12	29	6.69970000	2	3	8.5988392	0.00000001	0.0006671
3C 274	12	30	49.42341915	12	23	28.0421643	0.00030286	0.0040654
P 1244 - 255	12	46	46.80217065	-25	47	49.2896321	0.00004755	0.0009872
3C 279	12	56	11.16660613	-5	47	21.5265915	0.00002249	0.0008804
B2 1308 + 32	13	10	28.66380137	32	20	43.7811295	0.00003027	0.0004458
OP-322	13	16	7.98600558	-33	38	59.1718719	0.00013842	0.0015958
DW1335 - 12	13	37	39.78282936	-12	57	24.6935349	0.00002323	0.0008232
GC 1342 + 662	13	43	45.95944179	66	2	25.7429691	0.00011361	0.0007323
GC 1342 + 663	13	44	8.67949762	66	6	11.6417292	0.00008134	0.0004207
P 1349 - 439	13	52	56.53533280	-44	12	40.3888519	0.00019058	0.0015367
P 1354 + 19	13	57	4.43661980	19	19	7.3711230	0.00002427	0.0005085
OP-192	13	57	11.24503895	-15	27	28.7867430	0.00003011	0.0008814
OQ 208	14	7	0.39436580	28	27	14.6889018	0.00003172	0.0005476
GC 1418 + 54	14	19	46.59726856	54	23	14.7851817	0.00005389	0.0005101
OQ-151	14	32	57.69062126	-18	1	35.2493238	0.00027582	0.0038462
P 1445 - 16	14	48	15.05418706	-16	20	24.5487409	0.00003362	0.0008963
OR 103	15	4	24.97974565	10	29	39.1982308	0.00002066	0.0005423
P 1504 - 167	15	7	4.78702944	-16	52	30.2673226	0.00003536	0.0010599
P 1510 - 08	15	12	50.53293773	-9	5	59.8296275	0.00002424	0.0008074
P 1511 - 100	15	13	44.89334992	-10	12	0.2590722	0.00012513	0.0047020
P 1514 - 24	15	17	41.81336779	-24	22	19.4771955	0.00006876	0.0011368
P 1519 - 273	15	22	37.67599346	-27	30	10.7842172	0.00008423	0.0012447
P 1532 + 01	15	34	52.45359734	1	31	4.2075341	0.00002894	0.0007092
DW 1548 + 05	15	50	35.26917430	5	27	10.4485942	0.00002469	0.0006388
DW 1555 + 00	15	57	51.43393778	-0	1	50.4133609	0.00003074	0.0008196
DA 406	16	13	41.06412749	34	12	47.9077291	0.00003398	0.0005824
P 1614 + 051	16	16	37.55676335	4	59	32.7369010	0.00002445	0.0006393
GC 1633 + 38	16	35	15.49283800	38	8	4.4992151	0.00003668	0.0005022
NRAO 512	16	40	29.63260653	39	46	46.0280763	0.00004006	0.0005402
3C 345	16	42	58.80980000	39	48	36.9929752	0.00003562	0.0004583
OS 092	16	58	9.01137451	7	41	27.5425233	0.00004004	0.0009901

Table 1 (contd)

Name	Right Ascension			Declination			R.A. error	Dec. error
	h	m	s	°	'	"	s	"
DW 1656 + 05	16	58	33.44723638	5	15	16.4457268	0.00011546	0.0018022
P 1657 - 261	17	0	53.15421641	-26	10	51.7249256	0.00005629	0.0011049
OT-111	17	9	34.34542060	-17	28	53.3636920	0.00005597	0.0012250
GC 1717 + 17	17	19	13.04848761	17	45	6.4357837	0.00018765	0.0033162
NRAO 530	17	33	2.70579312	-13	4	49.5459665	0.00002905	0.0008033
OT 465	17	39	57.12882110	47	37	58.3606308	0.00004406	0.0006333
P 1741 - 038	17	43	58.85610106	-3	50	4.6152077	0.00002409	0.0006855
1749 + 701	17	48	32.83990518	70	5	50.7682029	0.00010839	0.0005487
OT 081	17	51	32.81849911	9	39	0.7294019	0.00002474	0.0007317
1803 + 784	18	0	45.68294237	78	28	4.0179481	0.00030408	0.0007011
3C 371	18	6	50.68020077	69	49	28.1086068	0.00008194	0.0004372
P 1821 + 10	18	24	2.85520298	10	44	23.7746943	0.00004006	0.0014917
OV-213	19	11	9.65297458	-20	6	55.1076354	0.00004556	0.0009498
OV-235	19	23	32.18985970	-21	4	33.3302802	0.00003863	0.0009280
OV-236	19	24	51.05617872	-29	14	30.1197921	0.00006617	0.0010194
OV 239.7	19	25	59.60523217	21	6	26.1630496	0.00002937	0.0005276
P 1936 - 15	19	39	26.65779536	-15	25	43.0564710	0.00003442	0.0008560
OV-198	20	0	57.09051289	-17	48	57.6703577	0.00003670	0.0008616
P 2008 - 159	20	11	15.71098648	-15	46	40.2508255	0.00003501	0.0008317
OW 637	20	22	6.68142238	61	36	58.8063153	0.00013980	0.0020267
OW 551	20	31	47.95843033	54	55	3.1481006	0.00032011	0.0036644
P 2029 + 121	20	31	54.99419610	12	19	41.3408484	0.00006625	0.0013675
B2 2113 + 29B	21	15	29.41332663	29	33	38.3684139	0.00003171	0.0006134
OX 036	21	23	44.51734537	5	35	22.0953806	0.00002753	0.0006091
P 2128 - 12	21	31	35.26172084	-12	7	4.7933609	0.00003429	0.0008040
P 2131 - 021	21	34	10.30957914	-1	53	17.2363608	0.00002918	0.0006818
P 2134 + 004	21	36	38.58638381	0	41	54.2145355	0.00002834	0.0006445
P 2145 + 06	21	48	5.45863130	6	57	38.6057827	0.00002729	0.0006223
OX 082	21	51	37.87535246	5	52	12.9567344	0.00004647	0.0010399
OX-192	21	58	6.28192609	-15	1	9.3260428	0.00014238	0.0020705
VRO 42.22.01	22	2	43.29124274	42	16	39.9817741	0.00003823	0.0004596
P 2216 - 03	22	18	52.03774097	-3	35	36.8783349	0.00003313	0.0008433
3C 446	22	25	47.25928239	-4	57	1.3886873	0.00003411	0.0008036
P 2227 - 08	22	29	40.08432367	-8	32	54.4322271	0.00003356	0.0007913
2229 + 695	22	30	36.46955494	69	46	28.0788256	0.00009468	0.0005378
CTA 102	22	32	36.40888182	11	43	50.9060834	0.00002704	0.0005475
GC 2234 + 28	22	36	22.47076608	28	28	57.4155420	0.00003138	0.0004683
OY-172.6	22	46	18.23197580	-12	6	51.2746746	0.00003672	0.0008401
P 2245 - 328	22	48	38.68575577	-32	35	52.1839028	0.00006021	0.0009986
3C 454.3	22	53	57.74787647	16	8	53.5628922	0.00003169	0.0006649
GC 2253 + 41	22	55	36.70769019	42	2	52.5347775	0.00005661	0.0010515
P 2254 + 024	22	57	17.56309267	2	43	17.5140935	0.00005252	0.0009038
GC 2318 + 04	23	20	44.85658844	5	13	49.9546884	0.00003016	0.0006805
P 2320 - 035	23	23	31.95371778	-3	17	5.0207566	0.00003543	0.0007714
P 2345 - 16	23	48	2.60855413	-16	31	12.0199908	0.00005820	0.0011185
P 2355 - 106	23	58	10.88243721	-10	20	8.6096896	0.00004540	0.0010260

Table 2. Correlated flux densities on Goldstone–Canberra baseline for all sources listed in Table 1 that have strengths determined from the experiments between 1986 and 1988^a

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
P 0008–264	0.50	0.13	0.71	0.32	17	0.30	0.05	0.38	0.22	9
P 0019+058	0.20	0.02	0.23	0.18	4	0.15	0.04	0.19	0.10	4
P 0048–09	0.80	0.01	0.82	0.79	4	1.04	0.15	1.20	0.84	4
P 0104–408	1.25	0.15	1.50	1.03	13	2.33	0.13	2.56	2.21	5
P 0106+01	2.26	0.82	3.49	0.82	20	0.66	0.25	1.00	0.22	12
P 0112–017	1.07	0.01	1.08	1.07	3	1.16	0.08	1.25	1.06	3
P 0113–118	1.39	0.07	1.46	1.31	4	0.74	0.10	0.90	0.65	4
P 0119+11	0.18	0.05	0.25	0.14	3	0.34	0.06	0.42	0.30	3
DA 55	0.40	0.15	0.64	0.21	6	0.98	0.01	0.99	0.98	2
0146+056	0.81	0.02	0.84	0.79	3	0.37	0.02	0.39	0.35	3
P 0202+14	0.86	0.20	1.23	0.47	20	1.05	0.50	1.90	0.58	12
GC 0221+06	0.33	0.01	0.34	0.31	2	0.73	0.00	0.73	0.73	2
P 0229+13	0.86	0.04	0.91	0.81	3	0.62	0.03	0.65	0.58	3
CTD 20	1.08	0.25	1.52	0.57	19	1.68	0.56	2.42	0.95	11
GC 0235+16	1.39	0.17	1.75	1.17	19	0.89	0.31	1.26	0.31	11
OD 166	1.18	0.07	1.34	1.07	19	0.46	0.14	0.60	0.23	11
OE 400	1.41	0.32	1.92	0.81	12	1.25	0.63	2.10	0.42	6
0306+102	0.20	0.08	0.28	0.12	2	0.21	0.02	0.23	0.19	2
0309+411	0.33	0.02	0.35	0.30	4	0.35	0.14	0.48	0.15	3
3C 84	1.45	0.50	2.24	0.74	9	0.18	0.08	0.32	0.11	4
0326+277	0.17	0.03	0.22	0.12	6	0.21	0.04	0.25	0.13	5
P 0332–403	0.97	0.03	0.99	0.93	4	0.63	0.08	0.72	0.54	4
NRAO 140	1.00	0.27	1.46	0.76	4	0.37	0.16	0.62	0.19	4
CTA 26	1.20	0.29	1.63	0.79	14	0.69	0.29	1.02	0.37	6
0342+147	0.27	0.05	0.38	0.21	21	0.23	0.06	0.31	0.15	12
P 0402–362	0.47	0.08	0.60	0.31	12	1.63	0.22	2.00	1.33	8
GC 0406+12	0.66	0.09	0.78	0.45	19	0.10	0.02	0.13	0.07	10
P 0420–01	2.91	0.41	3.55	2.33	22	1.71	0.65	2.44	0.45	13
P 0425+048	0.24	0.05	0.30	0.13	20	0.21	0.04	0.26	0.13	12
3C 120	0.50	0.10	0.59	0.37	5	0.36	0.11	0.50	0.25	5
P 0434–188	1.00	0.06	1.14	0.89	22	0.49	0.09	0.63	0.35	12
P 0438–43	1.67	0.08	1.81	1.59	5	0.80	0.03	0.83	0.75	3
0440+345	0.77	0.06	0.87	0.66	9	0.13	0.07	0.23	0.06	3
P 0446+11	0.36	0.07	0.43	0.30	2	0.34	0.04	0.38	0.29	2
0500+019	0.88	0.04	0.93	0.84	2	0.36	0.01	0.37	0.34	2
P 0506+101	0.56	0.14	0.90	0.39	8	0.44	0.13	0.62	0.31	3
P 0507+17	0.20	0.06	0.32	0.09	11	0.32	0.11	0.53	0.22	6
P 0528+134	1.21	0.12	1.47	1.01	12	0.56	0.21	0.87	0.22	7
0536+145	0.44	0.04	0.52	0.37	11	0.31	0.10	0.43	0.15	6
P 0537–441	2.62	0.04	2.67	2.56	4	1.95	0.11	2.04	1.77	4
0544+273	0.16	0.03	0.22	0.13	11	0.20	0.10	0.31	0.07	5
DA 193	2.54	0.16	2.86	2.30	18	1.43	0.34	1.78	0.41	12
0556+238	0.50	0.16	0.72	0.17	13	0.50	0.24	0.77	0.11	7
0600+177	0.37	0.07	0.50	0.23	11	0.22	0.04	0.29	0.15	7
3C 166	0.24	0.10	0.34	0.07	4	0.00	0.00	0.00	0.00	0
0657+172	0.57	0.08	0.70	0.46	9	0.50	0.12	0.70	0.43	4
P 0722+145	0.43	0.03	0.46	0.40	3	0.00	0.00	0.00	0.00	0
DW 072300	0.37	0.24	0.71	0.18	3	0.22	0.15	0.43	0.08	3
P 0727–11	2.28	0.66	3.49	0.59	25	1.72	0.59	2.58	0.37	12
P 0735+17	0.60	0.12	0.81	0.41	13	0.66	0.25	1.29	0.47	8
P 0736+01	1.26	0.02	1.28	1.24	3	0.00	0.00	0.00	0.00	0
OI 363	1.73	0.11	1.87	1.53	6	0.41	0.19	0.69	0.17	6
DW 0742+10	1.01	0.07	1.12	0.91	11	0.82	0.05	0.89	0.76	5
GC 0743+25	0.45	0.01	0.46	0.44	4	0.00	0.00	0.00	0.00	0
B2 0745+24	0.83	0.23	1.16	0.52	7	0.82	0.00	0.82	0.82	1
P 0748+126	0.44	0.10	0.61	0.33	6	0.43	0.11	0.59	0.30	6
P 0754+100	0.33	0.01	0.35	0.31	3	0.00	0.00	0.00	0.00	0

^a The “No.” columns give total number of observations in the database. Ranges in flux density may be caused by variability from experiment to experiment, structure-induced differences among observations in a single experiment, or both.

Table 2 (contd)

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
OJ 425	1.40	0.10	1.50	1.24	6	0.53	0.00	0.53	0.53	1
P 0823+033	0.78	0.12	0.94	0.58	10	0.63	0.19	0.92	0.35	8
OJ 287	1.84	0.53	2.85	1.01	25	2.85	0.91	4.30	1.15	15
OJ 499	0.12	0.02	0.17	0.10	5	0.17	0.06	0.24	0.10	5
P 0912+029	0.25	0.01	0.26	0.24	2	0.00	0.00	0.00	0.00	0
4C 39.25	2.61	0.66	3.89	1.57	31	1.19	0.38	2.13	0.37	18
1012+232	0.54	0.05	0.61	0.50	3	0.00	0.00	0.00	0.00	0
P 1034-293	0.90	0.20	1.26	0.58	20	1.16	0.30	1.76	0.77	13
OL 064.5	0.52	0.08	0.67	0.41	8	0.74	0.17	0.89	0.40	8
P 1055-01	0.90	0.10	1.12	0.68	28	1.97	0.34	2.68	1.24	17
P 1104-445	1.77	0.10	1.88	1.63	6	0.57	0.08	0.65	0.46	6
P 1123+26	0.76	0.07	0.90	0.62	29	0.80	0.14	1.04	0.57	18
P 1127-14	0.55	0.19	0.82	0.32	7	0.14	0.02	0.16	0.12	3
GC 1128+38	0.64	0.06	0.76	0.57	8	0.32	0.04	0.38	0.28	8
P 1144-379	0.63	0.26	0.96	0.20	22	0.99	0.32	1.53	0.44	16
P 1222+037	0.91	0.07	1.03	0.80	9	0.52	0.07	0.58	0.45	2
3C 273	1.71	0.69	3.02	0.41	39	1.81	0.44	2.74	0.89	22
P 1244-255	0.56	0.24	0.95	0.22	24	0.57	0.15	0.87	0.32	17
3C 279	1.21	0.44	1.87	0.26	11	0.96	0.24	1.32	0.64	6
B2 1308+32	0.95	0.29	1.42	0.37	37	1.17	0.70	2.34	0.33	22
OP-322	0.42	0.15	0.69	0.16	7	1.10	0.08	1.20	0.99	3
DW 1335-12	1.65	0.25	2.31	1.24	36	2.55	1.00	4.32	1.46	21
P 1349-439	0.38	0.05	0.44	0.31	5	0.33	0.03	0.36	0.31	2
OP-192	0.56	0.01	0.57	0.54	4	0.00	0.00	0.00	0.00	0
P 1354+19	0.88	0.15	1.18	0.59	34	1.02	0.32	1.45	0.52	21
OQ 208	0.22	0.05	0.31	0.16	10	0.13	0.01	0.14	0.12	3
OQ-151	0.61	0.08	0.72	0.45	8	0.18	0.04	0.23	0.13	7
OR 103	1.29	0.24	1.78	0.71	29	0.82	0.24	1.43	0.37	17
P 1504-167	1.25	0.24	1.72	0.85	11	0.48	0.11	0.65	0.23	8
P 1510-08	1.32	0.19	1.77	1.08	24	1.38	0.44	2.63	0.85	12
P 1511-100	0.80	0.03	0.82	0.76	3	0.10	0.03	0.13	0.07	2
P 1514-24	0.37	0.08	0.48	0.30	3	0.34	0.03	0.37	0.30	2
P 1519-273	1.38	0.26	1.74	0.87	24	1.28	0.17	1.49	0.81	16
P 1532+01	0.56	0.03	0.59	0.51	3	0.33	0.00	0.34	0.33	2
DW 1548+05	0.96	0.06	1.04	0.89	3	0.59	0.03	0.62	0.56	2
DW 1555+00	0.30	0.05	0.41	0.20	28	0.35	0.05	0.42	0.22	16
DA 406	1.00	0.20	1.38	0.76	16	0.59	0.12	0.81	0.43	8
P 1614+051	0.36	0.01	0.37	0.34	2	0.25	0.00	0.25	0.25	1
GC 1633+38	2.04	0.19	2.24	1.71	12	1.70	0.32	2.16	1.16	9
NRAO 512	0.87	0.11	1.04	0.67	14	0.64	0.12	0.75	0.43	7
3C 345	4.28	0.84	5.70	3.03	21	3.67	0.31	4.03	3.12	11
OS 092	0.61	0.04	0.66	0.54	8	1.12	0.13	1.32	0.93	8
P 1657-261	0.80	0.34	1.46	0.47	9	1.56	0.50	2.47	1.14	9
OT-111	0.20	0.06	0.31	0.08	11	0.18	0.04	0.25	0.13	5
NRAO 530	0.68	0.31	1.34	0.28	32	1.57	0.75	2.56	0.14	22
OT 465	0.68	0.15	0.88	0.52	3	0.08	0.00	0.08	0.08	1
P 1741-038	1.13	0.24	1.58	0.75	28	0.75	0.13	0.91	0.49	18
OT 081	1.07	0.26	1.36	0.67	18	1.59	0.53	2.25	0.97	9
P 1821+10	0.70	0.13	0.85	0.39	9	0.22	0.06	0.28	0.10	9
OV-213	0.59	0.11	0.69	0.42	4	0.38	0.11	0.55	0.24	6
OV-235	0.86	0.13	1.23	0.73	21	0.99	0.13	1.30	0.76	16
OV-236	2.60	0.73	3.98	1.49	26	3.99	0.90	5.72	1.90	18
OV 239.7	0.39	0.07	0.49	0.27	19	0.32	0.10	0.44	0.12	8
P 1936-15	0.12	0.01	0.13	0.11	3	0.23	0.01	0.24	0.22	3
OV-198	1.60	0.26	2.15	1.06	27	1.60	0.44	2.21	0.75	16
P 2008-159	0.45	0.06	0.50	0.36	3	0.37	0.05	0.42	0.31	3
P 2029+121	0.34	0.02	0.37	0.32	6	0.48	0.09	0.59	0.34	6
B2 2113+29B	0.46	0.08	0.67	0.40	9	0.66	0.15	0.82	0.44	6
OX 036	2.73	0.03	2.77	2.69	3	0.77	0.05	0.83	0.72	3
P 2128-12	0.27	0.02	0.28	0.24	3	0.67	0.07	0.77	0.60	3

Table 2 (contd)

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
P 2131-021	0.40	0.02	0.42	0.38	3	0.96	0.05	1.02	0.90	3
P 2134+004	0.65	0.20	0.97	0.32	16	1.20	0.29	1.73	0.79	12
P 2145+06	1.49	0.22	1.97	1.26	11	3.22	0.51	3.71	2.17	7
OX 082	0.59	0.06	0.69	0.45	13	0.25	0.02	0.27	0.19	7
OX-192	0.44	0.09	0.58	0.32	5	0.58	0.11	0.79	0.51	5
VRO 42.22.01	0.38	0.24	1.01	0.12	20	0.46	0.15	0.69	0.18	14
P 2216-03	1.37	0.19	1.62	1.08	9	0.50	0.14	0.70	0.38	3
3C 446	1.08	0.17	1.36	0.85	9	1.47	0.48	2.02	0.48	7
P 2227-08	0.40	0.05	0.44	0.33	3	1.13	0.10	1.24	1.00	3
CTA 102	1.89	0.20	2.46	1.54	19	0.75	0.14	1.10	0.52	16
GC 2234+28	1.55	0.23	1.96	1.11	20	0.41	0.13	0.73	0.23	16
OY-172.6	0.58	0.15	0.86	0.14	18	0.62	0.23	0.99	0.21	15
P 2245-328	0.52	0.20	0.86	0.17	25	0.32	0.06	0.39	0.14	14
3C 454.3	6.06	1.00	7.29	4.75	8	1.56	0.22	1.75	1.19	4
GC 2253+41	0.37	0.23	0.68	0.14	4	0.19	0.08	0.29	0.09	4
P 2254+024	0.24	0.02	0.27	0.23	3	0.30	0.02	0.32	0.27	3
GC 2318+04	0.79	0.01	0.81	0.78	3	0.61	0.04	0.66	0.57	3
P 2320-035	0.44	0.09	0.58	0.27	21	0.36	0.10	0.56	0.15	15
P 2345-16	1.36	0.47	1.88	0.28	16	0.56	0.29	0.94	0.22	8
P 2355-106	0.44	0.12	0.68	0.28	23	0.73	0.07	0.84	0.57	17

Table 3. Correlated flux densities on Goldstone-Madrid baseline for all sources listed in Table 1 that have strengths determined from the experiments between 1986 and 1988^a

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
P 0019+058	0.26	0.05	0.34	0.17	8	0.29	0.06	0.40	0.20	6
P 0048-09	0.73	0.07	0.80	0.57	6	0.75	0.09	0.86	0.63	5
P 0106+01	2.33	0.57	3.44	1.42	14	0.45	0.12	0.65	0.21	12
P 0111+021	0.17	0.04	0.23	0.11	5	0.14	0.02	0.17	0.11	5
P 0112-017	0.95	0.04	1.01	0.88	7	0.60	0.10	0.77	0.43	7
P 0113-118	1.22	0.14	1.35	0.89	7	0.22	0.07	0.28	0.10	5
P 0119+11	0.59	0.06	0.72	0.51	8	0.64	0.12	0.85	0.48	8
GC 0119+04	0.56	0.05	0.62	0.50	5	0.86	0.18	1.15	0.60	5
DA 55	0.89	0.33	1.51	0.43	12	0.80	0.31	1.36	0.48	8
0146+056	0.82	0.05	0.90	0.75	7	0.14	0.02	0.18	0.12	7
P 0201+113	0.75	0.04	0.79	0.68	5	0.32	0.05	0.41	0.25	5
P 0202+14	0.68	0.23	1.08	0.32	15	0.88	0.40	1.48	0.30	11
0212+735	0.64	0.44	1.73	0.06	41	0.52	0.36	1.33	0.07	26
GC 0221+06	0.14	0.04	0.21	0.11	5	0.63	0.09	0.71	0.45	5
DW 0224+67	0.54	0.33	1.30	0.09	15	0.73	0.42	1.66	0.29	8
P 0229+13	0.80	0.08	0.93	0.70	4	0.62	0.12	0.79	0.46	4
CTD 20	1.31	0.41	2.21	0.83	13	2.34	0.78	3.71	1.36	10
GC 0235+16	1.44	0.69	2.63	0.72	14	1.19	0.86	2.83	0.20	8
OD 166	0.81	0.12	0.92	0.50	15	0.17	0.09	0.34	0.07	11
OE 400	1.24	0.15	1.48	0.91	15	0.88	0.18	1.18	0.63	10

^a The "No." columns give total number of observations included in the database. Ranges in flux density may be caused by variability from experiment to experiment, structure-induced differences among observations in a single experiment, or both.

Table 3 (contd)

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
0306+102	0.19	0.07	0.26	0.08	5	0.21	0.03	0.24	0.16	5
0309+411	0.13	0.03	0.16	0.07	9	0.20	0.09	0.31	0.04	7
3C 84	1.80	0.91	2.88	0.44	6	0.54	0.10	0.68	0.42	4
0326+277	0.17	0.09	0.29	0.08	3	0.11	0.02	0.14	0.09	2
NRAO 140	0.42	0.04	0.47	0.38	3	0.29	0.00	0.29	0.29	1
CTA 26	1.45	0.20	1.70	1.18	8	0.56	0.13	0.71	0.32	6
0342+147	0.24	0.07	0.34	0.11	9	0.16	0.03	0.22	0.11	6
GC 0406+12	0.74	0.22	1.24	0.35	14	0.22	0.05	0.31	0.14	11
P 0420-01	3.32	0.37	3.84	2.68	13	2.03	0.63	3.05	1.10	7
P 0425+048	0.18	0.05	0.28	0.10	16	0.20	0.04	0.26	0.12	10
3C 120	0.29	0.05	0.33	0.24	2	0.00	0.00	0.00	0.00	0
P 0434-188	0.66	0.14	0.84	0.33	10	0.37	0.07	0.43	0.21	8
0440+345	0.82	0.07	0.90	0.75	2	0.35	0.00	0.35	0.35	1
P 0446+11	0.31	0.11	0.47	0.17	7	0.27	0.06	0.37	0.17	6
0454+844	0.32	0.08	0.47	0.16	33	0.13	0.06	0.25	0.04	22
0500+019	0.21	0.04	0.27	0.17	6	0.37	0.10	0.49	0.22	4
P 0506+101	0.55	0.11	0.70	0.38	7	0.36	0.07	0.48	0.27	6
P 0507+17	0.22	0.03	0.27	0.19	11	0.39	0.14	0.70	0.20	7
P 0528+134	1.65	0.21	2.01	1.37	9	0.98	0.47	1.70	0.44	4
0536+145	0.38	0.06	0.49	0.29	14	0.37	0.06	0.43	0.26	9
0544+273	0.20	0.07	0.32	0.08	15	0.29	0.11	0.54	0.14	10
DA 193	1.98	0.73	3.06	0.31	18	1.47	0.52	2.35	0.64	11
0556+238	0.65	0.13	0.81	0.49	11	0.59	0.05	0.65	0.52	7
0600+177	0.36	0.07	0.51	0.28	10	0.31	0.06	0.37	0.22	6
3C 166	0.30	0.07	0.38	0.16	11	0.23	0.05	0.29	0.16	7
0657+172	0.52	0.08	0.66	0.41	14	0.56	0.11	0.76	0.37	9
P 0722+145	0.29	0.06	0.35	0.20	3	0.18	0.05	0.23	0.14	2
DW 0723-00	0.29	0.00	0.29	0.29	1	0.20	0.00	0.20	0.20	1
P 0727-11	1.83	0.34	2.17	0.85	13	1.94	0.31	2.29	1.18	11
P 0735+17	0.81	0.15	1.09	0.58	7	0.86	0.32	1.24	0.46	3
P 0736+01	0.66	0.05	0.71	0.62	2	0.55	0.04	0.59	0.51	2
OI 363	1.71	0.15	2.00	1.50	6	0.37	0.21	0.58	0.16	2
DW 0742+10	2.71	0.24	3.19	2.31	10	0.45	0.06	0.53	0.34	6
GC 0743+25	0.42	0.02	0.45	0.40	4	0.20	0.02	0.22	0.19	2
B2 0745+24	0.39	0.20	0.71	0.04	8	0.44	0.16	0.63	0.23	4
P 0748+126	0.26	0.03	0.28	0.23	2	0.41	0.03	0.44	0.39	2
P 0754+100	0.50	0.06	0.56	0.43	3	0.29	0.01	0.30	0.27	2
OJ 425	1.12	0.11	1.33	0.93	13	0.41	0.08	0.53	0.27	7
P 0823+033	0.75	0.02	0.77	0.73	3	0.81	0.00	0.81	0.81	1
OJ 287	1.46	0.31	1.95	0.87	20	1.82	0.59	2.72	1.04	15
OJ 499	0.43	0.20	0.74	0.21	4	0.22	0.10	0.34	0.09	3
P 0912+029	0.43	0.02	0.45	0.40	3	0.36	0.02	0.38	0.34	2
4C 39.25	1.66	0.35	2.36	1.11	25	1.17	0.33	1.69	0.54	20
1012+232	0.20	0.07	0.28	0.10	4	0.50	0.10	0.63	0.38	3
OL 064.5	0.40	0.02	0.43	0.39	3	0.54	0.08	0.61	0.43	3
3C 245	0.24	0.01	0.26	0.23	3	0.06	0.00	0.07	0.06	3
1044+719	1.57	0.28	1.93	1.12	13	0.59	0.34	1.12	0.20	7
P 1055+01	0.61	0.10	0.83	0.41	17	1.41	0.28	2.12	1.01	12
P 1123+26	0.61	0.08	0.88	0.52	20	0.78	0.09	0.95	0.60	15
P 1127-14	1.13	0.23	1.37	0.74	7	0.15	0.02	0.18	0.12	4
GC 1128+38	0.57	0.11	0.69	0.32	7	0.28	0.10	0.39	0.08	6
P 1222+037	0.61	0.09	0.73	0.44	8	0.33	0.09	0.49	0.25	5
3C 273	3.79	0.60	4.53	2.07	21	1.58	0.64	2.60	0.83	13
P 1244-255	0.54	0.16	0.70	0.38	2	0.42	0.00	0.42	0.42	1
3C 279	3.23	0.24	3.51	2.83	6	1.81	0.52	2.56	1.22	5
B2 1308+32	0.60	0.23	0.94	0.27	27	0.40	0.15	0.80	0.16	18
DW 1335-12	2.17	0.63	3.16	0.89	12	2.20	0.83	3.87	1.11	9
GC 1342+663	0.48	0.07	0.64	0.39	32	0.14	0.06	0.25	0.04	22
OP-192	0.49	0.02	0.51	0.47	4	0.25	0.07	0.32	0.18	2
P 1354+19	0.65	0.13	0.90	0.40	26	0.52	0.19	0.89	0.25	17
OQ 208	0.63	0.03	0.68	0.58	9	0.35	0.06	0.45	0.28	6

Table 3 (contd)

Name	S-band Corr. flux density, Jy				No.	X-band Corr. flux density, Jy				No.
	Avg.	rms	Max	Min		Avg.	rms	Max	X _{min}	
GC 1418+54	0.75	0.19	1.18	0.38	21	0.43	0.23	0.81	0.05	13
P 1445-16	0.61	0.00	0.61	0.61	1	0.16	0.00	0.16	0.16	1
OR 103	1.23	0.26	1.84	0.90	25	0.43	0.12	0.77	0.23	17
P 1504-167	1.14	0.81	2.28	0.51	3	0.18	0.07	0.27	0.12	3
P 1510-08	0.98	0.52	2.19	0.44	10	1.78	0.19	2.02	1.58	4
P 1511-100	0.72	0.13	0.94	0.62	5	0.08	0.01	0.09	0.07	2
P 1532+01	0.47	0.04	0.54	0.42	7	0.21	0.07	0.29	0.08	5
DW 1548+05	1.49	0.17	1.82	1.30	6	0.77	0.19	0.98	0.50	4
DW 1555+00	0.26	0.04	0.35	0.19	19	0.25	0.07	0.38	0.13	15
DA 406	1.36	0.22	1.77	0.95	16	0.93	0.22	1.23	0.69	9
P 1614+051	0.52	0.02	0.56	0.48	6	0.24	0.03	0.28	0.20	5
GC 1633+38	1.72	0.25	2.13	1.33	20	0.88	0.38	1.80	0.31	11
NRAO 512	0.80	0.14	0.99	0.63	15	0.83	0.27	1.28	0.37	7
3C 345	3.20	0.48	4.61	2.24	41	2.17	0.65	4.13	1.36	16
OS 092	0.32	0.04	0.39	0.25	11	0.35	0.06	0.44	0.25	10
DW 1656+05	0.47	0.03	0.50	0.44	2	0.31	0.00	0.31	0.31	2
OT-111	0.21	0.09	0.35	0.10	5	0.17	0.04	0.23	0.14	3
NRAO 530	2.36	0.37	3.01	1.82	19	1.68	0.44	2.39	1.05	10
OT 465	0.67	0.14	0.81	0.24	17	0.30	0.17	0.53	0.05	14
P 1741-038	1.58	0.21	2.03	1.16	30	1.43	0.23	1.86	0.92	17
1749+701	0.20	0.06	0.33	0.13	10	0.11	0.02	0.15	0.08	8
OT 081	1.01	0.17	1.34	0.71	18	1.35	0.23	1.78	0.80	16
1803+784	1.24	0.09	1.33	1.15	2	1.71	0.15	1.86	1.57	2
3C 371	0.53	0.15	0.92	0.35	40	0.36	0.19	1.02	0.10	27
P 1821+10	0.83	0.10	0.95	0.60	16	0.32	0.04	0.40	0.28	7
OV-213	1.77	0.29	2.03	1.07	8	0.91	0.13	1.12	0.71	7
OV-235	1.36	0.41	1.81	0.75	7	1.05	0.35	1.46	0.58	6
OV 239.7	0.33	0.07	0.45	0.22	14	0.28	0.18	0.64	0.07	12
P 1936-15	0.27	0.03	0.32	0.22	7	0.22	0.05	0.26	0.13	5
OV-198	1.83	0.26	2.27	1.40	10	1.10	0.27	1.53	0.74	8
2008-159	0.81	0.04	0.88	0.76	5	0.66	0.05	0.71	0.58	5
OW 637	0.50	0.08	0.65	0.35	8	0.47	0.19	0.80	0.26	5
P 2029+121	0.58	0.06	0.66	0.49	10	0.19	0.06	0.29	0.13	8
OW 551	0.53	0.09	0.69	0.45	4	0.08	0.02	0.11	0.06	2
B2 2113+29B	0.75	0.15	0.96	0.41	12	0.56	0.12	0.80	0.36	9
OX 036	2.40	0.16	2.64	2.18	7	0.61	0.12	0.78	0.42	7
P 2128-12	1.08	0.07	1.17	0.95	5	0.55	0.13	0.77	0.39	5
P 2131-021	0.31	0.05	0.42	0.26	7	0.93	0.11	1.10	0.76	6
P 2134+004	2.34	0.33	2.92	1.83	17	1.12	0.34	1.84	0.48	14
P 2145+06	0.89	0.11	1.05	0.73	6	3.17	0.49	3.83	2.41	6
OX 082	0.60	0.06	0.68	0.47	9	0.21	0.03	0.26	0.15	7
VRO 42.22.01	1.68	0.81	3.21	0.84	21	1.19	0.48	2.05	0.59	14
P 2216-03	1.85	0.12	2.03	1.62	10	0.74	0.11	0.89	0.53	8
3C 446	1.37	0.03	1.41	1.34	2	1.76	0.19	1.95	1.57	2
P 2227-08	0.78	0.05	0.87	0.74	4	1.05	0.20	1.30	0.73	4
2229+695	1.05	0.12	1.25	0.86	13	0.35	0.14	0.62	0.13	9
CTA 102	0.80	0.57	2.19	0.13	22	0.42	0.17	0.74	0.06	17
GC 2234+28	1.22	0.22	1.58	0.38	24	0.42	0.18	0.71	0.06	19
OY-172.6	1.00	0.14	1.17	0.73	8	0.80	0.11	0.97	0.61	6
3C 454.3	4.98	1.62	6.97	1.94	12	1.57	1.04	3.10	0.16	9
GC 2253+41	1.06	0.11	1.15	0.91	3	0.31	0.07	0.39	0.24	2
P 2254+024	0.28	0.03	0.31	0.23	6	0.20	0.04	0.25	0.14	6
GC 2318+04	0.66	0.06	0.77	0.60	7	0.42	0.06	0.51	0.34	7
P 2320-035	0.55	0.04	0.63	0.44	15	0.35	0.10	0.56	0.18	13
P 2345-16	0.58	0.22	0.95	0.22	11	0.50	0.13	0.73	0.37	9
P 2355-106	0.59	0.18	0.86	0.30	18	0.56	0.09	0.74	0.42	12

Table 4. Correlated flux density classes for all sources listed in Table 1^a

Name	Class						Name	Class					
	Goldstone-Canberra			Goldstone-Madrid				Goldstone-Canberra			Goldstone-Madrid		
	S	X	S/X	S	X	S/X		S	X	S/X	S	X	S/X
P 0008-264	3	3	3			—	0657+172	3	2	3	3	2	3
P 0019+058	4	4	4	3	3	3	P 0722+145	3		(4)	3	4	4
P 0048-09	2	1	2	2	2	2	DW 0723-00	4	4	4	3	3	3
P 0104-408	1	1	1			—	P 0727-11	1	1	1	1	1	1
P 0106+01	1	2	2	1	2	2	P 0735+17	2	2	2	2	2	2
P 0111+021			(4)	4	4	4	P 0736+01	1		(2)	2	2	2
P 0112-017	1	1	1	2	2	2	OI 363	1	3	3	1	3	3
P 0113-118	1	2	2	1	4	4	DW 0742+10	2	2	2	1	2	2
P 0119+11	4	3	4	2	2	2	GC 0743+25	3		(3)	3	3	3
GC 0119+04			(3)	2	2	2	B2 0745+24	2		(2)	3	3	3
DA 55	3	1	3	2	2	2	P 0748+126	3	3	3	3	2	3
0146+056	2	3	3	2	4	4	P 0754+100	3		(3)	3	3	3
P 0201+113			(3)	2	3	3	OJ 425	1	2	2	1	3	3
P 0202+14			(3)	3	2	3	P 0823+033	2	2	2	2	1	2
0212+735			—	3	3	3	B2 0827+24			(3)			(3)
GC 0221+06	3	2	3	4	2	4	4C 71.07			—			(4)
DW 0224+67			—	3	2	3	OJ 287	1	1	1	1	1	1
P 0229+13	2	2	2	2	2	2	OJ 499	4	4	4	3	4	4
CTD 20	2	1	2	1	1	1	P 0912+029	3		(3)	3	3	3
GC 0235+16	1	2	2	2	2	2	4C 39.25	1	1	1	1	1	1
OD 166	1	3	3	2	4	4	1012+232	2		(3)	4	3	4
OD 094.7			(3)			(3)	P 1034-293	2	1	2			—
OE 400	1	2	2	1	2	2	OL 064.5	3	2	3	3	2	3
0306+102	4	4	4	4	3	4	3C 245			(3)	3		(3)
0309+411	3	4	4	4	4	4	1044+719			—	1	3	3
3C 84	1	4	4	1	2	2	P 1055+01	2	1	2	2	1	2
0326+277	4	4	4	4	4	4	P 1104-445	1	2	2			—
P 0332-403	2	2	2			—	GC 1111+14			(4)			(4)
NRAO 140	2	3	3	3	3	3	P 1123+26	2	2	2	2	2	2
CTA 26	2	2	2	1	2	2	P 1127-14	3	4	4	2	4	4
0342+147	3	4	4	3	4	4	GC 1128+38	2	3	3	2	3	3
P 0402-362	3	1	3			—	P 1144-379	3	2	3			—
GC 0406+12	2	4	4	2	4	4	P 1148-00			(4)			(4)
P 0420-01	1	1	1	1	1	1	P 1222+037	2	2	2	2	3	3
VRO 41.04.01			(3)			(3)	3C 273	1	1	1	1	1	1
P 0425+048	3	4	4	4	4	4	3C 274			(3)			(3)
3C 120		3	3	3	3	(4)	P 1244-255	3	2	3	3	2	3
P 0434-188	2	2	2	2	3	3	3C 279	1	2	2	1	1	1
P 0438-43	1	2	2			—	B2 1308+32	2	2	2	3	3	3
0440+345	2	4	4	2		(4)	OP-322	3	1	3			—
P 0446+11	3	3	3	3		(3)	DW 1335-12	1	1	1	1	1	1
P 0451-28			(4)			—	GC 1342+662			—			(4)
0500+019		3	(4)	4	3	4	GC 1342+663			—	3	4	4
0454+844			—	3	4	4	P 1349-439	3	3	3			—
P 0506+101	3	3	3	3	3	3	P 1354+19	2	2	2	2	3	3
P 0507+17	4		4	3	3	3	OP-192	2		(2)	3	4	4
P 0528+134	1	3	3	1	2	2	OQ 208	4	4	4	2	3	3
P 0537-441	1	1	1			—	GC 1418+54			—	2	3	3
0536+145	3	3	3	3	3	3	OQ-151	2	4	4			(4)
0544+273	4	4	4	4	3	4	P 1445-16			—	2	4	4
DA 193	1	1	1	1	1	1	OR 103	1	2	2	1	3	3
0556+238	3	2	3	2	2	2	P 1504-167	1	2	2	2	4	4
0600+177	3	4	4	3	3	3	P 1510-08	1	1	1	2	1	2
P 0607-15			(4)			(4)	P 1511-100	2	4	4	2	4	4
3C 166			(4)	3	4	4	P 1514-24	3	3	3			(3)

^a Class 1 sources are observable with two 34-m antennas, class 2 sources require at least one 70-m antenna, and class 3 sources require two 70-m antennas for observation with the navigation VLBI system. Class 4 sources are too weak even for two 70-m antennas. Values are given in parentheses for sources with insufficient data in the 1986-1988 period. Dashes indicate sources that are not visible on the specified baseline.

Table 4 (contd)

Name	Class						Name	Class					
	Goldstone-Canberra			Goldstone-Madrid				Goldstone-Canberra			Goldstone-Madrid		
	S	X	S/X	S	X	S/X		S	X	S/X	S	X	S/X
P 1519-273	1	1	1			—	P 2008-159	3	3	3	2	2	2
P 1532+01			(3)	3	3	3	OW 637			—	3	3	3
DW 1548+05	2	2	2	1	2	2	OW 551			—	3		(3)
DW 1555+00	3	3	3	3	3	3	P 2029+121	3	3	3	2	4	4
DA 406	2	2	2	1	2	2	B2 2113+29B	3	2	3	2	2	2
P 1614+051	3	3	3	2	3	3	OX 036	1	2	2	1	2	2
GC 1633+38	1	1	1	1	2	2	P 2128-12	3	2	3	1	2	2
NRAO 512	2	2	2	2	2	2	P 2131-021	3	1	3	3	1	3
3C 345	1	1	1	1	1	1	P 2134+004	2	1	2	1	1	1
OS 092	2	1	2	3	3	3	P 2145+06	1	1	1	2	1	2
DW 1656+05			(3)	3	3	3	OX 082	2	3	3	2	3	3
1657-261	2	1	2			—	OX-192	3	2	3			(3)
OT-111	4	4	4	4	4	4	VRO 42.22.01	3	3	3	1	2	2
GC 1717+17			(3)			(3)	P 2216-03	1	3	3	1	2	2
NRAO 530	3	1	3	1	1	1	3C 446	2	1	2	1	1	1
OT 465	2	4	4	2	4	4	P 2227-08	3	1	3	2	1	2
P 1741-038	1	2	2	1	1	1	2229+695			—	2	3	3
1749+701			—	4	4	4	CTA 102	1	3	3	3	3	3
OT 081	2	1	2	2	1	2	GC 2234+28	1	3	3	1	3	3
1803+784			—	1	1	1	OY-172.6	2	3	3	2	2	2
3C 371			—	3	3	3	P 2245-328	3	3	3			—
P 1821+10	2	3	3	2	3	3	3C 454.3	1	1	1	1	1	1
OV-213	2	3	3	1	2	2	GC 2253+41	4	4	4	2	3	3
OV-235	2	2	2	2	2	2	P 2254+024	3	3	3	3	4	4
OV-236	1	1	1			—	GC 2318+04	2	2	2	2	3	3
OV 239.7	3	3	3	3	4	4	P 2320-035	3	3	3	2	3	3
P 1936-15	4	3	4	3	3	3	P 2345-16	1	3	3	3	2	3
OV-198	1	1	1	1	1	1	P 2355-106	3	2	3	3	2	3

Table 5. Portion of global source-finding list for ecliptic sources. The most accurate positions available in the 1950 coordinate system are given.^a

Source	Right Ascension			Declination			Total Flux Density, Jy				Survey	VLBI	CEI	Notes
	h	m	s	°	'	"	5 GHz	2.7 GHz	1.4 GHz	0.4 GHz				
0000-177	00	00	48.6	-17	43	57.	0.95	1.37	2.2		6	N	N	
0001+128	00	01	07.06	12	49	58.1	0.41	0.66	1.12		20	N	Y	
0002+125	00	02	24.74	12	31	32.5	0.56		1.70		8	N	Y	
0003+158	00	03	26.7	15	52	38.	0.43	0.7	0.9	2.7	20	N	Y	
0003-066	00	03	40.293	-06	40	17.30	1.58	1.46	1.53		3	Y	Y	
0003-003	00	03	48.8	-00	21	06.	1.41	2.40	3.36		8	Y	Y	3C2
0005-199	00	05	43.8	-19	56	30.	0.25	0.45	0.53		3	N	N	
0005-062	00	05	55.8	-06	15	25.	0.33	0.68	1.2		6	N	N	
0007+124	00	07	18.5	12	28	40.	0.68	1.2	1.9		22	N	Y	
0007+106	00	07	57.9	10	41	30.	0.47		0.34			Y	Y	
0007+171	00	07	59.383	17	07	37.50	0.92	0.91	0.83		8	Y	Y	
0009+081	00	09	01.7	08	07	14.6	0.36	0.49			8	N	Y	
0010+005	00	10	37.40	00	35	09.2	0.50	0.95	1.53		10	Y	Y	3C5
0011-046	00	11	20.64	-04	40	32.7	0.41	0.21			6	Y	N	
0011-096	00	11	24.3	-09	36	38.	0.26	0.40	0.82		6	N	N	
0011+054	00	11	29.3	05	28	46.	0.47	0.70	1.29		20	N	N	
0012-184	00	12	30.4	-18	28	48.	0.42	0.48	0.52		6	N	N	
0013-197	00	13	28.4	-19	45	56.	0.27	0.48	0.69		6	N	N	
0013-005	00	13	37.347	-00	31	52.50	0.68	0.87	0.91		10	Y	Y	

^aVLBI and CEI indicate whether or not a source has been observed with VLBI or connected-element interferometry. List was last updated 14 July 1989. Selection criteria as of 18 January 1988:

- (1) All sources within 20 deg of ecliptic fulfilling the following:
 - (a) Five-GHz flux density above 250 mJy in MIT-GB survey.
 - (b) Five-GHz flux density above 250 mJy in Parkes surveys, or if no 5-GHz flux density is given in Parkes survey or other surveys such as MIT-GB survey, flux density above 250 mJy at 2.7 GHz.
 - (c) Any S-/X-band VLBI observations, regardless of total flux density.
 - (d) Five-GHz flux density above 250 mJy in other surveys (e.g., S surveys).
- (2) Exceptions:
 - (a) Declinations generally limited to +40 deg in north and to -35 deg in south.
 - (b) Some sources slightly more than 20 deg from ecliptic are included if they have lots of potential or are in otherwise empty regions (e.g., 0420-014). A few sources outside of +40-deg or -35-deg declination are also included.
 - (c) Sources observed specifically for ecliptic reasons (e.g., Phobos lander frame-tie candidates) are included even if their flux densities are below the selection limit.
 - (d) Sources in selected regions chosen from Condon-Broderick 1.4-GHz survey are included down to 200 mJy at 1.4 GHz. Even those sources which turn out to have 5-GHz flux densities below 250 or 300 mJy are still included.
 - (e) Of course, some other random sources are included for various reasons.
 - (f) The 0.4-GHz column is fairly incomplete, especially in cases where there are fluxes at two or three of the higher frequencies. In such cases, it was usually decided that it was not worth the trouble of finding a flux density at the lowest frequency, since it would add no useful information.

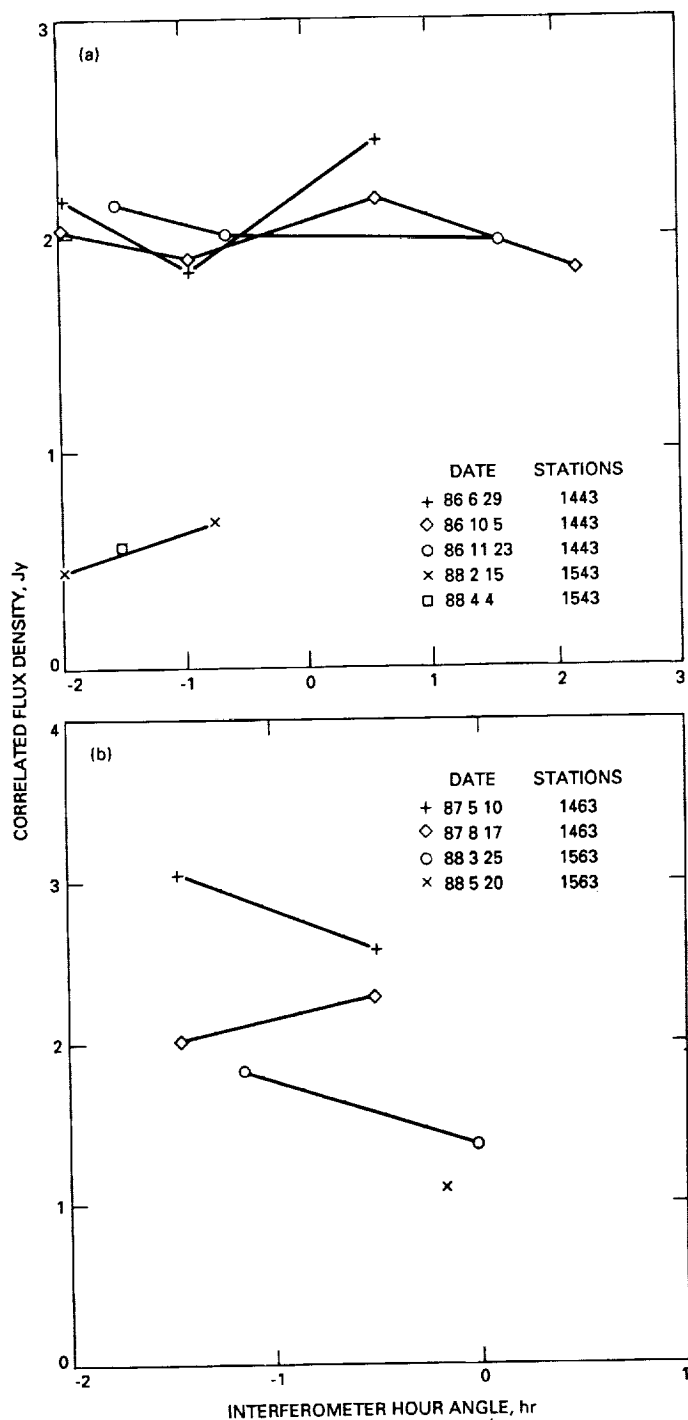


Fig. 1. Correlated flux density versus hour angle for P0420-014 at X-band: (a) on the California-Australia baseline, (b) on the California-Spain baseline.

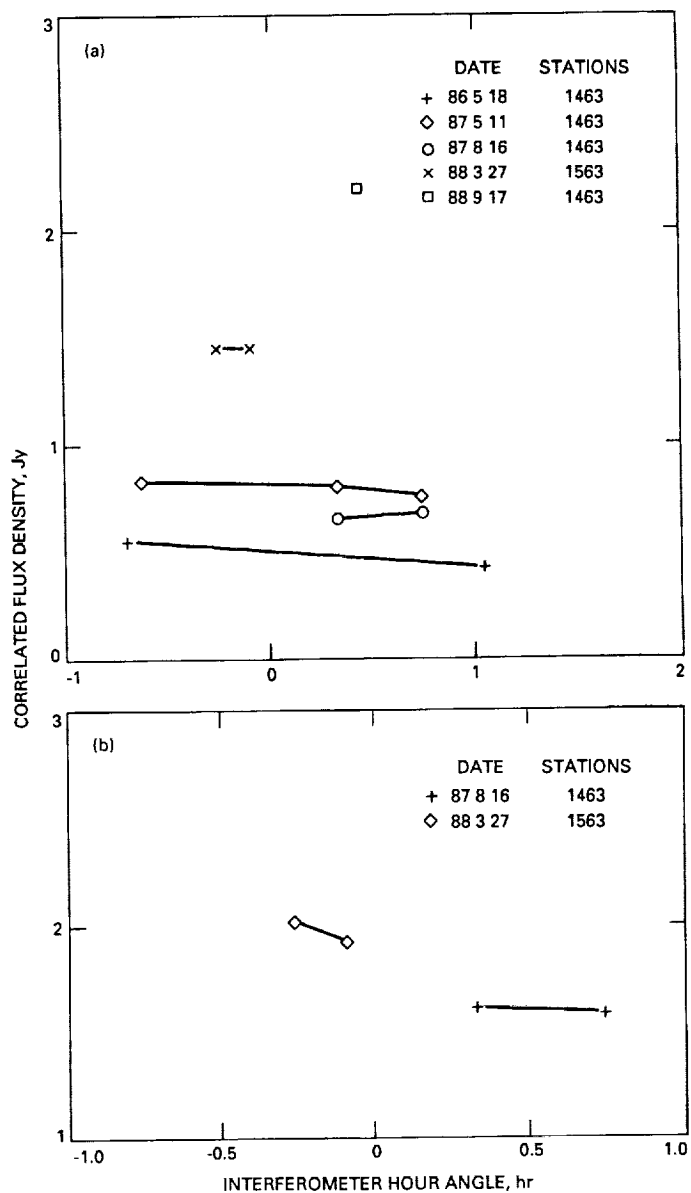


Fig. 2. Correlated flux density versus hour angle for P1510-08: (a) at S-band, (b) at X-band.

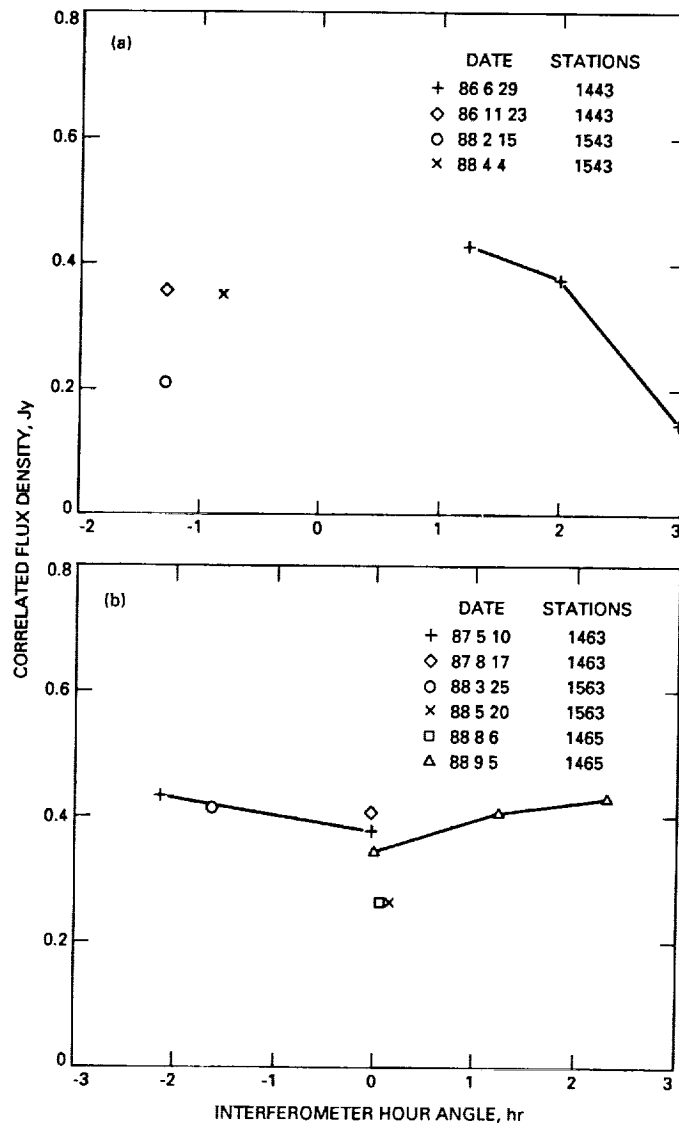


Fig. 3. Correlated flux density versus hour angle at X-band for 0536+145: (a) on the California-Australia baseline, (b) on the California-Spain baseline.

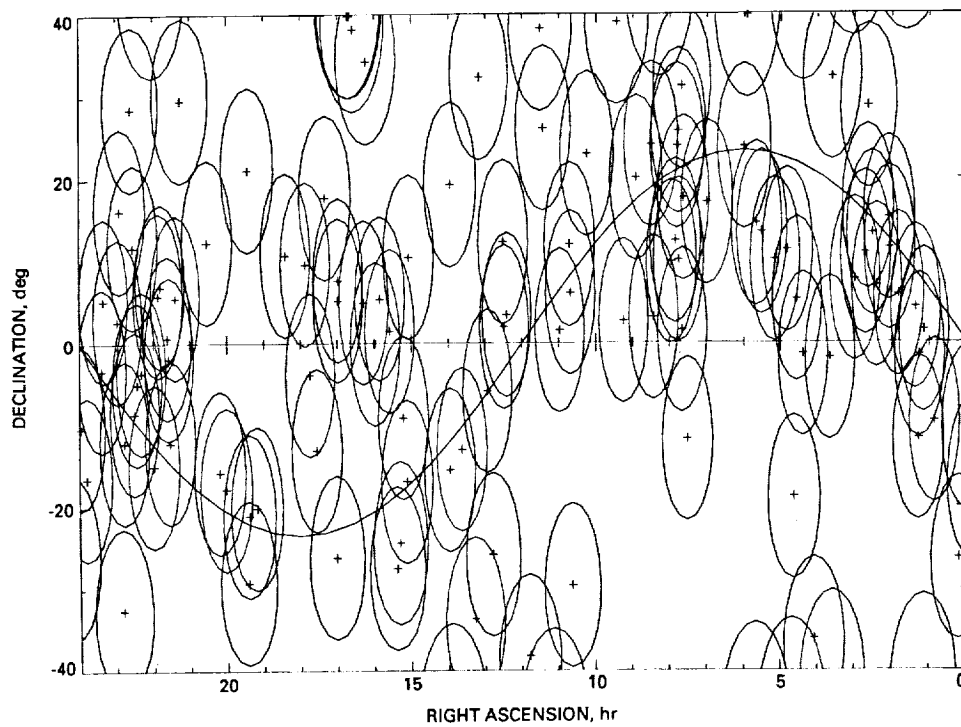


Fig. 4. Class 1, 2, and 3 sources on the California-Australia baseline are plotted on a projection of a portion of the celestial sphere. Circles of 10-degree radius are plotted at each source position, and the ecliptic plane is shown as the sinusoidal curve. These sources have correlated flux densities of at least 0.2 Jy at both S- and X-bands.

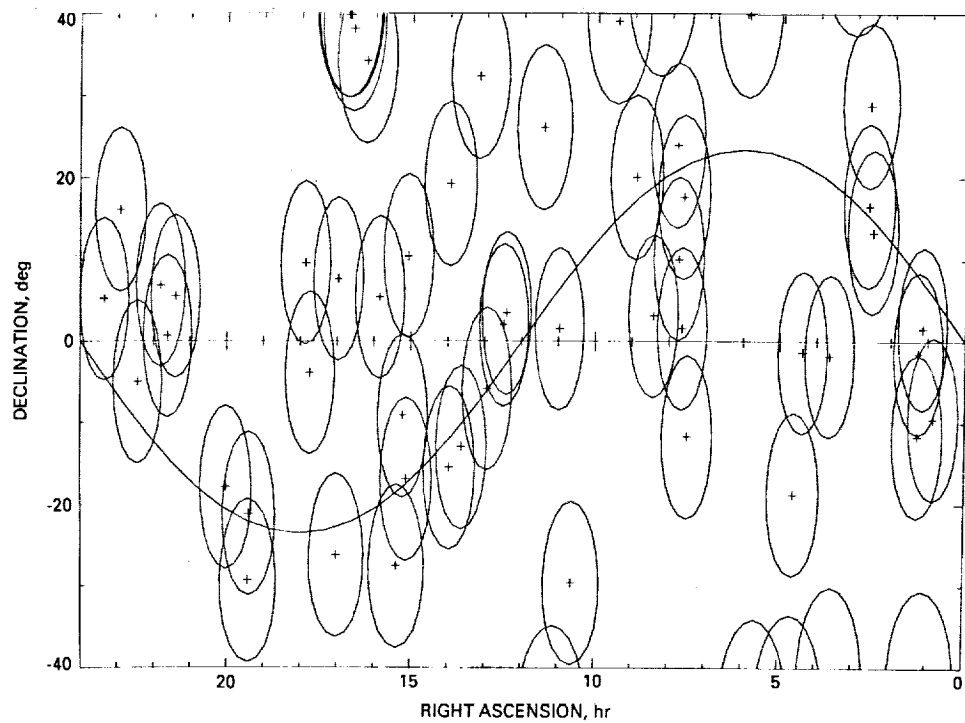


Fig. 5. Class 1 and class 2 sources on the California–Australia baseline are plotted on a projection of a portion of the celestial sphere. Circles of 10-degree radius are plotted at each source position. These sources have correlated flux densities of at least 0.5 Jy at S-band and 0.4 Jy at X-band. Similar to Fig. 4.

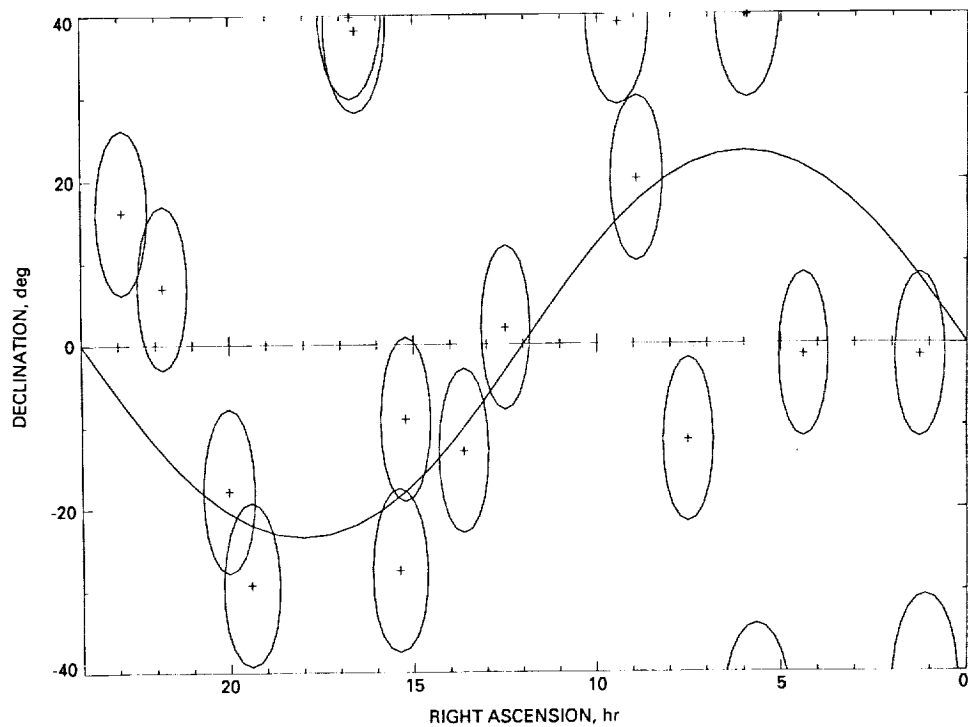


Fig. 6. Class 1 sources on the California-Australia baseline are plotted on a projection of a portion of the celestial sphere. These sources have correlated flux densities of at least 1.0 Jy at S band and 0.8 Jy at X-band. Similar to Fig. 4.

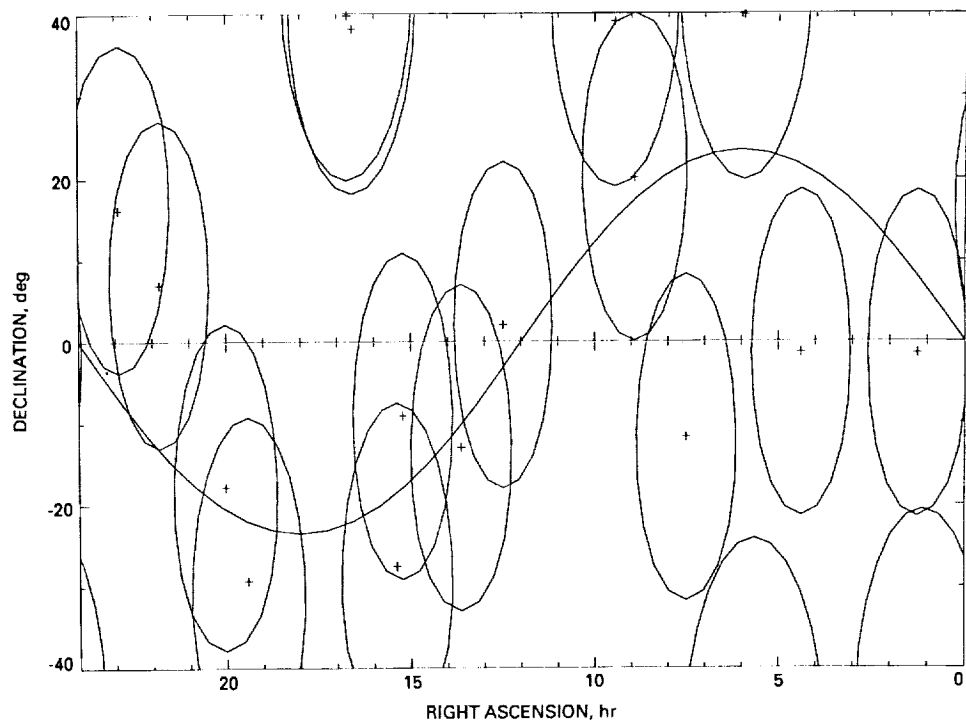


Fig. 7. Circles of radius 20 degrees are plotted at the position of each class-1 source on the California-Australia baseline. Similar to Fig. 4.

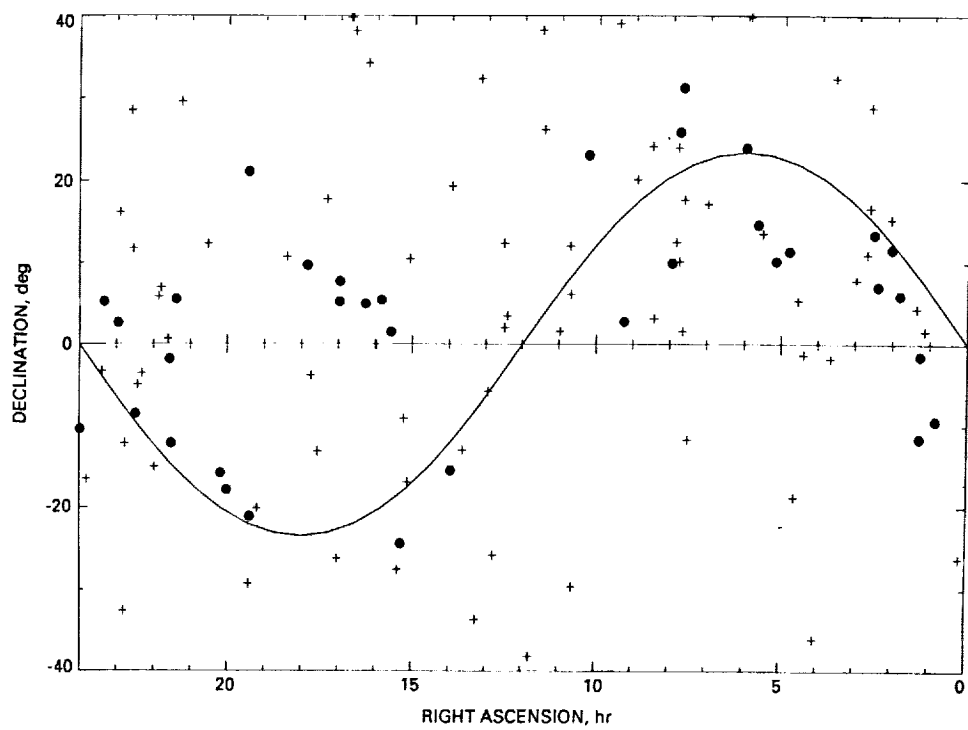


Fig. 8. Plot of each source in classes 1, 2, or 3 on the California-Australia baseline. Each new source is indicated by a \bullet and each old source is indicated by a $+$.

An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation

D. L. Trowbridge, J. R. Loreman, T. J. Brunzie, and R. Quinn
Radio Frequency and Microwave Subsystems Section

An 8.4-GHz front-end system consisting of a feedhorn, a waveguide feed assembly, dual masers, and downconverters was reimplemented at Parkes as part of the Parkes Canberra Telemetry Array for the Voyager Neptune encounter. The front-end system was originally assembled by the European Space Agency and installed on the Parkes antenna for the Giotto project. It was also used on a time-sharing basis by the Deep Space Network as part of the Parkes Canberra Telemetry Array to enhance the data return from the Voyager Uranus encounter. At the conclusion of these projects in 1986, part of the system was then shipped to JPL on loan for reimplementation at Parkes for the Voyager Neptune encounter. New design and implementation required to make the system operable at Parkes included new microwave front-end control cabinets, closed-cycle refrigeration monitor system, noise-adding radiometer system, front-end controller assembly, X81 local oscillator multiplier, and refurbishment of the original dual 8.4-GHz traveling-wave masers and waveguide feed system. The front-end system met all requirements during the encounter and was disassembled in October 1989 and returned to JPL.

I. Introduction

The front-end system was originally assembled by the European Space Agency (ESA) for installation on the Parkes antenna in support of the Giotto project. The radio frequency (RF) package, containing a waveguide feed system, dual traveling-wave maser and closed-cycle refrigerator assemblies (TWM/CCRs), and the TWM monitor and control instrumentation, was built by the Airborne Instrument Laboratories (AIL) division of Eaton Corporation. JPL provided AIL with the waveguide feed system, the TWM design based on the JPL Block II-A TWM [1], the TWM monitor and control instrumentation design, and technical consulting during the manufacturing and testing of the system. As a result of an agreement between

NASA/JPL and ESA, the front-end system was also used on a time-sharing basis by the Deep Space Network (DSN) as part of the Parkes Canberra Telemetry Array (PCTA) to enhance the data return from Voyager 2 at Uranus. At the conclusion of these projects in 1986, the front-end system was dismantled, packed, and shipped to Europe. Part of the system was then shipped from ESA to JPL on loan for reimplementation at Parkes for the Voyager Neptune encounter.

New design and implementation tasks required to make the system operable at Parkes for the Voyager Neptune encounter included new microwave front-end control cabinets, CCR monitor system, noise-adding radiometer

(NAR), front-end controller (FEC), X81 local oscillator multiplier, and refurbishment of the original dual 8.4-GHz TWMs and waveguide feed system.

During the Voyager Neptune encounter, Canberra Deep Space Communications Complex (CDSCC) personnel were responsible for the maintenance and operation of the Parkes antenna front-end microwave electronics. A decision was made not to replicate the ESA-designed monitor and control system, and a new system was built around an Intel Multibus computer similar to that used in the Parkes/CDSCC Telemetry Array (PCTA) equipment. Advantages of this approach included automated operation at the PCTA Parkes site, the availability of health and status information (and to some extent, control) at the Canberra site (SPC-40), and the ability to share common spares with PCTA equipment.

In addition to the front-end monitor and control functions, an NAR function was added to the system. Intended primarily to aid antenna pointing calibration procedures, the NAR was capable of monitoring system-noise temperature during either pre-pass or (in an optional mode) telemetry tracks.

The front-end system met all requirements during the encounter and was disassembled in October 1989 and returned to JPL.

II. Reimplementation for Parkes/Neptune

Reimplementation of the front-end system for the Neptune encounter required a substantial amount of design, fabrication, and procurement due to the absence of some key hardware components. Specifically, all Hewlett-Packard (HP) commercial equipment had been removed from the 1986 system with the following consequences:

- (1) There was no means of controlling or reading the position of the waveguide switches or polarizer (other than manually).
- (2) There was no means of monitoring the health and status of the CCRs and compressors other than by personal inspection.
- (3) There was no means of monitoring the health and status of the downconverters, monitor receiver, or upconverter other than by personal inspection.
- (4) There was no automation of system configuration or calibration.

Some consideration was given to purchasing all the missing hardware and software so that the system could

be rebuilt and operated exactly as it was in 1986; however, this idea was abandoned due to financial and practical considerations.

A number of fundamental constraints applied to this reimplementation task:

- (1) The equipment supplied by ESA was on loan to JPL and any modifications made to the equipment required prior approval from ESA or had to be reversible prior to returning the equipment to ESA.
- (2) There were a limited number of cables available through the wrap on the Parkes antenna so any new design was ideally to use no more cables.
- (3) The physical space for new equipment was limited and was not to exceed that used in 1986, if possible.
- (4) Equipment weight in the aerial cabin needed to be minimized to avoid damage to the focus drive gear. The total weight in 1986 was considered excessive.
- (5) There was no plenum air conditioning available in the aerial cabin or in the NASA trailer; however, a limited amount of room air conditioning was available.
- (6) Remote monitor and control of many parameters was mandatory because of the normally inaccessible location of the front end (at prime focus), and the fact that the equipment was unmanned during much of the pre-encounter period.
- (7) The usual constraints of operability, reliability, and maintainability also applied.

The preliminary system design has been previously reported [2].

III. Parkes Front-End System

A. General Description

The configuration of the front-end system, described below, is similar to that for the 1986 encounter, with the exception of pre-/post-TWM signal-coupling ports provided for the Commonwealth Science Industrial Research Organization (CSIRO) of Australia, who operate and maintain the Parkes antenna, and two new local oscillator frequency-multiplier chains (X81) that drive the two RF-IF (intermediate frequency) downconverters.

The front-end system configuration, shown in Fig. 1, consisted of the following:

- (1) Aerial cabin equipment, which included the feed and all microwave components, traveling-wave maser low-noise amplifiers and their closed-cycle refrigerators, downconverters, an upconverter, test-signal switching, noise-diode assemblies, and a monitor receiver.
- (2) NASA trailer control room equipment, which included all maser controls, switching and all monitor and control for the front-end system.
- (3) Antenna pedestal equipment, which consisted of the helium compressors for the CCRs, and an FEC remote terminal which supported antenna pointing calibrations.

B. Aerial Cabin Equipment

The RF package, shown in Figs. 2 and 3, contained the feed system and the dual-maser amplifiers. Three aerial cabin racks, A, B, and C (Fig. 4), contained the balance of front-end RF and TWM/CCR monitor and control equipment. This equipment must be in close proximity to the TWM/CCRs and waveguide feed system, and would require excessive cables to relocate the equipment. The total weight of the front-end equipment, including the RF package installed in the aerial cabin, was 630 kilograms.

A block diagram of the aerial cabin RF equipment is shown in Fig. 5. The feedhorn assembly was located at the primary focus of the antenna and was connected to a rotatable polarizer and orthomode transducer. The system provided two identical receive channels for redundancy, using TWMs based on the JPL Block II-A TWMs and operating at 8425 MHz (nominal) with 100-MHz bandwidth. Downconversion to 325 MHz (nominal) was accomplished with two identical downconverters, using fixed-frequency local oscillators at 8100 MHz phase-locked to the station 100-MHz timing. Either one of the 325-MHz downconverter outputs could be selected as the input to the short-loop telemetry receiver (not part of the front-end system).

A single monitor receiver with switchable input was provided for monitoring maser gain bandwidth response. The X-band input signal to the monitor receiver was obtained from couplers located in each of the downconverters. The local oscillator for the monitor receiver was also obtained from the downconverters, and hence the monitor receiver input switching selected an X-band input signal and local oscillator as a pair. Output was at 325 MHz (nominal).

The test-signal injection system consisted of a programmable synthesizer operating at 425 MHz (nominal), an upconverter known as the X-band test generator, and a test-signal switching and distribution network that allowed selection of the test signal to either maser input or output.

The synthesizer output could be phase modulated with high-rate data from an external source and was coherent with station timing. Upconversion to 8425 MHz was accomplished with an 8000-MHz fixed-frequency local oscillator, which was also coherent with station timing. A wideband output (for maser bandpass measurement) and a narrowband output (for telemetry testing) was provided on the upconverter.

The switching network consisted of a monitor/test-signal assembly and two TWM calibration assemblies based on JPL designs.

C. NASA Trailer Control Room

Two new RF front-end control (RFFEC) cabinets were installed in the NASA trailer. These cabinets are shown in Fig. 6 and a control flow diagram is shown in Fig. 7. The two new cabinets required new multi-conductor cables between the antenna pedestal control room and the NASA trailer. The two RFFEC cabinets were the maintenance points for tuning and adjusting the dual TWMs. The cabinets contained the monitor receiver display (an HP Spectrum Analyzer 8590A), the 420-MHz nominal test-signal source (an HP Synthesizer Signal Generator 8663A), the front-end controller including a local terminal, the CCR monitor system, and the TWM monitor and control equipment.

D. Detailed Functional Descriptions

1. Front-end controller. Reimplementation of the Parkes antenna front-end system centralized and automated the monitor and control of the microwave electronics. This monitor and control was provided by the front-end controller (FEC) assembly, a multibus-based computer mounted in control cabinet no. 2. The FEC tied together individual pieces of equipment in the trailer, some of which interfaced with hardware located in the antenna aerial cabin.

Control of the FEC was provided through a cathode-ray tube (CRT) terminal located in control cabinet no. 1. A second CRT was located in the antenna pedestal's 3rd floor, dedicated for use during antenna pointing calibration. A third CRT was located at CDSCC SPC-40 and communicated with the FEC through modem and data

link between Parkes and Canberra. All three of these terminals were able to send commands and to monitor front-end status at any time; however, only one terminal at a time could be used as the master terminal. The master terminal was able to lock the front-end configuration for security during actual tracking missions. Selection of the master terminal was through a switch mounted on the front of the FEC chassis. Terminal CRT displays could be printed out on inkjet printers, which provided a permanent record of all data, commands, and informational graphics.

Commands for the FEC were grouped according to function. Each group had a menu (one example is shown in Fig. 8), and one or more status displays. Commands were listed in the menus, along with a brief description of the function of the command. Syntax and range checking were enforced, with individual error statements for each command. Some commands also responded to a query; entering a question mark as a parameter would display the current variable values as a one-line status display.

The FEC provided cabling interfaces with the following front-end monitor and control instrumentation (Figs. 6 and 7):

- (1) IEEE-488 cabling from the FEC central processing unit (CPU) to an HP quartz thermometer, test-signal synthesizer, spectrum analyzer, and power meter, and (through an IEEE-488 extender to the aerial cabin) the noise-diode relay switch controller, the test oscillator upconverter, the ac power controller for the waveguide switches, the monitor receiver, and each of the two downconverters
- (2) Parallel transistor-transistor logic (TTL) control lines from the FEC CPU to the dual TWM control panel
- (3) RS-232 serial communication lines from the FEC serial ports to the following terminals:
 - (a) Communications modem (link with SPC-40 CRT)
 - (b) A local dumb terminal for front-end operation and maintenance
 - (c) A remote dumb terminal in the antenna to support NAR operation during antenna pointing calibration
 - (d) A CCR monitor (described in a following section)

- (4) 50-ohm coaxial cable from the PCTA receiver upconverter drawer to the power meter board in the FEC, and from the FEC to the noise-diode power supplies in the aerial cabin.

Front-end system tasks performed by the FEC included:

- (1) Monitor and control of the RF front-end waveguide switches and polarizer drive assembly via IEEE-488 using an HP 3488A remote switch/control unit with plug-in HP 44471A relay modules for power control and an HP 44474A digital input/output (I/O) module for telltale monitoring showing the true mechanical position of the devices (not provided in 1986). Manual control in the event of hardware failure was possible from the front panel of the HP 3488A in the aerial cabin. The status of these switches could be displayed on the terminal CRTs (see Fig. 9).
- (2) System temperature measurement (see Fig. 10) using the Y-factor technique through automation of the waveguide switches, quartz thermometer, dual TWM control assembly, and HP power meter (see Fig. 11 for output display).
- (3) Measurement and graphical display of TWM 1 and 2 gain versus frequency. The data was stored and could be retrieved as shown in Fig. 12.
- (4) Performance of NAR functions under remote control during antenna calibration via automation of the noise-diode assemblies and the digital power meter. This task included returning a time-varying analog signal to the pedestal, corresponding to measured system temperature, as well as periodic calibration of the noise diodes.
- (5) Providing status updates and charting performance history of the refrigerators and compressors, as requested. This task included generating alarm messages based on CCR out-of-limits conditions provided to the FEC by the CCR monitor.
- (6) Detection of alarms from the downconverters, monitor receiver, and test-signal upconverter. These included power supplies, local oscillator power levels, timing, and phase-lock conditions.

2. Noise-adding radiometer. Noise-adding radiometers (NARs) operate by periodically injecting small, known quantities of noise into the antenna front end, then measuring the resulting increase in system-noise power at the receiver. For the Parkes implementation, this additive noise was generated using a DSN noise-diode assembly; the power meter in the front-end controller performed

the noise-power measurements. During antenna-pointing NAR operations, system temperature calculations were carried out continually, with the FEC supplying a corresponding analog voltage to the antenna pedestal via a digital-to-analog (D/A) converter.

Two noise-diode assemblies were supplied for the Parkes antenna, one for each of the two X-band receive chains. Each assembly consisted of a noise-diode oven and associated power supply. Each oven contained three diodes, providing noise temperatures of 0.25 K, 0.5 K, 1 K, 2 K, 4 K, 8 K, and 50 K. The ovens were controlled through their power supply assemblies, each of which contained three independent power supplies, one for each diode. Three relays per supply were used to select the amount of diode current (allowing three noise levels per diode), while a fourth TTL signal modulated the diode on and off. Each power supply assembly was monitored and controlled through 21 digital I/O lines, consisting of relay closures, telltale sense, and diode-modulation input. Both assemblies were operated through an HP 3488A remote switch controller containing three HP 44474A digital I/O cards (16 channels per card); coaxial cable run directly from the FEC supplied the modulation control signals. Since an identical HP controller was used to operate the waveguide switches, common spares existed for both the switch controllers and the I/O cards.

At the other end of the receive chain, noise-power measurements were taken from the inputs of the Parkes telemetry receiver. Each of the two 285–360-MHz RF signals were split 3 dB in the signal-select drawer, and then one output arm was fed directly into the power meter in the FEC.

The NAR power meter was a DSN precision power meter (PPM) NAR RF assembly, containing an RF switch, a square-law diode, and a voltage-to-frequency converter. The output frequency was counted and averaged within the FEC, using a PPM frequency-counter board.

Diode control, noise measurement, system-temperature computation, and analog output programming were all handled by the FEC CPU, with results included in FEC status displays.

3. Closed-cycle refrigerator monitor. The closed-cycle refrigerator (CCR) monitor system consisted of three μ MAC 5000¹ data acquisition assemblies and an IBM PC-XT data acquisition unit (DAU), used for data display, processing, and storage [3]. The DAU was located

in front-end control cabinet no. 2 (Fig. 6) located in the NASA trailer. One μ MAC was located in the aerial cabin to monitor both CCRs, while the other two μ MACs were located in the compressor room in the antenna pedestal (Fig. 13) to monitor one compressor each. Data were sent from the μ MACs to the DAU at one-minute intervals.

Communication between the μ MACs and the DAU was serial RS-232 and utilized a coaxial cable daisy-chained to all devices. An additional RS-232 port on the DAU provided the interface to the FEC via the serial-communications card.

The DAU was dedicated solely to CCR system-monitor tasks. The DAU continuously displayed CCR and compressor parameters, updating the screen once a minute as data were received from the CCR system. These parameters were also echoed to the FEC, the most up-to-date values being used to generate routine PCTA status displays.

Monitored CCR/compressor parameters included:

- (1) CCR stage temperatures (nominally 4.5 K, 15 K, and 70 K)
- (2) Refrigerator reserve capacity (percentage of normal)
- (3) Helium pressures (supply, refrigeration return, Joule-Thomson return, tank, and oil stack differential)
- (4) Third-stage, i.e., Joule-Thomson helium flow
- (5) Compressor temperatures (first- and second-stage)
- (6) Vacuum pressure in CCR vessel
- (7) Compressor motor phase currents (A, B, C)

In addition to the CCR/compressor parameter values, each record sent by the DAU to the FEC was formatted with status flags indicating whether each value was within acceptable limits. (All CCR/compressor alarm limits were entered at the DAU keyboard.) Any out-of-limit conditions caused the FEC to immediately send a PCTA-compatible alarm message (including audible tone) to each FEC terminal.

In order to track performance history, the DAU time-tagged and logged the CCR/compressor data to its hard disk every 15 minutes. An entire year's worth of data (96 records/day) could be recorded. In addition to the long-term storage in the DAU, the FEC stored data received

¹ Trademark of Analog Devices, Inc.

during the preceding 72 hours (transmitted in one-minute intervals).

This performance history was made available in two forms; for both forms, one or more parameters were displayed in columns upon specification of a start date/time and a stop date/time. Short-term history (up to 72-hours old) had a maximum resolution of one minute; long-term history resolution was 15 minutes. Data records that were not found in the 72 hour FEC short-term cache could be automatically requested from the DAU hard disk log, indexed by day of year, time of day, and CCR/compressor system number.

4. Alarm monitoring. The downconverters, monitor receiver, and test-signal upconverter had built-in alarms to detect failures or performance degradation. These alarms were detected by polling the equipment over the IEEE-488, and were monitored by the FEC. Alarm conditions generated a PCTA-compatible alarm message and the condition of all monitored parameters was included in routine FEC status displays.

Alarms monitored included:

(1) Downconverters:

- (a) DC Power: Detects any power supply voltage drops of over 15 percent.
- (b) Phase lock: Detects X-band local oscillator PLL phase-lock.
- (c) 100-MHz standard: Detects low or missing 100-MHz reference from station frequency and timing subsystem (FTS).
- (d) Local Oscillator (LO) Power: Detects low local oscillator power.
- (e) Sum: Reports if any of the above alarms are true.

(2) Upconverter:

- (a) DC Power: Detects any power supply voltage drops of over 15 percent.
- (b) Phase lock: Detects the X-band local oscillator PLL phase-lock.
- (c) 100-MHz standard: Detects low or missing 100-MHz reference from station FTS.
- (d) LO power: Detects low local oscillator power.
- (e) Sum: Reports if any of the above alarms are true.

(3) Monitor Receiver:

- (a) DC Power: Detects any power supply voltage drops of over 15 percent.
- (b) Sum: Reports if any of the above alarms are true.

IV. Front-End System Refurbishment and TWM Performance

The two TWM/CCRs as received from ESA required some repair and upgrade. Both TWMs were tested when received at JPL; the measured gain/bandwidth performance fell short of Block II-A requirements. The length and thickness of the magnetic field shaping shim sets were modified to meet Block II-A maser 100-MHz 3-dB bandwidth specifications. The final gain/bandwidth performance and maser equivalent input noise-temperature measurements were performed on the two ESA TWMs before and after modifications. The results are shown in Figs. 14, 15, and 16. The performance of the two masers installed on the Parkes antenna is discussed below. The cooling capability of the CCR was verified, and a reserve refrigeration capacity of approximately 2.8 W was recorded for each CCR.

During the refurbishment period, a leaking waveguide vacuum window and faulty internal CCR wiring harness were also replaced. The new harness facilitated the installation of new cryogenic temperature-sensing diodes required by the new CCR monitor.

Due to the conditions under which the system had been previously stored and/or shipped, some corrosion had occurred inside the maser pump assemblies and at some of the waveguide flanges between the feedhorn and the TWM/CCRs. The pump assemblies were cleaned, and those components damaged by corrosion were replaced. New modulator/protect assemblies were installed to upgrade the pump source to meet Block II-A maser requirements. The input waveguide feed components were disassembled, inspected, cleaned, and the flanges lapped flat. To prevent further corrosion caused by moisture entering the feed system, a means of waveguide pressurization was investigated. There is no source of dry gas available in the aerial cabin on the Parkes antenna, so it was necessary to provide a lightweight source of gas that could be used during shipment and after installation on the antenna. Based on the very low leak-rate measurements made at JPL, a device supplying dry air at less than 47.2 cm³/min was required. Although the entire front-end assembly was wrapped in plastic sheeting while in shipment,

a commercial desiccant cartridge was enclosed to provide additional protection against moisture. The cartridge was attached to a tube fitting on an unused port of a waveguide switch. Where ac power was available at JPL, Tidbinbilla, and Parkes on the antenna, a Dielectric Communications Model 150 compressor dehydrator was used. This rack-mountable device, weighing 9 kg, was ordered with a special low-pressure option providing for operation at the desired pressure of approximately 0.04 kgs/cm². The dehydrator was mounted in the aerial cabin near the RF front-end assembly.

V. Front-End System Performance

The front-end system performance (gain, bandwidth, input noise temperature, and gain stability) was measured at JPL prior to shipment, at CDSCC after shipment, and at Parkes after installation on the antenna. The front-end system performance remained within project specifications in each case.

The noise-temperature performance of the RF package (in the absence of an antenna) was measured by attaching a high-quality feedhorn to the waveguide feed input flange (Fig. 3) and then measuring the Y-factor obtained when alternately switching (using the waveguide switch at the input of each maser) between the feedhorn looking at the "cold" sky and an ambient waveguide termination. The system temperature measurements were taken using the FEC to measure Y-factor at the output of the down-converter's 325-MHz IF. The system noise temperature of the front-end at JPL measured 15.4 K using TWM No. 1 and 15.6 K using TWM No. 2. At the CDSCC (Fig. 3), the same measurement resulted in 14.5 K and 15.7 K respectively.

The system was then installed on the Parkes antenna. The system-noise temperature measured 22 K using either maser channel (antenna at zenith, broken clouds). The gain/bandwidth curves measured on TWM 1 and TWM 2 after installation on the Parkes antenna are shown in Fig. 17. The system-noise temperature and maser gain/bandwidth data were within the specified requirements and indicated no change resulting from shipping.

The FEC software, CCR monitor software, and the front-end instrumentation control functions were tested after installation on the Parkes antenna. After some software corrections, all functions were verified to perform correctly, and all design parameters were met or exceeded.

VI. Conclusion

A PCTA design review for the Parkes front-end system was presented in November of 1987 and the majority of the front-end system requirements were established at that time. A previous TDA Progress Report [2] listed a number of requirements and constraints that affected the design approach selected. Later, frequency stability requirements were established for radio occultation experiments during the Voyager encounter with Neptune, and operational and maintenance requirements were requested by CDSCC engineering staff. The system with the newly designed FEC, NAR, CCR monitor, system alarm monitor, and refurbished RF front-end met all station operation and maintenance, Voyager Project, and radio occultation experiment requirements.

The design of the FEC provided a background of successful hardware and software development for the future design and development of controllers for low-noise amplifiers in the DSN.

Acknowledgments

The Parkes front-end system refurbishment and reimplementation is the product of the Parkes front-end team members L. Fowler, J. Kovatch, and T. Sweeney, in addition to the authors, who are members of the JPL Microwave Electronics Group, under the supervision of S. Petty. Also, the CDSCC engineering staff, under the direction of M. Dinn, contributed greatly to the success of this project from the design phase through installation, testing, and operation.

References

- [1] D. L. Trowbridge, "Block IIA Traveling-Wave Maser," *TDA Progress Report 42-87*, vol. July-September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 158-164, November, 15, 1986.
- [2] D. L. Trowbridge, J. R. Loreman, T. J. Brunzie, and B. Jenkins, "An 8.4-GHz Dual Maser Front-End for Parkes Reimplementation," *TDA Progress Report 42-93*, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 214-228, May 15, 1988.
- [3] L. Fowler and M. Britcliffe, "Traveling-Wave Maser Closed-Cycle Refrigerator Data Acquisition and Display System," *TDA Progress Report 42-91*, vol. July-September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 304-311, November 15, 1987.

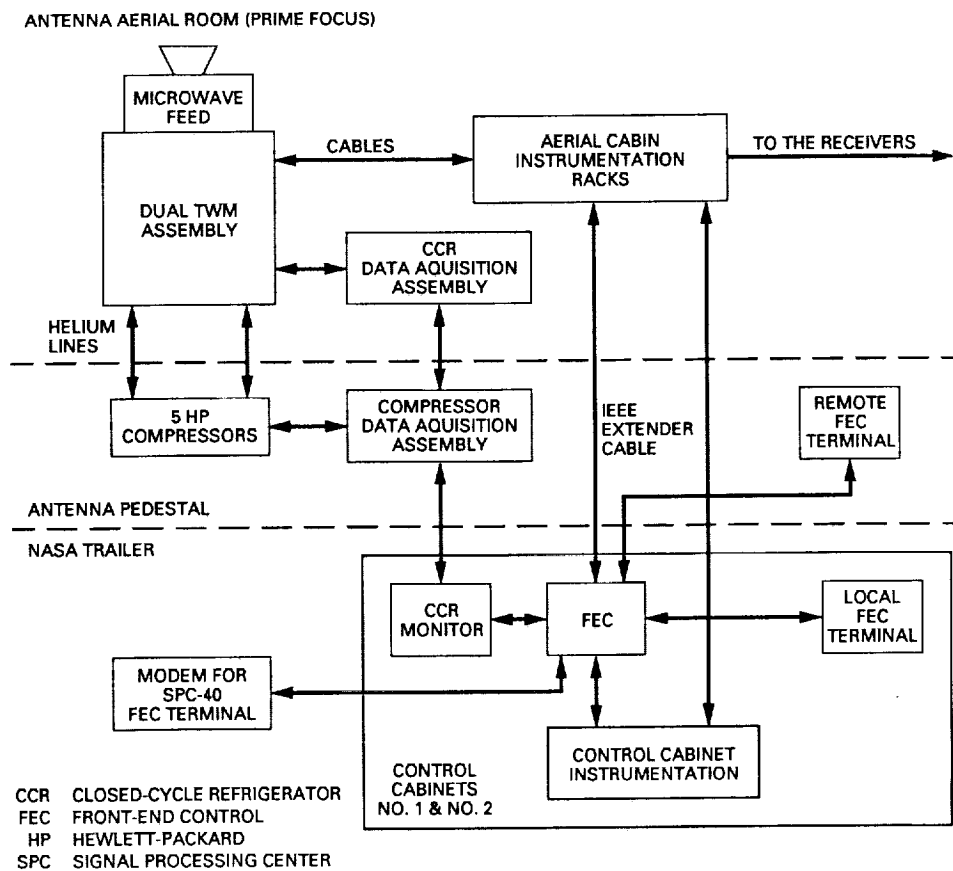


Fig. 1. Parkes front-end system.

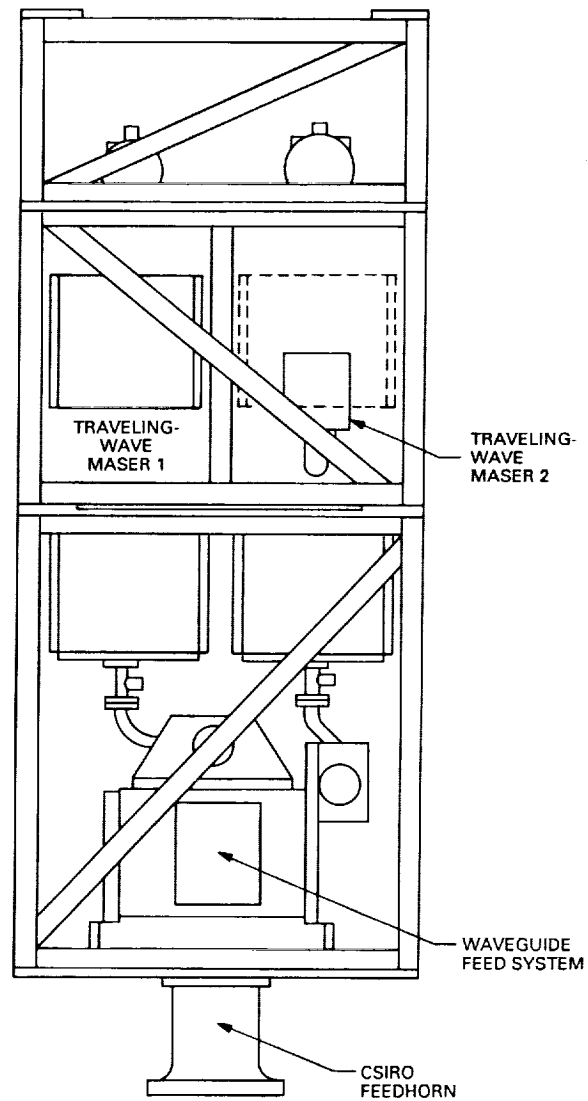


Fig. 2. Parkes RF package.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

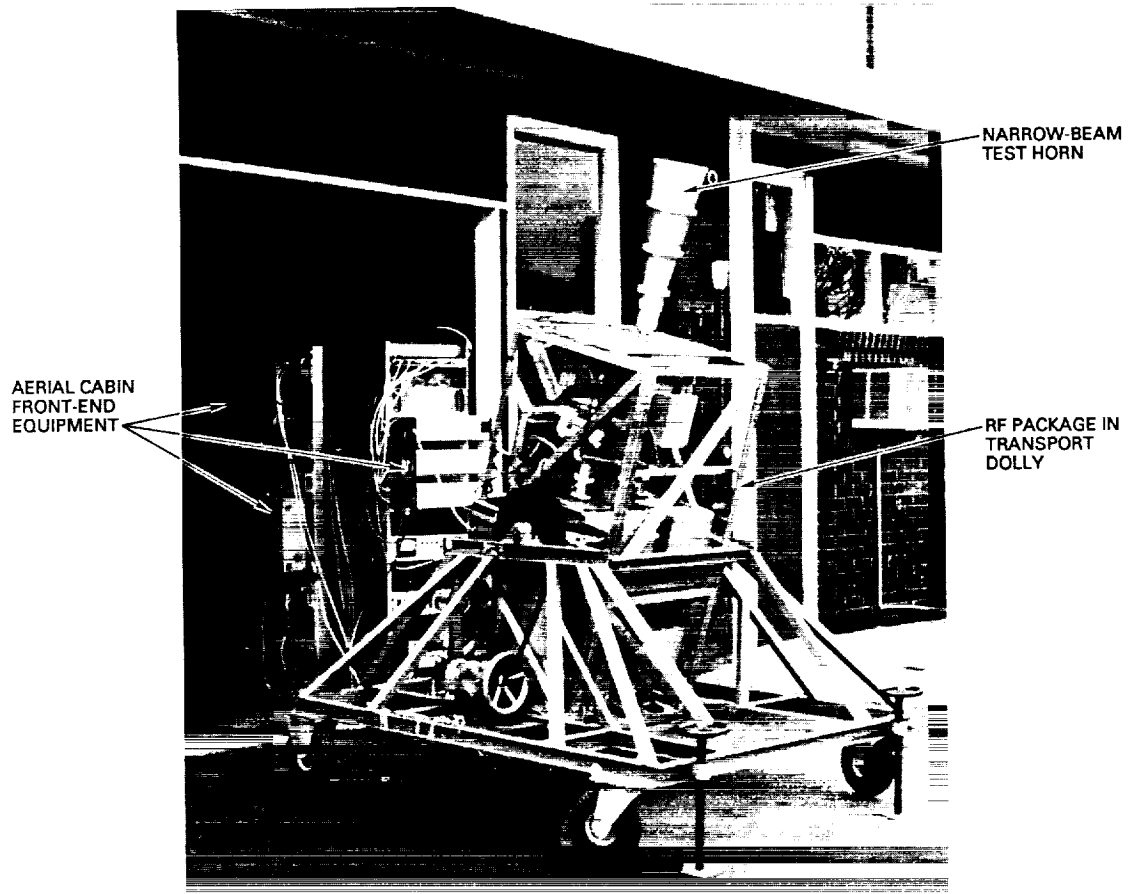


Fig. 3. Parkes front-end system under test.

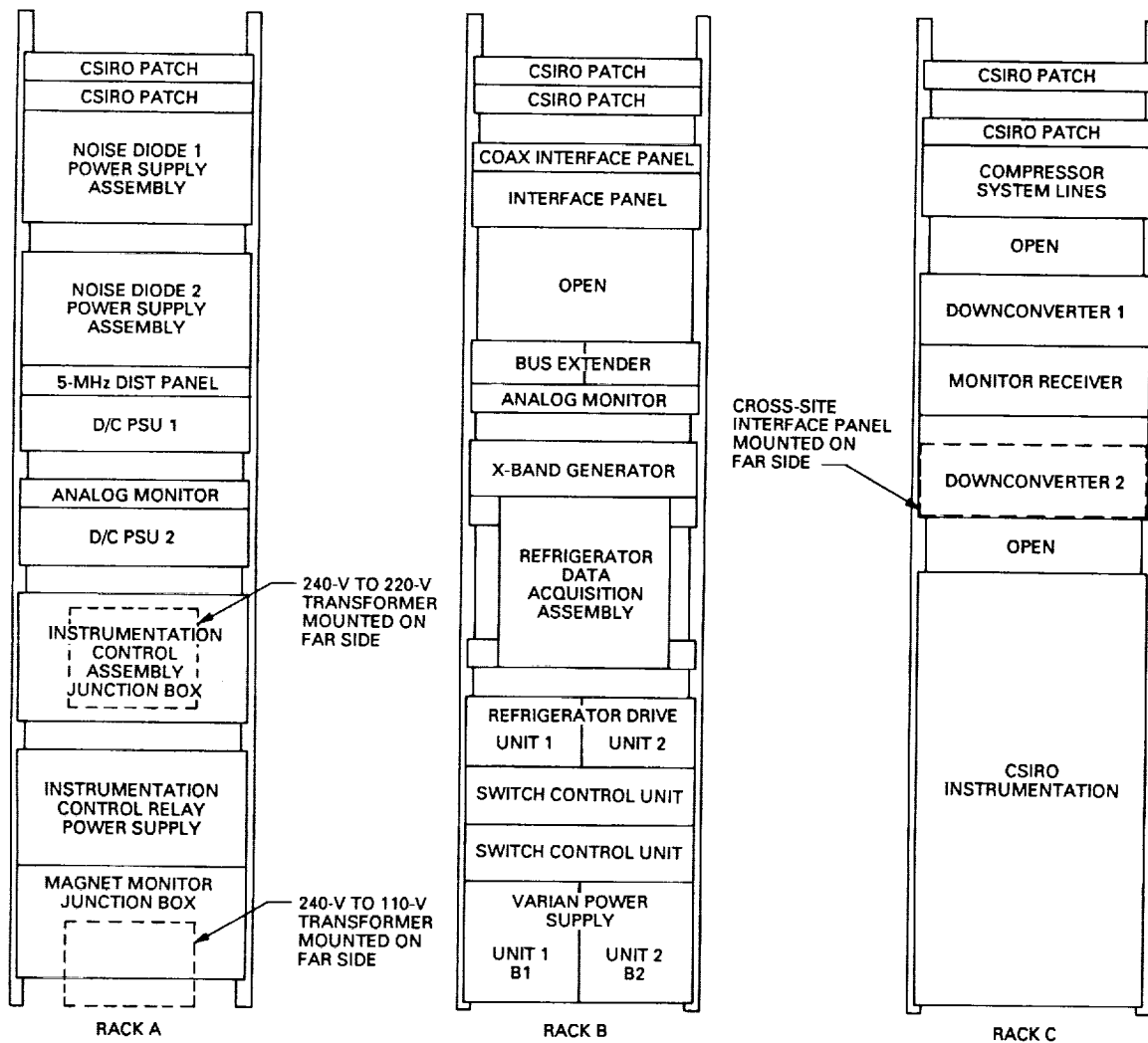


Fig. 4. Aerial cabin racks, 1989.

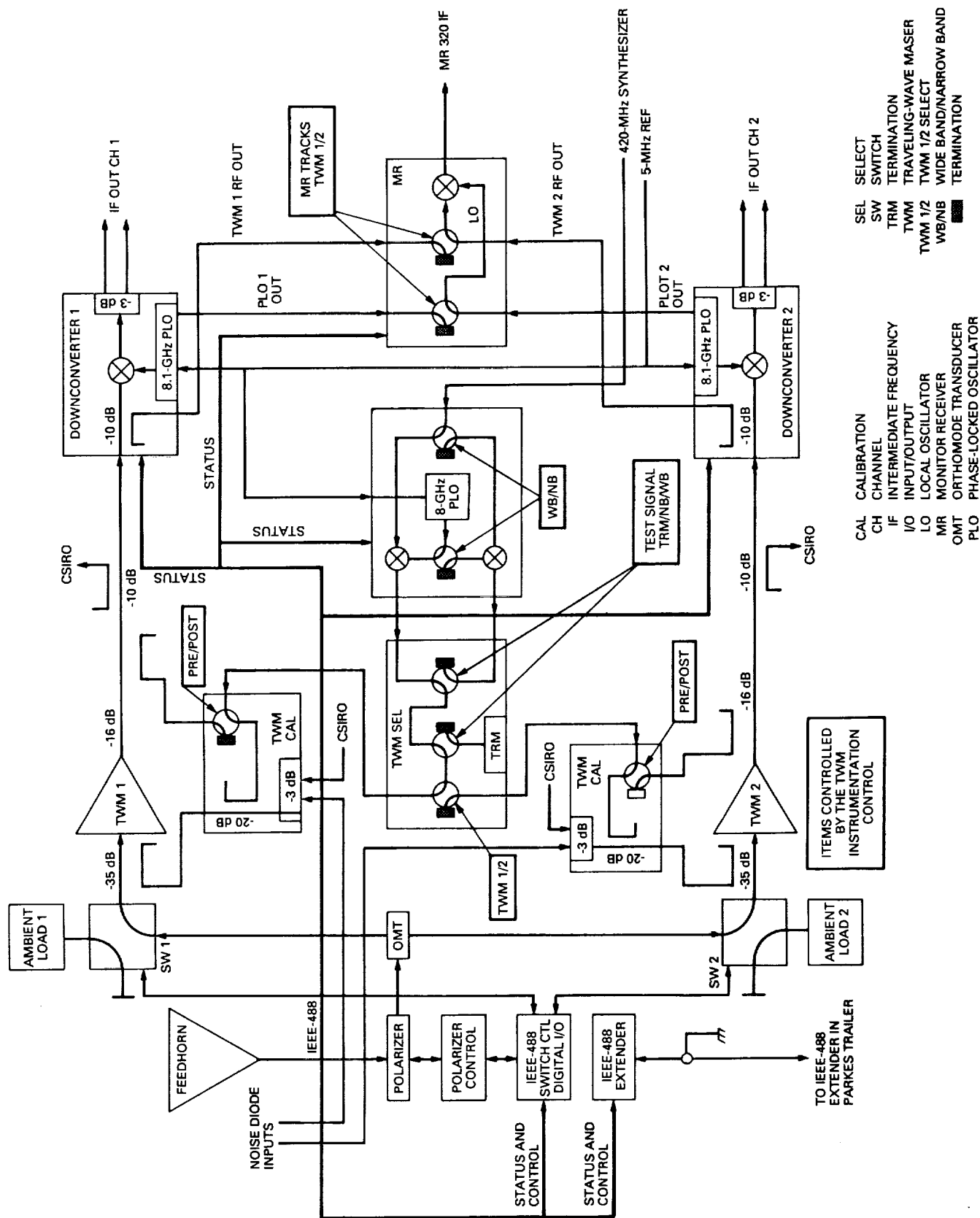


Fig. 5. RF equipment and control flow diagram: aerial cabin.

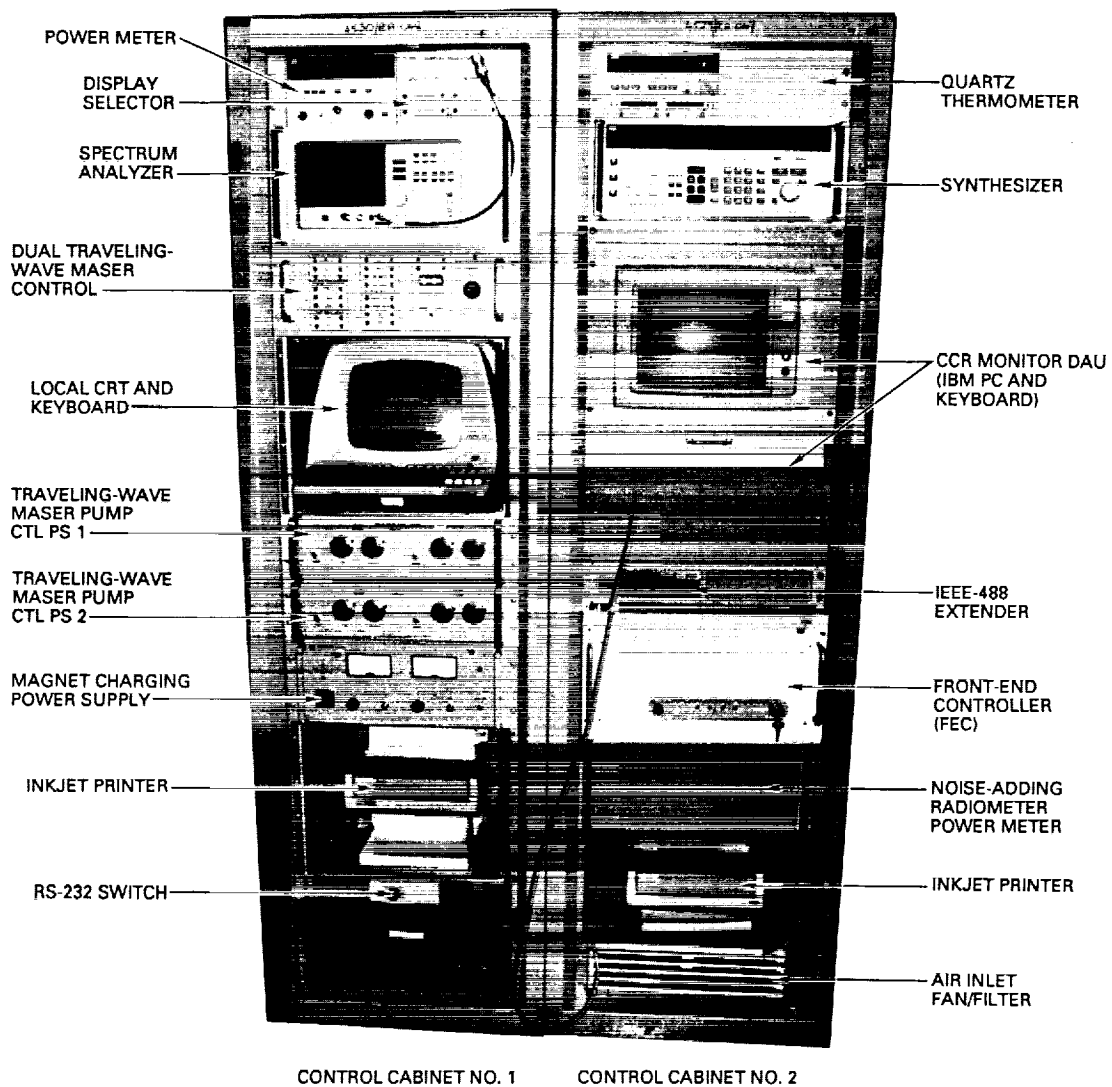


Fig. 6. RF front-end control cabinets.

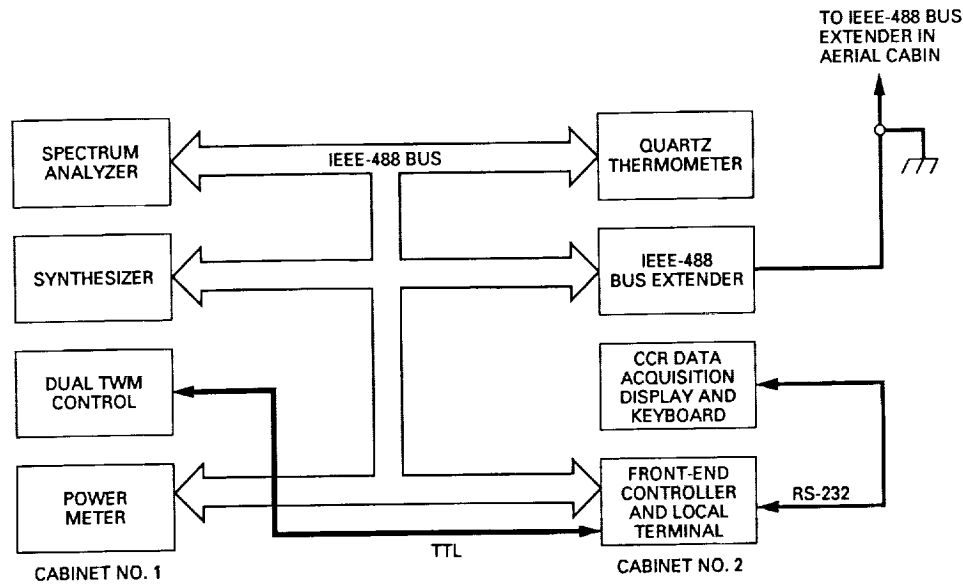


Fig. 7. Control flow diagram: control cabinets.

PFEC>help c

PFEC Front End Controller Configuration Help...
The format for user input is: COMMAND PARAMETER

Cmdnd	Description	Parameter
MINP	Set Maser Input Source.	[1],[2] [S]ky,[A]mbient Load
MPOL	Set Maser Input Polarization.	[1],[2] [L]eft,[R]ight
TSBW	Set Maser Test Signal Bandwidth.	[W]ide,[N]arrow,[T]erminate
TFRQ	Set Maser Test Signal Frequency.	[OFF],[n] MHz
TAMP	Set Maser Test Signal Amplitude.	[OFF],[n] dBm
TMOD	Set Maser Test Signal Modulation.	[Off],[n] deg
MRCV	Set Monitor Receiver Input Source.	[1],[2]
DSEL	Set Meter/Analyzer Display Select.	[M]easure,[C]alibrate
ITSG	Inject Maser Test Signal.	[OFF],[1],[2] [PRE],[PST]
TALK	Talk Directly to GP-IB Equipment.	[SC],[TH],[AN],[SG],
LSTN	Talk & Listen Directly to GP-IB Equip.	[PM],[D1],[D2],[MR]

PFEC>

Fig. 8. FEC configuration: help menu.

```

PFEC>stat c

PFEC Configuration Status (269 07:32:14)           Configuration is UNLOCKED

      53.36K/MOD----+                               8100 MHz      +----NAR
      |                                     |               |
Horn/RCP-----+-----Maser 1-----+-----X=====RF 1
      |                                     |               |
8425 MHz/WB-----+-----+-----+-----X=====Mon Rcvr
      |                                     |               |
Ambient Load-----+-----Maser 2-----+-----X=====RF 2
      |                                     |               |
ND OFF/OFF----+                               8100 MHz

Local CRT has command locking privileges.

PFEC>

```

Fig. 9. FEC configuration status display: signal path.

```

PFEC>help m

PFEC Front End Controller Measurement Help...
The format for user input is: COMMAND PARAMETER

  Cmnd  Description                               Parameter

YFAC    Perform Y-factor Measurement.             [1],[2] [L]eft,[R]ight
GAIN     Compute Maser Gain Profile.               [1],[2]
PMTR     Operate Power Meter.                     [M]easure,[C]alibrate
CCRH     Chart CCR Performance History.            [1],[2] [MMddHHmm] [MMddHHmm]
SYST     Measure System Temperature (TPR).         [],[OFF],[C]ont,[D]isp,[n]sec
NRES     Set NAR Sample Resolution.               [n] deg K
NRAT     Set NAR Sample Rate.                     [n] Hz
DRAT     Set NAR Diode Switching Rate.            [n] Hz (Debug only)
NSAM     Set Number of Resolution Samples.         [n]
CALG     Calibrate System Gain Factor.             [] (CFG only)
CALD     Calibrate Noise Diodes.                  [ALL],[n] deg K (CFG only)
DLTA     Measure System Linearity Delta.           [S]ky,[A]mbient,[F]ull Range

PFEC>

```

Fig. 10. FEC measurement: help menu.

```
PFEC>stat y 2
```

```
PFEC>STAT: Y-Factor Results for Maser 2, Right Polarization (271 10:28:39)
```

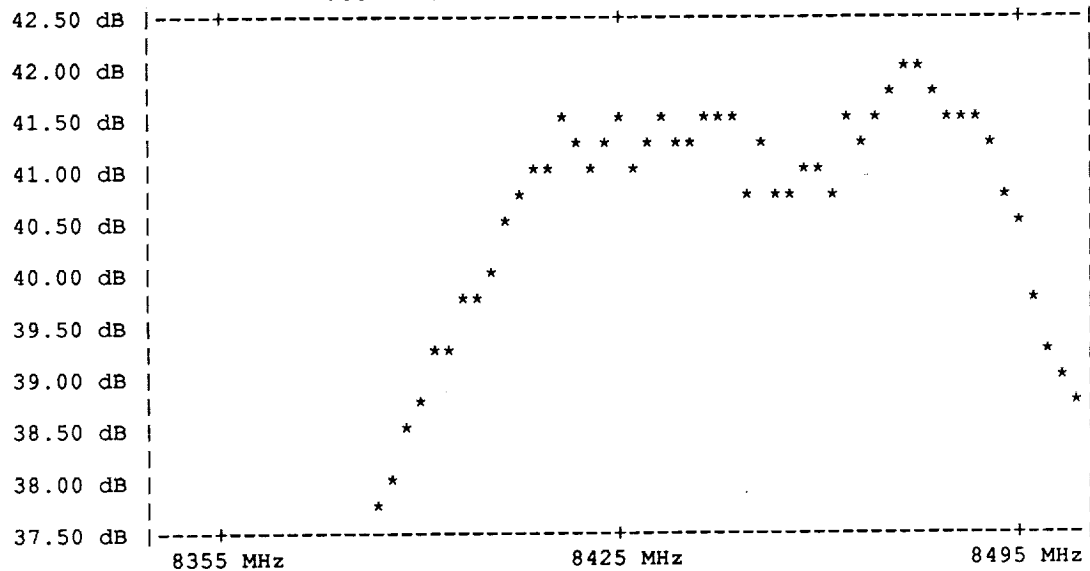
```
Average Ambient Load Power Level = -44.88 dBm
Average Ambient Load Temperature = 294.35 deg K
Average Antenna Power Level       = -57.67 dBm
Average Antenna Temperature      = 15.70 deg K
```

```
PFEC>
```

Fig. 11. FEC Y-factor measurement: status display.

```
PFEC>stat g 1
```

Maser 1 Gain Profile - Measured 271 12:38:07



```
PFEC>
```

Fig. 12. FEC maser gain profile: status display.

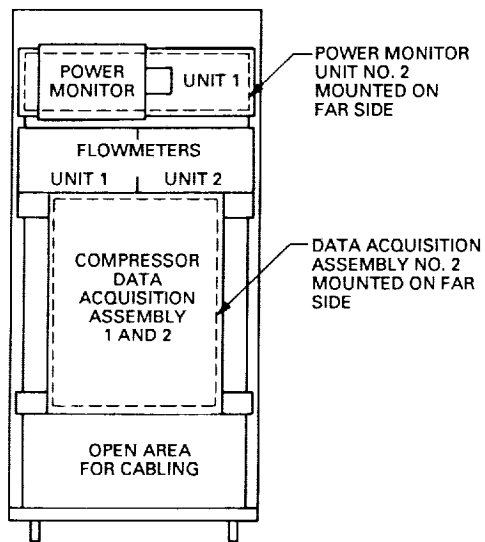


Fig. 13. Compressor data half-rack.

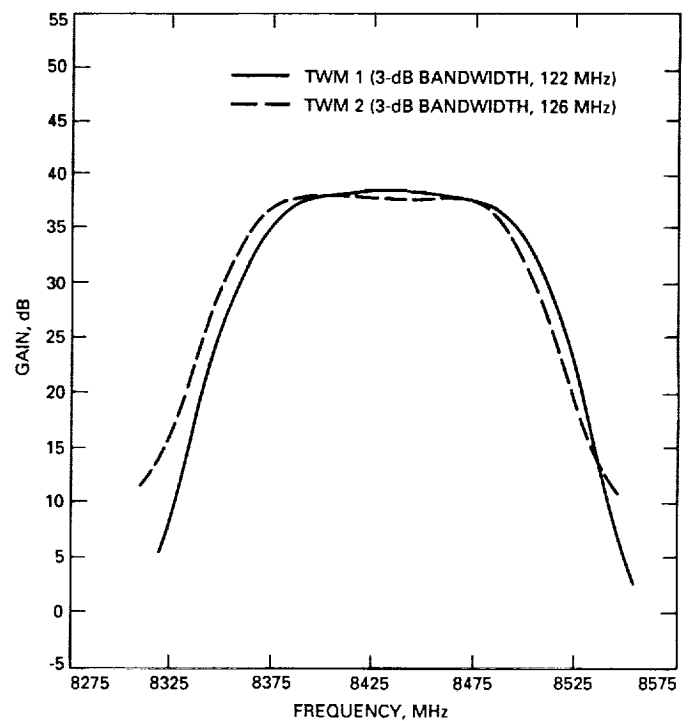


Fig. 15. TWMs 1 and 2 gain/bandwidth, after final adjustment.

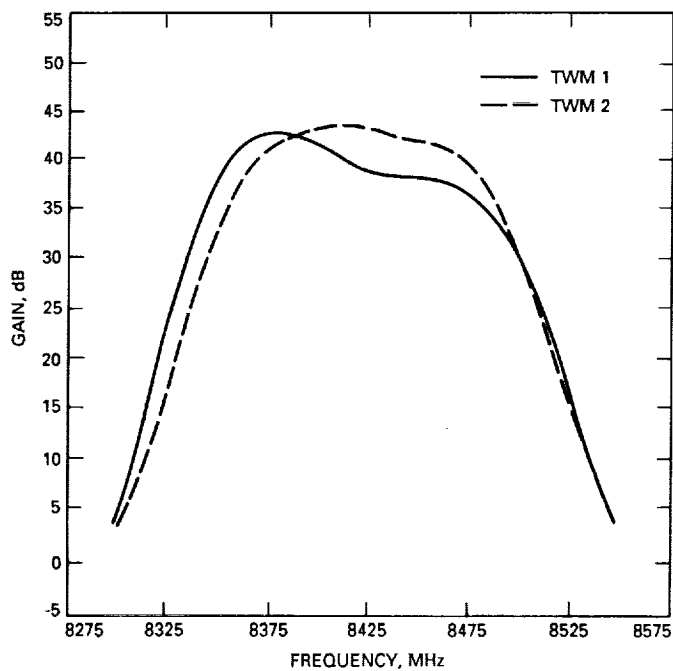


Fig. 14. TWMs 1 and 2 gain/bandwidth, prior to adjustment.

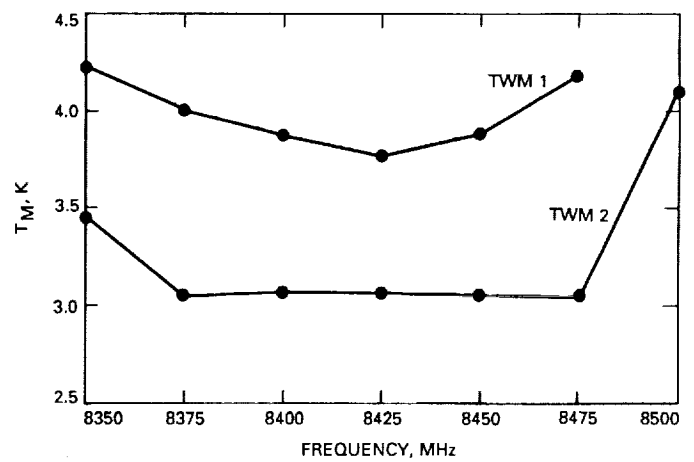


Fig. 16. TWMs 1 and 2 equivalent input noise temperature.

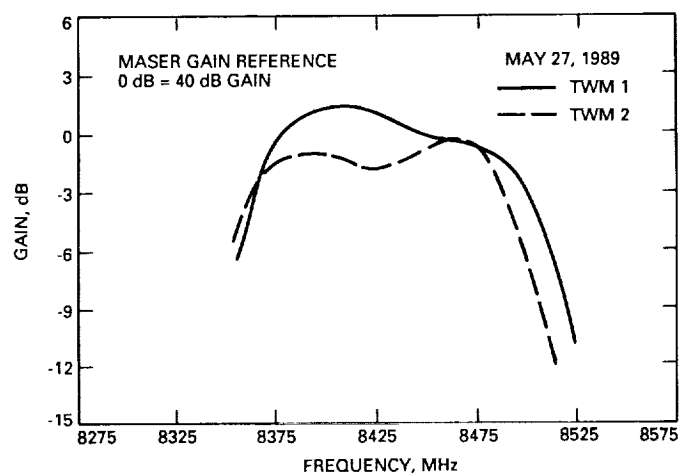


Fig. 17. Gain/bandwidth curves after installation on the Parkes antenna.

Author Index,¹ 1989

The Telecommunications and Data Acquisition Progress Report

42-97, January–March, May 15, 1989

42-98, April–June, August 15, 1989

42-99, July–September, November 15, 1989

42-100, October–December, February 15, 1990

Alvarez, L. S.

- 42-97 Application of Optimal Control Theory to the Design of the NASA/JPL 70-Meter Antenna Axis Servos, pp. 112–126.

L. S. Alvarez and J. Nickerson

Arnold, S.

- 42-100 Validity of the Two-Level Model for Viterbi Decoder Gap-Cycle Performance, pp. 191–202.

See Dolinar, S.

Badilla, G. L.

- 42-100 Using Connected-Element Interferometer Phase-Delay Data for Magellan Navigation in Venus Orbit, pp. 48–54.

See Thurman, S. W.

Bathker, D. A.

- 42-97 DSN 70-Meter Antenna Microwave Optics Design and Performance Improvements Part I: Design Optimization, pp. 306–313.

D. A. Bathker and S. D. Slobin

- 42-98 DSN 70-Meter Antenna Microwave Optics Design and Performance Improvements Part II: Comparison With Measurements, pp. 174–189.

D. A. Bathker and S. D. Slobin

Bevan, R.

- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160–173.

See Hinedi, S.

Britcliffe, M.

- 42-98 A Closed-Cycle Refrigerator for Cooling Maser Amplifiers Below 4 Kelvin, pp. 141–149.

Brunzie, T. J.

- 42-100 An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation, pp. 301–319.

See Trowbridge, D. L.

Cha, A. G.

- 42-97 A New Analysis of Beam-Waveguide Antennas Considering the Presence of the Enclosure, pp. 162–168.

Chen, C.-C.

- 42-97 Effective Amplifier Noise for an Optical Receiver Based on Linear Mode Avalanche Photodiodes, pp. 216–220.

Chen, J.

- 42-97 Microwave Time Delays for the X/S-Band Feed System at X-Band, pp. 248–251.

Cheung, K.-M.

- 42-97 Performance of Concatenated Codes Using 8-Bit and 10-Bit Reed-Solomon Codes, pp. 194–201.

See Pollara, F.

¹ In the case of joint authorship, the reader is referred to the citation under the first author where all authors of the article are listed.

- 42-97 The Weight Distribution and Randomness of Linear Codes, pp. 208–215.
- 42-98 A Performance Comparison Between Block Interleaved and Helically Interleaved Concatenated Coding Systems, pp. 95–103.
K.-M. Cheung and L. Swanson
- 42-98 Frame Synchronization Methods Based on Channel Symbol Measurements, pp. 121–137.
See Dolinar, S.
- 42-98 On the Decoder Error Probability of Linear Codes, pp. 104–109.
- 42-99 Fast Transform Decoding of Nonsystematic Reed-Solomon Codes, pp. 130–140.
See Truong, T. K.
- 42-99 Quantization Effects in Viterbi Decoding Rate 1/n Convolutional Codes, pp. 115–121.
See Onyszchuk, I. M.
- 42-100 An Adaptive Vector Quantization Scheme, pp. 214–220.
- Cheng, U.**
- 42-97 Software Package for Performing Experiments About the Convolutionally Encoded Voyager 1 Link, pp. 175–179.
- 42-99 Performance of the All-Digital Data-Transition Tracking Loop in the Advanced Receiver, pp. 60–71.
U. Cheng and S. Hinedi
- Chong, D.**
- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160–173.
See Hinedi, S.
- Christensen, C. S.**
- 42-97 High-Precision Radiometric Tracking for Planetary Approach and Encounter in the Inner Solar System, pp. 21–46.
C. S. Christensen, S. W. Thurman, J. M. Davidson, M. H. Finger, and W. M. Folkner
- Clements, P. A.**
- 42-98 Stable Kalman Filters for Processing Clock Measurement Data, pp. 190–201.
P. A. Clements, B. P. Gibbs, and J. S. Vandergraft
- Collins, O.**
- 42-99 Memory Management in Traceback Viterbi Decoders, pp. 98–104.
O. Collins and F. Pollara
- 42-99 Quantization Effects in Viterbi Decoding Rate 1/n Convolutional Codes, pp. 115–121.
See Onyszchuk, I. M.
- 42-100 A VLSI Decomposition of the deBruijn Graph, pp. 180–190.
O. Collins, S. Dolinar, R. McEliece, and F. Pollara
- Conroy, B.**
- 42-97 Multipurpose Exciter With Low Phase Noise, pp. 169–174.
B. Conroy and D. Le
- Cornier, R.**
- 42-100 Voltage Source AC-to-DC Converters for High-Power Transmitters, pp. 55–68.
- Cowles, K.**
- 42-97 A Visibility Characterization Program for Optical Communications Through the Atmosphere, pp. 221–225.
- 42-97 Options for Daytime Monitoring of Atmospheric Visibility in Optical Communications, pp. 226–234.
See Erickson, D.
- 42-97 Site Selection Criteria for the Optical Atmospheric Visibility Monitoring Telescopes, pp. 235–239.
- Cucchissi, J. J.**
- 42-98 Performance Effects of Tie-Truss Modifications for a 70-Meter Centerline Beam Waveguide Antenna, pp. 86–94.
- Davidson, J. M.**
- 42-97 High-Precision Radiometric Tracking for Planetary Approach and Encounter in the Inner Solar System, pp. 21–46.
See Christensen, C. S.
- Del Castillo, H.**
- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160–173.
See Hinedi, S.

- Dick, G. J.**
- 42-97 New Ion Trap for Atomic Frequency Standard Applications, pp. 58-63.
See Prestage, J. D.
- 42-97 Operational Parameters for the Superconducting Cavity Maser, pp. 64-71.
See Wang, R. T.
- 42-99 Microwave Oscillator With Reduced Phase Noise by Negative Feedback Incorporating Microwave Signals With Suppressed Carrier, pp. 20-33.
G. J. Dick and J. Saunders
- Diener, W.**
- 42-99 Thermal Coefficient of Delay for Various Coaxial and Fiber-Optic Cables, pp. 43-59.
See Lutes, G.
- Divsalar, D.**
- 42-97 The Use of Interleaving for Reducing Radio Loss in Trellis-Coded Modulation Systems, pp. 180-193.
D. Divsalar and M. K. Simon
- Dolinar, S.**
- 42-98 Frame Synchronization Methods Based on Channel Symbol Measurements, pp. 121-137.
S. Dolinar and K.-M. Cheung
- 42-99 Some Easily Analyzable Convolutional Codes, pp. 105-114.
See McEliece, R.
- 42-100 A VLSI Decomposition of the deBruijn Graph, pp. 180-190.
See Collins, O.
- 42-100 Exact Closed-Form Expressions for the Performance of the Split-Symbol Moments Estimator of Signal-to-Noise Ratio, pp. 174-179.
- 42-100 Validity of the Two-Level Model for Viterbi Decoder Gap-Cycle Performance, pp. 191-202.
S. Dolinar and S. Arnold
- Duffy, L.**
- 42-98 Ka-Band MMIC Beam Steered Transmitter Array, pp. 207-217.
See Rascoe, D. L.
- Edwards, C. D.**
- 42-97 The Effect of Spatial and Temporal Wet-Troposphere Fluctuations on Connected Element Interferometry, pp. 47-57.
- Erickson, D.**
- 42-97 Options for Daytime Monitoring of Atmospheric Visibility in Optical Communications, pp. 226-234.
D. Erickson and K. Cowles
- Estafan, J. A.**
- 42-100 The Determination of Maximum Deep Space Station Slew Rates for a High-Earth Orbiter, pp. 13-20.
- Finger, M. H.**
- 42-97 High-Precision Radiometric Tracking for Planetary Approach and Encounter in the Inner Solar System, pp. 21-46.
See Christensen, C. S.
- 42-99 Photon Statistical Limitations for Daytime Optical Tracking, pp. 90-97.
See Folkner, W. M.
- Folkner, W. M.**
- 42-97 High-Precision Radiometric Tracking for Planetary Approach and Encounter in the Inner Solar System, pp. 21-46.
See Christensen, C. S.
- 42-99 Photon Statistical Limitations for Daytime Optical Tracking, pp. 90-97.
W. M. Folkner and M. H. Finger
- Franco, M. M.**
- 42-100 Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels, pp. 111-130.
See Otoshi, T. Y.
- Freedman, A. P.**
- 42-99 Determination of Earth Orientation Using the Global Positioning System, pp. 1-11.

Galindo-Israel, V.

- 42-100 Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations, pp. 69–80.

V. Galindo-Israel, W. Imbriale, K. Shogen, and R. Mittra

Gatti, M. S.

- 42-99 32-GHz Performance of the DSS-14 70-Meter Antenna: 1989 Configuration, pp. 206–219.

M. S. Gatti, M. J. Klein, and T. B. H. Kuiper

Gibbs, B. P.

- 42-98 Stable Kalman Filters for Processing Clock Measurement Data, pp. 190–201.

See Clements, P. A.

Glass, G. W.

- 42-100 Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System, pp. 100–110.

See Johnson, D. L.

Greenhall, C. A.

- 42-97 Orthogonal Sets of Data Windows Constructed From Trigonometric Polynomials, pp. 300–305.

Gulkis, S.

- 42-98 The NASA SETI Sky Survey: Recent Developments, pp. 218–226.

See Klein, M. J.

Ham, N. C.

- 42-97 DSN Radio Science System Design and Testing for Voyager–Neptune Encounter, pp. 252–284.

N. C. Ham, T. A. Rebold, and J. F. Weese

Hill, R. E.

- 42-97 A New Method for Analysis of Limit Cycle Behavior of the NASA/JPL 70-Meter Antenna Axis Servos, pp. 98–111.

- 42-99 Disturbance Torque Rejection Properties of the NASA/JPL 70-Meter Antenna Axis Servos, pp. 170–188.

Hinedi, S.

- 42-98 Performance of the Split-Symbol Moments SNR Estimator in the Presence of Inter-Symbol Interference, pp. 157–173.

See Shah, B.

- 42-99 Performance of the All-Digital Data-Transition Tracking Loop in the Advanced Receiver, pp. 60–71.

See Cheng, U.

- 42-99 Costas Loop Lock Detection in the Advanced Receiver, pp. 72–89.

See Mileant, A.

- 42-100 A Functional Description of the Advanced Receiver, pp. 131–149.

- 42-100 QPSK Carrier-Acquisition Performance in the Advanced Receiver II, pp. 150–159.

S. Hinedi and B. Shah

- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160–173.

S. Hinedi, R. Bevan, H. Del Castillo, P. Kinman, D. Chong, and R. Labelle

Hoppe, D. J.

- 42-97 A Multiflare Horn With 1-Megawatt Power Handling Capability, pp. 149–155.

- 42-97 Thermal Measurements of Microwave Transmitter Feedhorn Window, pp. 156–161.

See Perez, R. M.

Howard, S. D.

- 42-100 Improving a Data-Acquisition Software System With Abstract Data Type Components, pp. 248–251.

Huang, J.

- 42-98 Ka-Band MMIC Beam Steered Transmitter Array, pp. 207–217.

See Rascoe, D. L.

Imbriale, W.

- 42-100 Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations, pp. 69–80.

See Galindo-Israel, V.

Jacobs, C. S.

- 42-100 A Higher Density VLBI Catalog for Navigating Magellan and Galileo, pp. 274-300.

See Ulvestad, J. S.

Janik, G. R.

- 42-98 An Apparatus for the Electrodynamical Containment of Charged Macroparticles, pp. 57-62.

See Williams, A.

- 42-99 Simple Analytic Potentials for Linear Ion Traps, pp. 12-19.

G. R. Janik, J. D. Prestage, and L. Maleki

Johnson, D. L.

- 42-100 Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System, pp. 100-110.

D. L. Johnson, S. M. Petty, J. J. Kovatch, and G. W. Glass

Kinman, P.

- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160-173.

See Hinedi, S.

Klein, M. J.

- 42-98 The NASA SETI Sky Survey: Recent Developments, pp. 218-226.

M. J. Klein, S. Gulkis, E. T. Olsen, and N. A. Renzetti

- 42-99 32-GHz Performance of the DSS-14 70-Meter Antenna: 1989 Configuration, pp. 206-219.

See Gatti, M. S.

Koerner, M. A.

- 42-97 Accuracy of Telemetry Signal Power Loss in a Filter as an Estimate for Telemetry Degradation, pp. 285-291.

Kovatch, J. J.

- 42-100 Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System, pp. 100-110.

See Johnson, D. L.

Kuiper, T. B. H.

- 42-99 32-GHz Performance of the DSS-14 70-Meter Antenna: 1989 Configuration, pp. 206-219.

See Gatti, M. S.

Kumar, R.

- 42-99 Application of Adaptive Least-Squares Algorithm to Multi-Element Array Signal Reconstruction, pp. 141-160.

Labelle, R.

- 42-100 Digital Doppler Extraction Demonstration With the Advanced Receiver, pp. 160-173.

See Hinedi, S.

Le, D.

- 42-97 Multipurpose Exciter With Low Phase Noise, pp. 169-174.

See Conroy, B.

Levy, R.

- 42-97 Reanalysis, Compatibility, and Correlation in Analysis of Modified Antenna Structures, pp. 367-381.

Lichten, S. M.

- 42-97 Estimation and Filtering Techniques for High-Accuracy GPS Applications, pp. 1-20.

- 42-100 Precise Estimation of Tropospheric Path Delays With GPS Techniques, pp. 1-12.

Logan, R. T., Jr.

- 42-99 Effect of Laser Frequency Noise on Fiber-Optic Frequency Reference Distribution, pp. 34-42.

R. T. Logan, Jr., G. F. Lutes, and L. Maleki

- 42-100 Microwave Analog Fiber-Optic Link for Use in the Deep Space Network, pp. 21-33.

R. T. Logan, Jr., G. Lutes, and L. Maleki

Loreman, J. R.

- 42-100 An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation, pp. 301-319.

See Trowbridge, D. L.

Lubecke, V.

- 42-98 Ka-Band MMIC Beam Steered Transmitter Array, pp. 207-217.

See Rascoe, D. L.

Lutes, G.

- 42-97 State-of-the-Art Fiber Optics for Short Distance Frequency Reference Distribution, pp. 81-87.

G. Lutes and L. Primas

- 42-99 Effect of Laser Frequency Noise on Fiber-Optic Frequency Reference Distribution, pp. 34-42.
See Logan, R. T., Jr.
- 42-99 Thermal Coefficient of Delay for Various Coaxial and Fiber-Optic Cables, pp. 43-59.
G. Lutes and W. Diener
- 42-100 Microwave Analog Fiber-Optic Link for Use in the Deep Space Network, pp. 21-33.
See Logan, R. T., Jr.
- Lyons, J. R.
- 42-98 Spin-Lattice Relaxation and the Calculation of Gain, Pump Power, and Noise Temperature in Ruby, pp. 63-85.
- Maleki, L.
- 42-97 New Ion Trap for Atomic Frequency Standard Applications, pp. 58-63.
See Prestage, J. D.
- 42-98 An Apparatus for the Electrodynamical Containment of Charged Macroparticles, pp. 57-62.
See Williams, A.
- 42-99 Effect of Laser Frequency Noise on Fiber-Optic Frequency Reference Distribution, pp. 34-42.
See Logan, R. T., Jr.
- 42-99 Simple Analytic Potentials for Linear Ion Traps, pp. 12-19.
See Janik, G. R.
- 42-100 Microwave Analog Fiber-Optic Link for Use in the Deep Space Network, pp. 21-33.
See Logan, R. T., Jr.
- McEliece, R.
- 42-100 A VLSI Decomposition of the deBruijn Graph, pp. 180-190.
See Collins, O.
- 42-99 Some Easily Analyzable Convolutional Codes, pp. 105-114.
R. McEliece, S. Dolinar, F. Pollara, and H. Van Tilborg
- Melbourne, R.
- 42-98 An Apparatus for the Electrodynamical Containment of Charged Macroparticles, pp. 57-62.
See Williams, A.
- Mileant, A.
- 42-99 Costas Loop Lock Detection in the Advanced Receiver, pp. 72-89.
A. Mileant and S. Hinedi
- Mitra, R.
- 42-100 Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations, pp. 69-80.
See Galindo-Israel, V.
- Nerheim, N.
- 42-97 The Effects of Atmospheric Turbulence on Precision Optical Measurements Used for Antenna-Pointing Compensation, pp. 141-148.
- Nguyen, L.
- 42-98 Wideband Phase-Locked Angular Modulator, pp. 150-156.
- Nickerson, J.
- 42-97 Application of Optimal Control Theory to the Design of the NASA/JPL 70-Meter Antenna Axis Servos, pp. 112-126.
See Alvarez, L. S.
- Olsen, E. T.
- 42-98 The NASA SETI Sky Survey: Recent Developments, pp. 218-226.
See Klein, M. J.
- Onyszchuk, I. M.
- 42-99 Quantization Effects in Viterbi Decoding Rate 1/n Convolutional Codes, pp. 115-121.
I. M. Onyszchuk, K.-M. Cheung, and O. Collins
- 42-100 Finding the Complete Path and Weight Enumerators of Convolutional Codes, pp. 203-213.
- Otoshi, T. Y.
- 42-100 Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels, pp. 111-130.
T. Y. Otoshi and M. M. Franco

- Peng, T. K.**
 42-100 Radar RFI at Goldstone DSS 12 and DSS 16, pp. 234-247.
 See Slobin, S. D.
- Perez, R. M.**
 42-97 Thermal Measurements of Microwave Transmitter Feedhorn Window, pp. 156-161.
 R. M. Perez and D. J. Hoppe
- Petty, S. M.**
 42-100 Ultralow Noise Performance of an 8.4-GHz Maser-Feedhorn System, pp. 100-110.
 See Johnson, D. L.
- Pitt, G. H., III**
 42-98 Performance of the Image Statistics Decoder in Conjunction With the Goldstone-VLA Array, pp. 138-140.
 See Wang, H. C.
- Pollara, F.**
 42-97 Performance of Concatenated Codes Using 8-Bit and 10-Bit Reed-Solomon Codes, pp. 194-201.
 F. Pollara and K.-M. Cheung
 42-99 Memory Management in Traceback Viterbi Decoders, pp. 98-104.
 See Collins, O.
 42-99 Some Easily Analyzable Convolutional Codes, pp. 105-114.
 See McEliece, R.
 42-100 A VLSI Decomposition of the deBruijn Graph, pp. 180-190.
 See Collins, O.
- Posner, E. C.**
 42-98 A Simple Model for DSS-14 Outage Times, pp. 202-206.
 See Rumsey, H. C.
- Prestage, J. D.**
 42-97 New Ion Trap for Atomic Frequency Standard Applications, pp. 58-63.
 J. D. Prestage, G. J. Dick, and L. Maleki
- 42-98** An Apparatus for the Electrodynamical Containment of Charged Macroparticles, pp. 57-62.
 See Williams, A.
- 42-99** Simple Analytic Potentials for Linear Ion Traps, pp. 12-19.
 See Janik, G. R.
- Primas, L. E.**
 42-97 Stabilized Fiber-Optic Frequency Distribution System, pp. 88-97.
 L. E. Primas, G. F. Lutes, and R. L. Sydnor
 42-97 State-of-the-Art Fiber Optics for Short Distance Frequency Reference Distribution, pp. 81-87.
 See Lutes, G.
- Quinn, R.**
 42-100 An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation, pp. 301-319.
 See Trowbridge, D. L.
- Rabkin, J.**
 42-99 Big Viterbi Decoder (BVD) Results for (7,1/2) Convolutional Code, pp. 122-129.
 See Statman, J.
- Rascoe, D. L.**
 42-98 Ka-Band MMIC Beam Steered Transmitter Array, pp. 207-217.
 D. L. Rascoe, A. L. Riley, J. Huang, V. Lubecke, and L. Duffy
- Rebold, T. A.**
 42-97 DSN Radio Science System Design and Testing for Voyager-Neptune Encounter, pp. 252-284.
 See Ham, N. C.
 42-99 Parkes Radio Science System Design and Testing for Voyager Neptune Encounter, pp. 189-205.
- Reed, I. S.**
 42-97 Decoding of 1/2-Rate (24,12) Golay Codes, pp. 202-207.
 See Truong, T. K.
 42-99 Fast Transform Decoding of Nonsystematic Reed-Solomon Codes, pp. 130-140.
 See Truong, T. K.

- 42-100 A VLSI Design for a Systolic Viterbi Decoder, pp. 221-233.
See Truong, T. K.
- Renzetti, N. A.**
- 42-98 The NASA SETI Sky Survey: Recent Developments, pp. 218-226.
See Klein, M. J.
- Richter, P. H.**
- 42-97 DSN 70-Meter Antenna X- and S-Band Calibration Part I: Gain Measurements, pp. 314-351.
P. H. Richter and S. D. Slobin
- 42-97 DSN 70-Meter Antenna X- and S-Band Calibration Part II: System Noise Temperature Measurements and Telecommunications Link Evaluation, pp. 352-366.
See Slobin, S. D.
- Riley, A. L.**
- 42-98 Ka-Band MMIC Beam Steered Transmitter Array, pp. 207-217.
See Rascoe, D. L.
- Robinson, D. L.**
- 42-97 A Novel Approach to a PPM-Modulated Frequency-Doubled Electro-Optic Cavity-Dumped Nd:YAG Laser, pp. 240-247.
- Rodemich, E. R.**
- 42-100 A Real-Time Signal Combining System for Ka-Band Feed Arrays Using Maximum-Likelihood Weight Estimates, pp. 81-99.
See Vilnrotter, V. A.
- Ross, D. L.**
- 42-100 Mark IV-A DSCC (Magellan-Era) Telemetry System Description, pp. 253-273.
- Rowan, D. R.**
- 42-97 The DSS-14 C-Band Exciter, pp. 292-299.
- Ruggier, C. J.**
- 42-99 The Effects of Sinusoidal Interference on the Second-Order Carrier Tracking Loop Preceded by a Bandpass Limiter in the Block IV Receiver, pp. 161-169.
- Rumsey, H. C.**
- 42-98 A Simple Model for DSS-14 Outage Times, pp. 202-206.
H. C. Rumsey, R. Stevens, and E. C. Posner
- Russell, R. K.**
- 42-98 An Analytic Development of Orbit Determination for a Distant, Planetary Orbiter, pp. 1-25.
R. K. Russell and S. W. Thurman
- Saldua, B. P.**
- 42-100 Structural Fatigue in the 34-Meter HA-Dec Antennas, pp. 252-262.
See Van Hek, R. A.
- Satorius, E.**
- 42-100 A VLSI Design for a Systolic Viterbi Decoder, pp. 221-232.
See Truong, T. K.
- Saunders, J.**
- 42-99 Microwave Oscillator With Reduced Phase Noise by Negative Feedback Incorporating Microwave Signals With Suppressed Carrier, pp. 20-33.
See Dick, G. J.
- Scheid, R. E.**
- 42-97 Precision Pointing Compensation for DSN Antennas With Optical Distance Measuring Sensors, pp. 127-140.
- Shah, B.**
- 42-98 Performance of the Split-Symbol Moments SNR Estimator in the Presence of Inter-Symbol Interference, pp. 157-173.
B. Shah and S. Hinedi
- 42-100 QPSK Carrier-Acquisition Performance in the Advanced Receiver II, pp. 150-159.
See Hinedi, S.
- Shih, M. T.**
- 42-100 A VLSI Design for a Systolic Viterbi Decoder, pp. 221-233.
See Truong, T. K.

Shiozaki, A.

- 42-99 Fast Transform Decoding of Nonsystematic Reed-Solomon Codes, pp. 130-140.

See Truong, T. K.

Shogen, K.

- 42-100 Dual-Shaped Offset Reflector Antenna Designs From Solutions of the Geometrical Optics First-Order Partial Differential Equations, pp. 69-80.

See Galindo-Israel, V.

Siev, B.

- 42-99 Big Viterbi Decoder (BVD) Results for (7,1/2) Convolutional Code, pp. 122-129.

See Statman, J.

Simon, M. K.

- 42-97 The Use of Interleaving for Reducing Radio Loss in Trellis-Coded Modulation Systems, pp. 180-193.

See Divsalar, D.

Slobin, S. D.

- 42-97 DSN 70-Meter Antenna Microwave Optics Design and Performance Improvements Part I: Design Optimization, pp. 306-313.

See Bathker, D. A.

- 42-97 DSN 70-Meter Antenna X- and S-Band Calibration Part I: Gain Measurements, pp. 314-351.

See Richter, P. H.

- 42-97 DSN 70-Meter Antenna X- and S-Band Calibration Part II: System Noise Temperature Measurements and Telecommunications Link Evaluation, pp. 352-366.

S. D. Slobin and P. H. Richter

- 42-98 DSN 70-Meter Antenna Microwave Optics Design and Performance Improvements Part II: Comparison With Measurements, pp. 174-189.

See Bathker, D. A.

- 42-100 Radar RFI at Goldstone DSS 12 and DSS 16, pp. 234-247.

S. D. Slobin and T. K. Peng

Smyth, P.

- 42-98 Automated Monitor and Control for Deep Space Network Subsystems, pp. 110-120.

Sovers, O. J.

- 42-100 A Higher Density VLBI Catalog for Navigating Magellan and Galileo, pp. 274-300.

See Ulvestad, J. S.

Statman, J.

- 42-99 Big Viterbi Decoder (BVD) Results for (7,1/2) Convolutional Code, pp. 122-129.

J. Statman, J. Rabkin, and B. Siev

Stevens, R.

- 42-98 A Simple Model for DSS-14 Outage Times, pp. 202-206.

See Rumsey, H. C.

Strayer, D. M.

- 42-97 Operational Parameters for the Superconducting Cavity Maser, pp. 64-71.

See Wang, R. T.

Swanson, L.

- 42-98 A Performance Comparison Between Block Interleaved and Helically Interleaved Concatenated Coding Systems, pp. 95-103.

See Cheung, K.-M.

Sydnor, R. L.

- 42-97 Stabilized Fiber-Optic Frequency Distribution System, pp. 88-97.

See Primas, L. E.

Thurman, S. W.

- 42-97 High-Precision Radiometric Tracking for Planetary Approach and Encounter in the Inner Solar System, pp. 21-46.

See Christensen, C. S.

- 42-98 An Analytic Development of Orbit Determination for a Distant, Planetary Orbiter, pp. 1-25.

See Russell, R. K.

- 42-100 Galileo Earth Approach Navigation Using Connected-Element Interferometer Phase-Delay Tracking, pp. 34-47.

- 42-100 Using Connected-Element Interferometer Phase-Delay Data for Magellan Navigation in Venus Orbit, pp. 48-54.

S. W. Thurman and G. L. Badilla

Truong, T. K.

- 42-97 Decoding of 1/2-Rate (24,12) Golay Codes, pp. 202-207.

T. K. Truong, I. S. Reed, and X. Yin

- 42-99 Fast Transform Decoding of Nonsystematic Reed-Solomon Codes, pp. 130-140.

T. K. Truong, K.-M. Cheung, I. S. Reed, and A. Shiozaki

- 42-100 A VLSI Design for a Systolic Viterbi Decoder, pp. 221-232.

T. K. Truong, E. Satorius, M. T. Shih, and I. S. Reed

Trowbridge, D. L.

- 42-100 An 8.4-GHz Dual-Maser Front-End System for Parkes Reimplementation, pp. 301-319.

D. L. Trowbridge, J. R. Loreman, T. J. Brunzie, and R. Quinn

Tucker, T. K.

- 42-97 Operating and Environmental Characteristics of Sigma Tau Hydrogen Masers, pp. 72-80.

Ulvestad, J. S.

- 42-98 A Statistical Study of Radio-Source Structure Effects on Astrometric Very Long Baseline Interferometry Observations, pp. 26-56.

- 42-100 A Higher Density VLBI Catalog for Navigating Magellan and Galileo, pp. 274-300.

J. S. Ulvestad, O. J. Sovers, and C. S. Jacobs

Vandergraft, J. S.

- 42-98 Stable Kalman Filters for Processing Clock Measurement Data, pp. 190-201.

See Clements, P. A.

Van Hek, R. A.

- 42-100 Structural Fatigue in the 34-Meter HA-Dec Antennas, pp. 252-262.

R. A. Van Hek and B. P. Saldua

Van Tilborg, H.

- 42-99 Some Easily Analyzable Convolutional Codes, pp. 105-114.

See McEliece, R.

Vilnrotter, V. A.

- 42-100 A Real-Time Signal Combining System for Ka-Band Feed Arrays Using Maximum-Likelihood Weight Estimates, pp. 81-99.

V. A. Vilnrotter and E. R. Rodemich

Wang, H. C.

- 42-98 Performance of the Image Statistics Decoder in Conjunction With the Goldstone-VLA Array, pp. 138-140.

H. C. Wang and G. H. Pitt III

Wang, R. T.

- 42-97 Operational Parameters for the Superconducting Cavity Maser, pp. 64-71.

R. T. Wang, G. J. Dick, and D. M. Strayer

Weese, J. F.

- 42-97 DSN Radio Science System Design and Testing for Voyager-Neptune Encounter, pp. 252-284.

See Ham, N. C.

Williams, A.

- 42-98 An Apparatus for the Electrodynamical Containment of Charged Macroparticles, pp. 57-62.

A. Williams, R. Melbourne, L. Maleki, G. Janik, and J. Prestage

Yin, X.

- 42-97 Decoding of 1/2-Rate (24,12) Golay Codes, pp. 202-207.

See Truong, T. K.

