

The Telecommunications and Data Acquisition Progress Report 42-119

July-September 1994

Joseph H. Yuen
Editor

(NASA-CR-197249) THE
TELECOMMUNICATIONS AND DATA
ACQUISITION REPORT Progress Report,
Jul. - Sep. 1994 (JPL) 300 p

N95-17279
--THRU--
N95-17297
Unclas

P.4 63/32 0033266

358 393

November 15, 1994



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

110864

The Telecommunications and Data Acquisition Progress Report 42-119

July–September 1994

Joseph H. Yuen
Editor

November 15, 1994



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 Telecommunications and Mission Operations Directorate (TMOD), which now includes the former Telecommunications and Data Acquisition (TDA) Office. In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, supporting research and technology, implementation, and operations. Also included are standards activity at JPL for space 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 Communications (OSC).

The TMOD also performs work funded by other NASA program offices through and with the cooperation of OSC. The first of these is the Orbital Debris Radar Program funded by the Office of Space Systems Development. It exists at Goldstone only and makes use of the planetary radar capability when the antennas are configured as science instruments making direct observations of the planets, their satellites, and asteroids of our solar system. The Office of Space Sciences funds the data reduction and science analyses of data obtained by the Goldstone Solar System Radar. The antennas at all three complexes are also configured for radio astronomy research and, as such, conduct experiments funded by the National Science Foundation in the U.S. and other agencies at the overseas complexes. These experiments are either in microwave spectroscopy or very long baseline interferometry.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TMOD are included.

This and each succeeding issue of *The Telecommunications and Data Acquisition Progress Report* will present material in some, but not necessarily all, of the aforementioned programs.

Contents

OSC TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

Evaluation of Current Tropospheric Mapping Functions by Deep Space Network Very Long Baseline Interferometry	1
O. J. Sovers and G. E. Lanyi NASA Code 314-30-56-28-50	
An Algorithm for Extraction of Periodic Signals From Sparse, Irregularly Sampled Data	12
J. Z. Wilcox NASA Code 314-40-22-12-03	
Space VLBI Telecommunication Characteristics, Protection Criteria, and Frequency Sharing	29
B. O. Gutierrez-Luaces and D. F. Bishop NASA Code 315-90-31-00-02	
Radio-Planetary Frame Tie From Phobos-2 VLBI Data	46
C. E. Hildebrand, B. A. Iijima, P. M. Kroger, W. M. Folkner, and C. D. Edwards NASA Code 310-10-60-50-00	
COMMUNICATIONS, SPACECRAFT-GROUND	
Closed-Loop Carrier Phase Synchronization Techniques Motivated by Likelihood Functions	83
H. Tsou, S. Hinedi, and M. Simon NASA Code 314-30-61-02-04	
On the Application of Under-Decimated Filter Banks	105
Y.-P. Lin and P. P. Vaidyanathan NASA Code 315-91-40-20-35	
A Simplified Integer Cosine Transform and Its Application in Image Compression	129
M. Costa and K. Tong NASA Code 310-30-71-83-04	
Adaptive Line Enhancers for Fast Acquisition	140
H.-G. Yeh and T. M. Nguyen NASA Code 315-91-10-13-08	
Antenna Noise Temperatures of the 34-Meter Beam-Waveguide Antenna With Horns of Different Gains Installed at F1	160
T. Y. Otoshi, P. R. Lee, and M. M. Franco NASA Code 310-20-65-86-08	
Low-Complexity Wavelet Filter Design for Image Compression	181
E. Majani NASA Code 315-91-40-20-36	
A Preliminary Optical Visibility Model	201
K. Cowles and B. M. Levine NASA Code 315-91-40-20-29	
Field Verification of the Wind Tunnel Coefficients	210
W. K. Gawronski and J. A. Mellstrom NASA Code 314-40-52-01-02	

DSN Systems Implementation CAPABILITIES FOR EXISTING PROJECTS

Tracking the Galileo Spacecraft With the DSCC Galileo Telemetry Prototype	221 -13
T. T. Pham, S. Shambayati, D. E. Hardi, and S. G. Finley	
NASA Code 314-30-61-02-10	
Computation of Reflected and Transmitted Horn Radiation Patterns for a Dichroic Plate	236 -14
J. C. Chen	
NASA Code 315-91-40-20-26	
X-/Ka-Band Dichroic Plate Noise Temperature Reduction	255 -15
W. Veruttipong and P. Lee	
NASA Code 315-91-40-20-26	

DSN Operations TDA PROGRAM MANAGEMENT AND ANALYSIS

A Model for Analysis of TDA Budget Allocation	262 -16
D. S. Remer and H. R. Buchanan	
NASA Code 314-40-31-30-03	

GROUND COMMUNICATIONS IMPLEMENTATION AND OPERATIONS

RF Optics Study for DSS-43 Ultracone Implementation	268 -17
P. Lee and W. Veruttipong	
NASA Code 314-30-61-02-17	
JPL-ANTOPT Antenna Structure Optimization Program	282 -18
D. M. Strain	
NASA Code 314-30-55-02-07	
Errata	293

33267

P-11

110865

35839B

Evaluation of Current Tropospheric Mapping Functions by Deep Space Network Very Long Baseline Interferometry

O. J. Sovers and G. E. Lanyi
Tracking Systems and Applications Section

To compare the validity of current algorithms that map zenith tropospheric delay to arbitrary elevation angles, 10 different tropospheric mapping functions are used to analyze the current data base of Deep Space Network Mark III intercontinental very long baseline interferometric (VLBI) data. This analysis serves as a stringent test because of the high proportion of low-elevation observations necessitated by the extremely long baselines. Postfit delay and delay-rate residuals are examined, as well as the scatter of baseline lengths about the time-linear model that characterizes tectonic motion. Among the functions that utilize surface meteorological data as input parameters, the Lanyi 1984 mapping shows the best performance both for residuals and baselines, though the 1985 Davis function is statistically nearly identical. The next best performance is shown by the recent function of Niell, which is based on an examination of global atmospheric characteristics as a function of season and uses no weather data at the time of the measurements. The Niell function shows a slight improvement in residuals relative to Lanyi, but also an increase in baseline scatter that is significant for the California-Spain baseline. Two variants of the Chao mapping function, as well as the Chao tables used with the interpolation algorithm employed in the Orbit Determination Program software, show substandard behavior for both VLBI residuals and baseline scatter. The length of the California-Australia baseline (10,600 km) in the VLBI solution can vary by as much as 5 to 10 cm for the 10 mapping functions.

I. Introduction

The tropospheric delay is one of the most poorly understood components in the theoretical model of very long baseline interferometric (VLBI) observables. This was first realized at least a decade ago, and numerous attempts have been made during the past 10 years to improve tropospheric modeling. These include the two high-quality mapping functions of Lanyi [1] and Davis et al. [2], the model of tropospheric turbulence by Treuhaft and Lanyi [3], and a plethora of mapping functions developed since the mid-1980s. Recent reviews by Gallini [4] and Estefan and Sovers [17] provide a good summary.

During the past decade at JPL, the Lanyi tropospheric mapping function has been part of the VLBI theoretical model in the MODEST software [5], the GPS models in the GIPSY software [6], and the 1992 International Earth Rotation Service (IERS) Standards [7]. The Orbit Determination Program (ODP)

still uses the much older Chao tables.¹ Other VLBI groups have employed, in succession, the Davis (CfA) [2], Herring (MTT) [8], and Niell (NMF)² [9] mapping functions. In view of the multitude of recent new mapping functions, it is of interest to determine whether the decade-old JPL VLBI tropospheric model is still state of the art, as it was when first introduced in 1984. To this end, some simple tests were devised to evaluate the performance of the newer mapping functions, as well as the older Chao algorithms, for a typical data set of DSN VLBI observations. The results also have bearing on the adequacy of the current ODP tropospheric model.

The next section describes the VLBI analyses that were carried out to evaluate the performance of 10 variants of tropospheric mapping algorithms. The data base consists of all currently processed DSN Mark III VLBI observations since their inception in late 1988. The quantities examined are the postfit residuals, as a function of elevation angle, and the scatter of daily baseline lengths about the linear model of tectonic plate motion. Conclusions and recommendations for theoretical modeling in VLBI and orbit analyses are presented in the last section.

II. Mapping Function Comparisons

Tests, using VLBI measurements with DSN antennas, were devised to evaluate the performance of a variety of tropospheric mapping functions carried out during the past decade. These were typically 24-hr observing sessions using two antennas separated by either of the two intercontinental baselines (approximately 8,400 or 10,600 km), looking over as large a region of the sky as permitted by mutual-source visibility. Because the extreme baseline length is a large fraction of an Earth diameter, the region of the sky accessible at either site was severely limited, and numerous observations were made at the lower elevation limits of the antennas (approximately 6 deg). These low-elevation observations help to decorrelate zenith delays at the two stations in data analyses, but require good accuracy in the tropospheric mapping at low-elevation angles. While tropospheric delay is also an important part of the theoretical model in orbit determination with the ODP software [17],³ similar tests for the Doppler and range observables in spacecraft tracking are precluded by the lack of accepted observations below 15-deg elevations, which is the region most sensitive to details of the mapping algorithm.

To avoid the possibility of drawing false conclusions induced by the much higher system noise level of the early observations (approximately 10 cm for Mark I, 2 cm for Mark III), only a subset of the VLBI data was selected for detailed analysis. It included all the newer data, recorded with Mark III data acquisition systems. The data spanned the time period 1988 to 1993; a total of 11,897 delay and delay-rate pairs was used. A standard VLBI parameter estimation fit was performed. The estimated parameters included positional coordinates of 283 radio sources, a pair of nutation angle offsets for each session (longitude and obliquity), and station coordinates for each overseas station (Australia and Spain) for each of the 61 sessions. The right ascension of one source, the nutation model on one day, and the Goldstone station coordinates were kept fixed. A new value of the zenith tropospheric value was estimated every 2 hr at each station. Correlations among the delay and delay-rate observables due to tropospheric fluctuations were ignored here, but will be considered in a future article. More details of both the DSN data-acquisition and parameter-estimation procedures have been previously published [10,11]. MacMillan and Ma [12] recently performed a similar evaluation for the Chao, Davis, Herring, and Ifadis tropospheric models, using the NASA Crustal Dynamics Project (CDP) and National Oceanic and Atmospheric Administration (NOAA) International Radio Interferometric Surveying (IRIS) data bases.

¹ C. C. Chao, "Improved Tropospheric Mapping Tables (Including Bending Effect) for SATODP," JPL Interoffice Memorandum 391.3-637 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 28, 1972.

² A. E. Niell, *Global Mapping Functions for the Atmospheric Delay at Radio Wavelengths*, Haystack Observatory, Westford, Massachusetts, a draft report received as a personal communication, March 31, 1994.

³ N. A. Mottinger, "Reflections on Refraction—A Historical Overview of the Tropospheric Refraction Model in the ODP," JPL Interoffice Memorandum 314.10-385 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 18, 1984.

A postfit analysis estimates time-linear variation of the station coordinates, as well as the Goldstone-overseas baseline lengths. Such fits and postfit baseline analyses were repeated for each of 10 tropospheric mapping functions, with all other aspects of modeling and parameter estimation (the latter with one exception) being identical. The mapping functions are listed in Table 1. For those functions requiring surface meteorological measurements (nos. 4 through 9), such data were taken from the Deep Space Communications Complex (DSCC) Media Calibration Subsystem data base.⁴ Temperature lapse rates had seasonal variations at the Australia and Spain stations,⁵ while the default -6.5 K/km was used for Goldstone. The two remaining atmospheric parameters (inversion and tropopause heights) were assigned default values of 1.25 and 12.2 km, respectively. The “Lanyi estimated” function requires some additional explanation. It uses the Lanyi 1984 function to accurately map the zenith delays, based on the available tabulated surface and balloon meteorological data. A crude single-parameter approximation is then used to account for the effect of deviations between the real and tabulated meteorological data on the mapping function [5]. One such parameter is estimated at each station for every observing session.

Table 1. Troposphere mapping functions used in fits to DSN VLBI data.

Mapping function	JPL VLBI archive ID	Reference	Comments
1 = Chao original	152	[13]	
2 = Chao revised	153	[14]	Revised constants
3 = Chao tables	154	Footnote 1	Tables used in ODP
4 = Lanyi standard	151	[1]	
5 = Lanyi updated	155	— ^a	Geometry and gravity curvature corrections
6 = Lanyi estimated	156	[5]	One parameter/station estimated per 24-hr experiment
7 = Davis	158	[2]	Alias CfA
8 = Ifadis	157	[15]	
9 = Herring	159	[8]	Alias MTT
10 = Niell	160	[9]	Alias NMF

^aG. E. Lanyi, “Tropospheric Propagation Delay Effects for Radio Waves,” JPL Interoffice Memorandum 335.1-156 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 15, 1983.

Our examination of tropospheric mapping functions includes

- (1) Comparison of the total tropospheric delays given by the various mapping functions for the DSN sites for the particular subset of meteorological conditions prevailing at the times of the VLBI measurements.
- (2) Comparison of VLBI delay and delay-rate residuals and baseline-length scatter resulting from multiparameter estimation in model fits to the VLBI observables.

The first category thus reflects the effect on the mapping functions of the elevation distributions and weather conditions at the times of the VLBI experiments, while the second category is a quantitative

⁴T. F. Runge, *Troposphere Calibration Software*, JPL SRD-NVI-5454-OP (internal document), Jet Propulsion Laboratory, Pasadena, California, August 31, 1993.

⁵M. A. Smith, “Analysis of Meteorological Balloon Data,” JPL Interoffice Memorandum 335.6-92-023 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 27, 1992.

assessment of how well each mapping function represents the data. These two categories of tests are described in turn in the two following subsections.

A. Statistics of Mapping Function Delay Values

Direct comparisons of mapping functions under varying atmospheric conditions yield voluminous data, most of which will not be presented here. They are in partial agreement with the results presented by Mendes and Langley [16].⁶ Discrepancies between tropospheric delays calculated with different mapping functions generally increase rapidly with decreasing elevation angles below 20 deg. In the extreme case, the difference reaches approximately 10 cm of tropospheric delay at the DSN lower observational limit of 6 deg. A histogram of the elevation angles of the DSN VLBI observations is shown in Fig. 1. The distribution is seen to peak between 10 and 15 deg, with more than half of the observations being below a 30-deg elevation at one or both ends of the baseline. This distribution of observations with elevation angle and the magnitude of the differences between the mapping functions are generally consistent with the postfit analyses.

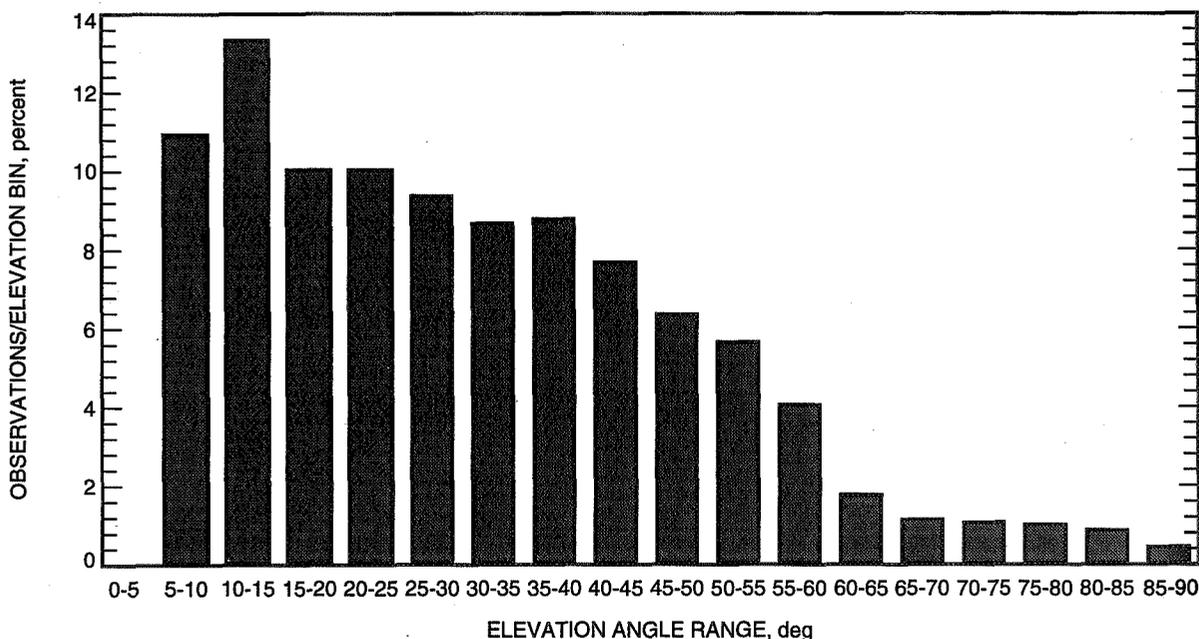


Fig. 1. Histogram of elevation angle distributions of DSN Mark III VLBI observations.

Figure 2 shows the total delay differences as a function of elevation angle for the Davis and Niell mapping functions. The Davis function is shown because it gives the closest fit results to Lanyi; the Niell function was not available for the initial work of Mendes and Langley and appears to show a performance nearly equivalent to that of the Lanyi mapping function. All tropospheric delay values were evaluated at each station for all 11,897 observations, using observed values for any required surface meteorological data. The results were placed in 0.5-deg bins for elevations lower than 20 deg and 1-deg bins above 20 deg; the error bars in Fig. 2 correspond to the standard deviation from the average in each bin. The differences are seen to be as large as 3 cm for Niell minus Lanyi just above a 6-deg elevation and to decrease rapidly toward zero when approaching 20 deg. The scatter within each bin approaches 1 cm below 10 deg for both functions. Note that the Davis-Lanyi differences have a negative region of 1-2 mm at mid-elevations. It should be noted that the plots in Fig. 2 represent statistical rather than exact mapping function comparisons. The Niell mapping function internally contains statistical

⁶ Better agreement was not attained because in some of the Mendes and Langley comparisons not all of the needed meteorological data were used.

averages of atmospheric parameters; for the Davis and Lanyi mapping functions, DSN statistical averages of atmospheric parameters were used, e.g., monthly averages of zenith delays. Since daily values of the surface pressure were used for the Davis function, there could be a small discrepancy (a few millimeters at 6 deg) in Fig. 2(a) due to the dependence of the bending effect on the zenith delay or, equivalently, on the surface pressure.

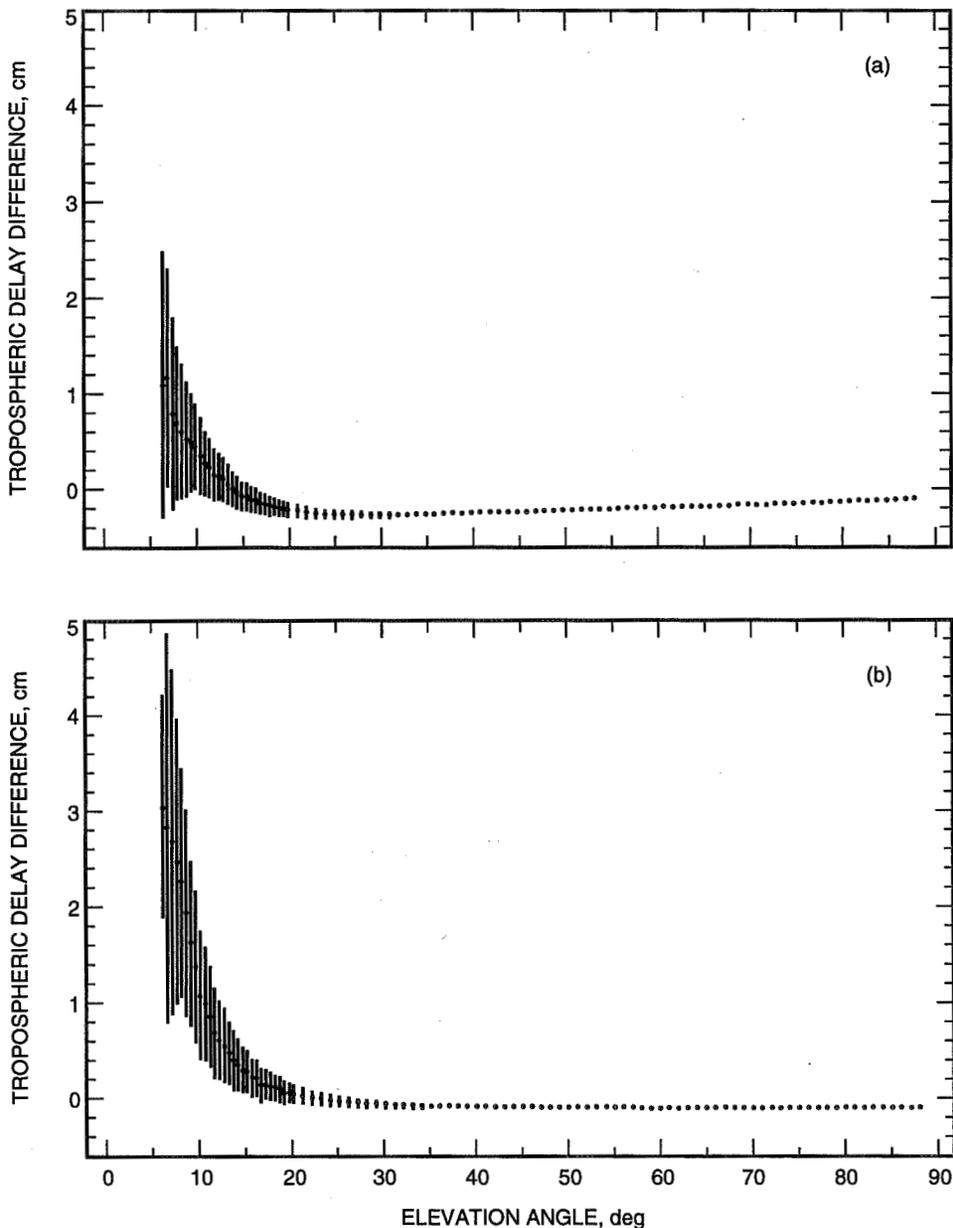


Fig. 2. Tropospheric delay difference versus elevation angle for DSN Mark III VLBI observations: (a) Davis minus Lanyi and (b) Niell minus Lanyi mapping functions. Each error bar is the standard deviation from the mean of all values in each elevation bin.

B. Statistics of Postfit Residuals

The postfit quantities that were examined included the delay and delay-rate residuals, the “baseline scatter” and χ^2 for a time-linear fit to baselines estimated independently for each observing session, and

the baseline length bias (change in baseline length at a given epoch induced by a change in the mapping function). Our criteria for a good mapping function are small values for the residuals and the baseline scatter. Differences in the average angular dependence of the mapping functions, as well as in the resulting baseline scales, can point to particular elevation ranges or baselines that are not properly modeled.

Table 2 shows the rms delay and delay-rate residuals and total χ^2 's for the 10 VLBI fits. Normalized χ^2 values are not shown because of the uncertainty in partitioning parameters between the two observable types. The “improvement” columns in this table (as well as in Tables 3 and 4) contain the signed rms difference between the residuals given by each mapping function and Lanyi standard. Focusing first on the residual results in Table 2, we see that the mapping functions fall into three broad groups. Both Chao functions and the Chao tables produce fits to VLBI data that are definitely inferior, especially for delay rates. All other functions do not differ by more than a few picoseconds in delay residuals from the Lanyi function and its variants. Only the Niell and Ifadis functions, and Lanyi with parameters estimated, improve both the delay and delay-rate residuals. As may be seen from the subsequent results as well, the Niell and Ifadis functions show very similar behavior.

Table 2. Mark III VLBI residuals.

Mapping function ^a	Delay (D)		Delay rate (DR)		Improvement ^b	
	rms, psec	χ^2	rms, fsec/sec	χ^2	D , psec	DR , fsec/sec
1	71.6655	9665	131.512	13,140	-19.4	-38.2
2	69.2404	8932	126.494	11,883	-5.9	-12.7
3	70.4522	9176	154.737	21,456	-14.3	-90.0
4	68.9867	8835	125.856	11,701	—	—
5	69.0001	8838	125.869	11,702	-1.4	-1.8
6	68.8222	8784	125.314	11,599	+4.8	+12.7
7	69.0492	8844	126.315	11,788	-2.9	-10.8
8	68.9374	8837	125.762	11,717	+2.6	+4.9
9	69.0331	8838	125.767	11,695	-2.5	+4.7
10	68.9273	8813	125.608	11,682	+2.9	+9.4

^a Mapping function definitions are provided in Table 1.

^b Relative improvement is based on mapping function 4.

To assess the behavior of each tropospheric mapping function in various elevation angle ranges, the delay residuals were divided into the six elevation bins shown in Table 3: observations below 10 deg, 5-deg bins up to 30 deg, and observations from 30 deg to zenith. Table 3 gives the raw results and relates them to the Lanyi standard function. It is seen that the Niell and Ifadis (and Lanyi with estimated parameters) improvements in residuals are not uniform across the range of elevations. They do well at very low angles, but not in the 10- to 15-deg and 25- to 30-deg ranges (this holds for most of the newer functions). No extraordinary elevation-partitioned results are seen, with one exception: It can be noted that the Lanyi parameters-estimated function apparently achieves reductions in residuals by improving the mapping function shape at very low elevations.

Table 4 shows baseline length results from the VLBI data analyses. It is seen that any departure from the standard Lanyi mapping function increases the baseline scatter and χ^2 per degree of freedom (χ^2_ν) (with the exception of Davis on the California–Spain [DSS 15–DSS 65] baseline), sometimes by substantial amounts. The statistical significance of these differences can be inferred from the formal uncertainties of χ^2_ν . The number of degrees of freedom N in the California–Australia (DSS 15–DSS 45) and California–Spain fits are 30 and 27, respectively. Thus, the standard deviation of the normalized χ^2 ,

$\sigma_{\chi^2_\nu}$, is approximately $(2/N)^{1/2} = 0.26$ and 0.27 . Even for the best fits, however, χ^2_ν shows substantial departures from unity. This originates both from model inadequacies and underestimated observable errors. We assume that the best fits correspond to the best modeling, and that in this case modeling errors and error underestimates contribute equally to the increased χ^2 . According to this assumption, the χ^2 value corresponding to the underestimated errors is $1 + (\chi^2 - 1)/2$. Consequently, to obtain a more realistic error for χ^2_ν , we multiply the formal $\sigma_{\chi^2_\nu}$ by this estimated quantity, 1.3 and 1.6 for the two baselines, respectively. This gives $\sigma_{\chi^2_\nu} = 0.34$ and 0.44 , which are the error bars used in Figs. 3(a) and (b). Since all fits are based on identical data, the statistical significance of the difference between different solutions is measured by the standard deviation of χ^2_ν itself, i.e., 0.34 or 0.44.

Table 3. Delay residuals by elevation range.

Mapping function ^a	Range of elevation angles, deg					
	0–10	10–15	15–20	20–25	25–30	30–90
	Delay residuals, psec					
1	98.993	76.114	66.665	64.213	65.090	66.386
2	94.562	73.217	65.754	62.232	62.562	64.346
3	97.358	74.320	66.923	62.694	62.537	65.477
4	94.117	72.692	65.883	62.025	61.914	64.221
5	94.134	72.708	65.904	62.036	61.911	64.235
6	93.679	72.807	65.770	61.903	61.938	64.000
7	94.419	72.716	65.873	61.990	61.900	64.267
8	94.049	72.782	65.632	61.965	62.171	64.121
9	94.174	72.918	65.777	62.027	62.173	64.210
10	93.994	72.772	65.741	61.938	62.151	64.104
	Improvement over Lanyi standard in quadrature, psec ^b					
1	−30.7	−22.6	−10.2	−16.6	−20.1	−16.8
2	−9.2	−8.8	+4.1	−5.1	−9.0	−4.0
3	−24.9	−15.5	−11.8	−9.1	−8.8	−12.8
4	—	—	—	—	—	—
5	−1.8	−1.5	−1.7	−1.2	+0.6	−1.3
6	+9.1	−4.1	+3.9	+3.9	−1.7	+5.3
7	−7.5	−1.9	+1.1	+2.1	+1.3	−2.4
8	+3.6	−3.6	+5.7	+2.7	−5.6	+3.6
9	−3.3	−5.7	+3.7	−0.5	−5.7	+1.2
10	+4.8	−3.4	+4.3	+3.3	−5.4	+3.9

^a Mapping function definitions are provided in Table 1.

^b Relative improvement is based on mapping function 4.

The differences are due to mapping errors and are seen to be highly significant for the group of three Chao functions versus Lanyi. They are marginally significant for the remaining functions on the California–Australia baseline, and $> 2\sigma_{\chi^2_\nu}$ significant for Ifadis, Herring, and Niell versus Lanyi on the California–Spain baseline. The Ifadis, Herring, and Niell mapping functions all worsen the scatter by at least 10 mm on both baselines. The Davis function resembles Lanyi most closely in regard to scatter.

Table 4. Baseline-length results.

Mapping function ^a	Baseline							
	DSS 15–DSS 45 ^b		DSS 15–DSS 65		DSS 15–DSS 45 ^{c,d}		DSS 15–DSS 65	
	δ_B	χ^2	δ_B	χ^2	$\Delta\delta_B^e$	β_B	$\Delta\delta_B$	β_B
1	49.4	5.8	24.6	8.0	–40.8	173	–20.6	172
2	42.1	4.2	24.3	7.8	–31.5	90	–20.3	112
3	38.0	3.1	24.8	7.9	–25.8	26	–20.9	9
4	27.9	1.7	13.4	2.2	—	—	—	—
5	27.9	1.7	13.4	2.2	0.0	–3	0.0	–2
6	38.4	1.7	19.9	3.0	–26.4	6	–14.7	6
7	28.1	1.6	13.9	2.3	–3.3	20	–3.7	15
8	29.9	2.0	17.3	3.8	–10.8	62	–10.9	41
9	32.5	2.3	17.1	3.7	–16.7	61	–10.6	40
10	30.3	2.0	16.2	3.3	–11.8	58	–9.1	45

^a Mapping function definitions are provided in Table 1.

^b δ_B = scatter (mm).

^c $\Delta\delta_B$ = improvement of scatter (mm).

^d β_B = bias (mm).

^e Relative improvement is based on mapping function 4.

The Lanyi map with parameters estimated achieves its residual improvement at the expense of increases in baseline scatter of 26 and 15 mm. The purpose of this fit is to model any remaining variation in the mapping function due to inaccurate atmospheric modeling; indeed, the rms observable residuals decrease. This improvement appears to be achieved by propagating the delay residual errors into systematic baseline errors. The mechanism is poorly understood and is currently under further study. Most mapping functions also yield substantial baseline length biases relative to Lanyi 1984, ranging from 20 to 62 mm for the California–Australia 10,600-km baseline (the Chao tables do very well in this regard, shifting the scale by only 26 mm).

III. Conclusions

Based on the comparisons discussed in the previous section, a number of conclusions can be drawn. In the absence of more generalized testing, these conclusions should be regarded as limited to mapping the tropospheric delays at the three DSN sites. The statistical significance of differences among the mapping functions is tied to the residual scatter of these particular DSN VLBI fits. First, all three variants of the Chao mappings are very inadequate. They produce postfit residual and baseline scatter values that are inferior by many centimeters relative to the more modern functions. Second, the 1984 Lanyi mapping function is still, by a small margin, equivalent or superior to all the newer algorithms developed during the intervening decade. Third, those mapping functions that employ no or limited surface meteorological data are either equivalent to or slightly worse than the Lanyi function. Some of their deficiencies may be due to the fact that their functional forms are based on atmospheric profiles measured at predominantly North American sites. Finally, of the functions using minimal or no surface data, either the Ifadis or Niell function produces tropospheric models of DSN VLBI measurements that are nearly equal in quality to those given by the Lanyi mapping function on the California–Australia baseline, but worse on the California–Spain baseline.

Of the three major JPL astrometric/geodetic/navigational software packages, MODEST and the GPS analysis software GIPSY have always used the Lanyi 1984 function for tropospheric delay calibration. The former also has options to employ any of the models discussed in the present article. The Chao table formulation in ODP is seen to be grossly inadequate, especially at low elevation angles, and should be updated to one of the modern functions.

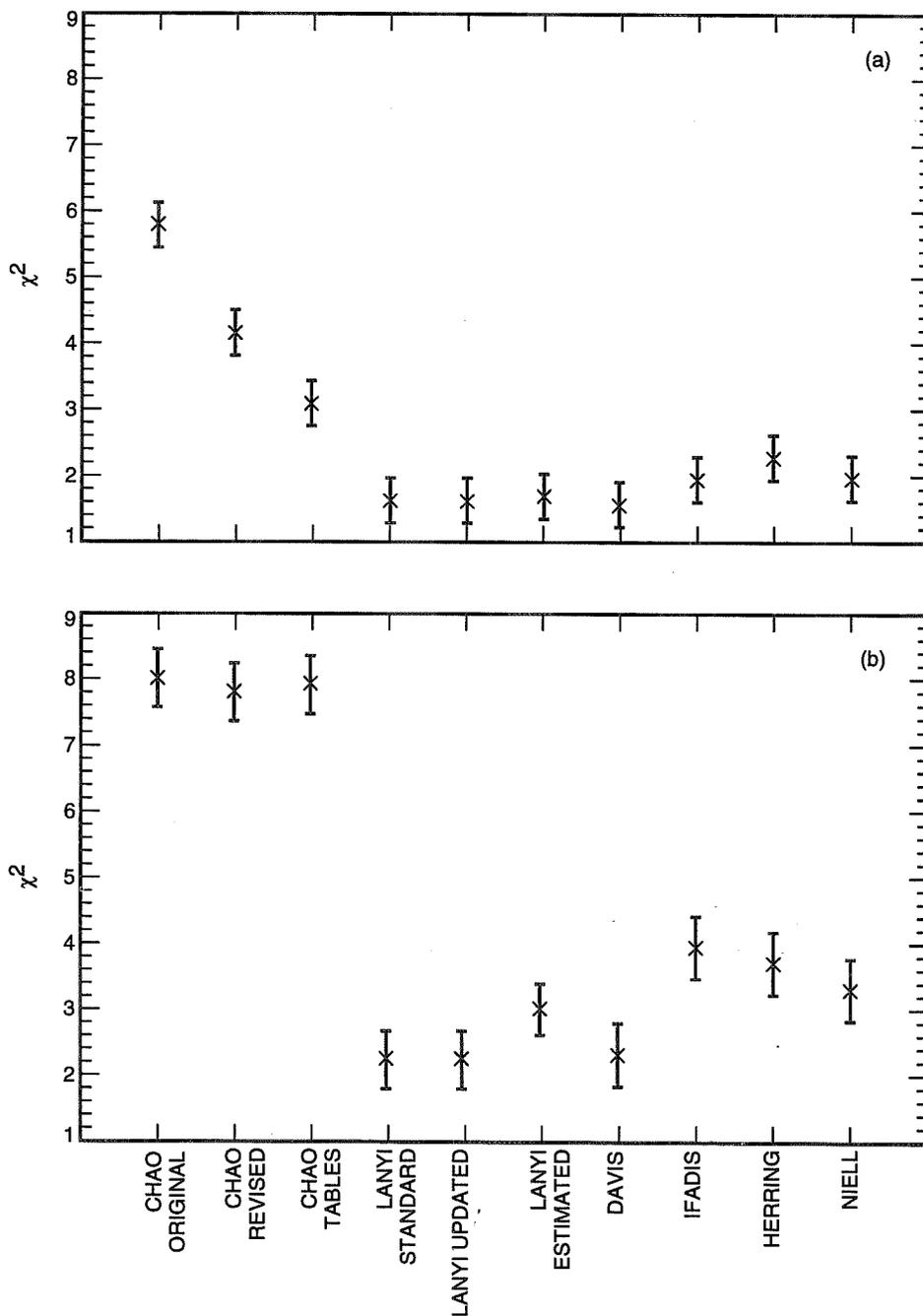


Fig. 3. Normalized χ^2 of residuals from a linear fit to baseline lengths versus time for (a) the California–Australia baseline and (b) the California–Spain baseline. See text for explanation of error bars.

Acknowledgments

We are grateful to A. E. Niell for supplying his new algorithms prior to publication. The present work originated in discussions of the merits of various tropospheric mapping functions with J. A. Estefan during the past 2 years.

References

- [1] G. E. Lanyi, "Tropospheric Delay Effects in Radio Interferometry," *The Telecommunications and Data Acquisition Progress Report 42-78*, vol. April-June 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 152-159, August 15, 1984.
- [2] J. L. Davis, T. A. Herring, I. I. Shapiro, A. E. E. Rogers, and G. Elgered, "Geodesy by Radio Interferometry: Effects of Atmospheric Modeling Errors on Estimated Baseline Length," *Radio Science*, vol. 20, pp. 1593-1607, 1985.
- [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, March-April 1987.
- [4] T. E. Gallini, *A Survey of Tropospheric Refraction Models*, Aerospace Report TOR-94(4488)-11, The Aerospace Corporation, El Segundo, California, April 20, 1994.
- [5] O. J. Sovers and C. S. Jacobs, *Observation Model and Parameter Partial for the JPL VLBI Parameter Estimation Software "MODEST"-1994*, JPL Publication 83-39, Rev. 5, Jet Propulsion Laboratory, Pasadena, California, August 1994.
- [6] O. J. Sovers and J. S. Border, *Observation Model and Parameter Partial for the JPL Geodetic GPS Modeling Software "GPSOMC"*, JPL Publication 87-21, Rev. 2, Jet Propulsion Laboratory, Pasadena, California, June 15, 1990.
- [7] *IERS Standards, IERS Technical Note 13*, edited by D. D. McCarthy, International Earth Rotation Service, Paris, France, 1992.
- [8] T. A. Herring, "Modeling Atmospheric Delays in the Analysis of Space Geodetic Data," *Refraction of Transatmospheric Signals in Geodesy*, edited by J. C. DeMunck and T. A. Th. Spoelstra, The Netherlands Commission on Geodesy, Delft, Netherlands, 1992.
- [9] A. E. Niell, "A New Approach for the Hydrostatic Mapping Function," *Proceedings of the International Workshop for Reference Frame Establishment and Technology Development in Space Geodesy*, Communications Research Laboratory, Koganei, Tokyo, Japan, pp. 61-68, January 1993.
- [10] 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," *Astron. J.*, vol. 95, pp. 1647-1658, 1988.
- [11] O. J. Sovers, "JPL 1990-3: A 5-nrad Extragalactic Source Catalog Based on Combined Radio Interferometric Observations," *The Telecommunications and Data Acquisition Progress Report 42-106*, vol. April-June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 364-383, August 15, 1991.

- [12] D. S. MacMillan and C. Ma, "Evaluation of Very Long Baseline Interferometry Atmospheric Modeling Improvements," *J. Geophys. Res.*, vol. 99, no. B1, pp. 637-651, 1994.
- [13] C. C. Chao, *The Troposphere Calibration Model for Mariner Mars 1971*, Technical Report 32-1587, Jet Propulsion Laboratory, Pasadena, California, pp. 61-76, March 1974.
- [14] C. C. Chao, "New Tropospheric Range Corrections With Seasonal Adjustment," *The DSN Progress Report for September and October 1971*, Technical Report 32-1526, vol. VI, Jet Propulsion Laboratory, Pasadena, California, pp. 67-82, December 15, 1971.
- [15] I. I. Ifadis, *The Atmospheric Delay of Radio Waves: Modeling the Elevation Dependence on a Global Scale*, Technical Report 38L, Chalmers University of Technology, Göteborg, Sweden, 1986.
- [16] V. B. Mendes and R. B. Langley, "Modeling the Tropospheric Delay From Meteorological Surface Measurements: Comparison of Models" (abstract), *EOS Transactions*, vol. 75, no. 16, p. 105, May 23-27, 1994.
- [17] J. A. Estefan and O. J. Sovers, *A Comparative Survey of Current and Proposed Tropospheric Refraction-Delay Models for DSN Radio Metric Data Calibration*, JPL Publication 94-24, Jet Propulsion Laboratory, Pasadena, California, October 1994.

An Algorithm for Extraction of Periodic Signals From Sparse, Irregularly Sampled Data

J. Z. Wilcox

Tracking Systems and Applications Section

Temporal gaps in discrete sampling sequences produce spurious Fourier components at the intermodulation frequencies of an oscillatory signal and the temporal gaps, thus significantly complicating spectral analysis of such sparsely sampled data. A new fast Fourier transform (FFT)-based algorithm has been developed, suitable for spectral analysis of sparsely sampled data with a relatively small number of oscillatory components buried in background noise. The algorithm's principal idea has its origin in the so-called "clean" algorithm used to sharpen images of scenes corrupted by atmospheric and sensor aperture effects. It identifies as the signal's "true" frequency that oscillatory component which, when passed through the same sampling sequence as the original data, produces a Fourier image that is the best match to the original Fourier map. Unlike the clean algorithm, it performs the search in the Fourier space. The algorithm has generally met with success on trials with simulated data with a low signal-to-noise ratio, including those of a type similar to hourly residuals for Earth orientation parameters extracted from VLBI data. For eight oscillatory components in the diurnal and semidiurnal bands, all components with an amplitude-noise ratio greater than 0.2 were successfully extracted for all sequences and duty cycles (greater than 0.1) tested; the amplitude-noise ratios of the extracted signals were as low as 0.05 for high duty cycles and long sampling sequences. When, in addition to these high frequencies, strong low-frequency components are present in the data, the low-frequency components are generally eliminated first, by employing a version of the algorithm that searches for noninteger multiples of the discrete FFT minimum frequency.

I. Introduction

In observational sciences like astronomy, it frequently happens that data are available at an unevenly spaced set of sampling times. The analyst may wish to extract from these data certain periodic components for which the frequencies are not precisely known. In fact, signal extraction from sparse or unevenly sampled data may prove crucial to the successful exploitation of large data sets assembled at considerable expense over many years. If the data acquisition process is characterized by a regular or nearly periodic succession of gaps during which no data are taken, sidelobes will occur at frequencies corresponding to the intermodulation products of the various signal and gap frequencies, giving rise to spurious signals in the Fourier map. The situation is aggravated by the inevitable presence of noise. In recent years, a number of authors have developed methods based on Lomb's normalized periodogram to deal with problems of this kind. A set of references and an implementation of Lomb's method are provided in [1]. A recent

article studied the effect of temporal windows on observed solar oscillation parameters [2]. However, there appears to be no currently available method to search for periodic signals with a low signal-to-noise ratio and unknown frequencies spanning a potentially broad spectral range; nor does there exist a clear understanding of the mathematical limits on the kinds of information that might be extracted from such data.

This article describes a new mathematical procedure for extraction of periodic signals from incomplete data sets. The procedure addresses a situation of particular interest in which only a relatively small number of periodic components are present within a noisy background. Although it may not be possible even in principle to extract a large array of broadband signals from sparsely sampled data, the fact that only a small number of components are present makes it possible to extract the components in a manner that effectively eliminates ambiguity. The main idea of the new procedure comes from the so-called “clean” algorithm [3], which has been extensively used within the astronomical community to sharpen images of scenes corrupted by atmospheric and sensor aperture effects. Clean algorithms have been most successful in cleaning up maps of images created by a small number of intense sources. The location of the most intense point source is identified with the highest peak in the scene’s “dirty” map. By modeling the effects of the atmosphere and sensor aperture (or other intermediate filters) on the image of a single localized source, the resulting point spread function is subtracted off, and a point reconstruction added to a clean map. The process is repeated for successively smaller peaks.

Analogous to the clean algorithm, the new algorithm identifies the strongest oscillatory component and produces a Fourier image of this signal as it would be registered if it were sampled in the same manner as the actual data. Unlike in the clean algorithm, a search for the component is performed in Fourier space. That frequency is identified as the frequency of the strongest oscillatory component whose Fourier image is the best match to the Fourier map of the actual data. The image is subtracted from the original Fourier map. The difference between the old map and the subtracted image is a new residual map, and the amplitude and frequency of the generating signal are recorded as contributors to the clean map.

The algorithm has generally met with success in trials with simulated data, including those of a type similar to hourly residuals for Earth orientation parameters extracted from VLBI data. Some eight signal components in the diurnal and semidiurnal bands with amplitudes varying by an order of magnitude were added to a background of Gaussian noise on the same order of magnitude as the oscillatory signal. For a 2-year-long sampling sequence and 20-percent duty cycle (the duty cycle is defined as the ratio of the length of time during which the data was taken to the sequence’s total length), all components with an amplitude–noise ratio greater than about 0.2 were successfully extracted; for a 100-percent duty cycle, the extracted amplitude–noise ratio was as low as 0.05. To minimize central processing unit (CPU) time (important for long data sequences), the algorithm uses fast Fourier transforms (FFTs) on uniformly spaced time markers. The data arrays have zeroes added into the temporal gaps and into the appended region (which is added to make the array size equal to a power of two). The CPU time scales as a constant $\times N \times \log N$, where the constant depends on the efficiency of the search for the signal oscillatory frequencies. For $N = 2^{18}$ time markers and 8 cleaning steps, the required CPU time on a VAX-class machine was on the order of several minutes. The frequency resolution is $1/T$, where T is the sampling sequence length. For cases when a more precise frequency determination is required, or when in addition to the high (i.e., diurnal and semidiurnal) frequencies, the data contains low-frequency components, a noninteger version of the algorithm has been developed. The noninteger version constructs the Fourier image as a function of frequencies that are not integer multiples of the discrete FFT minimum frequency. An additional search, performed by using a functional minimum-finding algorithm, extends the search to noninteger frequency values. Compared to the integer algorithm, the noninteger algorithm determines the signal frequencies with greatly improved accuracy. It also consumes more CPU time, depending on the desired accuracy and on the spectral content of the original data.

Section II describes the algorithm; Section III describes results of simulation experiments; and the Appendix describes the algorithm extension to frequencies that are not equal to discrete FFT frequencies.

II. The Algorithm

The data are sampled in a sequence that consists of discrete observation sessions (or windows, on the order of one day) between which there are large gaps (on the order of several days). Within each session, data are sampled at uniformly spaced discrete time intervals that are integers of some minimum δt (in our case, 1 hr). The set of hour markers at which the data are taken defines the sampling function $S(t_n)$, which assumes values of 1 or 0 depending upon whether or not a data value was acquired at that particular time. The data define the array $D(t_n)$, in which the entries where no data have been taken have been set to zero. The hour markers are $t_n = n \delta t$, $n = 0, 1, \dots, N-1$, where N is the total number of markers in the sampling sequence of duration $T = N \delta t$. Note that, to use the FFT, it is most convenient for N to be chosen equal to a power of 2.

Sampling at discrete intervals creates a situation whereby frequencies beyond the maximum resolved frequency (Nyquist, equal in our case to $1/2\delta t$) are aliased (or “folded back”) into the lowest frequency region. In addition, the presence of periodic or nearly periodic gaps in the data sequence leads to the appearance of sidelobes and spurious spectral peaks at the signal and gap intermodulation frequencies. The sidelobes are also aliased into the lowest spectral region. To remove the sidelobe effect, the algorithm produces a Fourier map of the actual data and identifies the strongest oscillatory component. It calculates a Fourier image of this signal filtered through the same sequence, $S(t_n)$, as the actual data, and subtracts it from the original Fourier map to produce a residual map. The most obvious candidate for the strongest component is a frequency with the biggest peak in the original Fourier map. Occasionally, however, the biggest peak occurs at one of the intermodulation frequencies. Therefore, the “true” modulation frequency is found to be that frequency which when passed through $S(t_n)$ is the best match (including the sidelobe structure) for the original Fourier map.

Let M_l designate the least squares function:

$$M_l \equiv \sum_k^N |D^S(k) - \sum_{\pm} a_{\pm l} f_{\pm l}^S(k)|^2 \quad (1)$$

where the sum over k is a sum over all (discrete) frequencies, $\omega_k \equiv 2\pi k/(N\delta t)$. The $D^S(k)$ is a discrete Fourier transform of $D(t_n)$, and the second term on the right-hand side of Eq. (1) is a Fourier representation of a candidate oscillatory signal of frequency ω_l taken over the sampling function $S(t_n)$. That is, $f_l^S(k)$ is a Fourier transform of $f_l(t_n) S(t_n)$, where the periodic function $f_l(t_n) = \exp(-2\pi i \frac{n l}{N})$. Note that if there were no gaps in the data (i.e., if $S(t_n)$ consisted of an uninterrupted series of data points), then $f_l^S(k)$ would consist of a single Kronecker delta, $\delta_{l,k}$. However, because of the window-gap structure of $S(t_n)$, the $f_l^S(k)$'s will have sideband components, similar to sideband components of $D^S(k)$.

The a_l 's specify amplitudes of the Fourier components $f_l^S(k)$. It is important that the Fourier representation of a candidate oscillatory signal sums over both the positive and negative frequency indexes to account for signal phase. To see this, assume that the contribution of the l th oscillatory component to data is

$$A_l \cos\left(2\pi \frac{n l}{N} + \phi_l\right) \quad (2)$$

where ϕ_l is phase. The contribution to the Fourier map is

$$\frac{A_l}{2} (e^{i\phi_l} f_{-l}^S(k) + e^{-i\phi_l} f_l^S(k)) \quad (3)$$

Because Eq. (3) consists of both a positive and its complex conjugate negative frequency components, so must the Fourier image, $\sum_{\pm} a_{\pm l} f_{\pm l}^S(k)$, in Eq. (1).

The index l that minimizes M_l as a function of l is identified as the signal's true frequency. To perform the minimization, $D^S(k)$'s and $f_l^S(k)$'s are computed (for all k 's and candidate l 's). The coefficients $a_{\pm l}$ are computed from the following complex conjugate relationships for spectral strengths at $\pm l$ locations in the Fourier map:

$$D^S(l) = a_{-l} f_{-l}^S(l) + a_l f_l^S(l) \quad (4a)$$

$$D^S(-l) = a_{-l} f_{-l}^S(-l) + a_l f_l^S(-l) \quad (4b)$$

By combining Eqs. (4a) and (4b), the $a_{\pm l}$'s are computed for the candidate l . The computed $a_{\pm l}$'s are substituted into Eq. (1), whose minimization yields the optimal index l for the frequency of the strongest oscillatory component. Note that the use of the positive and negative l allows for reconstruction of the signal phase. By comparing Eqs. (3) and (4), the signal phase ϕ_l is given by the following expression:

$$\tan \phi_l = \frac{\text{Im } a_{-l}}{\text{Re } a_{-l}} \quad (5)$$

Hence, the use of the above-described procedure allows one to determine signal phase.

The above relationships, Eqs. (4) and (5), were derived by relating the spectral strengths at $\pm l$ locations. It is possible to formally derive Eq. (4) by minimization of Eq. (1) as a function of a_l . The minimization is performed in the Appendix, together with the extension of the algorithm to noninteger l values. Note also that the use of the $a_{\pm l}$'s computed from Eq. (4) permits complete subtraction of the spectral strengths at the $\pm l$ locations; if the data consist of a single frequency, then one subtraction yields a residual at the level of the machine round-off error or background noise, whichever is bigger.

For data modulated with several frequencies, multiple applications of the algorithm are necessary. Equations (1)–(5) will apply, except that on the $(i+1)$ th iteration, $D^S(k)$ is replaced by a residual array ($y_i(k)$) that has resulted from the i th iteration, $y_i(k) \equiv y_{i-1}(k) - \sum_{\pm} a_{\pm l_i} f_{\pm l_i}^S(k)$, where l_i is the optimal index that minimizes M_l at the i th iteration, $y_{i-1}(k)$ is the residual array that has resulted from the $(i-1)$ th iteration, and $y_o(k) \equiv D^S(k)$ is the Fourier transform of the original data. With multiple iterations, it may sometimes be desirable to subtract the spectral peaks in fractions. That is (as is a common practice in the application of the clean algorithm), one subtracts at each iteration only a fraction of the image that has been constructed by using the peak's full strength. However, the phase relationship between the subtracted $a_{\pm l}$'s should be the same as that computed from Eq. (4). Sidelobes associated with one frequency can contribute signal at another frequency. Subtracting the peaks in fractions permits one to correct for sidelobe contributions associated with other signal frequencies. In any case, the algorithm first removes (or partially removes) the most intense component (including sidebands due to the window-gap structure), forms the residual array, and treats the residual as a new set of data. The process is iterated to yield a desired number of oscillatory components.

III. Results and Discussion

Two versions of the above-described general procedure have been implemented and tested on simulated data. In the first version, the oscillatory signal that minimizes Eq. (1) consists of two integers, $\pm l$, as in Eq. (3). This version works very well when all modulation frequencies are equal (or nearly equal) to discrete multiples of $1/T$. If T is the total length of the sampling sequence, then the spacing for the

resolved frequencies is $1/T$, corresponding to the resolution parameter of τ^2/T for an oscillatory signal with period τ . Thus, this version provides a sufficient resolution for relatively high frequencies, say, $\tau < 0.01 T$ (in which case, the resolution is better than 1 percent). Because the strongest components are extracted first, the accuracy with which these components are extracted impacts the extraction success of less intense components. The integer version works best when all (relatively rapid) modulation amplitudes are about the same order of magnitude. The second version of the algorithm searches for frequencies that are not integer multiples of $1/T$ (or, equivalently, the frequency indices are not integer multiples of $1/N$). This version allows one, at the expense of increased CPU time, to determine the oscillatory frequencies with greatly improved accuracy. This version must be used for analysis of data with strong low-frequency components.

A. Integer Frequency Index

The integer version was tested with success on many different types of simulated monochromatic as well as multifrequency signals; the modulations were imbedded in Gaussian noise. The success of frequency extraction depends on the number of modulation frequencies, the length and structure of the sampling sequence (including length, number, and regularity of individual observational sessions), and noise level.

For a monochromatic signal (data sampled hourly), the modulation frequencies ranged from those corresponding to periods of 4 hr to those corresponding to the length T of the data sequence. Table 1 summarizes the results of the simulations for $S(t_n)$ consisting of 20 and 200 1-day-long, periodically repeating sessions separated by several day-long gaps. The residual has been reduced to a background level by using a single subtraction step for signal-to-noise amplitude ratios as low as 0.17 and 0.04, respectively, when the duty cycle was 1; the minimum values of the successfully extracted ratios were 0.3 and 0.1, and 0.5 and 0.125, for the duty cycles of 0.2 and 0.1, respectively. Here, the duty cycle has been defined as the ratio of the number of days on which data were recorded to the total number of days in the observation period (e.g., for a 1-day-long observation session registered on every 5th day, the duty cycle is $1/5$). Results of simulations for nonperiodic sampling sequences were similar, except that the extraction tended to be successful even at somewhat higher noise levels. This is not unexpected, since constructive interference (leading to strong sidelobes) is less.

Results of simulation experiments for a signal consisting of several modulation frequencies are summarized in Tables 2 and 3 and Fig. 1. Eight oscillatory components (four in the semidiurnal band and four in the diurnal band) with amplitudes varying by an order of magnitude between the strongest and weakest component were added to a background of Gaussian noise. Typical $S(t_n)$'s consisted of 70, 140, and 740 1-day-long observation sessions (data sampled hourly) separated by 4-day- (duty cycle 0.2) and 9-day-long (duty cycle 0.1) gaps. The noise amplitude was as high as 100 (the amplitudes of the oscillatory components ranged from 3 to 23). The amplitudes and relative phase of the simulated signals are summarized in Table 2. To demonstrate the resolution for the extracted signals, the table also summarizes the results of the extraction process for the 140-data day, 0.2-duty cycle sampling sequence, with the noise level equal to the amplitude of the strongest component. All eight components were extracted within their resolution accuracy (for $T = 2^{16}$ hr, the resolution $\tau^2/T \simeq 0.002$ and 0.008 for semidiurnal and diurnal bands, respectively) in eight iteration steps. Figure 1(a) shows the original data, and Figs. 1(b), (c), and (d) show, respectively, the original, residual, and clean Fourier maps.

Results of simulations for the same oscillatory components but different observation parameters are summarized in Table 3. To investigate the effect of excess iterations, the number of cleaning steps used was 10. In Table 3, the symbol s designates an oscillatory component extracted within its expected resolution, c designates an extraction of a previously extracted component, i designates extraction at one of the intermodulation frequencies, and n designates noise (any frequency higher than $1/6 \text{ h}^{-1}$ was considered a noise). A component may be extracted twice (or several times) because of our use of the full spectral strength $D^S(\pm l)$ [Eq. (4)] for the subtraction of the Fourier image. The strength $D^S(\pm l)$ consists of contributions from the l th oscillatory component as well as from sidelobes produced by other frequencies

Table 1. Extraction results for simulated monochromatic data.^a

Duty cycle	Minimum extracted signal-noise amplitude ratio	
	<i>N</i> data days	
	20	200
1	≥0.17	≥0.04
0.2	≥0.3	≥0.1
0.1	≥0.5	≥0.125

^aThe listed numbers are the minimum values for the signal-noise amplitude ratio for which the extraction was successful. (The signal periodicity was $\tau = 11.9672$ hr, data sampled hourly.)

Table 2. Extraction results for simulated data with four oscillatory components in the diurnal and four in the semidiurnal band.^a

Periodicity, hr	Relative amplitude	Phase, deg
Simulated data		
11.9672	17.0	126
12.0000	8.6	92
12.4206	3.7	86
12.6583	3.2	103
23.9345	23.0	200
24.0659	7.0	240
25.8193	19.0	69
26.8684	5.3	230
Extracted signals		
11.9679	16.4	139
12.0007	9.2	104
12.4215	4.1	103
12.6591	4.2	106
23.9357	22.8	205
24.0676	6.0	243
25.8219	20.0	80
26.8700	6.2	243

^aThe background noise relative amplitude was 30. The sampling sequence consisted of 140 data days separated by 4-day-long gaps (duty cycle 0.2). Data were sampled hourly; total number of hour markers in the array was 2^{16} . The frequencies were extracted within the resolution of $12^2/2^{16}$ and $24^2/2^{16}$ for the semidiurnal and diurnal bands, respectively.

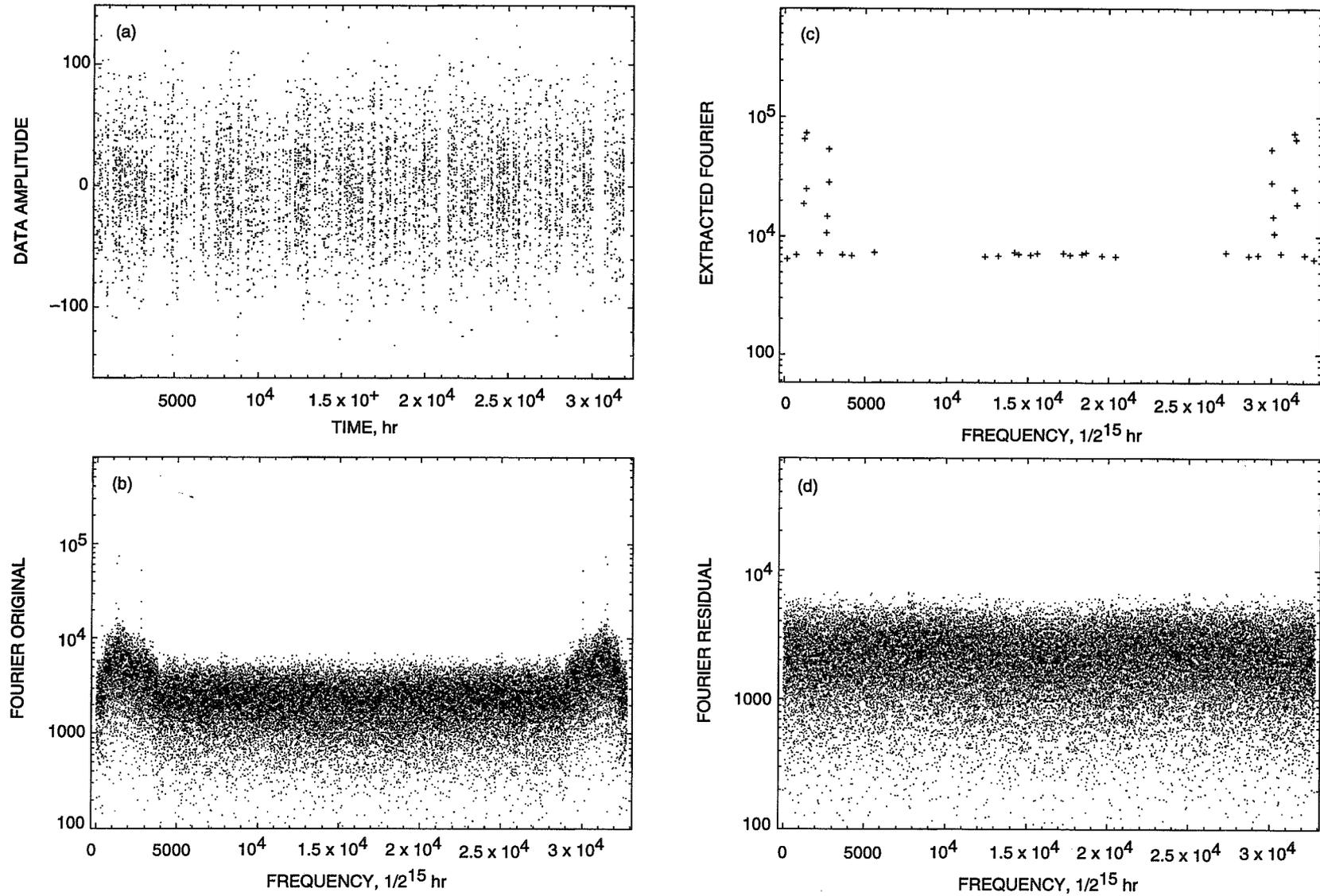


Fig. 1. Results of simulation experiments: (a) the original data, (b) the original Fourier map, (c) the clean Fourier map, (d) the residual Fourier map for simulated data with four diurnal and four semidiurnal signals (parameters of Table 2) buried in the background of Gaussian noise. The noise level was 30, the duty cycle was 0.2, the number of data days was 270 (data sampled hourly). The signals were extracted in eight iteration steps.

Table 3. Extraction results for the same eight oscillatory components as in Table 2 for periodically repeating observational sessions.

Noise level	Duty cycle	<i>N</i> data days		
		740	140	70
0	1	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>i</i>
	0.2	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>
	0.1	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>	7 <i>s</i> , 3 <i>c</i>
10	1	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>i</i>
	0.2	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 1 <i>c</i> , 1 <i>i</i>
	0.1	1 <i>i</i> , 3 <i>s</i> , 5 <i>i</i> , 1 <i>s</i> [5 <i>s</i> , 2 <i>c</i> , 3 <i>s</i>]	8 <i>s</i> , 2 <i>c</i> [8 <i>s</i> , 2 <i>c</i>]	8 <i>s</i> , 2 <i>i</i> [8 <i>s</i> , 1 <i>c</i> , 1 <i>i</i>]
30	1	8 <i>s</i> , 2 <i>c</i>	8 <i>s</i> , 2 <i>n</i>	6 <i>s</i> , 4 <i>n</i>
	0.2	8 <i>s</i> , 2 <i>c</i> [8 <i>s</i> , 2 <i>c</i>]	7 <i>s</i> , 1 <i>i</i> , 2 <i>n</i> [8 <i>s</i> , 2 <i>n</i>]	6 <i>s</i> , 1 <i>i</i> , 3 <i>n</i> [7 <i>s</i> , 3 <i>n</i>]
	0.1	1 <i>i</i> , 3 <i>s</i> , 5 <i>i</i> , 1 <i>s</i> [3 <i>s</i> , 1 <i>c</i> , 5 <i>s</i> , 1 <i>c</i>]	6 <i>s</i> , 2 <i>i</i> , 2 <i>n</i> [8 <i>s</i> , 2 <i>n</i>]	5 <i>s</i> , 1 <i>i</i> , 4 <i>n</i> [6 <i>s</i> , 1 <i>i</i> , 3 <i>n</i>]
100	1	6 <i>s</i> , 2 <i>i</i> , 2 <i>n</i>	4 <i>s</i> , 6 <i>n</i>	3 <i>s</i> , 7 <i>n</i>
	0.2	5 <i>s</i> , 1 <i>c</i> , 3 <i>i</i> , 1 <i>n</i> [6 <i>s</i> , 2 <i>n</i> , 2 <i>i</i>]	4 <i>s</i> , 1 <i>n</i> , 5 <i>i</i> [4 <i>s</i> , 1 <i>i</i> , 5 <i>n</i>]	3 <i>s</i> , 7 <i>n</i> [3 <i>s</i> , 7 <i>n</i>]
	0.1	3 <i>s</i> , 1 <i>n</i> , 6 <i>i</i> [5 <i>s</i> , 2 <i>c</i> , 3 <i>n</i>]	3 <i>s</i> , 2 <i>n</i> , 5 <i>i</i> [3 <i>s</i> , 7 <i>n</i>]	2 <i>s</i> , 3 <i>i</i> , 5 <i>n</i> [4 <i>s</i> , 2 <i>i</i> , 4 <i>n</i>]

(sidelobes produced by one frequency contribute spurious strength at other frequencies). Therefore, the use of the full $D^S(\pm l)$ may add (or subtract) signal strength at the l th or other frequency. The appearance of c 's (c stands for correction) in Table 3 signifies that some of the excess strength has been recovered on subsequent iteration of the algorithm (other side effects may be errors at other frequencies or noise). The number of successful extractions decreases with decreasing T , decreasing duty cycle, and increasing noise. In Table 3, successful extractions range from all eight (low noise, long T case) to two components (70 data days, a 0.1 duty cycle, noise amplitude of 100, corresponding to the minimum extracted amplitude–noise ratio of 0.2). For 140 and 740 data days, the critical amplitude–noise ratio is 0.1 and 0.07, respectively, when the duty cycle is 0.2; the ratio is somewhat less (higher) for higher (lower) duty cycles. Note that in case of zero noise, the extraction was complete in most cases.

Because of high sidelobes, the most unfavorable situation for extraction is a periodic sampling sequence. In Table 3, results are also listed (in square brackets) for nonperiodic sequences characterized by the nonperiodicity parameter σ_p of about 1 (corresponding to randomly distributed observation sessions). Here, σ_p is the normalized root-mean-square deviation of the separation p_i between the beginning times for subsequent sessions,

$$\sigma_p \equiv \sqrt{\frac{1}{N_d} \sum_i^{N_d} \left(\frac{p_i}{\langle p \rangle} - 1 \right)^2} \quad (6)$$

where $\langle p \rangle$ is the average separation, $\langle p \rangle = \sum_i^{N_d} p_i / N_d$, where N_d is the total number of (1-day-long) observation sessions. The main effect of a finite (nonzero) σ_p is to somewhat suppress the height of the spectral peaks at the signal and gap intermodulation frequencies. The suppression is most effective for

long sampling sequences, because (for periodically repeating sessions) the sidelobes are biggest there. For example, for 740 data days, a 10-percent duty cycle, and a noise level of 100, the randomization increased (Table 3) the number of extracted frequencies from 3 to 5. The randomization makes the extraction easier to obtain if the sampling sequence is long and random.

In the majority of the simulations, the optimal frequency index was found to coincide with the frequency of the highest spectral peak in the current Fourier map. Occasionally, however, the second or third (or tenth) strongest peak actually produced the smallest quadratic residual function M_l , suggesting that during the search for the optimal frequency, one should test some finite number (N_p) of the biggest spectral peaks. For the sequences tested, we have found that $N_p \leq 50$ was quite sufficient. The fact that only a relatively small number of candidate l 's need to be inspected makes the minimization of the M_l rather efficient. For N_p inspected peaks, the CPU time scales as $N_c \times N_p \times N \times \log N$, where N_c is the number of iterations (typically not higher than several tens), and where $N \times \log N$ is the scale factor for CPU time for FFT of data arrays of dimension N . (If all indices l were to be tested, the CPU time would scale as $N_c N^2 \log N$, which for large arrays would be excessive.) For $N_c = 8$, $N_p = 10$, and $N = 10^{18}$, the CPU time was about 1 minute on a VAX 3000/400 and several minutes on a VAX 4000/90.

B. Noninteger Frequency Index

The precision of the extracted frequencies affects the accuracy of the extracted phase and, ultimately, the success of retrieving oscillatory components with lower amplitudes. To improve the extraction accuracy, the total length of the sampling sequence T can be increased (by padding it with zeroes) to a higher power of two. Or, if the array size is limited, a second version of the algorithm can be used that involves nonintegral l 's. This noninteger version is especially useful if, in addition to high frequencies, the data contain slowly varying signals (for τ on the order of or longer than T , the resolution τ^2/T resulting from the use of the standard discrete FFT would be on the order of or greater than τ itself). The use of the noninteger algorithm allows for elimination of strong low-frequency components before the extraction of high-frequency components.

If the data include frequencies not equal to an integer multiple of $1/N$, the residual is minimum for a set $\pm(l + \Delta l)$, where $0 \leq \Delta l \leq 1$. The noninteger algorithm searches for the optimal Δl by using a noninteger FFT together with a minimization scheme that does not require construction of derivatives. The main mathematical basis for the algorithm is described in the Appendix. The noninteger algorithm was tested by using the same simulated data as for the integer algorithm. The results were similar in that when one algorithm met with success, so did the other algorithm. The trade-off was a bigger array size for the integer algorithm versus longer CPU time for the noninteger algorithm. To achieve a fractional accuracy of $\delta\tau/\tau$ for an oscillatory signal with the period τ , the array size (i.e., the total length of the sampling sequence T) used by the integer algorithm must be greater than $\tau(\tau/\delta\tau)$. For the noninteger algorithm, the required array size is less; however, the CPU time is increased due to the use of the search and minimization procedure that optimizes Δl . For the eight oscillatory components in Table 2, CPU time for the noninteger algorithm was typically a factor of 10 bigger, while the frequency accuracy increased by one significant place for the same array size in both algorithms.

The principal application of the noninteger algorithm is for spectral analysis of data with both rapid and slow modulations present. To simulate a train of hourly residuals of Earth orientation parameters obtained from reduced VLBI data, the simulated data consisted of the same four near-diurnal and four near-semidiurnal oscillatory frequencies as in Table 2, superimposed on a slowly varying signal that was represented by a sum of 1-month, 1-year, and 18-year periodic components. The signal was imbedded in a background of Gaussian noise on the same order of magnitude as the strongest diurnal component. The extraction success depends on the length and structure of the sampling sequence, noise level, and the number and amplitudes of low-frequency components. A typical sampling sequence consisted of 270 data days (data sampled hourly) separated by 4-day-long gaps (with a corresponding duty cycle of 0.2 and a total sampling sequence length of $T = 3.5$ yr). Table 4 summarizes the simulation parameters for

three cases: In the first case, the amplitudes of the 1-month, 1-year, and 18-year components were 500, 2000, and 0, respectively; in the second case, the amplitudes were 500, 2000, and 200,000, respectively; in the third case, the amplitudes were 500, 20,000, and 500,000, respectively. The noise level was 30, and the eight diurnal and semidiurnal signals were the same as in Table 2 (amplitudes between 3.2 and 23). Figures 2, 3, and 4 show the corresponding original data and the original, residual, and clean Fourier maps. On comparison, the figures illustrate the effect of the number and amplitudes of low-frequency components on the success of extraction of high-frequency components.

Table 4. Extraction results when the data also contained strong low-frequency components (corresponding to Figs. 1, 2, 3, and 4).^a

Relative amplitudes of low-frequency components		Iterations for extraction of high-frequency components
(18 yr):	0	
(1 yr):	0	1, 2, 3, 4, 5, 6, 7, 8
(1 mo):	0	
(18 yr):	0	
(1 yr):	2,000	4, 5, 6, 7, 8, 9, 10, 11
(1 mo):	500	
(18 yr):	200,000	
(1 yr):	2,000	11, 14, 15, 18, 20, 22, 25, 31
(1 mo):	500	
(18 yr):	500,000	
(1 yr):	20,000	26, 28, 31, 47, 53, 65, 72, 86
(1 mo):	500	

^a The listed numbers designate iterations that extracted one of the high-frequency components. The high-frequency components are the same as in Table 2, the number of data days was 270 (data sampled hourly), the duty cycle was 0.2, and the noise level was 30.

In Figs. 2–4, iterative application of the algorithm has resulted in successful extractions of the diurnal and semidiurnal frequencies in all cases; however, the extraction process was lengthier (resulting in richer recovered spectral content) when the number and amplitudes of the original low-frequency components were higher. In general, the algorithm finds the strongest (in these cases, the low-frequency) components first. However, since the values of the extracted frequencies and phase are always somewhat in error (the accuracy will eventually be limited by the available numerical accuracy), the effect of errors in the values of the extracted parameters is made up by retrieving, in successive iterations, additional (false) frequencies, until the slow variations are matched sufficiently well for the high-frequency modulations to become the strongest Fourier components in the current residual map. This process of matching the time sequence with a richer Fourier spectrum becomes more intricate with increasing amplitudes and complexity of the original spectrum. In Fig. 2, the original spectrum contains no 18-year component. The high-frequency signals were found immediately after the low (1-year and 1-month) frequency signals, except that (due to a finite length T of the data array) the algorithm also found (as it usually does when variations on a scale comparable to or longer than T are present) a frequency $\simeq 1/T$. In Figs. 3 and 4, the original spectrum includes an 18-year-long component with the amplitude four orders of magnitude higher than

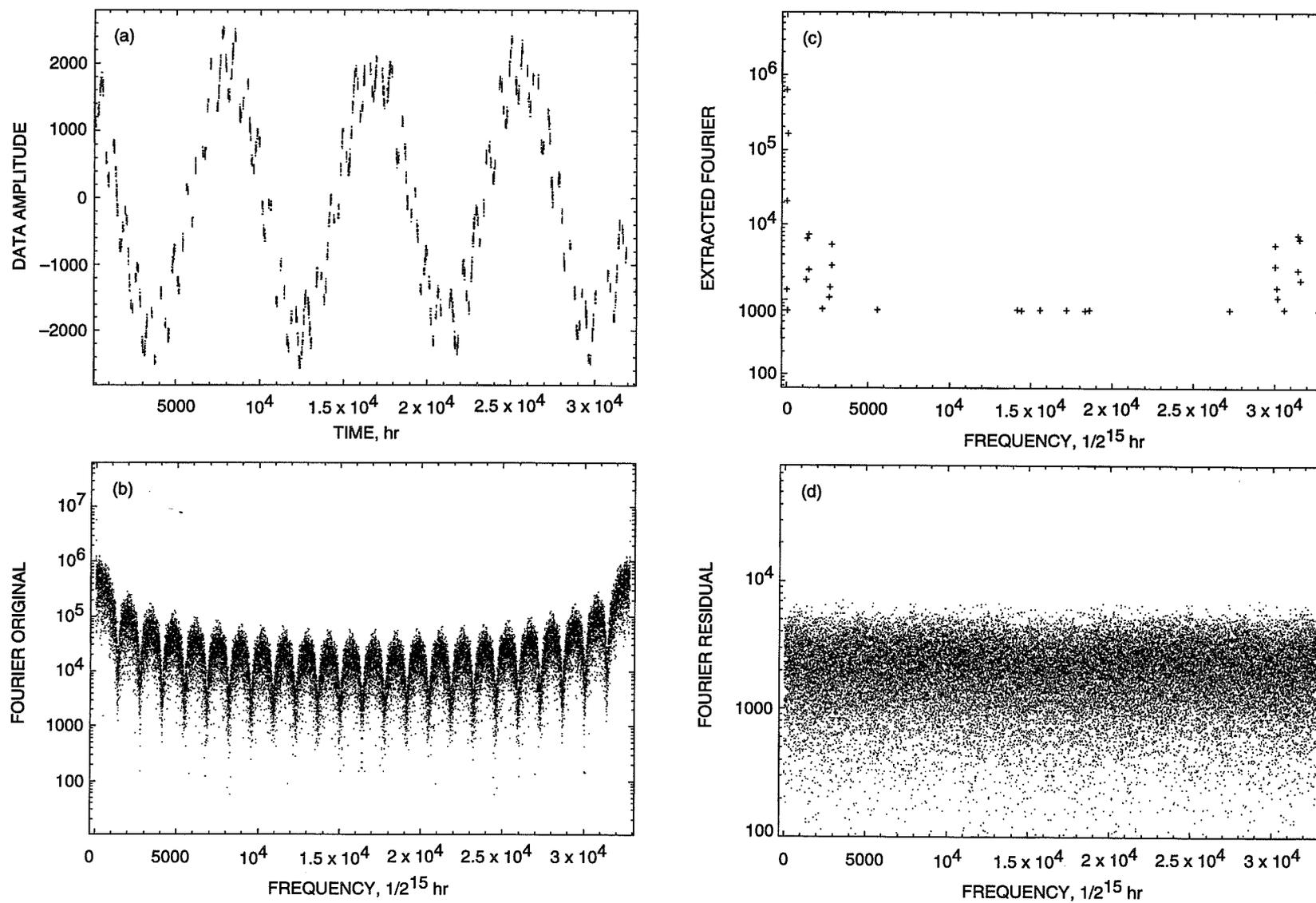


Fig. 2. Results for the same simulated data as in Fig. 1, but with the diurnal and semidiurnal components (Table 2) superimposed on a slowly varying signal with the amplitudes of 500 and 2000 for 1-mo and 1-yr periodic components respectively: (a) the original data, (b) the original Fourier map, (c) the clean Fourier map, (d) the residual Fourier maps.

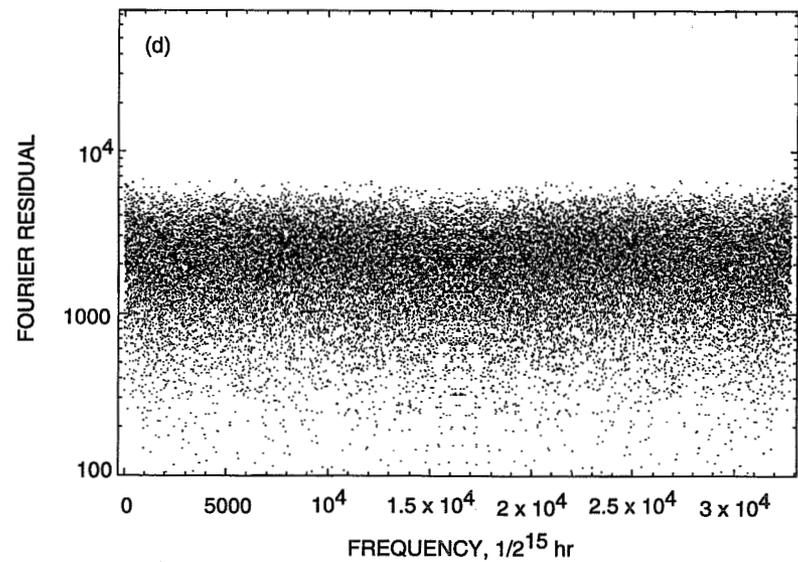
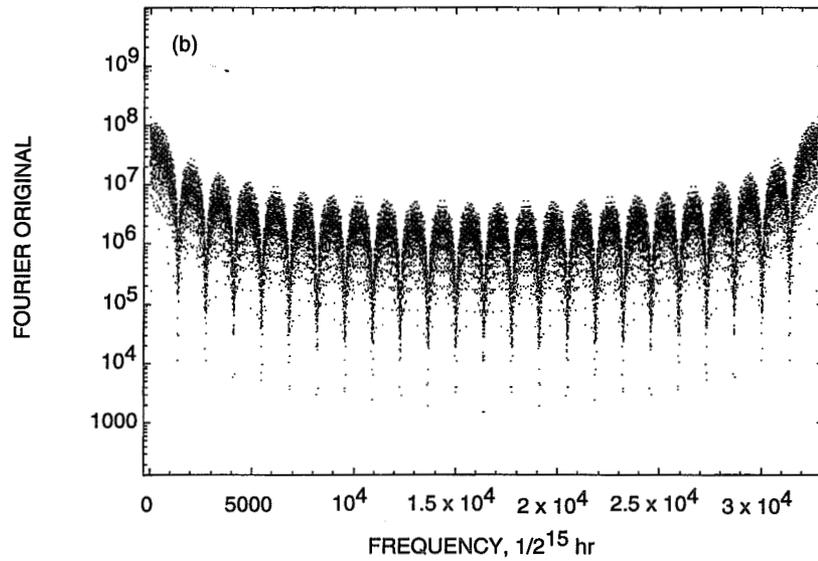
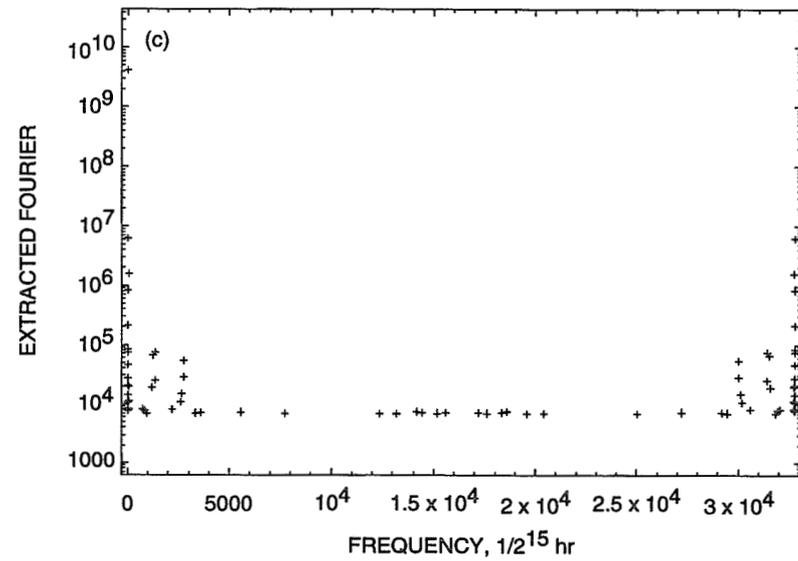
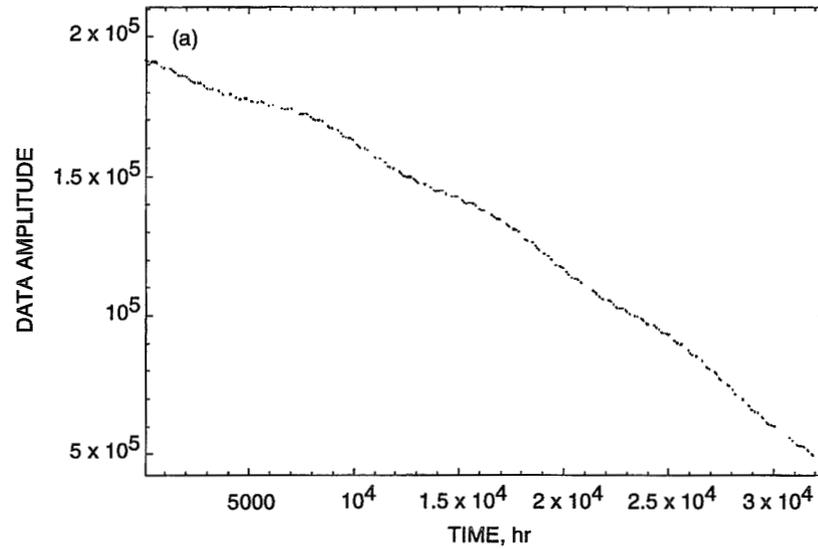


Fig. 3. Results for the same simulated data as in Figs. 1 and 2, but with the diurnal and semidiurnal components (Table 2) superimposed on a slowly varying signal with the amplitudes of 500, 2000, and 200,000 for a 1-mo, 1-yr, and 18-yr periodic components respectively: (a) the original data, (b) the original Fourier map, (c) the clean Fourier map, and (d) the residual Fourier map.

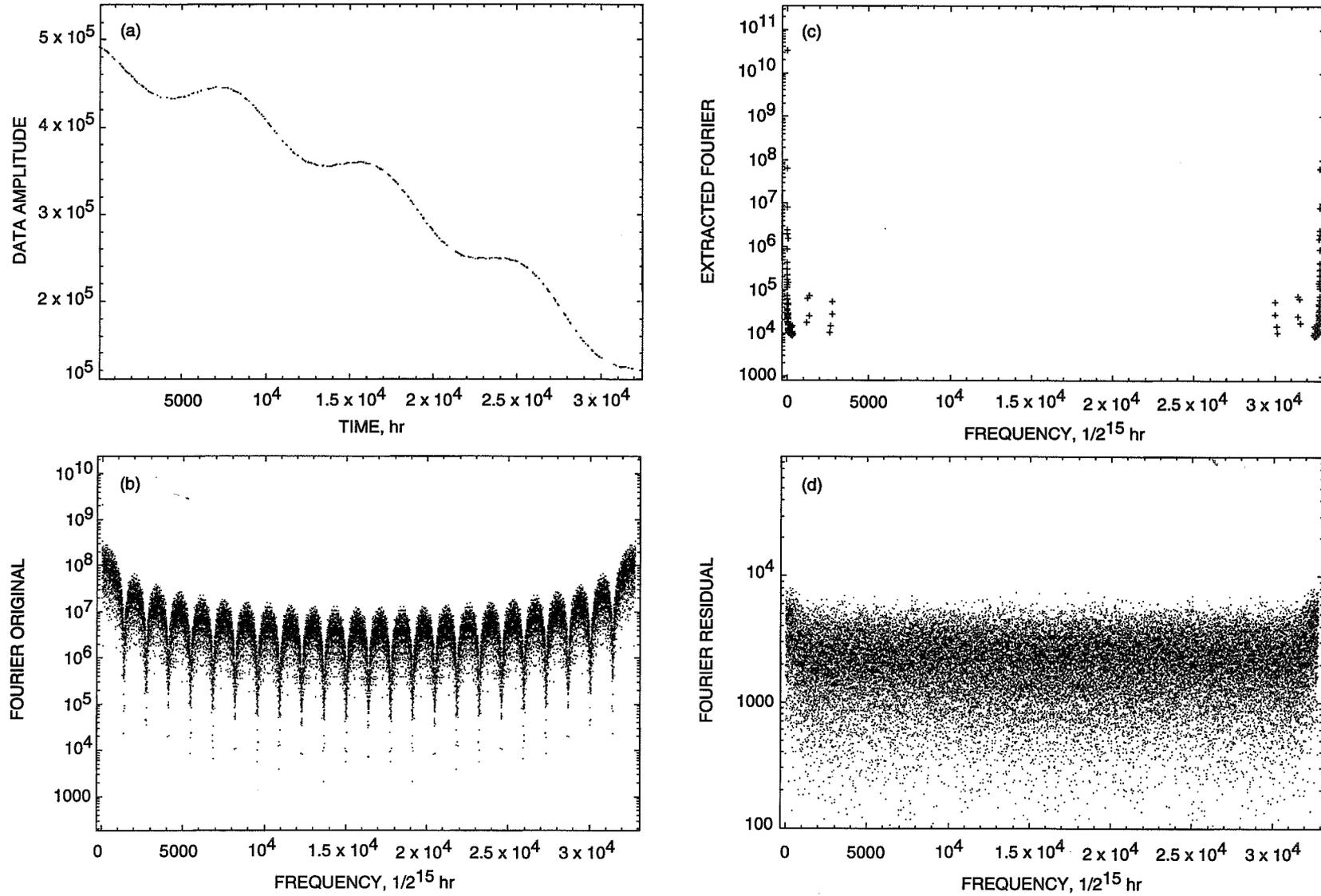


Fig. 4. Results for the same data as in Figs. 1, 2, and 3, but with the diurnal and semidiurnal components (Table 2) superimposed on a slowly varying signal with amplitudes of 500, 20,000, and 500,000 for 1-mo, 1-yr, and 18-yr periodic components, respectively: (a) the original data, (b) the original Fourier map, (c) the clean Fourier map, and (d) the residual Fourier map.

the strongest diurnal component. The algorithm also fits the slow variations first. However, because the variations are not fit exactly (the 18-year period was found with 2- and 3-percent accuracy in Figs. 3 and 4, respectively, and the 1-year and 1-month periods were found with 0.7- and 0.1-percent accuracy, respectively), the algorithm continues to search for additional frequencies until the fit is sufficiently tight. In Figs. 3 and 4, the high frequencies were extracted at the 11th, 14th, 15th, 18th, 20th, 22nd, 25th, and 31st iterations, and at the 26th, 28th, 31st, 47th, 53rd, 65th, 72nd, and 86th iterations, respectively. Thus, the complexity of the low-frequency content will play a significant role in establishing limits on the algorithm applicability.

IV. Summary and Conclusions

A new algorithm has been developed for spectral analysis of sparse, irregularly sampled data with low signal-to-noise ratios; the extracted parameters include frequency, phase, and amplitude of the signal components. The algorithm identifies the strongest component as that component which when passed through the same sampling sequence as the original data produces a Fourier image that is best matched to the current residual map. The algorithm has met with success in trials with simulated data, including those of a type similar to hourly residuals for Earth orientation parameters. The simulated data consisted of four near-diurnal and four semidiurnal oscillatory components with amplitudes varying by an order of magnitude. The background noise was on the same order of magnitude as the strongest high-frequency component. The number of data days ranged from 70 to 740 (data sampled hourly), and the duty cycle for the observation sessions varied between 1 and 0.1. Two versions of the algorithm were developed. In the integer version, the frequency solution is limited to $1/T$, where T is the sampling sequence length. With only the high-frequency components present in the data, the integer version was successful in extracting all components with an amplitude-noise ratio greater than about 0.2; the extraction success was higher (less) for longer (shorter) data sequences and higher (lower) duty cycles. Sessions that are not periodic make the extraction easier to achieve since the sidelobes (produced by the interference with the average periodicity of the sessions) are less than for the periodic case. Sidelobes associated with one frequency can contribute signal at other frequencies. If the residual map has been constructed by completely subtracting the identified spectral peak, the subtracted sidelobes can produce errors in this and other estimates. The solution to this problem is to subtract the peak in fractions. That is, at multiple applications of the algorithm, one subtracts from the current residual map only a fraction of the image that has been constructed by using the peak full strength. The effect of subtracted fractions on the extraction results could be used to achieve confidence in the found solutions. Another test of confidence can be obtained by modifying the observation strategy, e.g., by changing the duty cycle, the length of the sampling sequence, and the gap randomness. The frequency resolution can be increased by padding the data array with zeroes to a higher power of two (higher T).

For signal frequencies that are not integer multiples of the minimum discrete FFT frequency, the residual map will be minimized by a noninteger frequency value. We have implemented (and tested) an algorithm version that searches for a noninteger frequency index. The resolution is greatly improved over $1/T$ without requiring an increase in the array length. However, the CPU time is increased. In simulations with diurnal and semidiurnal bands, the typical CPU time was longer by a factor of 10, while the extracted frequency accuracy increased by one significant place relative to results obtained with the integer algorithm. Many real data also contain (in addition to the high frequencies) very strong low-frequency components. Unless some other scheme is used to eliminate the slow variations first, when applied, the algorithm fits the slowly varying signal with low frequencies before it searches for the high frequencies. The increase in the extraction accuracy comes at the expense of increased CPU time. The CPU time is increased depending on the signal frequencies, strength, and number of low-frequency components. Several types of simulated data with strong 18-year, 1-year, and 1-month periodic components were analyzed in detail in Section III.B (Figs. 2, 3, and 4). The accuracy of the extracted frequencies influences the efficiency of retrieval of components with lower amplitudes.

For uniformly spaced time markers, the CPU time scales as a constant $\times N_c \times N_p \times N \times \log N$, where N_c is a number of iterations of the algorithm (to find multiple oscillatory components), N_p is a number of candidate spectral peaks that must be inspected to minimize the residual sum, and N is the array total size chosen as a power of 2. For 2^{18} time markers in the sampling sequence and eight cleaning steps, typical CPU time for the integer algorithm was on the order of several minutes on a VAX-class machine. For the noninteger algorithm, the CPU time can exceed several hours, depending on the complexity of the overall frequency content.

The algorithm is currently applied to actual hourly Earth orientation parameters for the available International Radio Interferometric Surveying (IRIS), Crustal Dynamics Project (CDP), and DSN VLBI data.

Acknowledgments

The author is grateful to O. Sovers and C. Jacobs for encouragement and to R. Gross for critical reading of the manuscript.

References

- [1] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed., Cambridge: Cambridge University Press, Ch. 13, pp. 530–602, 1992, and references therein.
- [2] M. Lazrek and F. Hill, “Temporal Window Effects and Their Deconvolution From Solar Oscillation Spectra,” *Astron. Astrophys.*, vol. 280, pp. 704–714, 1993.
- [3] J. A. Hogbom, “Aperture Synthesis With a Non-Regular Distribution of Interferometer Baselines,” *Astron. and Astrophys. Supplement*, vol. 15, pp. 417–426, 1974.

Appendix

Spectral Analysis for a Noninteger Frequency Index

The method outlined in Section II is frequently sufficient provided that the periodic components are sufficiently few in number, well separated in frequency, not too disparate in the relative magnitude of their amplitudes, and not too corrupted by the presence of noise. If stressing conditions are present, the method can be enhanced by introducing fractional frequencies. The Fourier map becomes a function of the fractional value Δl . An additional search is performed on Δl to identify the residual minimum by using some standard function minimum-finding algorithm.

Similarly to Eq. (1) of the main text, define the minimal function

$$M_{l+\Delta l} \equiv \sum_k^N |D^S(k) - \sum_{\pm} a_{\pm(l+\Delta l)} f_{\pm(l+\Delta l)}^S(k)|^2 \quad (\text{A-1})$$

where all terms have the same meaning as in Eq. (1) except that the second term on the right-hand side of Eq. (A-1) represents a signal at the candidate frequency $\omega_{l+\Delta l} = 2\pi(l + \Delta l)/(N\delta t)$, and $f_{l+\Delta l}^S(k)$ is the Fourier transform of $f_{l+\Delta l}(t_n) S(t_n)$, where the periodic function $f_{l+\Delta l}(t_n) \equiv \exp(-2\pi i(n(l + \Delta l)/N)) S(t_n)$.

As in the main text, the optimal value of $l + \Delta l$ is found as that frequency which when filtered through the same sampling sequence $S(t_n)$ as the original data provides the best match to the original Fourier map. The algorithm computes the Fourier map, $D^S(k)$, selects a value of Δl , $0 \leq \Delta l \leq 1$, and, for the selected Δl , computes the functions $f_{l+\Delta l}^S(k)$. To evaluate $M_{l+\Delta l}$, the values of the amplitudes $a_{l+\Delta l}$ are determined from the following least-square minimization principle:

$$\frac{\partial M_l}{\partial a_m^*} = \sum_k^N \left[D^S(k) - \sum_{\pm} a_{\pm(l+\Delta l)} f_{\pm(l+\Delta l)}^S(k) \right] f_m^{S*}(k) = 0 \quad (\text{A-2})$$

For $m = \pm(l + \Delta l)$, Eq. (A-2) defines a linear system of two equations from which the optimal magnitudes of the unknown $a_{\pm(l+\Delta l)}$'s are determined. By using these optimal $a_{\pm(l+\Delta l)}$'s in Eq. (A-1), the algorithm searches for an optimal l that minimizes M_l . (Similar to that in the main text, this minimization is most efficiently achieved by inspecting several of the highest peaks in the original Fourier map.) The above-described procedure yields an optimal l for any given Δl . The algorithm performs a systematic search for a Δl that minimizes $M_{l+\Delta l}$ as a function of Δl by using Brent's method (it does not require construction of derivatives; see [1]).

To compute the various functions required in the evaluation of Eqs. (A-1) and (A-2), the following relationships are used to minimize the number of the computations. Let F denote the discrete Fourier transform operator. Then the Fourier transform of $f_{l+\Delta l}(t_n)S(t_n)$ is

$$\begin{aligned} f_{l+\Delta l}^S(k) &\equiv F_k \left[e^{-2\pi i \frac{n(l+\Delta l)}{N}} S(t_n) \right] = \sum_n^N e^{2\pi i \frac{n(k-l)}{N}} \left[S(t_n) e^{-2\pi i \frac{n\Delta l}{N}} \right] \\ &= F_{k-l} \left[S(t_n) e^{-\frac{2\pi i n \Delta l}{N}} \right] \equiv S(k-l-\Delta l) \end{aligned} \quad (\text{A-3})$$

That is to say, the k th Fourier component of $f_{l+\Delta l}(t_n)S(t_n)$ is the $(k-l)$ th Fourier component of $S(t_n) \exp^{-2\pi i(n\Delta l/N)}$.

By using relationships similar to those used to derive Eq. (A-3), Eq. (A-2) is simplified:

$$\begin{aligned} \sum_k^N f_{l+\Delta l}^S(k) f_m^{S*}(k) &= \sum_{k,n,n'}^N S(t_n) e^{2\pi i \frac{n(k-l)}{N}} S(t'_n) e^{-2\pi i \frac{n'(k-m)}{N}} e^{-2\pi i \frac{n\Delta l}{N}} \\ &= \sum_{n,n'}^N \delta_{n,n'} S^2(t_n) e^{-2\pi i \frac{n\Delta l}{N}} = N S(m-l-\Delta l) \end{aligned} \quad (\text{A-4})$$

and, similarly,

$$\sum_k^N D^S(k) f_m^{S*}(k) = N D^S(m) \quad (\text{A-5})$$

By using Eqs. (A-4) and (A-5), Eq. (A-2) simplifies to

$$\sum_{\pm} S(m \mp (l + \Delta l)) a_{\pm(l+\Delta l)} = D^S(m) \quad (\text{A-6})$$

For $m = \pm(l + \Delta l)$, Eq. (A-6) is solved for a pair of coefficients $a_{\pm(l+\Delta l)}$. Note that for $\Delta l = 0$, Eq. (A-6) reduces to Eqs. (4a) and (4b) of the main text.

Other useful relationships involve the evaluation of the noninteger Fourier components of the sampling function $S(t_n)$:

$$S(k + \Delta l) = F_k \left[S(t_n) e^{2\pi i \frac{n\Delta l}{N}} \right] \quad (\text{A-7})$$

$$S(2(l + \Delta l)) = F_{2l} \left[S(t_n) e^{4\pi i \frac{n\Delta l}{N}} \right] \quad (\text{A-8})$$

$$D^S(l + \Delta l) = F_l \left[F_n^{-1} [D^S(k)] e^{-2\pi i \frac{n\Delta l}{N}} \right] \quad (\text{A-9})$$

The use of Eqs. (A-3)–(A-5) and (A-7)–(A-9) minimizes the number of computations required to recompute Eqs. (A-1) and (A-6) at each iteration step to minimize M_l .

33269

P-17

November 15, 1994

110867

358414

Space VLBI Telecommunication Characteristics, Protection Criteria, and Frequency Sharing

B. O. Gutierrez-Luaces
DSN Operations Office

D. F. Bishop
Telecommunications Systems Section

A brief description of the technical characteristics of space VLBI is made, emphasizing the VLBI cross-correlation process. The signal-to-noise ratio of the cross-correlation process should be maintained as large as possible for the duration of the observation. Protection of this process from unwanted interference is a primary objective. The telecommunication radio links required in a space VLBI system are identified and characterized. Maximum bandwidths are suggested, as well as the minimum carrier frequencies required for the telemetering and the phase-transfer radio links. Planned space VLBI system models—Radioastron (Russia), VLBI Space Observatory Project (VSOP) (Japan), and the DSN orbiting VLBI subnet (United States)—are taken as a baseline to determine the interference criteria. It is concluded that existing interference criteria for near-Earth research satellites are suitable for the protection of the space VLBI systems planned.

I. Introduction

Very long baseline interferometry (VLBI) is a technique that allows experimenters to achieve angular resolution of observed radio sources that cannot be approached by other radio or optical methods. VLBI has a wide variety of scientific and engineering uses [1,2]. Observations of distant radio sources with two or more VLBI stations are combined to determine the structure of extragalactic radio sources, determine geodynamical characteristics of the Earth, study the Moon's libration and tidal response, determine orientation of the solar system with respect to the extragalactic inertial frame, determine vector separation between antenna sites, and provide navigation and tracking of spacecraft.

II. Technical Characteristics

The operating approach of the most simple VLBI system, composed of two VLBI Earth stations, may be summarized as follows. The VLBI Earth antennas will point to the radio source, common to both antennas, for the planned experiment. Because of engineering limitations, the resulting observed frequency spectrum is usually translated down to a lower frequency. The amplitude and phase characteristics of this observed spectrum are maintained by using a highly stable reference frequency, a local oscillator (LO). The observed spectrum at each antenna is recorded independently in some supported medium (e.g., magnetic tape). In the case of space VLBI, one of the antennas used for the radio source observation

is space borne, and the downconverted spectrum is transmitted and recorded at a space VLBI Earth station (Fig. 1).

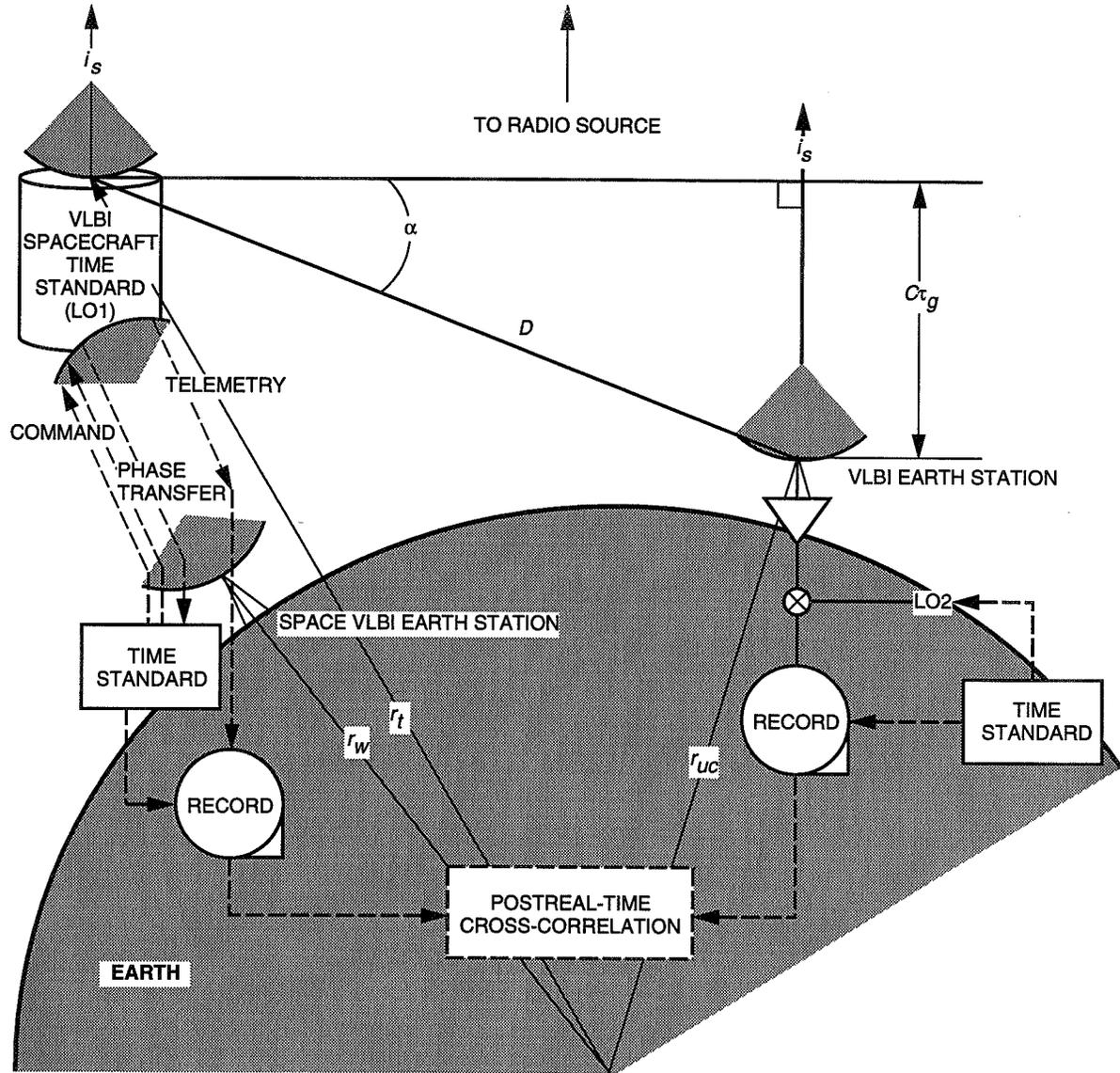


Fig. 1. Space VLBI geometry.

The basic observables in radio interferometry are the amplitude and relative phase of the cross-correlation of the two observed spectra. This cross-correlation process usually is performed in nonreal time and may be expressed as

$$R_{xy}(\tau_g) = \langle x(t)y(t - \tau_g) \rangle \quad (1)$$

where

R_{xy} = cross-correlation function

$\langle \rangle$ = estimated mean for the observation period

$x(t)$ = recorded signal at site 1

$y(t)$ = recorded signal at site 2

τ_g = wave front time delay

In the cross-correlation function of Eq. (1), the prerecorded signals will be contaminated with noise from the receiving systems. It has been shown that the cross-correlation signal-to-noise ratio, SNR_{cross} , may be expressed as a function of the two observing signal-to-noise ratios, SNR_{obs1} and SNR_{obs2} , as

$$SNR_{cross} = (SNR_{obs1}SNR_{obs2}BT)^{1/2} \quad (2)$$

where B is the observing bandwidth and T is the integration time of each observation. The SNR_{cross} should be maintained as large as possible to decrease the error in the τ_g measurement in Eq. (1).

The sensitivity (signal-to-noise ratio = 1) of this two-element VLBI interferometer may be determined [3]:¹

$$S_d = 4(2)^{1/2}10^{26}k(T_1T_2)^{1/2}(\pi gD_1D_2)^{-1}(\eta_1\eta_2)^{-1/2}(BT)^{-1/2} \text{ (Janskys)} \quad (3)$$

where

k = Boltzmann's constant = 1.38×10^{-23} (W/K-Hz)

T_1, T_2 = system temperatures

D_1, D_2 = antenna diameters

η_1, η_2 = antenna aperture efficiencies

g = coherence of the VLBI system

T = integration time

B = noise bandwidth

This is equivalent to the root-mean-square (rms) noise divided by the coherence, g .

In VLBI, a "quasi-common" time reference frame at both observing stations is essential because of the need for precise knowledge of the signal frequency and phase. Also, precise time information is needed for the postreal-time cross-correlation. These requirements are met with high-stability oscillators often referred to as "atomic clocks." It is desirable to provide the space VLBI spacecraft with a space-qualified atomic clock in the future. For the time being, an Earth-to-space (E-S) phase-transfer radio link will be needed to impart the required timing or phase reference to the spacecraft's onboard clock.

A. Telecommunication Links for Space VLBI

The telecommunication radio links to be considered in a space VLBI system are represented in Fig. 1 by the four dashed lines between the space VLBI spacecraft telecommunication antenna and the space VLBI Earth station. A description of the radio links follows.

¹ J. Ulvestad, R. Freeland, G. Levy, D. Meier, D. Murphy, and R. Preston, "Future Space VLBI Options," draft midyear report (internal document), Jet Propulsion Laboratory, Pasadena, California, May 4, 1992.

1. E–S Telecommand Radio Link. This radio link is used for reliable transmission of telecommands required for operation and correction of possible malfunctions of spacecraft behavior.

2. E–S Phase-Transfer Radio Link. The main use of this link will be for translation to the spacecraft of the phase and frequency stability of the atomic clock located at the space VLBI Earth station. This high stability is needed for the duration of the observation time and should be of the same order of magnitude as the one in the atomic clock at the space VLBI Earth station.

3. Space-to-Earth (S–E) Telemetry Radio Link. The space VLBI spacecraft observes the radio source over a selected bandwidth. This observed spectrum is transmitted to the space VLBI Earth station for recording and future cross-correlation with the observed spectrum from one or more VLBI Earth stations.

4. S–E Phase-Transfer Radio Link. This radio link will be a coherent frequency translation of the E–S phase-transfer radio link described above and will be used to calibrate the phase errors introduced in the E–S phase-transfer radio link by various causes. This radio link may be dedicated to this phase-transfer operation or may simultaneously be used to transfer the observing spectra from the spacecraft, as described in Section II.A.3.

B. Telemetry Link Characterization

The space VLBI spacecraft receives the radio source frequency spectrum contaminated with noise (background, system, etc.) in a selected observing bandwidth, B , at a given observing SNR, SNR_{obs1} . This observed spectrum has to be transmitted to the space VLBI Earth station to be recorded and further processed (cross-correlated). This transmission may be an analog transmission or the observed analog signal may be converted to a digital format and transmitted to the space VLBI Earth station for recording.

The transmission through space of a telemetry signal implies some signal degradation when detected at the intended receiver. In digital transmissions, this degradation is due to the probability of information bits being in error and is dependent on the received symbol signal-to-noise ratio (SSNR). This link degradation will affect the final process of the space VLBI experiment, i.e., the cross-correlation function in Eq. (1). Figure 2 shows typical results of the SNR_{cros} degradation [Eq. (2)] as a function of the space VLBI Earth station telemetry link performance. Results for analog, 1-bit, and 2-bit binary representations have been included. Note the inherent degradation introduced by the digital conversion. This degradation is a function of the quantization levels utilized in the analog-to-digital conversion.

1. Required Telemetry Channel Bandwidth. Phase modulation has been shown to attain optimum performance on satellite telecommunications links. Therefore, binary phase shift keying (BPSK) or quadriphase shift keying (QPSK) will be considered the preferred digital modulation schemes.

When digitizing the observing bandwidth of B Hz, the required Nyquist sampling rate will be twice the bandwidth, or $2B$ samples per sec. Each observed voltage sample is quantized at either two levels (1-bit representation), four levels (2-bit representation), or eight levels (3-bit representation), etc. The total telemetry channel symbol rate required will, therefore, follow

$$SR = 2B \log_2 (L) \quad (4)$$

with

SR = total data rate (symbols/sec)

B = observed bandwidth (Hz)

L = total number of quantization levels

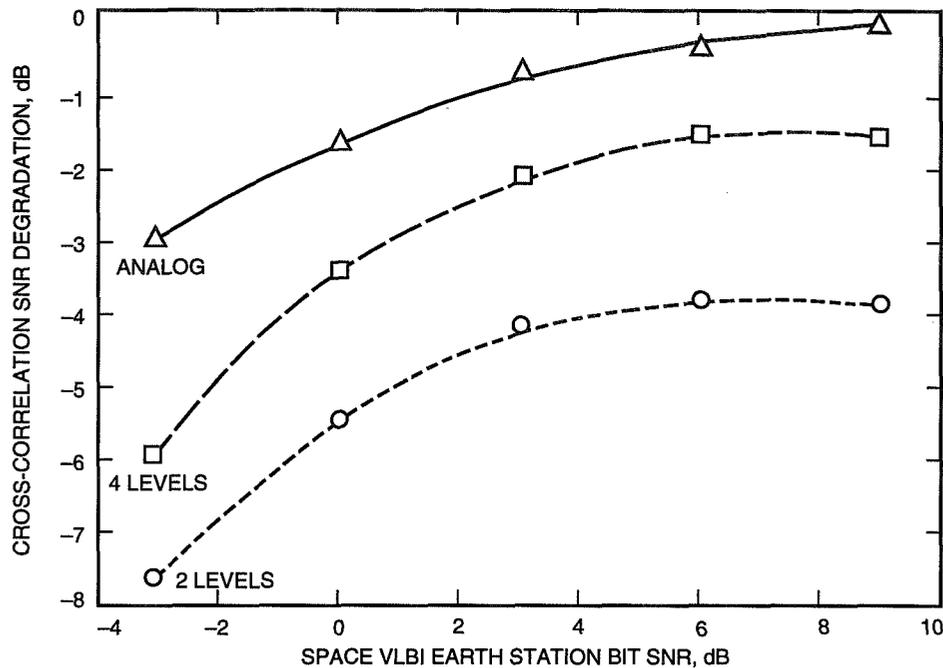


Fig. 2. Cross-correlation SNR degradation as a function of binary representation and telemetry SNR.

The radio frequency bandwidth, BW , required for the transmission of BPSK (telemetry losses less than 0.3 dB) has been recommended by the International Radio Consultative Committee (CCIR) as

$$BW = 5SR \quad (5)$$

If QPSK is used, the same bandwidth can accommodate twice the symbol rate with approximately the same performance as the BPSK case. Table 1 presents a summary of all the above considerations, showing the required radio frequency (RF) bandwidths as a function of observation bandwidth, B . Note that smaller bandwidths than those recommended may be used at a cost of link performance.

Planned space VLBI systems (see Table 2) typically use data rates on the order of 72 megasymbols/sec and QPSK modulation.² The maximum RF bandwidth required would, therefore, be on the order of 360 MHz (from Table 1). Theoretical studies of propagation effects on wide bandwidth transmissions have indicated that the atmosphere can support several gigahertz of bandwidth at carrier frequencies above 10 GHz. Therefore, transmission bandwidths on the order of 3–4 GHz may very well be envisioned for future VLBI systems.

2. Required S-E Telemetry Carrier Frequencies. Planned space VLBI systems with maximum RF transmission bandwidth requirements of less than 500 MHz will be very well allocated at carrier frequencies larger than 3 GHz. Future RF bandwidth requirements (4 GHz) indicate the need for carrier frequencies larger than 20 GHz.

C. Phase-Transfer Link Characterization

A prime requirement of a space VLBI spacecraft's onboard clock is that its frequency/phase stability be nearly as good as that of a VLBI Earth station's atomic clock. No space-qualified atomic clocks

² J. Ovnick, "Design Requirements for DSN Orbiting VLBI Subnet," DM515606A, Rev. A (internal document), Jet Propulsion Laboratory, Pasadena, California, March 31, 1992.

exist today; therefore, the required stability will be transferred to the space VLBI spacecraft via an E-S radio link. The carrier frequency of this radio link, f_{up} , is recovered at the spacecraft to generate the onboard reference frequencies to be used in the radio source observing process. In order to calibrate all the unknown phase errors introduced in this E-S phase-transfer radio link, this carrier frequency is coherently downconverted and transmitted back to the space VLBI Earth station, f_{down} . In this two-way phase calibration transfer system, phase errors are mainly introduced by the propagation medium and the receiving systems (spacecraft and space VLBI Earth station). These phase errors will contribute to the uncertainty in the determination of the amplitude and relative phase of the nonreal-time cross-correlation process of Eq. (1), effectively lowering the SNR_{cross} of Eq. (2).

Table 1. Required radio frequency bandwidth.

Signal presentation	Quantization levels	Bits, number	Symbol rate, symbols/sec	RF bandwidth, Hz		
				Analog	BPSK	QPSK
Analog	—	—	—	2B (minimum)	—	—
[0,1]	2	1	2B	—	10B	5B
[00,01,10,11]	4	2	4B	—	20B	10B
[000,001,010,011,100,101,110,111]	8	3	6B	—	30B	15B

Table 2. Characteristics of planned space VLBI systems.

Parameter	Radioastron	VSOP	IVS
Observing antenna diameter, m	10	10	20
Observing frequency and system temperature, GHz; K	0.3; 90 1.6; 60 5.0; 70 22.0; 135	1.6; 40 5.0; 60 22.0; 110 —	4.5; 8.5 15; 23 42; 63 86; 120
Nominal integration time, sec	300	300	—
Space-to-Earth			
Frequency, GHz	14–15.35	14–15.35	—
Modulation type	QPSK	QPSK	—
Maximum bit rate, MB/sec	144	128	—
Quantization, levels	2, 4	2, 4	—
RF bandwidth, MHz	500	500	—
Minimum E_b/N_o , dB	11.2	9.1	—
Earth-to-space (phase transfer)			
Frequency, GHz	7.145–7.235	15.25–15.35	—
Modulation type	None	None	—
Maximum bit rate, MB/sec	—	—	—
RF bandwidth, MHz	50	100	—
PLL bandwidth, Hz	1000	1000	—
Minimum P_c/N_o , dB/Hz	63	60	—
Orbital characteristics			
Inclination, deg	51.5–65.0	31	63
Height at perigee, km	2000	1000	5000
Height at apogee, km	78,980	20,000	150,000
Period, hr	28	6.06	67.14

1. Radio Frequency Propagation-Induced Phase Noise. The phase, ϕ_{up} , of the onboard reference frequency, f_{up} , is retrieved from the measured round-trip phase, ϕ_{round} , measured at the ground station through the following:

$$\phi_{up} = \frac{f_{up}}{f_{up} + f_{down}} \phi_{round} \quad (6)$$

There exists frequency-dependent path delay, τ_i , in the propagation of an electromagnetic wave through the ionosphere. Therefore, Eq. (6) should be modified to

$$\phi_{up} = \frac{f_{up}}{f_{up} + f_{down}} \phi_{round} + 2\pi \frac{f_{up}f_{down}}{f_{up} + f_{down}} [\tau_i(f_{up}) - \tau_i(f_{down})] \quad (7)$$

with

$$\tau_i(f) = \frac{40.3}{cf^2} TEC_i \text{ (sec)}$$

with

c = velocity of light (m/sec)

f = propagation frequency (Hz)

TEC_i = total electron content (electrons/m²)

The second term in the right side of Eq. (7) is an error term due to a frequency-dependent ionospheric delay. Unless additional information about the total electron content, TEC_i , in the ionosphere is provided, a proper correction for this error cannot be made. Nevertheless, this error becomes smaller if frequencies of both f_{up} and f_{down} are made higher and closer to each other. Table 3 gives the calculated results of this error in units of picosecond (psec) time delay, i.e., $\phi_{up}/2\pi f_{up}$, for two frequency pairs (7.2–8.46 GHz and 15.3–14.2 GHz). A total electron content of 8×10^{17} electrons/m² has been assumed.

Table 3. Ionospheric propagation effects.

Link frequencies, GHz		Absolute value of ionospheric error, psec (for $TEC = 8 \times 10^{17}$ electrons/m ²)	Coherence factor scintillation index, S	
F_{up}	F_{down}		$S = 0.1$	$S = 0.5$
7.2	8.46	308.8	0.867	0.028
15.3	14.2	35.2	0.998	0.954

From Table 3, it is concluded that the phase transfer at higher frequencies is much better than at lower frequencies. Note that, in this particular case, the lower-frequency phase error introduced is approximately a wavelength of the highest observing band planned for VSOP and Radioastron (22 GHz with a period of 45.4 psec). If the ionospheric delay fluctuates rapidly (ionospheric scintillation), the phase error introduced cannot be removed in the postreal-time cross-correlation process. The optimum coherence factor of 1.0 is reduced. Note that, at the higher pair of frequencies shown, almost optimum coherency is kept even for a scintillation index of 0.5. For the lower band, coherency is almost completely lost. The coherence factor

is inversely proportional to the sensitivity [Eq. (3)] of the interferometer. When the coherence factor is 1, the sensitivity is equal to the rms noise. When the coherence factor is below 1, the sensitivity rises above the rms noise, a nonideal situation.

2. Carrier Recovery Phase Noise. At the space VLBI spacecraft receiver of the E-S radio link, as well as at the space VLBI Earth station's receiver, the carrier recovery process considered may be the result of any combination of the following modulation schemes: an unmodulated carrier, a spread spectrum phase modulation (BPSK), or a quadriphase modulation (QPSK).

It has been shown that the phase error variance for carrier recovery processes, σ_{rcvr}^2 , may be expressed as a function of the symbol signal-to-noise ratio, $SSNR$, the phase lock loop (PLL) receiver closed-loop bandwidth, B_l , and the symbol period, T_s , as

$$\frac{\sigma_{rcvr}^2}{B_l T_s} = \frac{1}{SSNR} \quad (8)$$

for an unmodulated carrier, as

$$\frac{\sigma_{rcvr}^2}{B_l T_s} = \frac{1}{SSNR} + \frac{1}{SSNR^2} \quad (9)$$

for BPSK modulation, and as

$$\frac{\sigma_{rcvr}^2}{B_l T_s} = \frac{1}{SSNR} + \frac{9}{2SSNR^2} + \frac{6}{SSNR^3} + \frac{3}{2SSNR^4} \quad (10)$$

for QPSK modulation. For very strong $SSNR$, the three cases converge to $1/SSNR$.

3. Required Phase-Transfer Link Frequencies. In choosing the phase-transfer frequencies, it seems that the most important consideration should be given to the ionospheric propagation effects. Therefore, frequencies approximately 14 GHz or higher are the most suitable for phase transfers for space VLBI missions. Also, the uplink and downlink frequencies should be kept as close as possible.

D. Space VLBI System Characteristics

Table 2 is a summary of the salient radio link and orbital characteristics of Radioastron (Russia) and VSOP (Japan) [3,4,7].^{3,4} These are space VLBI spacecraft to be launched in 1996. Many telemetry receiving stations spread around the Earth will be used. An example of these is the DSN orbiting VLBI subnet (United States), whose main characteristics are summarized in Table 4. A next-generation space VLBI mission being considered, the International VLBI Satellite (IVS), has also been included in Table 2 [5].

III. Interference Criteria

A. S-E Telemetry Link

A computer simulation was used to determine the cross-correlation SNR degradation versus the interference to the system noise ratio of the space VLBI Earth station. The VLBI Earth station is assumed to be interference free. Three types of interference were used: wideband noise, continuous-wave (CW) worst case (constant phase), and continuous-wave most-likely case (random phase). Figure 3 contains

³ Ibid.

⁴ D. W. Murphy, personal correspondence with D. Murphy, Jet Propulsion Laboratory, Pasadena, California, 1992.

the results of the simulation for the analog system and for the two-level (1-bit) and four-level (2-bit) quantization schemes. In all cases, the telemetry SSNR of the space VLBI Earth station was set to a baseline value (no interference) of $SSNR = 6$ dB.

Table 4. Summary of DSN space VLBI Earth station characteristics.

Parameter	X-band	Ku-band
Receive frequency tuning, GHz	8.025–8.5	14.0–15.35
–1-dB receive bandwidths, MHz	50	500
Receive zenith G/T , dB/K	33.7	37.3
Transmit frequency tuning, GHz	7.145–7.235	15.25–15.35
Transmit antenna gain, dB	54.7	61.0
Transmit power levels, W	5	0.5
–1-dB transmit bandwidths, MHz	50	100
Receive or transmit polarizations	RHCP or LHCP	RHCP or LHCP
Telemetry receiver capability, MB/sec	144	144
Antenna diameter, m	11	11

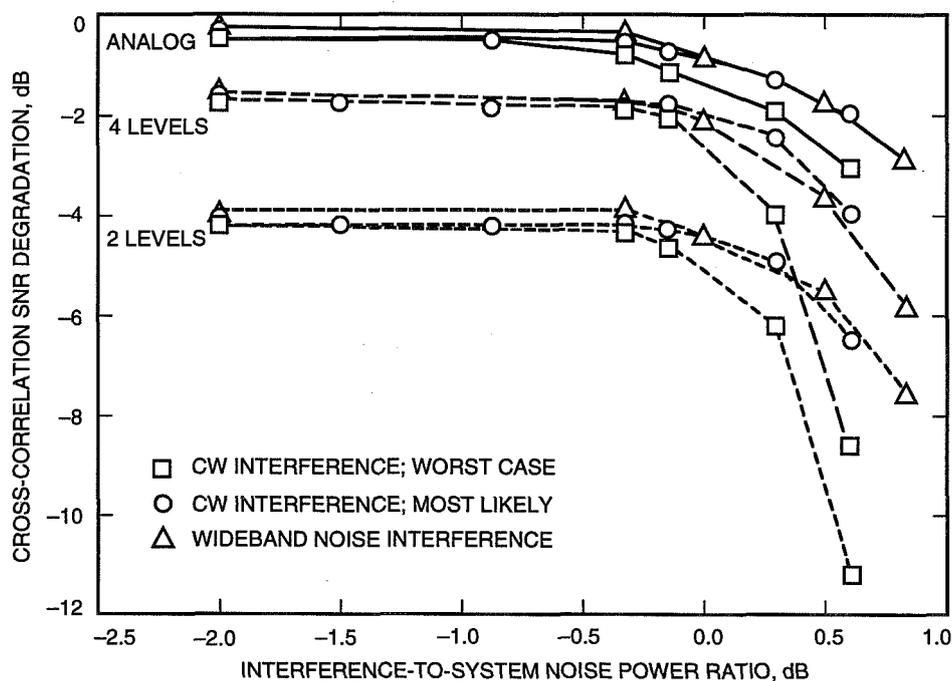


Fig. 3. Cross-correlation SNR degradation as a function of interference-to-system noise power ratio.

Protection criteria for telecommunication links for near-Earth research satellites [6] suggest that interference is harmful if the link threshold performance is decreased by more than 1 dB. From Fig. 3, the 1-dB degradation threshold for the three cases being considered may be set to a system noise power ratio to interference power, N/I , of approximately 0 dB. A typical system noise temperature for space VLBI Earth stations is 250 K. This corresponds to a system noise spectral density of -204 dB (W/Hz). Therefore, the maximum allowable interference power would be -204 dB (W/Hz).

B. Phase-Transfer Link

This link has been discussed in Sections II.A.2 and II.C. Phase errors are introduced primarily by the propagation medium at 8 GHz. A phase coherence error budget is contained in [7].

1. **E–S Link.** The system noise temperature of planned receiving systems on board space VLBI spacecraft is approximately 600 K. This corresponds to a system noise spectral density of -171 dB (W/kHz). Using the protection criteria of $N/I = 6$ dB [6], the maximum allowable interference noise power, therefore, would be -177 dB (W/kHz). The satellite circuit random phase-noise estimation of 2.56-deg rms [7] would be increased to 3.2-deg rms by applying Eq. (8).

2. **S–E Link.** The system noise temperature of planned space VLBI Earth station receiving systems is approximately 250 K. This corresponds to a system noise spectral density of -204 dB (W/Hz). Using the protection criteria of $N/I = 6$ dB [6], the maximum allowable interference noise power would be -210 dB (W/Hz). The estimated space VLBI Earth station receiver phase-noise contribution, the receiver phase-lock-loop SNR [7], of 2.56-deg rms would be increased to 3.2-deg rms by applying Eq. (8).

C. Interference Criteria Summary

Maximum allowable interference power at the input terminals of the space VLBI Earth station receiver has been found to be -204 dB (W/Hz) for the telemetering link and -210 dB (W/Hz) for the phase-transfer link. Maximum allowable interference power at the input terminals of the space VLBI spacecraft receiver has been found to be -177 dB (W/kHz). According to the recommendations of [6], there should be maximum interferences of -216 dB (W/Hz) at the input terminals of the Earth station receiver and -177 dB (W/kHz) at the input terminals of the space research space station, for 0.1 percent of the time in both cases. Therefore, the levels recommended in [6] are suitable for the protection of planned space VLBI systems.

IV. Sharing and Protection Criteria of Space VLBI Services in Bands Near 8, 15, and 40 GHz

There are many radio services that use the same frequency bands that are planned for space VLBI. The frequency bands and the radio services that have allocations [8] are listed in Table 5. These frequency bands will be used for the VSOP and Radioastron missions. There are no current missions that plan to use the frequency bands shown in Table 6 and allocated to the space research services. These frequencies may be used for space VLBI in the future. The proposed IVS mission is used as an example of a space VLBI mission that may use the near-40 GHz band.

The space VLBI communications links are susceptible to interference from other radio services that use the same frequency bands. Also, the other radio services are susceptible to interference from the space VLBI communications links. Therefore, a compatibility study was performed on these frequency bands. No study of the 74- to 84-GHz band is presented at this time.

A. Methodology

The initial approach is to calculate a worst-case interference value. If the interference paths are line of sight, then it is assumed that the boresights of the transmitting and receiving antennas are aligned and there are minimum distances between the transmitter and receiver. If the interference paths are over a great circle, then the interference margin is set equal to 0 dB and the coordination distance between the interfering transmitter and the receiver is computed. Interference margin is defined as the difference between the interference criterion at a receiver and the computed level of interference at a receiver.

Table 5. Frequency bands for space VLBI.

Frequency band, GHz	Direction	Existing services
8.025–8.5	Space-to-Earth	Fixed Mobile Fixed satellite (E–S) Earth exploration satellite (EES) (S–E) Meteorological satellite (E–S) Space research (S–E) Radio location Space research
7.145–7.235	Earth-to-space	Fixed Mobile EES (passive) Space research (passive) Space operations (E–S) Space research (E–S)
14–15.35	Space-to-Earth	Fixed satellite (E–S) Radio navigation Space research Fixed Mobile Radio navigation satellite Radio astronomy Land mobile satellite (E–S) Space research (passive) EES (passive)
15.25–15.35	Earth-to-space	Fixed Mobile Space research Space research (passive) EES (passive)

Table 6. Frequency bands for future space VLBI.

Frequency band, GHz	Direction	Existing services
37–38	Space-to-Earth	Fixed Mobile Fixed satellite (S–E)
40–40.5	Earth-to-space	Fixed Fixed satellite (S–E) Mobile Mobile satellite (S–E)
74–84	Space-to-Earth	Fixed Mobile Fixed satellite (E–S) Amateur Amateur satellite Radio location Fixed satellite (S–E) Mobile satellite (S–E)

A more realistic representation of interference is made when the space VLBI spacecraft orbits are generated from the orbital characteristics planned (Table 2). Computer simulations of interference are performed for each of the worst cases for which the interference margin is less than zero.⁵ Tables 7–19 are summaries of the computer simulations involving the different interference scenarios.

B. Sharing and Protection Summary

The worst-case interference scenario results shown in Tables 7–19 (“worst-case interference margin” column) indicate that, for line-of-sight paths, the interference margins are negative for nearly all cases. For great-circle interference paths, the coordination distances shown in Tables 7–19 are large in many cases. Therefore, assuming these worst-case conditions, band sharing between space VLBI and other services is difficult. However, when the most realistic approach of computing the interference levels as a function of orbital position of the space VLBI spacecraft is taken, the percentage of interfering time results shown in Tables 7–19 (“Time” column) indicates that space VLBI communication links are very compatible with other services. It should be noted that those interfering times that are different from zero happen at the space VLBI spacecraft perigee and that they can be eliminated by appropriate mission operations design, e.g., not pointing the space VLBI spacecraft antenna to the Earth while in closest approach. Therefore, it is recommended that space VLBI can share the frequency bands in Tables 5 and 6 with existing radio services. Careful coordination may be required in some cases.

V. Conclusions

The technical characteristics and telecommunications requirements for space VLBI have been provided. It has been shown that the use of high, closely spaced frequencies for the phase-transfer uplink and downlink produces lower ionospheric errors and larger VLBI coherence factors and that the protection criteria for near-Earth research satellite links [6] are sufficient to protect the space VLBI links.

A frequency-sharing analysis between space VLBI radio links and existing radio services has been performed. The analysis shows that the space VLBI links are compatible with these other services.

Table 7. Interference from Radioastron space VLBI phase reference downlink (8.025–8.5 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion	Orbit or location	Antenna and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	–150 dBW/m ² in 4 kHz	Canberra	—	3	—	–46.6
Fixed satellite (E–S)	–138.6 dBW in 50 MHz	343.8 deg east	36.5 dBi, Earth center	0	—	–12.9
Earth-exploration satellite (S–E)	–214.0 dBW/Hz	500-km altitude, 35-deg inclination	55.5 dBi, EES spacecraft	0	—	–53.7
Meteorological satellite (E–S)	–189.9 dBW in 960 kHz	343.8 deg east	27.0 dBi, Earth center	10.7	—	–54.7
Radio location	–166 dBW in 3 MHz	51.7 deg south, 148.98 deg east	34.0 dBi, east horizon	0	—	–56.8

⁵ D. F. Bishop, “Frequency Band Compatibility Between Space VLBI and Other Radio Services,” JPL Interoffice Memorandum 3396-92-33 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 10, 1992.

Table 8. Interference to Radioastron space VLBI phase reference Earth station receiver (8.025–8.5 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-216.0	—	38 dBi, horizon	—	311	0.0
Fixed satellite (E-S)	-216.0	—	20.1 dBi, 3-deg elevation	—	225	0.0
Earth-exploration satellite (S-E)	-216.0	500-km altitude, 35-deg inclination	6 dBi, EES Earth station	0	—	-52.9
Meteorological satellite (E-S)	-216.0	—	21.9 dBi, 3-deg elevation	—	452	0.0
Radio location	-216.0	—	34.0 dBi, east horizon	—	481	0.0

Table 9. Interference from Radioastron space VLBI phase reference uplink (7.145–7.235 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-150.4 in 4 kHz	—	38.0 dBi, horizon	—	531	0.0
Passive sensors	-158 in 400 MHz	500-km altitude, 35-deg inclination	34.5 dBi, Earth center	0.12	—	-90.7
Space operations (E-S)	-113.5 in 100 kHz	500-km altitude, 35-deg inclination	0 dBi	0	—	-11.7

Table 10. Interference to Radioastron space VLBI phase reference spacecraft receiver (7.145–7.235 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-207.0	51.7 deg south, 163.8 deg east	38.0 dBi, east horizon	0	—	-31.7
Space operations (E-S)	-207.0	Canberra	49.5 dBi, space operations spacecraft	2	—	-100.2

Table 11. Interference from VSOP space VLBI downlink (14–15.35 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion	Orbit or location	Antenna and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed satellite (E-S)	-211.3 dBW/Hz	81 deg east	40.2 dBi, Earth center	0	—	-16.6
Radio navigation	-209.0 dBW/Hz	31.3 deg north, 261 deg east	50 dBi, east horizon	0	—	-48.2
Fixed and mobile	-148.0 dBW/m ² in 4 kHz	Goldstone	—	6	—	-17.6
Radio navigation satellite	-136.0 dBW in 34 MHz	31.3 deg north, 261 deg east	0 dBi	0	—	-0.4
Radio astronomy	-221.0 dBW/m ² -Hz	30 deg north, 260 deg east	—	0	—	-54.6
Land mobile satellite (E-S)	-211.8 dBW/Hz	Geostationary, 81 deg east	33.0 dB, Earth center	0	—	-9.9
Passive sensors	-160.0 dBW in 200 MHz	500-km altitude, 35-deg inclination	41.8 dBi, Earth center	0	2.793 × 10 ⁶	0.0

Table 12. Interference from Radioastron space VLBI downlink (14–15.35 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion	Orbit or location	Antenna and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed satellite (E-S)	-211.3 dBW/Hz	343 deg east	40.2 dBi, Earth center	0	—	-23.2
Radio navigation	-209.0 dBW/Hz	51.7 deg south, 148.98 deg east	50.0 dBi, east horizon	0	—	-50.1
Fixed and mobile	-148.0 dBW/m ² in 4 kHz	38.47 deg south, 148.98 deg east	—	0	—	-19.5
Radio navigation satellite	-136 dBW in 34 MHz	51.7 deg south, 163.8 deg east	0 dBi	0	—	-2.2
Radio astronomy	-195.6 dBW/Hz	51.7 deg south, 163.8 deg east	27 m, zenith	0.02	—	-56.5
Land mobile satellite (E-S)	-211.8 dBW/Hz	Geostationary, 343.8 deg east	33.0 dBi, Earth center	0	—	-16.5
Passive sensors	-160.0 dBW in 200 MHz	500-km altitude, 35-deg inclination	41.8 dBi, Earth center	0	3.48 × 10 ⁶	0.0

Table 13. Interference to space VLBI Earth station receiver (14–15.35 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed satellite (E-S)	-216.0	—	20.1 dBi, 3-deg elevation	—	172	0.0
Radio navigation	-216.0	—	0 dBi	—	65.9	0.0
Fixed and mobile	-216.0	—	35.0 dBi, horizon	—	268	0.0
Radio navigation satellite	-216.0	20,200-km altitude, 55-deg inclination	—	0	—	-32.0
Land mobile satellite (E-S)	-216.0	—	20.2 dBi, 3-deg elevation	—	216	0.0

Table 14. Interference from space VLBI Earth station (15.25–15.35 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-156.0 in 4 kHz	—	35.0, horizon	—	341	0.0
Passive sensor	-160.0 in 200 MHz	500-km altitude, 35-deg inclination	41.0, toward 38.37 deg north 244.15 deg east	0	—	-88.9

Table 15. Interference to space VLBI spacecraft receiver (15.25–15.35 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-207.0	31.3 deg north, 261 deg east	38.0 dBi, east horizon	0	—	-33.6

Table 16. Interference from IVS space VLBI spacecraft (37–38 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed satellite (E–S)	28.3 dB <i>C/I</i>	Geostationary, 210 deg east	36.5 dBi, Earth center	0	—	–22.5
Fixed and mobile	–203.0 dBW/Hz	38.47 deg south, 148.98 deg east	38.5 dBi, east horizon	0	—	–26.1
Fixed satellite (S–E)	–215.8 dBW/Hz	Canberra	58.4 dBi to geostationary at 148.98 deg east	0	—	–58.8

Table 17. Interference to IVS space VLBI Earth station receiver (37–38 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed satellite (E–S)	–216.0	—	14.5 dBi, 5–deg elevation	—	98	0.0
Fixed and mobile	–216.0	—	55 dBW EIRP, horizon	—	131	0.0
Fixed satellite (S–E)	–216.0	Geostationary, 148.98 deg east	36.5 dBi, Canberra	0	—	–66.6

Table 18. Interference from IVS space VLBI Earth station transmitter (39.5–40.5 GHz) to other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	–133	—	38.5 dBi, horizon	—	110	0.0
Fixed and mobile satellite (S–E)	–134.1	—	14.5 dBi, horizon	—	60	0.0

Table 19. Interference to IVS space VLBI spacecraft receiver (39.5–40.5 GHz) from other radio services.

Existing services	Service parameters			Interference results		
	Interference criterion, dBW/Hz	Orbit or location	Antenna gain and pointing	Time, percent	Coordination distance, km	Worst-case interference margin, dB
Fixed and mobile	-207	38.47 deg south, 148.98 deg east	55 dBW EIRP, horizon	0.023	—	-32.5
Fixed and mobile satellite (S-E)	-207	Geostationary, 210 deg east	36.5 dBi, Earth center	0	—	-27.9

References

- [1] R. A. Preston, *Accurate Satellite Tracking With Very Long Baseline Interferometry*, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, July 1972.
- [2] H. F. Hinteregger, *Geodetic and Astrometric Applications of Very Long Baseline Interferometry*, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, October 20, 1972.
- [3] B. F. Burke, D. H. Roberts, C. R. Lawrence, C. S. Readhead, A. E. E. Rogers, A. R. Whitney, D. L. Jones, J. F. Jordan, G. S. Levy, R. P. Linfield, D. W. Murphy, R. A. Preston, L. R. D'Addario, J. D. Romney, K. J. Johnston, R. S. Simon, J. H. Spencer, J. M. Moran, M. J. Reid, and I. I. Shapiro, "US Participation in VSOP and Radioastron," OVLBI Science Consulting Group, August 1, 1989.
- [4] G. S. Levy, R. P. Linfield, C. D. Edwards, J. S. Ulvestad, J. F. Jordan, Jr., S. J. DiNardo, C. S. Christensen, R. A. Preston, L. J. Skjerve, L. R. Stavert, B. F. Burke, A. R. Whitney, R. J. Cappallo, A. E. E. Rogers, K. B. Blaney, M. J. Mather, C. H. Ottenhoff, D. L. Jauncey, W. L. Peters, J. Reynolds, T. Nishimura, T. Hayashi, T. Takano, T. Yamada, H. Hirabayashi, M. Morimoto, M. Inoue, T. Shiomi, N. Kawaguchi, H. Kunimori, M. Tokumaru, and F. Takahashi, "VLBI Using a Telescope in Earth Orbit: I. The Observations," *The Astrophysical Journal*, vol. 336, pp. 1098–1104, January 15, 1989.
- [5] J. A. Estefan, "Precise Orbit Determination of High-Earth Elliptical Orbiters Using Differenced Doppler and Range Measurements," *The Telecommunications and Data Acquisition Progress Report*, vol. April–June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 1–22, August 15, 1991.
- [6] "Protection Criteria for Telecommunication Links for Manned and Unmanned Near-Earth Research Satellites," *CCIR Recommendations, RTF, RSA, and RRA Series, Science Services*, CCIR Recommendation 609-1, International Radio Consultative Committee, Geneva, Switzerland, pp. 39–42, 1992.
- [7] J. C. Springett, "Phase Transfer Phase Components and Phase Compensation Using Two-Way Phase Transfer," Technical Memorandum, NeoComm Systems, La Cañada, California, October 1990.
- [8] *Manual of Regulations and Procedures for Federal Radio Frequency Management*, National Telecommunications and Information Administration, U.S. Department of Commerce, Washington, DC, May 1989.

54-32

33270

110868

P-39

TDA Progress Report 42-119

November 15, 1994

358418

Radio-Planetary Frame Tie From Phobos-2 VLBI Data

C. E. Hildebrand, B. A. Iijima, P. M. Kroger, and W. M. Folkner
Tracking Systems and Applications Section

C. D. Edwards
DSN Technology and Science Office

In an ongoing effort to improve the knowledge of the relative orientation (the "frame tie") of the planetary ephemeris reference frame used in deep space navigation and a second reference frame that is defined by the coordinates of a set of extragalactic radio sources, VLBI observations of the Soviet Phobos-2 spacecraft and nearby (in angle) radio sources were obtained at two epochs in 1989, shortly after the spacecraft entered orbit about Mars. The frame tie is an important systematic error source affecting both interplanetary navigation and the process of improving the theory of the Earth's orientation. The data from a single Phobos-2 VLBI session measure one component of the direction vector from Earth to Mars in the frame of the extragalactic radio sources (the "radio frame"). The radio frame has been shown to be stable and internally consistent with an accuracy of 5 nrad. The planetary ephemeris reference frame has an internal consistency of approximately 15 nrad. The planetary and radio source reference frames were aligned prior to 1989 with an accuracy of approximately 250 nrad, using Earth-based optical data and measurements of occultations of the radio source 3C273 by the Moon. The Phobos-2 VLBI measurements provide improvement in the accuracy of two of the three angles describing a general rotation between the planetary and radio reference frames. A complete set of measurements is not available because data acquisition was terminated prematurely by loss of the spacecraft. The analysis of the two Phobos-2 VLBI data sets indicates that, in the directions of the two rotation components determined by these data, the JPL planetary ephemeris DE200 is aligned with the radio frame as adopted by the International Earth Rotation Service within an accuracy of 20-40 nrad, depending on direction. The limiting errors in the solutions for these offsets are spacecraft trajectory (20 nrad), instrumental biases (19 nrad), and dependence of quasar coordinates on observing frequency (24 nrad).

I. Introduction

Planetary approach navigation is crucial to the success of many interplanetary missions. Whether it be for targeting a planetary flyby to carry out a gravity assist trajectory or for planning a maneuver to enter into a closed orbit about a target planet, the planetary approach phase almost always imposes key navigation requirements. Minimizing target-relative trajectory errors results in fuel savings that can

extend mission lifetimes and enhance the total scientific return. Future missions may employ aerobraking and aerocapture maneuvers to provide more efficient orbit insertion. These techniques can impose stringent constraints on navigation accuracy, as the spacecraft trajectory must pass through a narrow atmospheric entry corridor.

Earth-based tracking of radio signals from interplanetary spacecraft provides essential information for navigating deep space missions to their targets. The radio metric data types acquired by stations on the Earth include range from the station to the spacecraft, range rate via the Doppler shift of the spacecraft signal frequency received at the station, and interferometric delay and delay rate, in which very long baseline interferometry (VLBI) is used to obtain the angular position and velocity of the spacecraft in a reference frame of quasars (the "radio frame"). For many missions (e.g., Magellan, Ulysses), Earth-based radio metric observations are the only data available for orbit determination.

Spacecraft-based optical observations of a target body against the stellar background can provide information on the spacecraft-target relative position in the two directions orthogonal to the spacecraft-target line of sight. Since the position accuracy scales directly with distance from the target body, spacecraft imaging data are useful only during the latter stages of planetary approach. Also, the line-of-sight component of relative position is typically not well determined by onboard optical data. Thus, even for spacecraft equipped with onboard imaging systems, Earth-based radio data play an important role in determining the full spacecraft target-relative state vector.

The concept of a reference frame is central to navigation. An inertial reference frame is assumed in modeling the spacecraft trajectory and in the analysis of spacecraft position measurements. There are two main reference frames involved in the interplanetary navigation process. One frame is defined by the positions of distant radio sources; the other is defined in constructing the planetary ephemeris. The ephemeris gives the positions of the gravitational bodies affecting the spacecraft trajectory as well as the location of the target body. The radio reference frame is important in the reduction of spacecraft position measurements because the rotational orientation of the Earth, which has random components, is monitored and reported routinely in the radio reference frame. Thus, the inertial locations of Earth tracking stations are implicitly referred to the radio frame. With range and Doppler observables, the radio frame position of a spacecraft during interplanetary cruise can be determined with an accuracy of about 100–200 nrad. More precise spacecraft positions (5–10 nrad) can be obtained with differential spacecraft-quasar VLBI observations that directly measure the angular position of the spacecraft relative to selected quasars.^{1,2,3} To most effectively exploit the accuracy of spacecraft interplanetary cruise trajectories given by Earth-based radio metric tracking, it is necessary that the target body position also be accurately determined in the radio frame.

The orientation offset between the planetary and radio reference frames, which may be parameterized as three small rotation angles collectively referred to as the "frame tie," was known to an accuracy of 100–200 nrad at the time the observations described herein were planned. This uncertainty greatly exceeds the accuracy attainable with Earth-based radio tracking techniques and is a limiting error for planetary approach navigation. As a result, several efforts have been initiated in recent years to improve the tie between the planetary and radio frames. Observations of the millisecond pulsar PSR 1937+21 [1] and comparisons of VLBI and lunar laser ranging (LLR) observations [2] have provided tentative

¹ J. S. Border, "Analysis of Δ DOR and Δ DOD Measurement Errors for Mars Observer Using the DSN Narrow Channel Bandwidth VLBI System," JPL Interoffice Memorandum 335.1-90-026 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 15, 1990.

² W. M. Folkner, P. M. Kroger, and C. Hildebrand, "Preliminary Results From VLBI Measurement of Venus on September 12, 1990," JPL Interoffice Memorandum 335.1-92-25 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1992.

³ W. M. Folkner, P. M. Kroger, and B. A. Iijima, "Results From VLBI Measurement of Venus on March 29, 1992," JPL Interoffice Memorandum 335.1-93-22 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 1993.

measurements of the frame tie to an accuracy of 25–50 nrad. Several other, earlier determinations are discussed in [3].

Differential VLBI observations of a spacecraft in orbit about a planet provide a direct means of measuring the frame tie, as depicted in Fig. 1. Conventional Doppler tracking of a planetary orbiter typically can provide planet-relative spacecraft positions with kilometer- or sub-kilometer-level accuracy. With VLBI time delay measurements, differenced between the spacecraft and angularly nearby quasars to cancel common measurement errors, the spacecraft angular position (and thus the planet angular position as well) can be precisely determined in the quasar reference frame. The positions of the remaining bodies of the solar system are then also tied to the quasar frame to the accuracy of the relative angular position errors of the ephemerides. Newhall et al. [4] and Niell et al. [3] used a similar technique to estimate corrections to Mars' and Venus' right ascensions and declinations with data from the Viking and Pioneer Venus orbiters. These data were not VLBI time delays, but rather phase delay rates. Differential delay measurements from the Soviet Venus–Halley (VEGA) spacecraft as they flew past Venus were analyzed in [5].

In 1988, the Soviet Union launched two spacecraft, Phobos-1 and -2, to explore Mars and its moon Phobos. One of the scientific goals of the mission was a 5-nrad frame tie using VLBI measurements from a lander on the moon's surface.⁴ Although both spacecraft were lost before completing their missions, Doppler and VLBI tracking data acquired from Phobos-2 in orbit about Mars have provided for an improved frame-tie determination. This article describes the analysis of the Phobos Mission (PHOBOS) frame-tie VLBI data and presents the resulting estimates of two components of the rotational offset between the frame of JPL planetary ephemeris DE200 [6,7] and the radio frame adopted and maintained by the International Earth Rotation Service (IERS) [8].

The following section contains a more detailed description of the radio and planetary reference frames within which the data presented here were analyzed. The PHOBOS VLBI observations are summarized in Section III. Section IV contains the analysis of measurement information content and error sources. The strategy for estimating the frame-tie parameters is given in Section V, and a reference solution is obtained. Consistency of solutions with variations of data set and fitting strategy is also examined. The PHOBOS result is compared with a frame tie derived from lunar laser ranging data and VLBI observations of natural sources. Section VI gives a brief discussion of further opportunities to improve the frame tie.

II. Reference Frames

A. IERS Reference Frames and Earth Orientation Parameters

The radio reference frame is based on the observed positions of extragalactic radio sources. The source positions are measured using VLBI, as shown in Fig. 2. Two widely separated radio antennas record the signal from the radio source. The recorded signals are processed to find the time delay, τ , which is the difference in arrival times at two stations. The time delay is approximately given by

$$\tau = \frac{1}{c} \left(\vec{B} \cdot \hat{S} \right)$$

where \vec{B} is the vector pointing from (the inertial location of) station 1 to station 2 and \hat{S} is a unit vector in the direction of the source. The observation of a single source thus gives the angle between the baseline vector and the direction to the source. By observing a number of sources widely scattered over the sky, the relative positions of the sources can be inferred along with the length and direction of the baseline vector.

⁴R. A. Preston, personal communication, Tracking Systems and Applications Section, Jet Propulsion Laboratory, Pasadena, California, 1991.

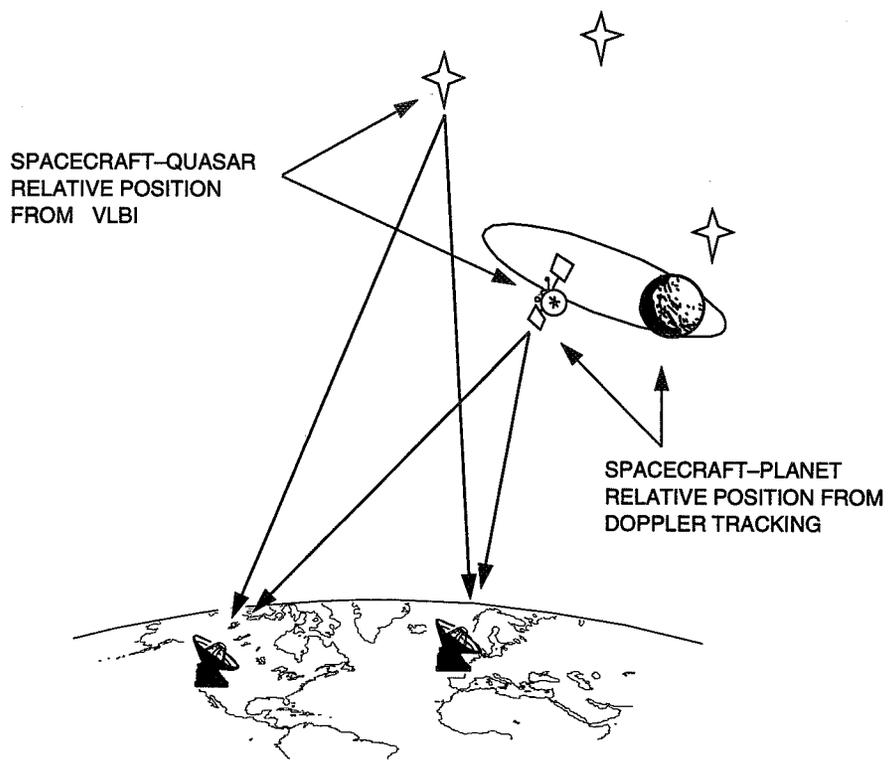


Fig. 1. Spacecraft-quasar differential VLBI geometry.

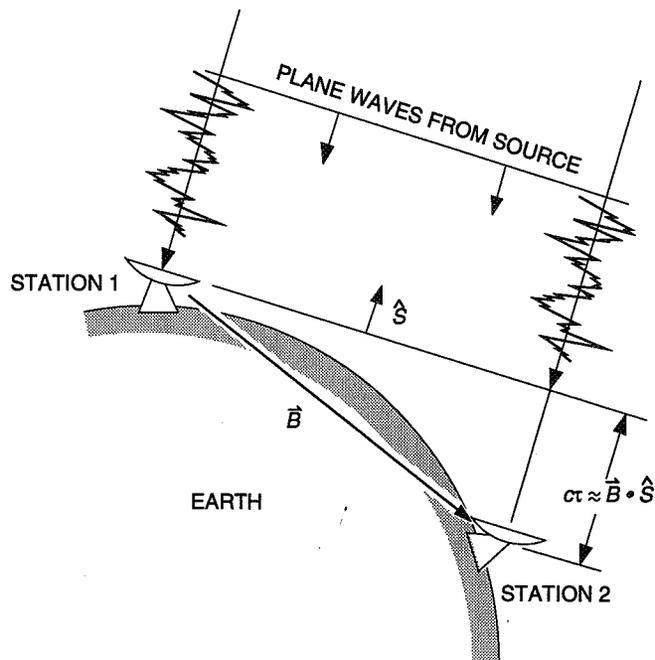


Fig. 2. VLBI geometry.

Since angles are measured only between the baseline vector and radio source directions, there is some arbitrariness in the definition of the direction of the coordinate axes. The celestial axes are traditionally defined in terms of the normal to the Earth's equator and the intersection of the equator and the ecliptic at some reference epoch (such as J2000). Because the Earth's rotation changes the baseline orientation with a diurnal signature, the VLBI data are sensitive to the equator of date. But these VLBI data are not sensitive to the position of the ecliptic. Before 1984, the orientation of the radio frame catalog about the equatorial pole was set by assigning a value to the right ascension of the source 3C273B based on observations of the occultation of this source by the Moon [9]. This choice of reference frame definition was found to be inadequate at the 5-nrad level of accuracy typical of modern VLBI measurements for two reasons: (1) The extrapolation of the equator of date to the equator of the reference epoch proved to be problematical, in that the standard theory for the motion of the equatorial pole is not accurate at the 5-nrad level, and (2) the source 3C273B has an apparent radio position that changes with time due to variations in source structure.

To avoid problems with changes in radio frame definition caused by variations in the model for motion of the Earth's pole and in the defined right ascension, the IERS in 1984 established a celestial reference frame by adopting coordinates for more than 20 commonly observed quasars with relative positions consistent at the 5-nrad level [10]. This celestial reference frame has been stable to better than 5 nrad since 1984 and is in approximate agreement with the Earth's equator and equinox of J2000 [8].

Given this definition of the celestial frame, it is necessary to define the zero point of baseline orientation. The IERS has specified positions (in terms of overall orientation) of a network of Earth tracking stations that define an "Earth-fixed" or terrestrial frame. This network includes VLBI antennas as well as sites for other techniques that can measure Earth orientation, including satellite laser ranging, lunar laser ranging, and Global Positioning System sites [11]. The network positions are continually updated as relative position information is improved. However, the overall orientation of the network is fixed so that the orientation of the Earth can be routinely monitored and reported in terms of the angles between the defined terrestrial network and the defined celestial frame.

B. Planetary Ephemeris Reference Frames

Range observations to planetary orbiters and landers, and radar ranges to the planetary surfaces, have provided a wealth of data on the motions of the planets in the solar system. Relative positions of Earth and Mars during the Viking lander ranging data arc (1976–1982) were known to an accuracy of a few nanoradians. Since then, the errors in relative longitudes have increased with time at a rate on the order of 1 nrad/yr [12]. The positions of Mercury and Venus relative to Earth have been determined to an accuracy of tens of nanoradians from planetary ranging data. Outer planets are known to lesser accuracy, based on spacecraft radio metric data during encounters, optical observations, and limited interferometric observations. For example, Jupiter's position is currently known to an accuracy of about 200 km, or 250 nrad, relative to Earth's orbit.

Beginning with JPL ephemeris DE200, the origin of right ascension for the planetary frame has been the dynamical equinox as determined from lunar laser ranging (LLR) data. LLR data are sensitive to the Earth's equator through the diurnal signature due to Earth rotation, and to the ecliptic through the effect of the luni-solar precession on the Earth–Moon distance. The LLR data determine the dynamical equinox at a mean data epoch of 1980 with an accuracy of approximately 5 nrad [12]. The ecliptic longitude of the Earth is determined to 15 nrad. The distributed ephemerides, however, are referred to the epoch J2000 rather than 1980. The uncertainty in predicting the difference in the equator of J2000 and the equator of 1980 is approximately 50 nrad. In general, each successive distributed ephemeris uses a longer arc of LLR data, which changes the estimate of the equator of J2000, so that the reference frames of two ephemerides are slightly different.

The distributed ephemerides do not contain information on the model for the motion of the Earth's pole that was used in defining the frame. Thus, angular offsets between different ephemerides cannot be

ascertained by comparing the orientations of the equator and equinox of date, but only by examining the tabulated positions of the planets. The relative positions of Earth and Mars are known to approximately 15 nrad, which is at least a factor of five better than the relative position accuracy of any other pair of bodies from the planetary ephemerides. Consequently, frame-tie comparisons between ephemerides, or between an ephemeris and the radio frame, have the most consistency and accuracy when based on the relative orientation of the orbit of the Earth and/or Mars in the two frames.

III. Data Acquisition

A. Observation Summary

The PHOBOS frame-tie data consist of two passes of VLBI observations of the Phobos-2 spacecraft and angularly nearby quasars. The PHOBOS data were recorded with the DSN wide-channel bandwidth (WCB) VLBI system, which is the DSN realization of the Mark III system developed for radio astronomy applications. Angular accuracy in the range of 1–5 nrad has been achieved with WCB observations of quasars [13,14]. The PHOBOS frame-tie data are the first interplanetary spacecraft–quasar differential VLBI measurements obtained with the WCB system.

The observing time in both passes was 1 hr. During the first pass on February 17, 1989, data were acquired by the DSN stations at Goldstone and Madrid and by the 76-m Lovell telescope at Jodrell Bank in England. Although the addition of Lovell improved the baseline geometry somewhat, the main reason for its inclusion was to enhance reliability. The second pass occurred on March 25, 1989, on the Goldstone–Canberra baseline. The DSN 70-m antennas were used for both passes, as they were the only DSN stations equipped to receive at the Phobos-2 spacecraft transmit frequency of approximately 1.7 GHz.

The Phobos-2 spacecraft and three extragalactic radio sources (see Fig. 3 for source geometry) were observed sequentially, not simultaneously, during each pass, according to the schedule in Fig. 4. Since uncalibrated path delay changes due to variations in solar plasma, troposphere, or ionosphere between spacecraft and radio source observations can degrade the accuracy of the measured differential delay, the observation sequences provided for sampling the temporal variation of measurement errors via multiple observations of a given source. Similarly, observations of quasars at various separation angles from the spacecraft were included to yield some indication of the spatial variation of media delays. An additional factor affecting the PHOBOS observation sequences was that the quasars with the smallest angular separations from Mars were relatively weak. Observations of well-known sources with high correlated flux density were included to provide a margin of reliability in case these weak sources could not be detected.

B. PHOBOS Frame-Tie Radio Sources

Of the six sources selected for the PHOBOS frame-tie experiments, the position of only one (0235+164) was known to an accuracy of 10 nrad or less. The positions of the remaining sources were determined with data from four DSN WCB VLBI passes in July and September of 1989—three passes on the Goldstone–Madrid baseline and one on the Goldstone–Canberra baseline. All data were acquired at 34-m stations.

During the astrometric passes, the PHOBOS sources and several quasars whose positions were precisely known were observed repeatedly over periods of about 8 hr. Data were acquired at both S-band (2.3 GHz) and X-band (8.4 GHz) so that signal delays due to charged particles of the ionosphere and the solar plasma could be calibrated. These observations were then combined with other DSN VLBI astrometric data. The entire data set was processed through the VLBI parameter estimation program, MODEST [15], to provide a quasar catalog, as well as Earth orientation parameters for each VLBI pass [16]. These results were then incorporated into the radio source catalog RSC(IERS) 91 C 01 of the International Earth Rotation Service. The $1\text{-}\sigma$ r.s.s. error in the estimated angular positions of the PHOBOS sources is about $15\sqrt{2}$ nrad (see Table 1 and the discussion in Section IV.B).

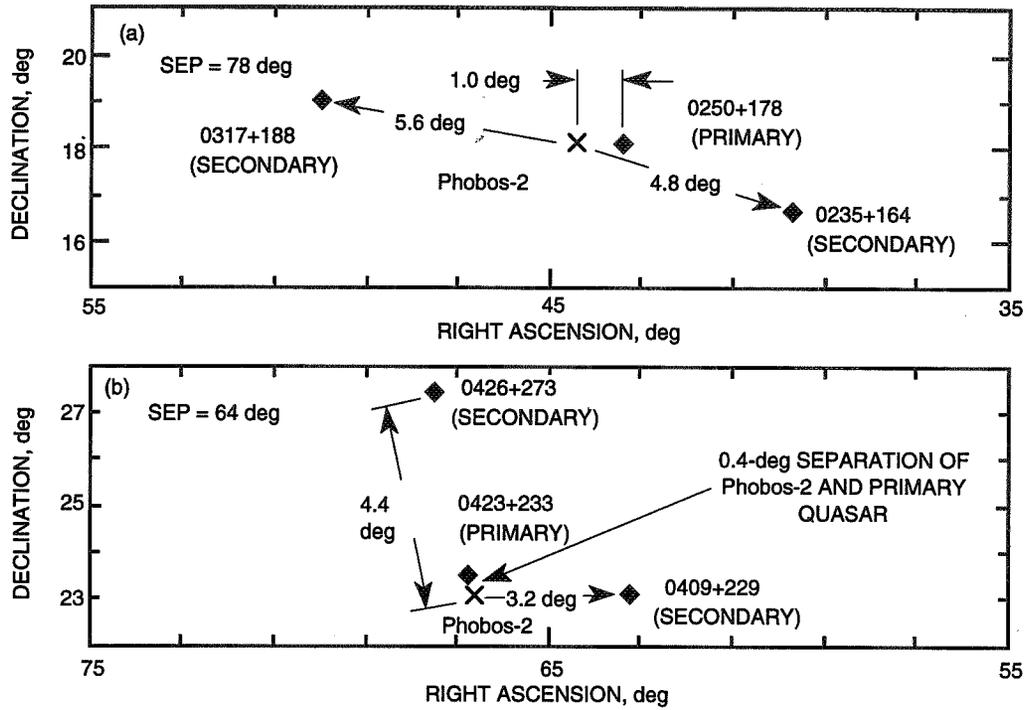


Fig. 3. Angular separation of the Phobos-2 spacecraft (X) from the radio sources (◆): (a) February 17, 1989, and (b) March 25, 1989. The source with the smallest angular offset from the spacecraft on each day is denoted as the primary source. SEP is the angular separation of the spacecraft and the Sun as viewed from Earth.

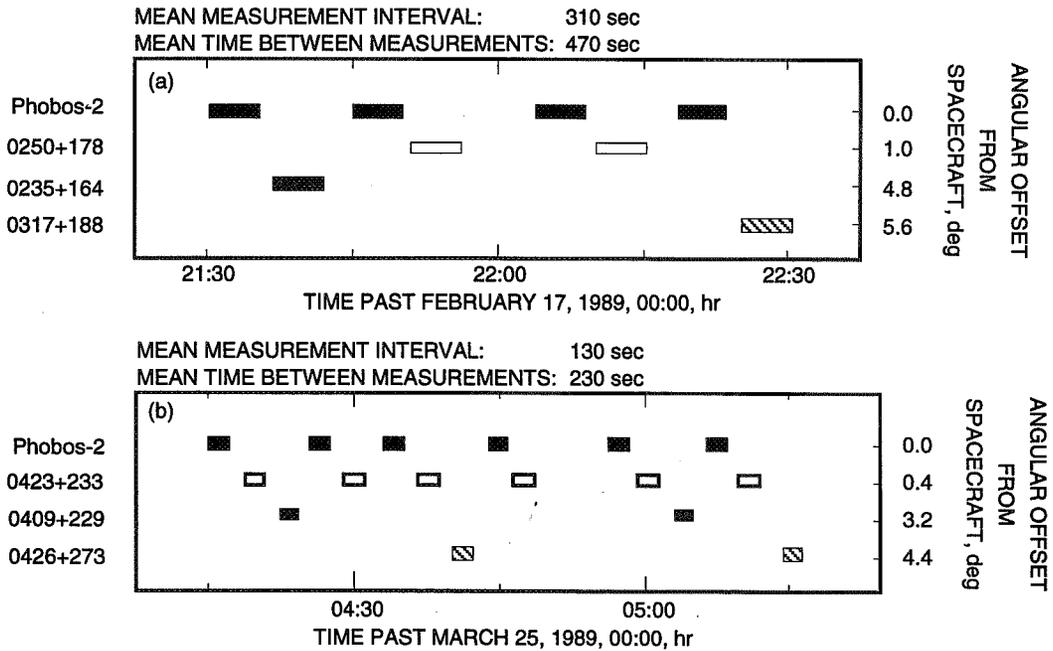


Fig. 4. Observation sequence: (a) February 17, 1989, and (b) March 25, 1989. Intervals of VLBI data acquisition for each source are shown by the horizontal bars.

Table 1. PHOBOS frame-tie error sources.

Error source	Magnitude	Notes
Thermal noise	Spacecraft, 0.18 nsec Quasar Wide-spanned bandwidth, 0.05 nsec Narrow-spanned bandwidth, 0.15 nsec	Approximate—sigmas are computed for each point during observable processing (see Appendix A)
Spacecraft orbit	February 17, 5.5 km March 25, 0.6 km	Spacecraft relative to Mars (See Footnote 7)
Planetary ephemeris	≈ 3 km	Mars relative to Earth orbit (See [6])
Zenith troposphere		
Constant	4 cm	(See [18] and Footnote 10)
Fluctuation at 1000 sec	1 cm	(See [19])
Ionosphere (Faraday calibration error)		
Time dependence	$0.5\sqrt{\Delta t/1000}$ nsec	See Appendix B
Bias	10 percent of calibration	
Solar plasma	~ 0.1 nsec	(See [21])
Instrumental bias	Goldstone–Madrid, 0.39 nsec Goldstone–Canberra, 0.10 nsec	(See Appendix A)
Quasar position	15 nrad (each component)	Source structure
Inertial station locations	20 cm (each component)	

IV. VLBI Measurement Sensitivity and Error Sources

A. Data Sensitivity

Following Finger and Folkner [2], the relation between the planetary and radio-frame positions of a body is assumed to be a rotation of coordinates having the form

$$\vec{r}_R = \vec{r}_P - \vec{A} \times \vec{r}_P$$

where \vec{r}_P is the position in the planetary frame, \vec{r}_R is the position in the radio frame, and \vec{A} is a vector of small rotations ($|\vec{A}| \ll 1$). If the positions of a body are independently determined in both frames, the frame-tie vector \vec{A} can be computed.

In the present case, the position of the Phobos-2 spacecraft is known in the planetary frame, and the PHOBOS VLBI data measure components of its angular position in the radio frame. The sensitivity of the VLBI delays to the frame-tie rotation vector is described below.

For the purposes of this discussion, the VLBI delay for the spacecraft is approximated by the time required for an electromagnetic wave transmitted by the spacecraft to traverse a distance in free space equal to the instantaneous offset in the locations of the two stations in the direction of the spacecraft geocentric position vector. That is, if \hat{S}_R is a unit vector along the Earth-to-spacecraft line in the radio frame, the observed delay is

$$\tau_{obs} = \frac{\vec{B}}{c} \cdot \left(\frac{\vec{r}_R}{r} \right) = \frac{\vec{B}}{c} \cdot \hat{S}_R$$

where \vec{B} is the baseline (i.e., vector separation of the two stations) in the radio frame and c is the speed of light.

The modeled delay is obtained from the previous equation by writing \hat{S}_R as a function of the rotation vector \vec{A} and the unit vector \hat{S}_P but expressed in planetary frame coordinates:

$$\tau_{model} = \frac{\vec{B}}{c} \cdot \left(\hat{S}_P - \vec{A} \times \hat{S}_P \right)$$

Nominally, the rotation vector is assumed to be zero, so the modeled delay is

$$\tau_{model} = \frac{\vec{B}}{c} \cdot \hat{S}_P$$

and the delay residual (observed minus modeled) can be written as

$$\Delta\tau = \frac{\vec{B}}{c} \cdot (\hat{S}_R - \hat{S}_P) = -\frac{\vec{B}}{c} \cdot (\vec{A} \times \hat{S}_P)$$

Since it is assumed that $|\vec{A}| \ll 1$, \hat{S}_P can be replaced by \hat{S}_R , so to first order,

$$\Delta\tau = -\frac{\vec{B}}{c} \cdot (\vec{A} \times \hat{S}_R) \quad (1)$$

The cross-product is given in terms of the components of the rotation vector \vec{A} as

$$\vec{A} \times \hat{S}_R = \begin{bmatrix} 0 & -A_3 & A_2 \\ A_3 & 0 & -A_1 \\ -A_2 & A_1 & 0 \end{bmatrix} \hat{S}_R$$

From Eq. (1), it is clear that the delay residual depends only on the rotation component which is in the plane of the sky (POS). The POS is orthogonal to \hat{S}_R (the Earth-spacecraft line of sight), so the relation of the delay residual to the rotation vector takes the form

$$\Delta\tau = -\vec{\tau}_{POS} \cdot (\vec{A}_{POS} \times \hat{S}_R) \quad (2)$$

where \vec{A}_{POS} is the POS component of \vec{A} . The vector $\vec{\tau}_{POS}$ is also in the POS and is in the direction of the projection of the station baseline, i.e.,

$$\vec{\tau}_{POS} = \frac{1}{c} \left[\vec{B} - (\vec{B} \cdot \hat{S}_R) \hat{S}_R \right]$$

The rotation vector can be represented in terms of small increments in right ascension and declination as

$$\vec{A}_{POS} = -\Delta\delta \hat{u} + \Delta\alpha \cos \delta \hat{v}$$

The quantities α and δ are the right ascension and declination of the source, and the unit vectors \hat{u} and \hat{v} lie in the POS, with \hat{u} in the direction of increasing right ascension (RA) and \hat{v} in the direction of increasing declination (DEC) (Fig. 5).

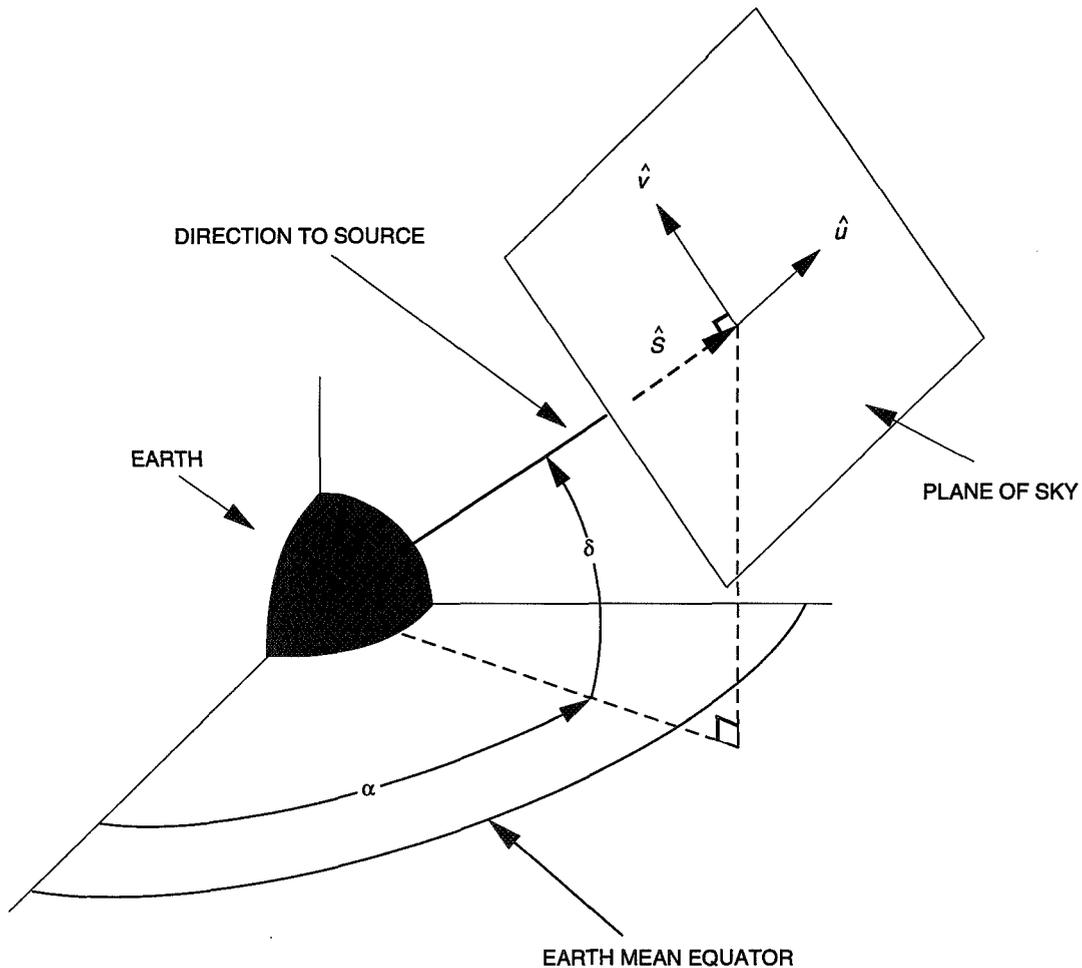


Fig. 5. Plane-of-the-sky coordinates.

A single delay measurement on a single baseline is sensitive only to that portion of \vec{A}_{POS} which is orthogonal to $\vec{\tau}_{POS}$. With VLBI measurements on one baseline, some sensitivity to the second component of \vec{A}_{POS} (in the POS and parallel to the projected baseline) is obtained with delay measurements separated in time so that Earth rotation causes the orientation of the projected baseline $\vec{\tau}_{POS}$ to change. The changes in projected baselines during the PHOBOS passes were mainly in length, not orientation (see Fig. 6), and the sensitivities to this rotation component are 30 to 40 times weaker than for the orthogonal

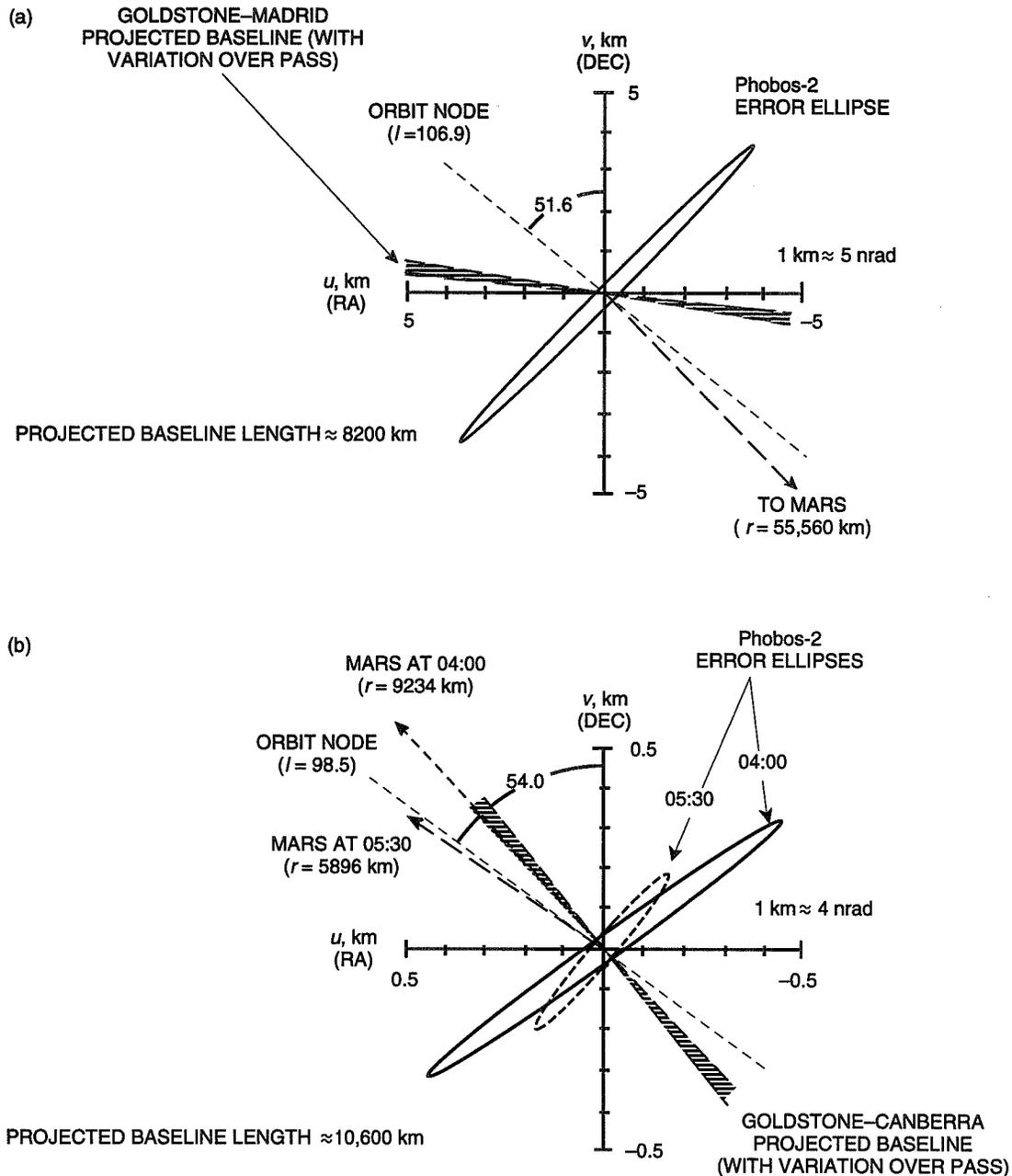


Fig. 6. Phobos-2 orbit errors in the plane of sky are represented by 1- σ error ellipses: (a) February 17, 1989, and (b) March 25, 1989. The projection of the station baseline is also shown. Only the component of the orbit error in the direction of the projected baseline affects the frame-tie parameter estimates. The orbit error for the March 25 pass is an order of magnitude smaller than for the February 17 pass.

component. A more accurate determination of the collinear component can be made by using data from a second baseline for which the projection on the POS has a large component orthogonal to the original $\vec{\tau}_{POS}$. A precise determination of the third component of \vec{A} , the rotation about the normal to the POS, would require an additional observation when the Earth-spacecraft line of sight was offset in angle by about 90 deg from its original direction.

The longest baselines for the first PHOBOS pass are approximately 8000 km in length, and mostly in the equatorial plane. The baseline for the second pass is about 10,000-km long, and has a large component

perpendicular to the equatorial plane. Together, the two passes provide baseline projections on the POS ($\overline{\tau}_{POS}$) which are large and have a large orthogonal component. But the motions of Earth and Mars between passes changed the orientation of the Earth-spacecraft line of sight by only 20 deg. Consequently, the PHOBOS data can provide accurate estimates of only two of the three rotation components.

B. Measurement Error Sources

The VLBI time delays are not just the differential signal arrival times at the two stations plus thermal noise, but include path delays due to media (ionosphere, troposphere, and solar plasma), instrumental phase response, and clock offsets, which are imperfectly known a priori. Furthermore, the delay models contain errors due to inaccuracies in spacecraft trajectory, Earth and Mars ephemerides, station locations, Earth orientation, and radio source positions. All of these effects have the potential to cause significant deviations in the residual delay, which in turn could cause errors in the frame-tie parameter estimates. The treatment of thermal noise, media, instrument, and clock effects in the process of obtaining the observed delay from the raw data is described in Appendix A. The important characteristics of both measurement and model errors as they pertain to the determination of the frame-tie parameters are summarized in Table 1.

1. Thermal Noise. Data weighting was based on estimates of the delay uncertainty from system noise obtained for each data point during postcorrelation processing (there is one data point for each observation interval on each source for a total of 22 delay measurements [see Fig. 4 and Section IV in Appendix A]). System noise effects in delay depend both on the signal strength and the range of frequencies that are sampled (the spanned bandwidth). For the spacecraft, with a usable bandwidth of about 30 MHz, the 1- σ delay uncertainty was typically about 0.2 nsec. Quasar delays were obtained for two different bandwidths—one set included measurements over the total spanned bandwidth of about 70 MHz and the other used essentially the same bandwidth as the spacecraft. The delay standard deviation from system noise for the widest spanned bandwidth was on the order of 0.05 nsec. For the narrow bandwidth, the delay uncertainty increased to 0.2 nsec. Solutions were obtained for each set of quasar delays separately, in combination with spacecraft delays, as a check on instrumental phase effects. (Note that a measurement error of 0.1 nsec corresponds roughly to an angular error of 3–4 nrad.)

2. Spacecraft Orbit. Since the VLBI data actually measure the angular offset between the quasars and the spacecraft, the frame tie can be achieved only if the a priori knowledge of the spacecraft position relative to Mars is sufficiently accurate. Phobos-2 spacecraft orbit estimates were obtained from the Navigation Systems Section. T. McElrath provided preliminary spacecraft trajectories in EME50 Cartesian coordinates relative to the DE118 planetary ephemeris, as well as covariances of errors in the spacecraft state estimates at the beginning of each pass.⁵ Refined trajectories and covariances from orbit determination processing incorporating additional data were provided by J. P. Berthias.⁶ The trajectories and covariances were transformed from EME50 coordinates to the Earth mean equator and equinox of J2000 (the DE200 system) by applying the rotation matrix given by E. M. Standish [6].

For the February 17, 1989, pass, the spacecraft was in a highly elliptic orbit with eccentricity of 0.8 and a period of 3.6 days. On March 21, 1989, it entered a nearly circular orbit with a period of 7.66 hr, which kept it close to Phobos (the mean distance from the satellite on March 25 was about 275 km). The data sets used for trajectory determination included instantaneous range rates derived from two-way C-band Doppler acquired by the Soviet network, two passes of two-way Doppler (C-band up,

⁵T. McElrath, personal communication, Navigation Systems Section, Jet Propulsion Laboratory, Pasadena, California, 1989.

⁶J.-P. Berthias, personal communication, Navigation Systems Section, Jet propulsion Laboratory, Pasadena, California, October 1990.

L-band down) from the DSN,⁷ and, for the March 25 orbit, measurements of the relative angular position of the spacecraft and Phobos derived from onboard imaging.⁸

The accuracy of the estimated orbits is represented by the $1\text{-}\sigma$ error ellipses for the projection of the Phobos-2 Mars-relative position onto the plane of sky (Fig. 6). The shape and orientation of the error ellipses indicate that the dominant uncertainty is in the orbit angular orientation about the Earth-spacecraft line of sight, which is typical of orbits determined primarily from Doppler data.

The magnitude of the orbit uncertainty depends on several factors, including the accuracy of the data, the length of the data arc, and the orbit shape and orientation. The orbit accuracy for the pass on March 25, about 500 m, is an order of magnitude smaller than the several-kilometer uncertainty for the February 17 pass. Favorable orbit geometry for the second pass further reduces the effect of orbit error on the frame-tie estimates. These trajectory uncertainties correspond roughly to 15 nrad and <1 nrad in geocentric angular position for the first and second pass, respectively.

3. Planetary Ephemeris. Uncertainty in the DE200 positions of Mars and the Earth are described by a covariance matrix on the errors in the estimated orbital elements.⁹ Included in this covariance are contributions due to a common angular offset of the orbits of Earth and Mars relative to the actual equator and equinox of the year 2000. It is this common angular orientation offset of the planetary and radio reference frames that is to be determined by the PHOBOS VLBI data, so the planetary ephemeris error model should reflect only the relative error in the positions of Earth and Mars. A covariance matrix for only this component of the error was obtained by statistically combining the a priori DE200 covariance with another covariance in which the sigmas on the parameters describing the orientation of the orbit of the Earth were essentially zero, and other orbit parameters were unconstrained. The $1\text{-}\sigma$ error ellipses for the resulting Earth-Mars direction uncertainty are shown in Fig. 7.

The ecliptic longitude of Mars relative to Earth's perihelion, and the inclination of the orbit of Mars to the ecliptic, have been determined from Viking lander range data [17]. The direction of the Earth-Mars line referenced to the orbit of the Earth is known to an uncertainty of 3 km (15 nrad) or better.

4. Troposphere. Tropospheric delay calibrations are computed using a seasonal model [18]. The calibration error is modeled as a constant offset in zenith troposphere delay at each site with a standard deviation of 4 cm.¹⁰ The zenith uncertainty is mapped to the proper elevation angle using the standard mapping function (approximately $1/\sin \gamma$, where γ is the elevation angle). At the lowest elevation angles ($\gamma \approx 22$ deg), the delay uncertainty at one station is about 11 cm or 0.4 nsec. The troposphere error in the delay differenced between sources separated in elevation angle by 1 deg is nearly 60 times smaller (0.2 cm, or less than 0.3 nrad).

The magnitudes of the delay error to be expected from fluctuations in the wet troposphere delay on time scales corresponding to the time between measurements ($\approx 200\text{--}500$ sec) and the duration of a pass (≈ 1 hr) are taken from Fig. 2 of [19], which is based on the assumption of Kolmogorov turbulence, with parameters adjusted to agree with variations measured by water vapor radiometers at the DSN sites. This model predicts variations of several millimeters on time scales of a few minutes, and about 1 cm over the pass at an elevation of 20 deg.

⁷ J.-P. Berthias, "Analysis of the Phobos-2 Radiometric Data Set," JPL Interoffice Memorandum 314-6 1129 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 1990.

⁸ J.-P. Berthias, October 1990, op. cit.

⁹ E. M. Standish, personal communication, Navigation Systems Section, Jet Propulsion Laboratory, Pasadena, California, 1991.

¹⁰ S. E. Robinson, "Errors in Surface Model Estimates of Zenith Wet Path Delays Near DSN Stations," JPL Interoffice Memorandum 335.4-594 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 1986.

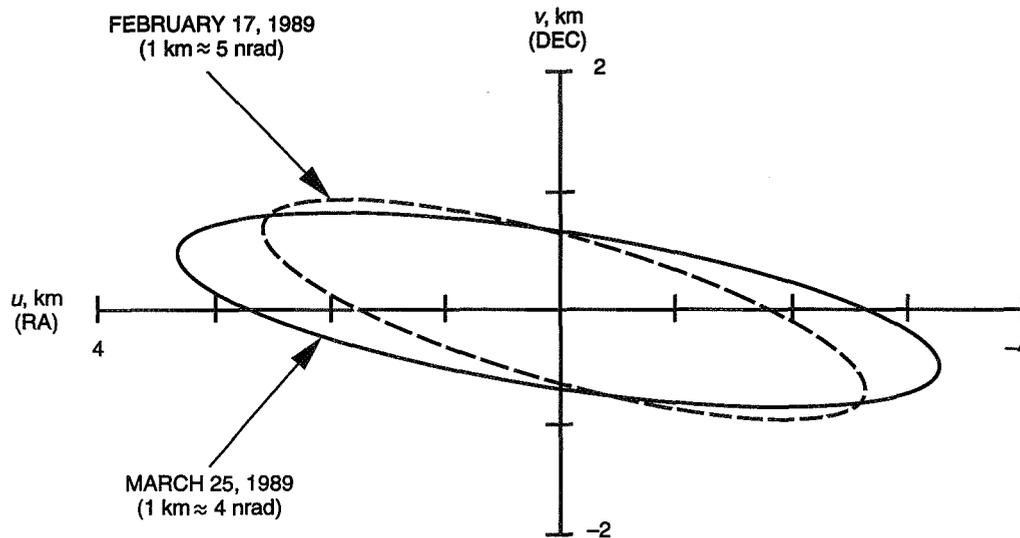


Fig. 7. Ephemeris DE200 error in the plane of sky (Mars relative to Earth-Moon barycenter orbit).

5. Ionosphere. Estimates of signal delays due to propagation through charged particles in the ionosphere were provided by the Tracking System Analytic Calibration Group. These estimates were based on measurements at each DSN site of the ionospheric delay of the signal from a geosynchronous Faraday beacon satellite [20]. To obtain delays along other ray paths, it is necessary to use a model for the temporal and spatial distribution of the ionosphere to map the measured ionosphere delay to the direction of other radio sources. The ionosphere mapping error is a major source of uncertainty in the ionosphere delay calibrations.

It is expected that the error in the ionosphere delay calibrations differenced between the spacecraft and quasar will depend to some extent on the apparent angular separation—the smaller the angle between the lines of sight, the closer together the ray paths through the ionosphere. For the small spacecraft-quasar separation angles of the PHOBOS data (in the range 0.4–5.6 deg), the time interval between measurements is also important. Earth rotation causes the position at which the ray path traverses the ionosphere to move in longitude at a rate of about 100 km in 5 min, whereas an angular source separation of 1 deg is equivalent to a position offset of approximately 10 km.

Because of uncertainties in assessing the spatial variability of the ionosphere over angular scales of many degrees, a subset of the quasar delay data was used to obtain the nominal frame-tie solution. The reduced quasar data set contains observations only of the quasar with the smallest angular separation from the spacecraft on each day (1.0 and 0.4 deg for the February 17 and March 25 passes, respectively). The main calibration errors were then modeled as temporal rather than spatial variations. In determining the frame-tie parameters, offsets of station clock epochs and rates were estimated for each pass with large a priori uncertainties ($1\text{-}\sigma$ of 1 msec in epoch and 10^{-9} in rate). This essentially eliminated estimate errors due to linear time variations of the calibration error. An analysis of Magellan S- and X-band data (Appendix B) showed that the deviation of the calibration error from a linear model had the form of a random walk. In the estimation process, the amplitude of the random walk was adjusted to yield a reduced χ^2 of 1.

An indication of the remaining sensitivity to calibration error was obtained by assessing the effect on the frame-tie parameter estimates of an error in Faraday calibration scaling. The a priori uncertainty for the calibration scaling was set to 10 percent of the calibration. This translates to approximately 2 cm or less in the delay difference between the spacecraft and nearest quasar.

6. Space Plasma. Plasma-induced variances of delay measurement errors are derived from the structure function of plasma phase variations. With the assumption that the plasma is concentrated in a thin screen that passes through the center of the Sun and is perpendicular to the line of sight from the Earth to the radio source, the structure function of phase for a homogeneous and isotropic plasma takes the form

$$D_\phi(b) \equiv \langle [\Delta\phi_{plasma}(x+b) - \Delta\phi_{plasma}(x)]^2 \rangle$$

where $\Delta\phi_{plasma}(x)$ is the plasma phase delay for a ray path that crosses the thin screen at a distance x from the center of the Sun [23]. From the analysis of [21], the structure function $D(b)$ of quasar group delay fluctuations due to the charged particles of the solar plasma is

$$D(b) = \frac{2.5 \times 10^4}{f^4} \left(\frac{b}{V_{SW}} \right)^{1.65} (\sin SEQ)^{-2.45} \text{ nsec}^2$$

SEQ is the Sun–Earth–quasar angle in deg, b is the length of the projection of the baseline on the plane of sky (or equivalently the thin screen) in km, V_{SW} is solar wind speed in km/sec, and f is signal frequency in GHz. This structure function is based on the phase scintillation spectrum of [22], which is specified to be valid in the range of 20 to 100 solar radii, or about 5–30 deg in SEQ . For SEQ angles of 30–90 deg, the model slightly overestimates the phase scintillation spectrum. For the spacecraft delay, the structure function has the same form, but the value of b is taken as the separation between the lines of sight at the points of closest approach to the Sun, which is smaller than the projected baseline due to parallax.

Using the approach given in [23], it can be shown that the standard deviation $\sigma_{\Delta\tau}$ of the plasma error in delay averaged over an interval T in sec is given by

$$\sigma_{\Delta\tau} = \frac{\sqrt{2.5 \times 10^{-4}}}{f^2} \left(\frac{b}{V_{SW}} \right)^{0.825} (\sin SEQ)^{-1.225} G(\kappa) \text{ nsec}$$

where

$$\kappa \equiv T \left(\frac{b}{V_{SW}} \right)^{-1}$$

$$G(\kappa) = \left\{ (2.65 \times 3.65 \times \kappa^2)^{-1} [(\kappa + 1)^{3.65} + |\kappa - 1|^{3.65} - 2(\kappa^{3.65} + 1)] \right\}^{1/2}$$

For the March 25 geometry (projected baseline of 10,600 km and $SEQ = 64$ deg) with a solar wind speed of 400 km/sec, the delay standard deviation for an averaging time of 130 sec is 70 psec. The corresponding value for February 17 (8200-km baseline, $SEQ = 78$ deg, and $T = 310$ sec) is 42 psec.

The 1- σ error in the difference of two averaged delays separated in time by Δt is

$$\sigma_{\Delta^2\tau} = \frac{\sqrt{2.5 \times 10^{-4}}}{f^2} \left(\frac{b}{V_{SW}} \right)^{0.825} (\sin SEQ)^{-1.225} F(\kappa, \Delta\kappa) \text{ nsec}$$

where

$$\begin{aligned}
F^2(\kappa, \Delta\kappa) &= 2G^2(\kappa) + (2.65 \times 3.65 \times \kappa^2)^{-1} \left\{ 2 [|\Delta\kappa + 1|^{3.65} + |\Delta\kappa - 1|^{3.65} - 2|\Delta\kappa|^{3.65}] \right. \\
&\quad - [|\kappa + \Delta\kappa + 1|^{3.65} + |\kappa + \Delta\kappa - 1|^{3.65} - 2|\kappa + \Delta\kappa|^{3.65}] \\
&\quad \left. - [|\kappa - \Delta\kappa + 1|^{3.65} + |\kappa - \Delta\kappa - 1|^{3.65} - 2|\kappa - \Delta\kappa|^{3.65}] \right\} \\
\Delta\kappa &\equiv |t_2 - t_1| \left(\frac{b}{V_{SW}} \right)^{-1}
\end{aligned}$$

For time intervals up to 1 hr, the model standard deviations for the difference of averaged plasma delays are less than 0.1 nsec.

For *SEQ* near 90 deg, measured spectral densities vary by up to 2 orders of magnitude [22], so that the plasma structure function given above may be too large or too small by a factor of 10. The corresponding upper limit for the PHOBOS plasma delays is about 0.3 nsec.

7. Instrumental Bias. The difference between the recorded spectra of the spacecraft and quasars may cause spacecraft delays to be biased relative to those of the quasars. The sampled spacecraft signal consisted of a series of discrete, coherent tones. For the quasars, the recorded spectrum was composed of several narrow “windows” (commonly referred to as “channels”), each consisting of white noise over a 2-MHz bandwidth (Fig. 8). The overall bandwidth of the sampled quasar spectra was much larger than that of the spacecraft signal—approximately 70 MHz versus 30 MHz. Estimates of the resulting delay bias errors are 0.39 nsec on the Goldstone–Madrid baseline and 0.10 nsec for Goldstone–Canberra. These values for the delay bias uncertainties are obtained from postcorrelation processing, as described in Appendix A.

8. Quasar Position. Source coordinates were taken from the IERS celestial reference frame RSC(IERS) 91 C 01 [8]. The accuracy with which the position of any given source is tied to this frame depends on several factors, including the number of observations available, observation frequencies, signal strength, and source structure. Generally, quasar positions are known with an accuracy of <5 nrad. The accuracy for the PHOBOS sources is not this good. With only a single exception (0235+164), the PHOBOS sources have only recently been added to the catalog, and repetitive observations that would provide a check on the consistency of the position determinations are not available.

Although the formal uncertainties in PHOBOS source positions are 6 nrad or less, the actual errors are almost certainly larger. The signal strengths of the quasars within 1 deg of the Phobos-2 spacecraft are fairly low. This could be an indication that the sources are not compact but have significant structure that might affect repeatability of the position estimates. Furthermore, the astrometric measurements were made at S- and X-bands while the frame-tie data were taken at L-band, so the position used in the frame-tie determination may be in error due to differences in structure at the different observing frequencies. For these reasons, the standard deviation of the quasar positions is conservatively taken to be 15 nrad in both right ascension and declination.

9. Inertial Station Locations. Locations of the radio antennas are given in the IERS Terrestrial Reference Frame (TRF) 1988.0 [8]. Continental drift was not modeled, so this error is included in the station location standard deviations. The size of the continental drift error is taken as the maximum value given by the AM0-2 plate motion model (7 cm or less [24]). Uncertainties of the crust-fixed locations (10 cm) are based on the accuracy of the determination of the geocenter [25,26]. The orientation of IERS TRF 1988.0 in inertial space is described by Earth polar motion, universal time, and the direction of the celestial pole. The polar motion and universal time calibrations used were consistent with the

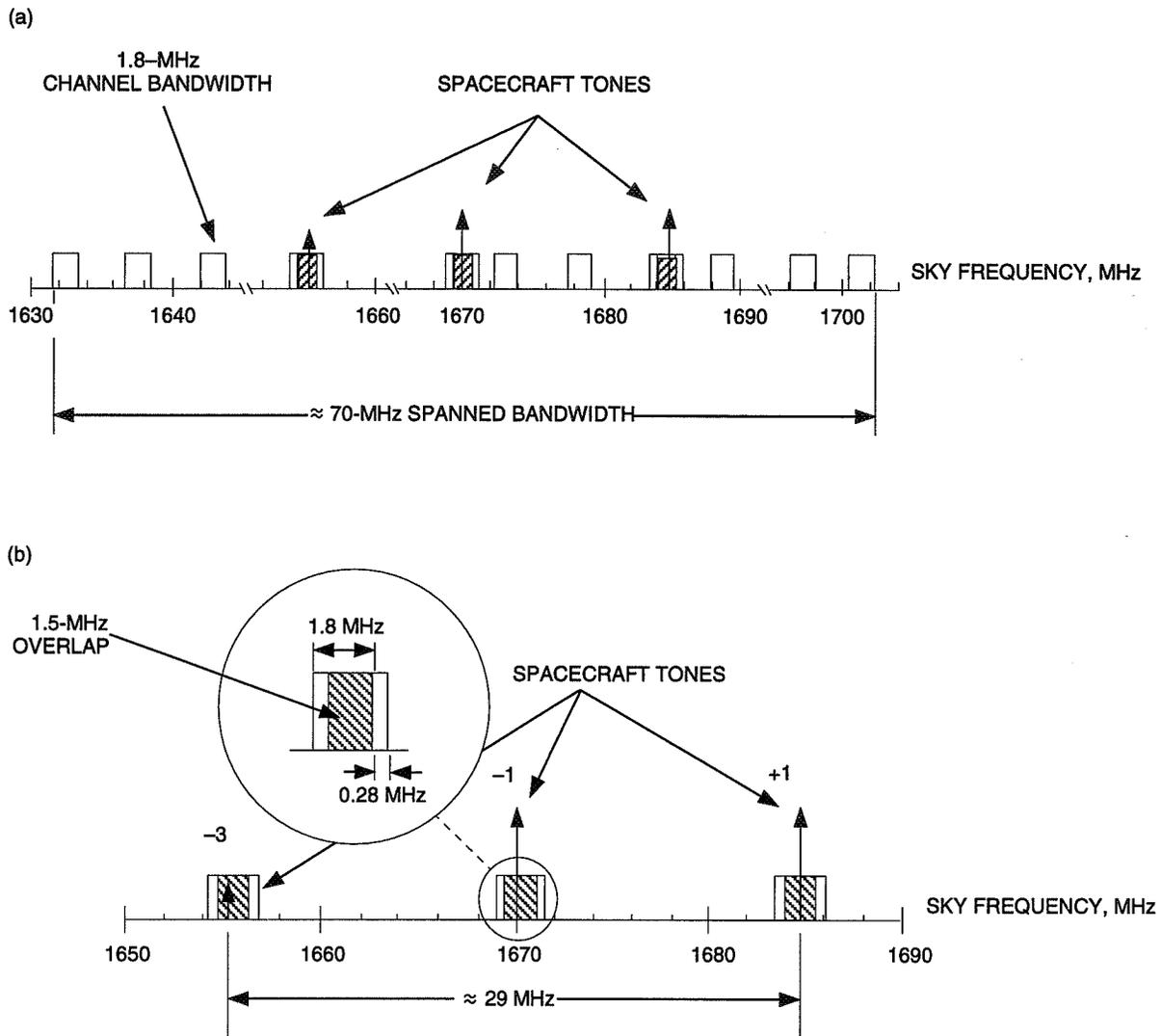


Fig. 8. VLBI channel configuration: (a) all channels and (b) spacecraft channels.

IERS celestial and terrestrial frame definitions at the 6-cm level [8]. The direction of the celestial pole was modeled by the 1980 IAU nutation model, which is consistent with the IERS frame definitions at the 6-cm level [8]. The uncertainty in inertial station locations due to errors in crust-fixed position and Earth orientation were combined and treated as “equivalent” errors in station locations with a standard deviation of 20 cm per component.

V. Results

A. Reference Solution

Figure 9 gives delay residuals relative to an a priori model. Drifts on the order of a few parts in 10^{13} caused by station clock rate offsets and uncalibrated ionosphere have been removed by fitting all residuals from a given baseline to a linear polynomial. The spacecraft delay residuals are indicated by filled circles, those for the primary quasars by open circles, and for the secondary quasars by triangles. The “primary quasar” is the one with the smallest angular separation from the spacecraft for each pass. The other

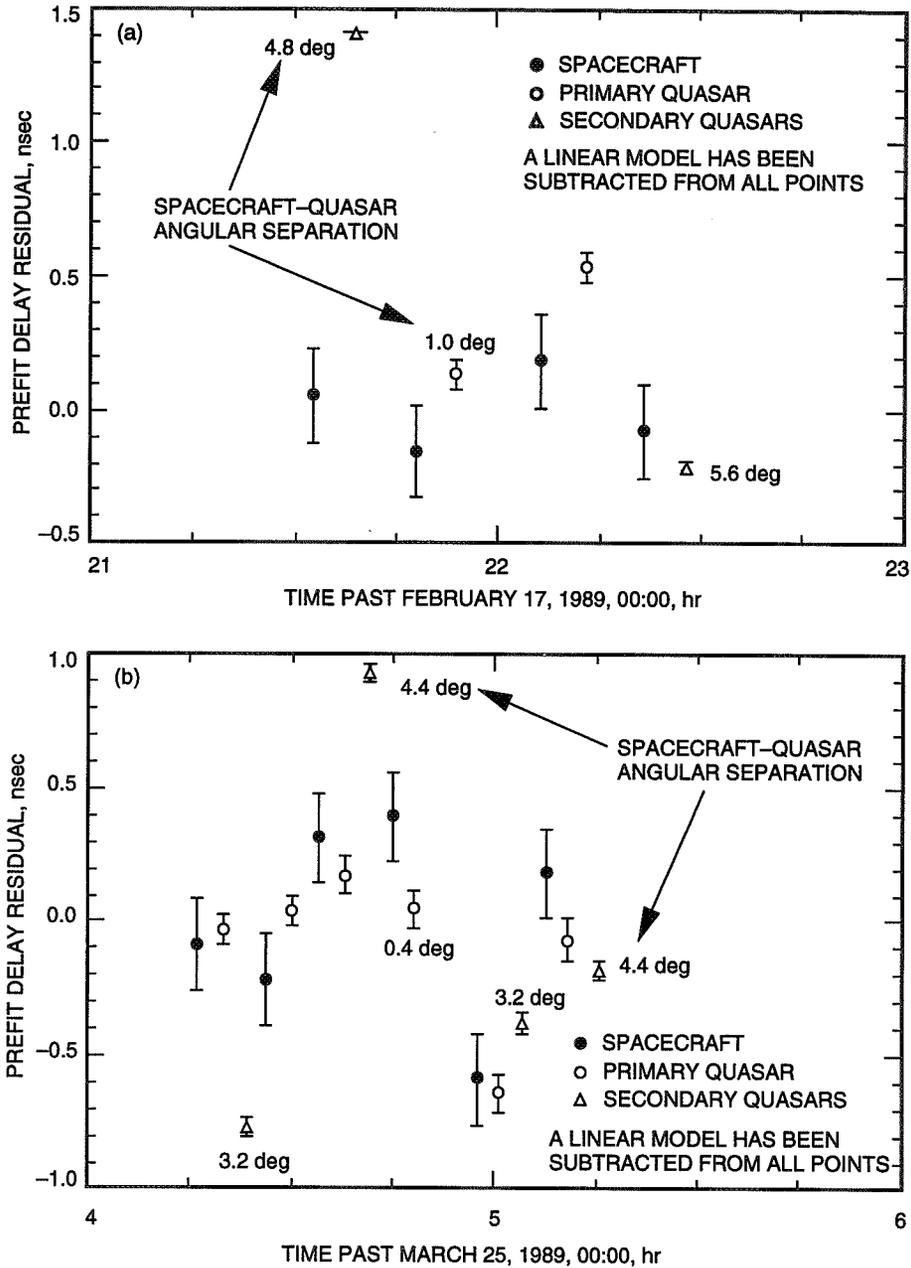


Fig. 9. Prefit delay residuals: (a) Goldstone-Madrid and (b) Goldstone-Canberra.

sources are referred to as "secondary quasars." The a priori model assumes the offsets of the planetary frame relative to the radio frame are zero; a nonzero offset causes a bias in the difference of the spacecraft and quasar delays.

The maximum difference between the spacecraft and quasar residuals is about 1.6 nsec for both passes. Residuals of this magnitude could be produced by errors of about 60 nrad in spacecraft angular position along the projection of the baselines on the plane of sky. However, for those cases in which a spacecraft observation was immediately followed by an observation of the primary quasar, the difference in delay is much smaller—always less than 0.5 nsec. This is consistent with the hypothesis that the 1.6-nsec offsets are due to model errors that are not canceled in the difference between the spacecraft and secondary quasars. Furthermore, the scatter in the residuals is larger than the $1\text{-}\sigma$ system noise error (indicated

by error bars), and significant changes in delay residuals are seen between successive points on a single source—the largest is 1.2 nsec between the spacecraft points at 4:45 and 4:58 for the March 25 pass. These effects are most likely due to error in the Faraday calibration of the ionospheric delays.

A noise model for the measurement error due to uncalibrated ionosphere was devised as described in Appendix B. The resulting structure function for the calibration error relative to a linear variation has the form of a random walk. That is, the error for measurement $i + 1$ is related to that of measurement i by

$$\tau_{i+1}^I = \tau_i^I + w_i$$

and the a priori standard deviations of the w_i have the form

$$\sigma_w(i) = C \left(\frac{t_{i+1} - t_i}{1000} \right)^{1/2}$$

Here, C is the a priori standard deviation for points separated in time by 1000 sec. Analysis of Magellan S-/X-band (2.3 and 8.4 GHz, respectively) data yielded an initial value for C of 0.26 nsec at the PHOBOS frequency.

The w_i ($i = 1$ to $N-1$ for N data points) were then taken to be additional (correlated) measurement noise. The value of C was adjusted to 0.5 nsec to give a weighted sum of squares of postfit residuals near the expected value. The initial error τ_1^I was not included, since this constant delay offset can be absorbed into the clock epoch correction (see below in this section).

Measurements from the secondary quasars were not included in the reference fit because the errors in the differenced ionosphere delay calibrations were expected to be several times larger at separation angles of 3.2–5.6 deg than at 1 deg or less. Although the random walk model strictly applies only to successive observations of the same source, it was assumed to accurately represent the differential delay error between the spacecraft and primary sources as well.

Data for baselines including the Jodrell station were not used in the reference fit. Jodrell is at a significantly higher latitude than the DSN stations, and the mapping of the Faraday calibration to the ionospheric pierce point may not be as accurate as for the DSN data.

Clock epoch offset and rate for each baseline were estimated to account not only for actual clock offsets, but also for any errors that are common to the spacecraft and quasar delays and which vary over the pass approximately linearly in time. Solving for these clock model parameters provides error cancellation even though the spacecraft and quasar delays were not explicitly differenced.

The delays were weighted with the inverse of the delay variances obtained during postcorrelation processing (Appendix A). These statistics mainly reflect the delay uncertainties due to system noise. For the reference solution, the quasar delays were produced from phases measured in 14 channels, each with a bandwidth of approximately 2 MHz, spread over 70 MHz (Fig. 8). The spacecraft delays utilize four channels for the February 17 pass and six for the pass on March 25, with a total spanned bandwidth of 29.4 MHz.

The remaining error sources were included via “consider” analysis [27]. That is, the errors in the estimates of the frame-tie parameters produced by $1-\sigma$ errors in the a priori value of each consider parameter were determined. The considered parameters were as follows:

- (1) Spacecraft position and velocity at a specified epoch.
- (2) Ephemeris correction parameters for the orbit of Mars relative to the orbit of the Earth (at the epoch of the ephemeris).
- (3) Constant zenith troposphere calibration error for each station.
- (4) Constant Faraday ionosphere calibration error in scale (each station).
- (5) Spacecraft versus quasar instrumental biases (constants, modeled as a bias in spacecraft delay for each baseline).
- (6) “Equivalent” station locations (includes crust fixed positions, continental drift, and Earth orientation).
- (7) Constant errors in quasar positions.

The PHOBOS data are sensitive to only two of the three components of the frame-tie rotation vector, so the solutions are best presented in a plane orthogonal to the major principal axis of the error ellipsoid. Orthogonal axes P_1, P_2 in this plane are defined such that P_1 is along the intersection with the nominal inertial reference plane (DE200 Earth mean equator of J2000). This coordinate system is shown in Fig. 10, and the two well-determined components of the rotation vector along with their $1\text{-}\sigma$ error ellipse are given in Fig. 11(a). The P_1, P_2 axes are closely aligned with the POS vectors $-\hat{u}$ and $+\hat{v}$ for some effective Mars right ascension and declination determined by the data. The right ascension of Mars for the two passes differed by about 22 deg, and the declination by 5 deg, so this alignment of axes is only approximate. The minor axis of the error ellipsoid is nearly orthogonal to the projection of the Goldstone–Canberra baseline on the POS.

For convenience of comparison with other frame-tie determinations, the solutions and error ellipses are also presented in terms of spacecraft right ascension and declination [Fig. 11(b)]. Again, these values are only approximate, for the reasons given above. The complete consider covariance on the components of the frame-tie vector \underline{A} and the corresponding error ellipsoid are given in Tables 2 and 3. (Note that the consider covariance is the covariance on the estimates accounting for all error sources, including thermal noise and random walk calibration error.)

The reference solution is consistent with a frame offset of zero in the P_1, P_2 plane. A breakdown of the contribution of the individual error sources to the overall accuracy of the determination of these rotation components is shown in Fig. 12. The labels “February 17 Orbit” and “March 25 Orbit” denote Phobos-2 orbit errors, and “G/C” and “G/M” refer to the Goldstone–Canberra and Goldstone–Madrid baselines, respectively. The consider errors tend to align with either the semimajor or semiminor axes of the total error ellipse. Most of the error is due to the instrumental delay biases for the February 17 pass and to the uncertainty in the Phobos-2 orbit on this day. The uncertainties in the minor axis direction have a large contribution from errors in the position of Mars relative to Earth as given by DE200. There is also a significant error due to uncertainty in the quasar positions in the radio frame.

B. Solution Consistency

Several variations on the basic estimation scheme were tried to determine if the scatter in the solutions was consistent with the expected errors. The results below are for the following fits:

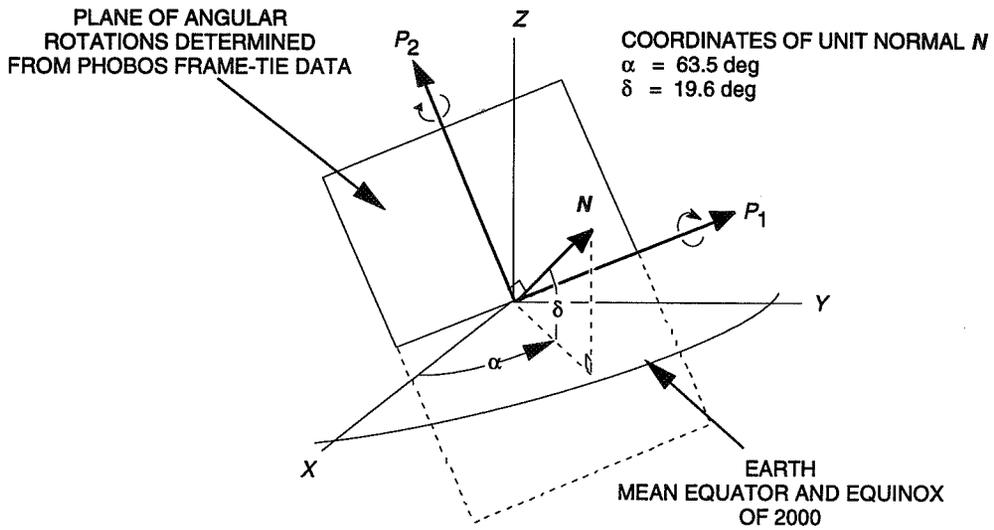


Fig. 10. PHOBOS frame-tie coordinates.

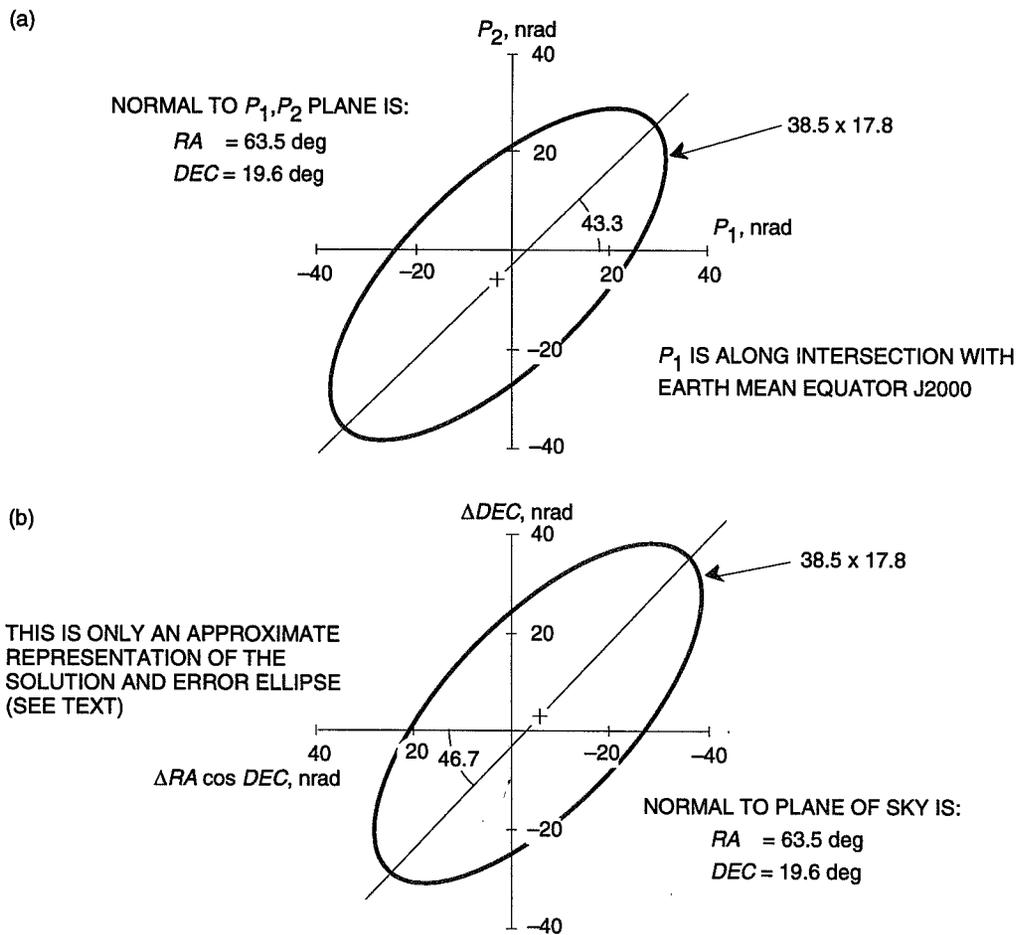


Fig. 11. PHOBOS frame-tie reference solution and error ellipse: (a) P_1, P_2 coordinates and (b) right ascension and declination.

- (1) Madrid (DSS 63) replaced by Jodrell
- (2) Quasar delays computed from phases in spacecraft channels only
- (3) Delays from all quasars included in solution
- (4) Case (3) without random walk ionosphere measurement noise

Solutions for these cases are shown with the reference solution in Fig. 13.

For fit (1), the data set is the same as for the reference solution except that Goldstone–Jodrell delays are used instead of those from the Goldstone–Madrid baseline. Comparison of these estimates with the reference solution potentially could reveal delay measurement errors due to differences in instrumentation between the DSN stations and the Jodrell station. Instrumental errors may occur if the manual phase calibration procedure fails, or if there are significant shifts in phase in spacecraft tones due to small scale phase variations with frequency. But the observables are also different because of other errors that are not the same at Jodrell and Madrid. For example, differences on the order of 0.1 nsec (about 4 nrad) are expected from system noise, and differences of several nanoradians could be caused by ionosphere calibration error in scale. The actual difference in solutions is 7 nrad, so there is no evidence of significant instrumental error.

For case (2), instrumental errors could cause the quasar delays to depend upon the specific channels used in obtaining the measured delays. Also, reducing the number of channels increases the system noise error in the delays. Again the estimates for this case are in good agreement with the reference solution.

Fits for cases (3) and (4) show the effect on the solutions of including the delays from the secondary sources, which are expected to have nanosecond-level errors due to uncalibrated ionosphere. For case (3), the a priori sigmas on the random walk ionosphere parameters were doubled (to 1 nsec in 1000 sec) to yield an acceptable weighted sum-of-squares (WSOS) for the postfit residuals. Doubling the sigmas produced only small increases in the uncertainties of the estimates. It is probably fortuitous that the

Table 2. Estimated J2000 Cartesian frame-tie vector \vec{A} and consider covariance matrix.

Component	Solution, nrad	Consider covariance, nrad ²		
		A_1	A_2	A_3
A_1	61.1	89,285.29	180,220.16	72,333.97
A_2	131.1	—	368,556.02	148,959.86
A_3	58.8	—	—	60,993.31

Table 3. Frame-tie estimate and covariance referred to consider covariance principal axes.

Solution, nrad	Semiaxes, nrad	Unit principal axes in J2000 frame		
		a	b	c
-156.1	719.05091	0.41338401	-0.75676500	0.50637970
5.8	38.53070	0.84417073	0.11006405	-0.52465386
2.3	17.77020	0.34130548	0.64435444	0.68433758

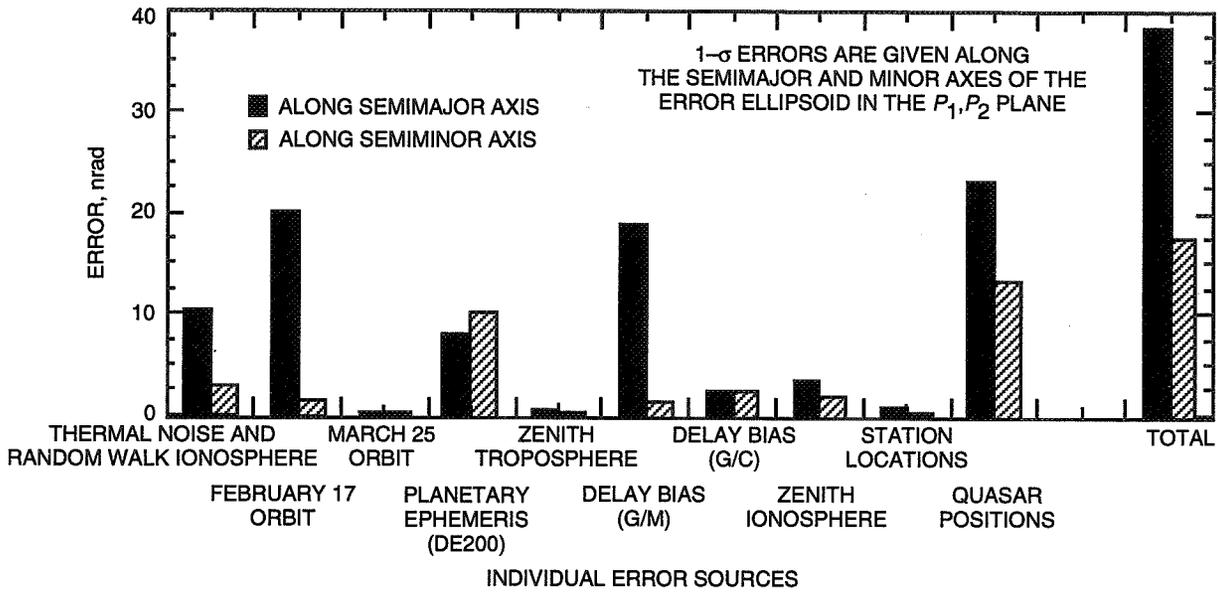


Fig. 12. PHOBOS frame-tie error summary.

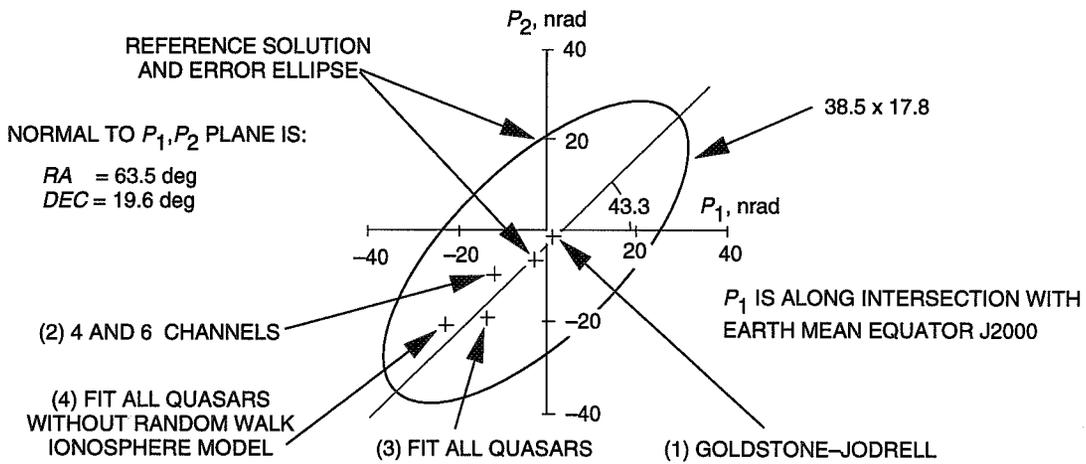


Fig. 13. PHOBOS frame-tie solution consistency.

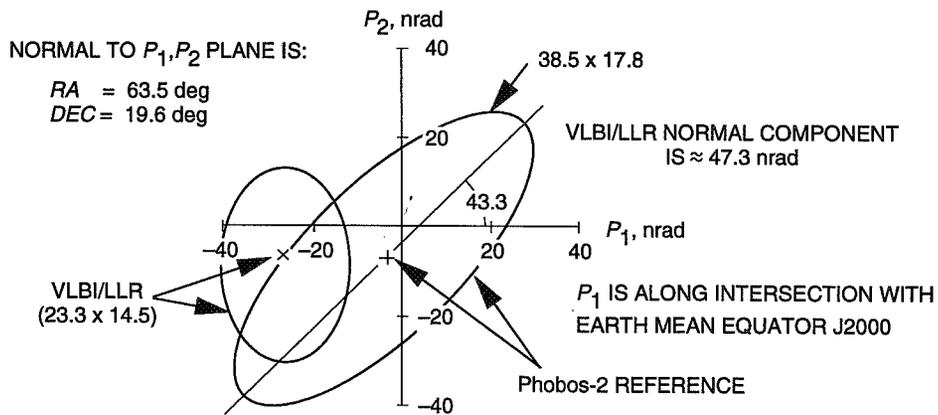


Fig. 14. PHOBOS frame tie versus VLBI/LLR.

case (4) result is consistent with other solutions; without the ionosphere contribution to the measurement noise, the WSOS of the postfit residuals was extremely large.

C. Comparison With Other Frame-Tie Estimates

Since the PHOBOS data provide accurate estimates of only two components of the small rotation vector relating the planetary and radio frames, these results cannot in general be compared with estimates presented as corrections to right ascension and declination. It is possible, however, to check the agreement with the three-component VLBI-LLR frame tie [2].

The VLBI-LLR frame tie was derived from a comparison of station location and Earth orientation parameter values based on DSN VLBI data with those obtained from lunar laser ranging data. Ground ties between the DSN stations, VLBI stations of the Crustal Dynamics Project, satellite laser ranging sites, and lunar laser ranging sites were also employed. The LLR data provide the tie to the planetary frame, as these data provide the location of the dynamical equinox to which the positions of the planetary ephemeris are referenced. The accuracy of the VLBI/LLR frame tie is about 15 to 20 nrad (Table 4).

The VLBI-LLR components in the PHOBOS P_1, P_2 frame with their corresponding error ellipse are shown in Fig. 14. The difference between the two solutions corresponds to about 0.7σ . The PHOBOS data are not sensitive to the third component, which, with a value of 47.3 nrad, is the largest of the three VLBI-LLR rotation parameter estimates.

Table 4. VLBI/LLR frame-tie referenced to Earth mean equator and equinox of J2000.

Component	Value, nrad	σ , nrad
A_1	5	15
A_2	-49	15
A_3	-19	25

VI. Conclusions and Future Activities

Two passes of VLBI data were acquired at the DSN from the Soviet Phobos-2 spacecraft while it was in Mars orbit. VLBI observations of quasars at small angular separations (1 deg or less) from the spacecraft were also included. These data are sensitive to the orientation of the Earth-Mars line of sight relative to the reference frame defined by the coordinates of selected extragalactic radio sources. Because the positions of Earth and Mars are well known within the frame of the planetary ephemeris, analysis of the VLBI measurements has yielded estimates of two components of the angular rotation relating the frame of JPL planetary ephemeris DE200 to the IERS radio frame. The $1-\sigma$ uncertainty in the frame-tie rotation components is 20-40 nrad, depending on direction.

The accuracy of the frame tie cannot be better than the internal consistency of either the planetary ephemerides or the radio source catalog. Relative positions of catalogued sources are stable to the order of 5 nrad [10]. The internal consistency of the planetary ephemeris is not as good. For example, Earth and Mars have uncertainties in their mean motions on the order of 1 nrad/yr [12].¹¹

In the same way that the PHOBOS data measured the Earth-Mars direction in the radio frame, VLBI data acquired from the Magellan spacecraft are being used to measure the Earth-Venus direction with

¹¹ J. G. Williams, personal communication, Tracking and Applications Section, Jet Propulsion Laboratory, Pasadena, California, 1993.

an accuracy of 5 nrad.¹² The Magellan VLBI data cannot directly improve the frame-tie determination because the uncertainty in the position of Venus in the planetary frame is about 100 nrad. But these data can give precise information on the motions of Earth and Venus in the radio (inertial) frame. Future missions using planetary orbiters or landers with the capability of providing ranging data with meter-level accuracy, in addition to VLBI delay measurements, would enable additional opportunities to improve the frame tie. With differential VLBI delay accuracy comparable to the Magellan data, a frame-tie accuracy of 10 nrad or less should be achievable.

Acknowledgments

The Phobos-2 VLBI measurements were acquired in preparation for a high-accuracy frame-tie determination that would have employed VLBI data from a lander on the surface of the Mars' satellite Phobos. The principal investigators for the PHOBOS lander frame-tie experiment and their affiliations were Jacques Blamont (CNES), Vyacheslav Linkin (IKI) and Robert Preston (JPL/NASA). The authors would like to thank the following: Roger Linfield for his careful reading of several versions of this article and his helpful suggestions; Dave Fort for providing valuable consultations in processing the data on the DSN Block II correlator; Jim Border for help in generating spacecraft and correlator predicts; Tim McElrath and Jean-Paul Berthias for spacecraft trajectory estimates and the associated covariances; Brian Wilson and Titus Roth for providing ionosphere calibrations; the TEMPO and TSAC teams for providing Earth-orientation, troposphere calibrations and Faraday calibrations of ionosphere delays; and Jim Williams and Mark Finger for many helpful discussions and suggestions along the way.

References

- [1] D. Jones, R. Dewey, C. Gwinn, and M. Davis, "Mark III VLBI Astrometry of Pulsars," IAU Colloquium 131, *Radio Interferometry: Theory, Techniques, and Applications*, T. Cornwell, ed., Siorro, New Mexico, 1990.
- [2] M. H. Finger and W. M. Folkner, "A Determination of the Radio-Planetary Frame Tie From Comparison of Earth Orientation Parameters," *The Telecommunications and Data Acquisition Progress Report 42-109*, vol. January-March 1992, Jet Propulsion Laboratory, Pasadena, California, pp. 1-21, May 15, 1992.
- [3] A. E. Niell, XX Newhall, R.A. Preston, G. L. Berge, D. O. Muhleman, D. J. Rudy, J. K. Campbell, P. B. Esposito, and E. M. Standish, "Relating the Planetary Ephemerides and the Radio Reference Frame," *The Telecommunications and Data Acquisition Progress Report 42-81*, vol. January-March 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 1-8, May 15, 1985.

¹² Folkner et al., op. cit.

- [4] XX Newhall, R. A. Preston, and P. B. Esposito, "Relating the JPL VLBI Reference Frame and the Planetary Ephemerides," *Proceedings of IAU Symposium 109, Astrometric Techniques*, H. K. Eichhorn and R. J. Leacock, eds., D. Riedel, pp. 789–794, 1986.
- [5] T. P. McElrath and R. S. Bhat, "Determination of the Inner Planet Frame Tie Using VLBI Data," paper 88-4234, AIAA/AAS Astrodynamics Conference, Minneapolis, Minnesota, 1988.
- [6] E. M. Standish, "Orientation of the JPL Ephemerides, DE200/LE200, to the Dynamical Equinox of J2000," *Astron. Astrophys.*, vol. 114, pp. 297–302, 1982.
- [7] E. M. Standish, "The Observational Basis for JPL's DE200, the Planetary Ephemeris of the Astronomical Almanac," *Astron. Astrophys.*, vol. 223, pp. 252–271, 1990.
- [8] *Annual Report for 1990*, International Earth Rotation Service, Observatoire de Paris, Paris, France, 1991.
- [9] C. Hazard, J. Sutton, A. N. Argue, C. M. Kenworthy, L. V. Morrison, and C. A. Murray, "Accurate Radio and Optical Positions of 3C273b," *Natural Physical Science*, vol. 233, pp. 89–91, 1971.
- [10] E. F. Arias, M. Feissel, and J.-F. Lestrade, "An Extragalactic Celestial Reference Frame Consistent With BIH Terrestrial System (1987)," *BIH Annual Report for 1987*, Observatoire de Paris, Paris, France, 1988.
- [11] C. Boucher and Z. Altamimi, "The Initial IERS Terrestrial Reference Frame," *IERS Technical Note 1*, Observatoire de Paris, Paris, France, 1989.
- [12] E. M. Standish and J. G. Williams, "Dynamical Reference Frames in the Planetary and Earth–Moon Systems," *Inertial Coordinate Systems on the Sky*, Dordrecht, Holland: Kluwer Academic Publishers, pp.173–181, 1990.
- [13] O. J. Sovers, "JPL 1990-3: A 5-nrad Extragalactic Source Catalog Based on Combined Radio Interferometric Observations," *The Telecommunications and Data Acquisition Progress Report 42-106, vol. April–June 1991*, Jet Propulsion Laboratory, Pasadena, California, pp. 364–383, August 15, 1991.
- [14] R. N. Treuhaft and S. T. Lowe, "A Measurement of Planetary Relativistic Deflection," *Astron. J.*, vol. 102, no. 5, pp. 1879–1889, 1991.
- [15] O. J. Sovers, *Observation Model and Parameter Partial for the JPL VLBI Parameter Estimation Software MODEST—1991*, JPL Publication 83-39, Rev. 4, Jet Propulsion Laboratory, Pasadena, California, 1991.
- [16] J. A. Steppe, O. J. Sovers, and S. H. Oliveau, "Smoothed Standard-Coordinate Earth Rotation From Deep Space Network VLBI," *IERS Technical Note 11: Earth Orientation, Reference Frames, and Atmospheric Excitation Functions*, submitted for the *1991 IERS Annual Report*, ed. P. Charlot, Observatoire de Paris, Paris, France, June 1992.
- [17] J. G. Williams and E. M. Standish, "Dynamical Reference Frames in the Planetary and Earth–Moon Systems," *Reference Frames in Astronomy and Geophysics*, J. Kovalevsky et al., eds., Dordrecht, Holland: Kluwer Academic Publishers, pp. 67–90, 1989.
- [18] C. C. Chao, *The Tropospheric Calibration Model for Mariner Mars 1971*, JPL Technical Report 32-1587, Jet Propulsion Laboratory, Pasadena, California, pp. 61–76, 1974.

- [19] R. N. Treuhaft and G. E. Lanyi, "The Effect of the Dynamic Wet Troposphere on Radio Interferometric Measurements," *Radio Science*, vol. 22, pp. 251–275, 1987.
- [20] H. N. Royden, D. W. Green, and G. R. Walson, "Use of Faraday-Rotation Data From Beacon Satellites to Determine Ionospheric Corrections for Interplanetary Spacecraft Navigation," *Proc. of the Beacon Satellite Symposium, COSPAR/URSI*, Warsaw, Poland, 1980.
- [21] R. D. Kahn and J. S. Border, "Precise Interferometric Tracking of Spacecraft at Low Sun–Earth–Probe Angles," AIAA-88-0572, 26th Aerospace Sciences Meeting, Reno, Nevada, 1988.
- [22] R. Woo and J. W. Armstrong, "Spacecraft Radio Scattering Observations of the Power Spectrum of Electron Density Fluctuations in the Solar Wind," *J. Geophys. Res.*, vol. 84, no. A12, pp. 7288–7296, 1979.
- [23] J. S. Border, W. M. Folkner, R. D. Kahn, and K. S. Zukor, "Precise Tracking of the Magellan and Pioneer Venus Orbiters by Same-Beam Interferometry," AAS/AIAA Spaceflight Mechanics Meeting, Houston, Texas, 1991.
- [24] J. B. Minster and T. H. Jordan, "Present-Day Plate Motions," *J. Geophys. Res.*, vol. 83, pp. 5331–5354, 1978.
- [25] M. H. Finger and W. M. Folkner, "A Determination of the Radio-Planetary Frame Tie and the DSN Tracking Station Locations," AIAA-90-2905, AIAA/AAS Astrodynamics Conference, Portland, Oregon, 1990.
- [26] Y. Vigue, S. M. Lichten, G. Blewitt, M. B. Heflin, and R. P. Malla, "Precise Determination of Earth's Center of Mass Using Measurements From the Global Positioning System," *Geophysical Research Letters*, vol. 19, no. 14, pp. 1487–1490, 1992.
- [27] G. Bierman, *Factorization Methods for Discrete Sequential Estimation*, vol. 128, New York: Academic Press, 1977.
- [28] J. B. Thomas, *Interferometry Theory for the Block II Processor*, JPL Publication 87-29, Jet Propulsion Laboratory, Pasadena, California, October 1987.
- [29] B. D. Mulhall, V. J. Ondrasik, and K. L. Thuleen, *The Ionosphere*, JPL Technical Report 32-1499, Jet Propulsion Laboratory, Pasadena, California, pp. 45–67, 1970.

Appendix A

PHOBOS Observable Processing

I. Computing Observables

In a VLBI measurement, the signal from a celestial source is recorded simultaneously by two stations. The recorded signals from the two stations are processed to obtain the signal delay between them. This interferometric delay can be written as a sum of a geometric delay plus propagation media and instrumentation effects and clock synchronization error:

$$\tau = \frac{1}{c} \vec{B} \cdot \hat{S} + \tau_{media} + \tau_{instr} + \tau_{clock}$$

where \vec{B} is the baseline vector, c is the speed of light, and \hat{S} is the direction to the source. Measurement of the delay is thus a measurement of the component of \hat{S} along the direction of the baseline. The astrometric accuracy of a VLBI measurement depends on the ability to calibrate propagation media and instrumentation effects and synchronize clocks, and on the accuracy of the baseline and Earth orientation information used.

In a differential spacecraft–quasar VLBI measurement, the signals of a spacecraft and an angularly nearby reference quasar are recorded, and the interferometric delays for the two sources are differenced. The differenced delay can be written

$$\Delta\tau = \frac{1}{c} \vec{B} \cdot \Delta\hat{S} + \Delta\tau_{media} + \Delta\tau_{instr} + \Delta\tau_{clock}$$

where $|\Delta\hat{S}|$ is equal to the angle between the two sources. Errors in the calibration of media and instrumental effects and clock synchronization, as well as errors in the modeling of baseline and Earth orientation, largely cancel in the difference. The astrometric accuracy of this measurement of the angular separation between the two sources is thus much better than for a VLBI measurement of a single source. (In the frame-tie estimation in Section V, the delays are not explicitly differenced, but only the delay differences contribute to the frame-tie solution.) The closer together the sources, the better the common mode cancellation. Very rarely will a spacecraft be angularly close enough to an appropriate quasar for the two sources to fit into an antenna beamwidth. Therefore, typically the spacecraft and quasar are observed sequentially rather than simultaneously. The spacecraft and quasar signals are recorded alternately for intervals typically of a few minutes. Each recording interval is called a “scan.”

In a VLBI measurement, the spacecraft and quasar signals are recorded at multiple frequencies at each station. Differential phase between the stations is measured at each frequency. The interferometric delay is the slope of the differential phase with frequency, so the delay precision is roughly inversely proportional to the separation of the highest and lowest frequencies.

The recorded differential phase between the two stations for a spacecraft or quasar signal of frequency ω is

$$\phi(\omega, t) = \omega\tau(t) + \phi_I(\omega)$$

where $\tau(t)$ is the interferometric delay and $\phi_I(\omega)$ is instrumentation phase, which includes an arbitrary local oscillator startup phase, filter response, and any other nonlinear phase shifts that occur in the recording system. (The instrumentation phase is generally also a function of time, but extant phase calibration

tone data indicate that instrumentation phase temporal behavior was sufficiently stable over a pass to be neglected.) The interferometric delay also has a frequency dependence due to signal propagation through the ionospheric and solar plasma which has been suppressed here for simplicity.¹³ The instrumentation phase must be calibrated to get interferometric delays. Also, the cycle ambiguity of the signal phase must be resolved to get meaningful delays. The observation of the quasar enables the resolution of the spacecraft phase ambiguity.

II. Station Configuration

An example of the WCB VLBI channel configuration is shown in Fig. 8. Fourteen channels of 2-MHz bandwidth were distributed over the L-band receiver bandwidth of approximately 70 MHz. This arrangement of channels is not optimal for measuring quasar delays because these passes were originally intended as preparation for future acquisitions of VLBI data from the Phobos-2 lander. The configuration was designed to sample the bandpass of the receiver and to provide data to be used in verifying procedures for phase ambiguity resolution. Each of the 14 channels was sampled at the Nyquist rate, 4 Mbits/sec, for a total data rate of 56 Mbits/sec.

The figure also illustrates the recorded spectrum for the spacecraft. The spacecraft signal consisted of odd harmonics of a 7.4-MHz square-wave phase modulated onto the carrier. Because only the first harmonic and the -3rd harmonic provided an adequate SNR, the spanned bandwidth of the spacecraft signal was limited to 29.4 MHz. The SNR of the first harmonic was 23 dB-Hz; that of the -3rd harmonic was 12 dB-Hz.

Each spacecraft tone was typically recorded by a pair of channels offset in frequency. This provided redundancy in case of channel failure. Also, comparison of the tone phases from each channel in a pair yielded information on instrumentation phase.

It had been planned to use phase calibration tones to determine and remove instrumental errors in delay, but due to equipment failures, phase calibration tones were present at only a single station for each pass. Comparison of phase differences of single tones that appeared in a pair of channels showed no significant drift between channels.

III. Correlation

The WCB VLBI data were recorded on tape at a data rate of 56 Mbits/sec at each station. This large volume of data was processed to produce a VLBI group delay observable for each scan. The processing was done in two steps: First, the WCB tapes were correlated using model delays to compress the data to 2-sec integrations. Second, the compressed data were processed by postcorrelation software to produce group delays and ancillary data.

The WCB data were correlated using the JPL/Caltech Block II VLBI Processor [28], henceforth called "the correlator." Because the spacecraft signal consisted of discrete frequency sinusoidal "tones" and the quasar signal is broadband noise, the spacecraft and quasar scans were processed rather differently in the correlator, but the end results are similar in form.

For the spacecraft data, an a priori spacecraft trajectory was used to produce a spacecraft range model for each station, which was in turn used to produce a model for the phase of each tone at each station. The correlator correlates the recorded spacecraft signal with the phase model and integrates for 2-sec intervals. The correlator output is a time series of correlation amplitudes and residual phases

¹³ B. Iijima, "Block II Processor Delta-DOR Observable Extraction," in preparation, Jet Propulsion Laboratory, Pasadena, California, 1994.

(the difference between the spacecraft phase and phase model) for each tone in each channel at each station. Differencing the correlator output phase from the two stations gives residual phase for each tone at frequency ω_i :

$$\phi_r(\omega_i, t_j) = \omega_i \tau_r(t_j) + \phi_I(\omega_i) \quad (\text{A-1})$$

where t_j is the center of the j th 2-sec integration interval and the residual delay τ_r is the difference between the interferometric delay and the difference of the range models used to produce the phase models.

The correlator processes quasar scans by synchronizing the quasar signals in two stations' bit streams using model delays, then multiplying the two streams, correlating the resultant interferometric fringes with a differential phase model, and integrating for 2-sec intervals. This is done for each channel on each baseline. The correlator also forms a cross-power spectrum, effectively dividing each 2-MHz channel into 8 "bins" 0.25-MHz wide. The output of the correlator is a time series of 2-sec correlation amplitudes and residual phases for each bin. The residual phases for bin i can be written as

$$\phi_r(\omega_i, t_j) = \langle \omega_i \tau_r(t_j) + \phi_I(\omega_i) \rangle_{2 \text{ sec time avg}} = \omega_i \tau_r(t_j) + \phi_I(\omega_i) + \phi_{avg}(\omega_i) \quad (\text{A-2})$$

where ω_i is the RF centroid frequency of the bin, t_j is the center time of the 2-sec integration, and the residual delay, τ_r , is the interferometric delay minus the synchronization delay model. The correlator output phase is a complicated average of the quasar signal phase over the 0.25-MHz-wide bin, and is approximately equal to the quasar signal phase at the center of the bin. The phase ϕ_{avg} represents the small discrepancy between the average phase over the bin and the phase at the center of the bin. Because of the proximity of all the sources and the relative brevity of each pass, ϕ_{avg} is approximately the same in all quasar scans during a pass.

IV. Postcorrelation Processing

The PHOBOS frame-tie experiment was the first spacecraft-quasar differential VLBI measurement using the WCB data acquisition terminals and Block II processor. New postcorrelation software¹⁴ was developed to produce observables from the correlator output. The software compressed each scan of VLBI data to give eight numbers: a time tag and reference frequency, two VLBI observables (group delay and phase delay rate) and the system noise uncertainty on each, the instrumentation phase calibration error on the group delay, and an effective frequency for ionospheric calibration. This section briefly describes how the postcorrelation software produced these eight numbers from the array of phases in Eqs. (A-1) and (A-2).

The group delays for each scan were computed from the phases in Eqs. (A-1) and (A-2) by a three-step process: First, the time series at each frequency was compressed to a single point for each frequency. Second, the instrumentation phase was calibrated out. Third, the residual group delay was computed, and the total group delay observable was formed by restoring the delay models used in the correlator. Phase delay rates were also produced in this process. These three steps are described in some detail below.

Step 1: Phase Compression. The correlator station-differenced output residual phase can be rewritten as

$$\phi_r(\omega_i, t_j) = \Phi(\omega_i) + \omega_i f(t_j)$$

¹⁴ Ibid.

where

$$\Phi(\omega_i) = \omega_i \tau_r(t_0) + \phi_I(\omega_i) + [\phi_{avg}(\omega_i)]$$

$$f(t_j) = \tau_r(t_j) - \tau_r(t_0)$$

The quantities $\Phi(\omega_i)$, one for each frequency, depend only on frequency, and $f(t_j)$ contains all temporal delay variations. The brackets enclosing ϕ_{avg} indicate that it is zero for spacecraft data. One can estimate Φ and f by doing a least-squares fit to the phases $\phi_r(\omega_i, t_j)$ weighted by the system noise uncertainty (which is approximately inversely proportional to the correlation amplitude). In order to make the least-squares fit well defined, it was required that a linear fit to $f(t_j)$ be zero at the scan reference time t_0 . This ensured that the slope of Φ would be the group delay at the reference time. The scan reference time tag t_0 , and a reference frequency ω_0 (to be used below), were computed by doing a weighted average of time and frequency over all the data for the scan.

Step 2: Instrumentation Phase Calibration. Since phase calibration tone data were not available from all stations, a technique was used to estimate the nonlinear component of the instrumental phase directly from the data. This technique is called manual phase calibration (MPC) and works as follows: A strong quasar scan is chosen to be the reference scan. The compressed phase of this reference scan is defined to be the manual phase calibration phase:

$$\Phi_{MPC}(\omega_i) = \omega_i \tau_{MPC} + \phi_I(\omega_i) + \phi_{avg}(\omega_i)$$

where τ_{MPC} is the residual delay of the reference scan.

Instrumentation phase calibration is performed on quasar scans by subtracting Φ_{MPC} from Φ to get the calibrated phase Φ_{cal} :

$$\Phi_{cal}(\omega_i) = \Phi(\omega_i) - \Phi_{MPC}(\omega_i) = \omega_i(\tau_r - \tau_{MPC})$$

To calibrate a spacecraft scan, Φ_{MPC} is interpolated to the spacecraft frequency and then subtracted from the spacecraft phase. The result is

$$\Phi_{cal}(\omega_i) = \Phi(\omega_i) - \Phi_{MPC}^{interpolated}(\omega_i) = \omega_i(\tau_r - \tau_{MPC}) + \varepsilon_{MPC}$$

where the spacecraft phase calibration error ε_{MPC} contains ϕ_{avg} interpolated to the spacecraft frequency and the discrepancy between ϕ_I and the instrumental phase interpolated between quasar bin frequencies. Because the spacecraft frequencies received at the station vary less than 1 kHz over each pass and the phase response is expected to be fairly smooth on that scale, ε_{MPC} is common to all spacecraft scans in a pass.

Analysis of a limited data set measuring the phase responses of several of the WCB video converters suggests that ε_{MPC} will be around 4 millicycles (mcyc) for most WCB video converters. However, a few of the WCB video converters deviate 10 mcyc or more from the nominal WCB nonlinear phase response, and much larger calibration errors are expected in those converters. A way of determining the size of ε_{MPC} directly from the data is discussed in Step 3.

Step 3: Computing the Group Delay Observable. A priori delay information was used to resolve the relative phase cycle ambiguity of $\Phi_{cal}(\omega_i)$ between frequencies, enabling the residual group delay,

which is the slope of Φ_{cal} , to be computed. Doing a linear least-squares fit of a line to Φ_{cal} weighted by the sum of the system noise uncertainty (including ϵ_{MPC} for the spacecraft) gives the measured (residual) group delay

$$\tau_r^{meas} = \tau_r(t_0) - \tau_{MPC} + [\delta_{MPC}]$$

where δ_{MPC} is the contribution of ϵ_{MPC} to the slope. The brackets around δ_{MPC} indicate that it is zero for quasar scans. There is a bias $-\tau_{MPC}$ in all delays. In the frame-tie parameter estimation, this will be absorbed into the clock bias parameter and have no effect on the frame-tie parameters. There is, in addition, an extra bias δ_{MPC} common to all spacecraft delays, but absent from the quasar delays, that will contribute an error to the frame tie.

An estimate of ϵ_{MPC} can be obtained by observing the scatter in the data in excess of the system noise uncertainty in the linear fit to Φ_{cal} for the spacecraft scans. An example of the residuals to the linear fit is shown in Fig. A-1. The data shown are from the March 25 pass, and the error bars displayed are system noise uncertainty. The magnitude of the instrumentation calibration error on each spacecraft channel phase was determined empirically by forcing the reduced chi-squared of the linear fit to be 1. This gave an instrumentation calibration error of about 8 mcyc in each channel phase on February 17 and 3 mcyc in each channel phase on March 25, roughly in agreement with the expected size given at the end of the discussion of Step 2.

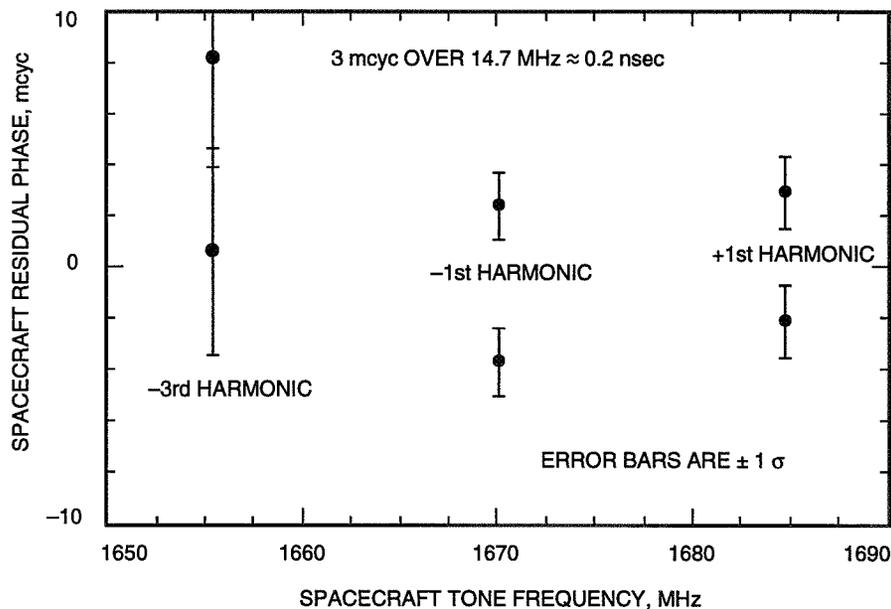


Fig. A-1. Phobos-2 instrumental phase error for Goldstone–Canberra baseline.

The system noise and δ_{MPC} both contribute to the spacecraft delay observable uncertainty, and their contributions to the uncertainty are separately computed and passed to the frame-tie estimation process. The contribution of δ_{MPC} to the delay uncertainty is a measure of the instrumental phase bias between the quasar and spacecraft scans. For the quasar scans, only the system noise uncertainty of the quasar delays is passed on to the frame-tie estimation process.

The result of the linear fit above is the residual group delay. The total group delay observable is obtained by evaluating the delay models used in correlation at the scan reference time and adding the delay model to the delay residual to get the total group delay observable.

The other observable extracted from the data, the phase delay rate, was computed by taking the slope of $f(t_j)$ from Step 1 and dividing by the reference frequency ω_0 to get the residual phase delay rate. The correlator delay rate model was then restored to get the total phase delay rate observable.

Appendix B

Faraday Calibration Error Model

Because of the relatively low frequency of the PHOBOS frame-tie observations (L-band is approximately 1.7 GHz), dispersive delays from charged particles represent a significant source of measurement error. Charged particles are encountered in the interplanetary solar plasma as well as in the Earth's ionosphere. It is the charged particle delay, differenced between the spacecraft and quasars, that corrupts the frame-tie determination. The analysis of Section IV indicates that the differential solar plasma delay should be at or below the 0.1-nsec level. Ionospheric errors are expected to be significantly larger. The total zenith daytime ionospheric delay can be 25 nsec or more at one station, with low-elevation delays approaching 100 nsec.

The general characteristics of the signal delays due to the charged particles of the ionosphere are well known. The DSN routinely monitors the ionosphere's charged-particle content along a fixed ray path by acquiring Faraday rotation data from a geosynchronous satellite at each site [20,29]. Although the calibrations from Faraday data have been applied to the PHOBOS delays, the calibrations are not perfect and the effect of these errors on the frame-tie parameter estimates must be included in assessing estimate accuracy. The use of "consider" analysis to get the sensitivity of the frame-tie solution to a scaling of the Faraday calibrations was described in Section V. However, it is also important to assess the point-to-point statistics of the ionospheric delay errors in weighting the data. Because the parameters of a linear clock model on each baseline are adjusted in the solution process, the calibration error model must specify the error statistics for differential delay after a linear drift has been removed.

A good data set that illustrates both spatial and temporal ionospheric delay variations was obtained during an experiment to detect the relativistic deflection of an electromagnetic wave by the mass of Jupiter [14]. The "planetary deflection" data consist of two VLBI passes, 3 to 4 hr in duration, from March and April of 1988. Delays were obtained for a number of quasars at a range of separation angles at both S- and X-band (2.3 and 8.4 GHz). Two sources separated by 3.7 deg were observed repeatedly. Because the ionospheric delay scales as f^{-2} , where f is signal frequency, a linear combination of the dual-frequency delays provides a measure of the station-differenced ionospheric delay. These are given in terms of nanoseconds at L-band in Fig. B-1. Note that the resulting ionospheric delay for each day contains an overall bias due to uncalibrated delay biases in the S- and X-band instrumentation. Nonetheless, temporal changes in the ionosphere during each day, as well as differences in the ionosphere between sources, are well determined from these data. Since the minimum elevation for the PHOBOS observations was 22 deg, only data above a 20-deg elevation are shown in the figure.

The major characteristics of the ionosphere observed in this data set are the increase in delay of 4-5 nsec over the pass and the offset of about 2 nsec between sources. The delays show a fairly linear variation. For the analysis of the PHOBOS VLBI data, a linear clock model is removed; as a result, any linear variation common to the spacecraft and quasar will be absorbed by the clock solution. The remaining point-to-point scatter, which is due to small-scale ionospheric irregularities, will manifest itself as an additional noise source for the observed delays. In addition, any bias between the spacecraft and quasar ionospheric delays will directly corrupt the frame-tie estimate.

In Fig. B-2, the same planetary deflection data are shown after Faraday calibrations have been applied. The calibrations have reduced the delay drift over each pass by a factor of 2, but drifts of 0.6–1.5 nsec per hour remain. The bias in differential delay between sources has decreased, to 0.2 nsec for the first pass and 0.7 nsec for the second. The latter bias roughly corresponds to an error in relative position of 30 nrad.

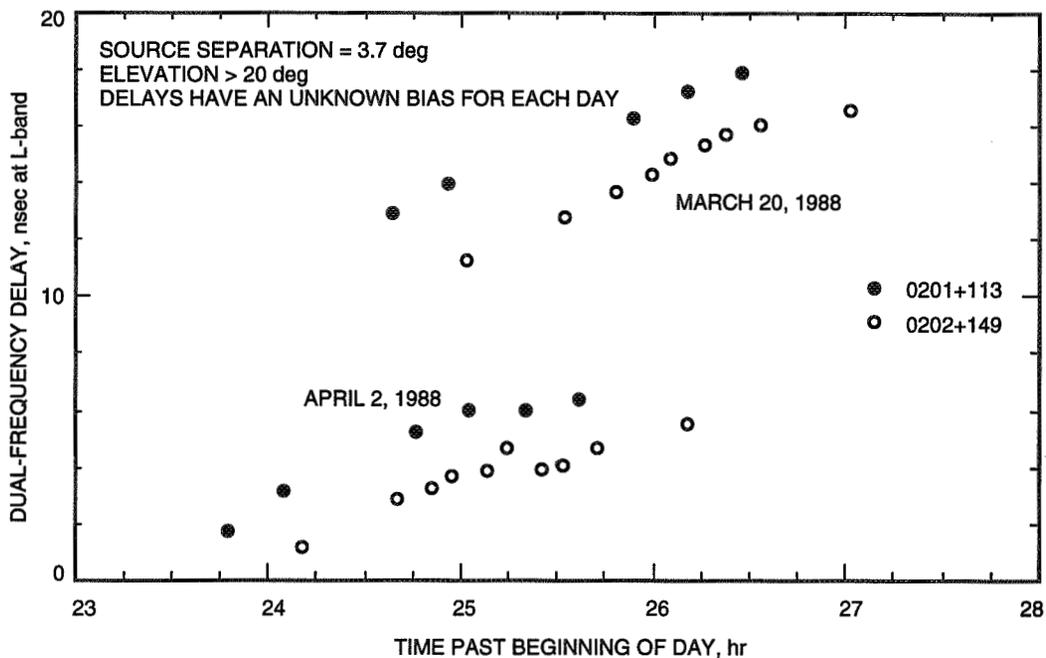


Fig. B-1. Ionosphere delay from dual-frequency VLBI.

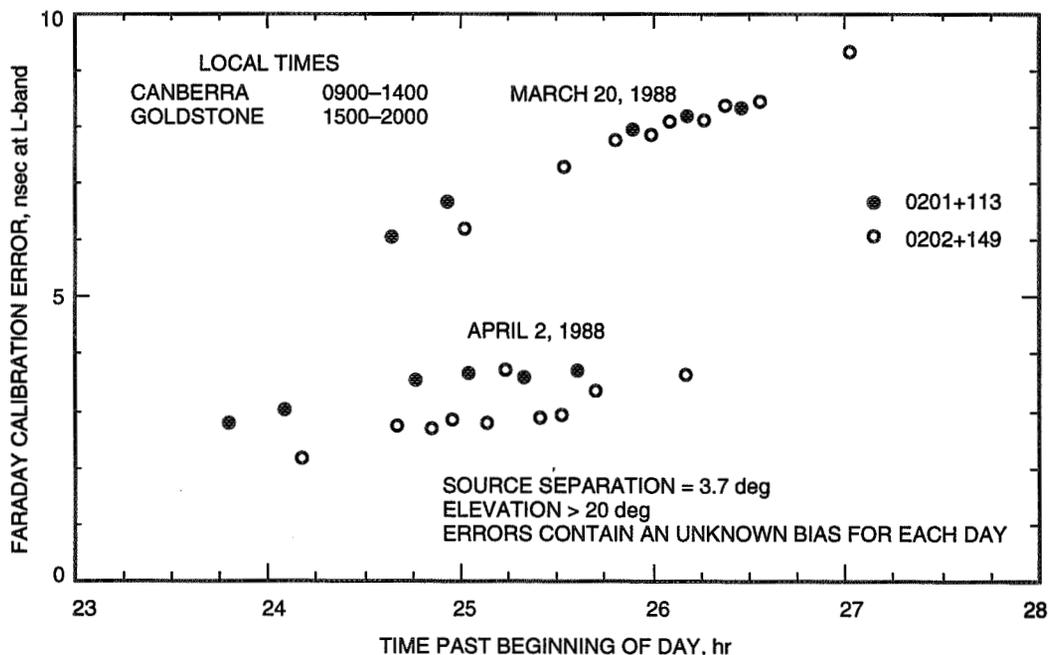


Fig. B-2. Error in Faraday calibration of ionosphere delay.

Given that nanosecond-level differential ionospheric delay biases can remain after Faraday calibration for angular separations of 3–4 deg, it is reasonable to include delays only for the quasar closest to the Phobos-2 spacecraft on each day in the nominal frame-tie solution. This corresponds to spacecraft-quasar angular separations of 1.0 deg for the February 17 pass and 0.4 deg for the March 25 pass. The angular separations from the Phobos-2 spacecraft of other quasars observed in these two experiments ranged from 3.2–5.6 deg, and hence were expected to have significantly larger ionospheric delay biases relative to the spacecraft scans.

Another benefit of using only the closest quasar is that it simplifies the modeling of the point-to-point error statistics. The spatial separation at the peak ionospheric altitude of 350 km for a pair of ray paths separated by 1 deg will be much smaller than the temporal displacement of the ray path due to Earth rotation during the 5-min time interval between adjacent scans. Thus, to a good approximation, we can represent the point-to-point scatter of the spacecraft and primary quasar scans based on statistics derived for ionospheric fluctuations along a single line of sight.

To quantify these fluctuation statistics, several days of dual-frequency (S-/X-band) observations of the Magellan spacecraft in orbit about Venus have been examined. The data were collected simultaneously at California and Australia during February 1991. Phase observations of the S-band and X-band carrier signals were made continuously at 2-sec intervals for a period of several hours. The spacecraft elevations ranged from 20–50 deg at the two stations. The passes occurred at midafternoon in California and midmorning in Australia.

Dual-frequency phases from each station were linearly combined to track variations in electron (el) density integrated along the line of sight (TEC_{los}) according to

$$\phi_S(t) - \frac{3}{11}\phi_X(t) = -0.54 \times TEC_{los}(t) + bias + solar\ plasma$$

where phase is in cycles and TEC_{los} is in units of 10^{16} el/m² (one TEC is about a 0.5-nsec delay at L-band). The single-station data show significant solar plasma fluctuations. Differencing this quantity between the California and Australia stations greatly reduced the solar plasma error, leaving a precise measure of variations in the station-differenced line-of-sight electron density ΔTEC_{los} . Figure B-3 shows an example of the variability of ΔTEC_{los} over about a 1-hr time span, after removal of a straight line. The ΔTEC_{los} is observed to vary by about $\pm 1 \times 10^{16}$ el/m² over the 1-hr span. From these data, the structure function (the expected mean-square change in ΔTEC_{los} over an interval T) was calculated as

$$D_{TEC}(T) \equiv \left\langle [\Delta TEC_{los}(t+T) - \Delta TEC_{los}(t)]^2 \right\rangle$$

Figure B-4 depicts four structure functions derived from Magellan S-/X-band data collected on two different days of the year (DOY). All the structure functions exhibit power-law slopes close to 1, characteristic of a random walk process. The straight line in the figure is the result of a fit of the four structure functions to a random walk model of the form $D_{TEC}(T) = C^2 T$. The random walk amplitude C is found to be

$$C = 1.80 \times 10^{14} \frac{\text{el}}{\text{m}^2 \sqrt{\text{sec}}}$$

Converting to the delay variation at the 1.7-GHz PHOBOS frequency yields $C = 8.3 \times 10^{-3}$ nsec/ $\sqrt{\text{sec}}$, or about a 0.26-nsec variation in 1000 sec. Although the examined data set is limited, it nonetheless serves

to guide the choice of random walk process noise and sets the a priori value of the random walk amplitude. The actual amplitude was adjusted in the course of parameter estimation to yield a satisfactory value of χ^2 . For the reference solution, the adjusted amplitude was equivalent to a 1- σ error of 0.5 nsec in 1000 sec.

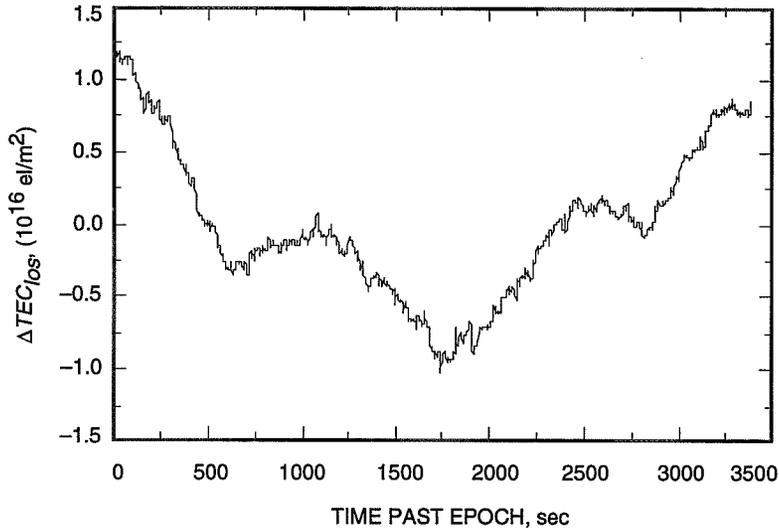


Fig. B-3. Line-of-sight station-differenced integrated electron content variations derived from S-X-band carrier phase observations of the Magellan spacecraft after removal of a linear fit (DOY 048, 1991, Goldstone-Canberra).

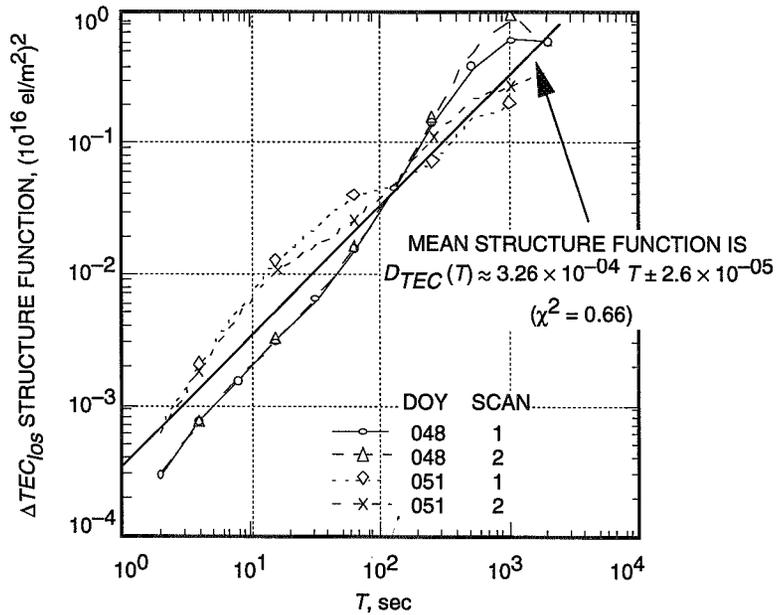


Fig. B-4. The ΔTEC_{los} structure functions derived from dual-frequency observations of the Magellan spacecraft (Goldstone-Canberra, 1991). A fit of the four structure functions to a random walk model yields a random walk amplitude C of 1.8×10^{14} el/m²/sec (= 0.26 nsec at 1000 sec at L-band).

In summary, the nominal PHOBOS frame-tie result was based on quasar delays only for the sources closest to the Phobos-2 spacecraft on each day because of significant uncertainty in the statistical characterization of postcalibration ionospheric delay errors for angular separations of a few degrees or more. Examination of several time series of station-differenced dual-frequency spacecraft data yielded a random walk model for the point-to-point ionospheric delay fluctuations, which was used in the final frame-tie estimation. The sensitivity of the frame-tie solution to an overall scaling error in the Faraday calibrations was also determined. Several alternative solutions were also computed, including two using all of the observed quasars (one with random walk ionosphere noise and one without). The differences from the reference solution are all within the stated accuracy. In no case was the ionosphere the dominant error in the final frame-tie estimate.

110869

33271

p. 22

358422

Closed-Loop Carrier Phase Synchronization Techniques Motivated by Likelihood Functions

H. Tsou and S. Hinedi

Communications Systems Research Section

M. Simon

Telecommunications Systems Section

This article reexamines the notion of closed-loop carrier phase synchronization motivated by the theory of maximum a posteriori phase estimation with emphasis on the development of new structures based on both maximum-likelihood and average-likelihood functions. The criterion of performance used for comparison of all the closed-loop structures discussed is the mean-squared phase error for a fixed-loop bandwidth.

I. Introduction

It is well known [1] that estimation of an unknown parameter based on a likelihood function approach is optimum in the sense of maximizing the a posteriori probability of the parameter given the observation. For the case where the unknown parameter is the random phase of a carrier received in a background of additive white Gaussian noise (AWGN), optimum open-loop structures have been derived for implementing the resulting phase estimate [2,3]. Herein, these structures are referred to as "open-loop carrier phase estimators."

When the carrier is data modulated, the conditional probability density function (pdf) of the observation—given the carrier phase—depends on the data sequence that exists during the interval of observation for the received signal. Hence, before maximizing this function with respect to the carrier phase, one has to choose how to eliminate its dependence on the unknown data sequence. If one is interested in determining only the optimum carrier phase estimate, the appropriate choice is to *average* the conditional pdf over the unknown data sequence. We shall refer to the phase estimate obtained by this process as the "average-likelihood" (AL) estimate. If, however, one is interested in joint phase estimation and data detection, the appropriate choice is to first *maximize* the conditional pdf with respect to the data sequence (resulting in the most probable sequence), and then to maximize it with respect to the carrier phase.¹ We shall refer to the phase estimate obtained by this process as the "maximum-likelihood"

¹In principle, the order of maximization operations could be reversed.

(ML) estimate.² It has often been conjectured, although never proven, that from the standpoint of phase estimation alone, the ML phase estimate is suboptimum to the AL estimate. Because of this, what is typically done in practice is to derive the AL carrier phase estimate and then use this estimate as the phase of a demodulation reference signal for performing bit-by-bit data detection. However, it should be understood that, from the standpoint of joint estimation of data and carrier phase, this sequential operation of *first* deriving the carrier phase estimate in the absence of any knowledge of the data (the AL approach) and *then* detecting the ensuing data using the phase estimate so derived is, in general, suboptimum.

Aside from the optimality of the AL and ML approaches to open-loop estimation of carrier phase, likelihood functions have also been used as *motivation* for closed-loop carrier phase synchronization. Emphasis is placed on the word “motivation” since, indeed, there is no guarantee that the resulting closed-loop schemes are optimal; nor can one guarantee that those schemes motivated by the AL approach will outperform those motivated by the ML approach (although typically this turns out to be the case, as we shall show.) Nonetheless, as we shall see, closed-loop carrier phase estimation schemes motivated by likelihood functions do indeed yield good tracking performance (as measured by the mean-squared value of the loop phase error). In fact, under suitable assumptions, many of them are synonymous with well-known carrier tracking loops, e.g., the I-Q Costas loop and the I-Q decision feedback or polarity-type Costas loop [4,5] that have been around for many decades.

It is the intent of this article to explore in more detail the structure and performance of closed-loop carrier phase synchronization loops motivated by likelihood functions, i.e., those in which the derivative (or some monotonic function of the derivative) of the conditional pdf of the observation given the carrier phase is used as an *error signal* in a closed-loop phase estimation scheme. Herein, for the purpose of abbreviated notation, we shall refer to such loops as AL and ML closed loops depending on the particular likelihood function used to define the error signal.

It is important at this point to mention that the notion of closed loops based on likelihood functions according to the above definition is indeed not new, and one should not attribute its originality to the authors of this article. Rather, the purpose of this article is to expand upon this notion and present some new loops motivated by likelihood functions along with their tracking performances. As such, we are not reinventing the wheel but, rather, adding some more spokes to it. Our specific motivation for reexamining this problem comes from a deep-space communication application involving the Galileo S-band (2.3 GHz) mission, which employs low-rate ($r = 1/4$) concatenated Reed–Solomon/convolutionally encoded binary phase-shift keying (BPSK) [6]. Because of a malfunctioning high-gain X-band (8.4 GHz) antenna, the mission must rely on a low-gain S-band antenna (and, thus, much reduced link margin) for data transmission back to Earth. At Jupiter encounter, the symbol energy-to-noise spectral density ratio, E_s/N_0 , could be as low as -11 dB. One technique for improving this situation is to use antenna array combining [7] wherein the signals from multiple antennas, either collocated or at distant geographical locations, are combined to build E_s/N_0 . Even then the equivalent E_s/N_0 could still be as low as -5 dB. Thus, in our application, there is a serious need to find as efficient a carrier tracking loop as possible in the sense of producing minimum phase jitter at very low E_s/N_0 . In the more general context, it is important to point out that, in coded systems, the carrier-loop performance is dependent on the symbol energy-to-noise ratio E_s/N_0 rather than the bit energy-to-noise ratio E_b/N_0 and, thus, becomes critical when E_s/N_0 becomes small, despite the fact that E_b/N_0 might be large. In uncoded systems where $E_s/N_0 = E_b/N_0$ and is large, the search for a more efficient carrier tracking loop is somewhat academic since the known configurations perform quite well and are virtually identical to one another.

² In the strictest of parlance, both the AL and the ML phase estimates are maximum-likelihood estimates since the term “maximum-likelihood estimation” is typically reserved for estimating a purely unknown (uniformly distributed) random parameter. However, to allow for distinguishing between the two different ways in which the data sequence is handled, we shall use the above terminology.

The hierarchical structure of the problem and also the way in which it is addressed in this article is illustrated by the tree diagram of Fig. 1. We have already discussed the first level of the overall dichotomy in terms of the ML and AL approaches. This level of the chart as well as those below it will take on more meaning as soon as we develop a mathematical formulation of the problem in Section III.

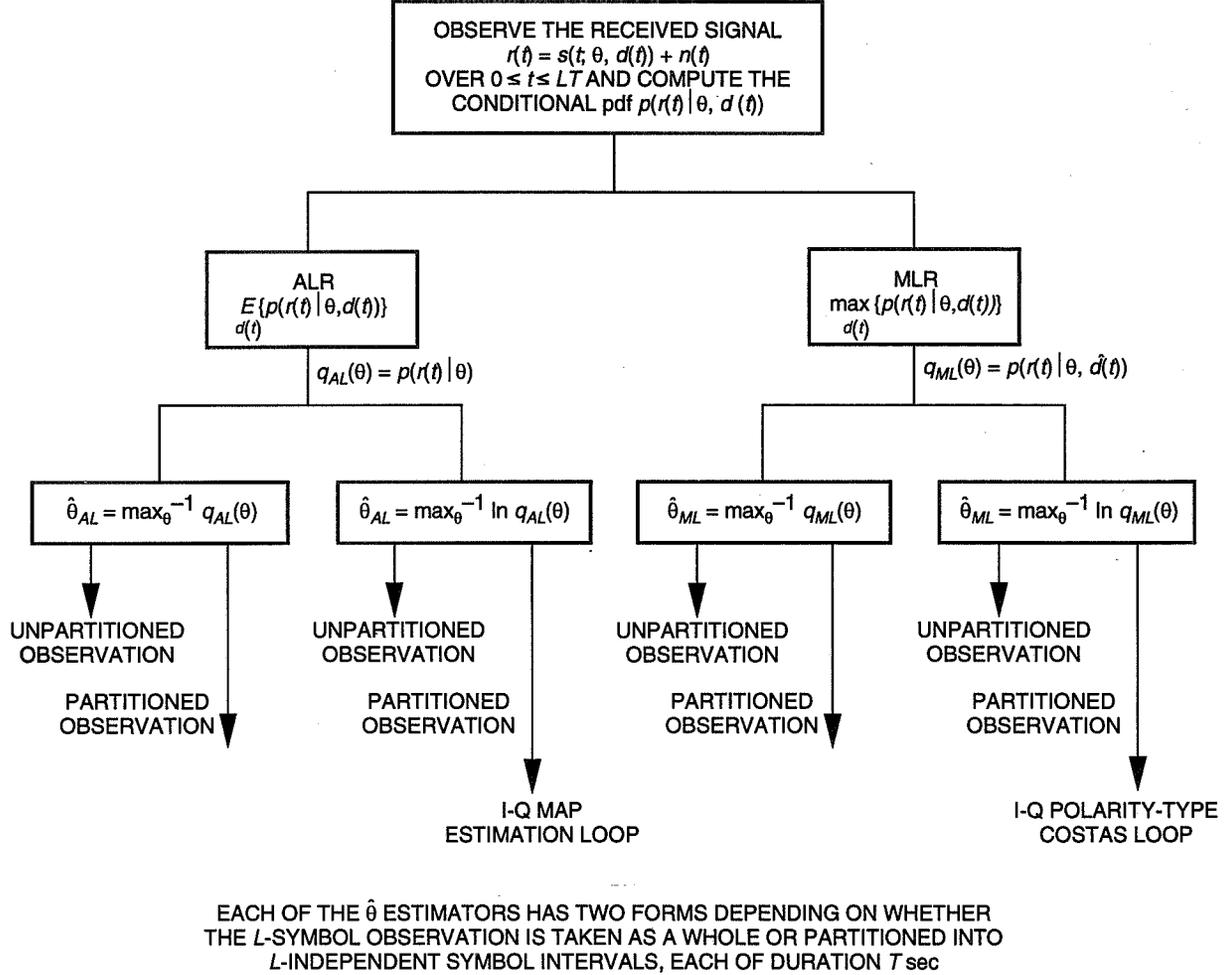


Fig. 1. A hierarchical structure of the open/closed-loop carrier phase estimation problem for data-modulated signals.

II. System Model

Consider a system that transmits BPSK³ modulation over an AWGN channel. As such, the received signal takes the form

$$r(t) = \sqrt{2S}d(t) \sin(\omega_c t + \theta) + n(t) = s(t; \theta, d(t)) + n(t) \quad (1)$$

where S denotes the received power, ω_c is the carrier frequency in rad/sec, θ is the unknown phase assumed to be uniformly distributed in the interval $(-\pi, \pi)$, $n(t)$ is an AWGN with single-sided power

³ We restrict ourselves to the case of binary modulation. By a straightforward extension of the procedures discussed, the results can easily be extended to M -ary modulation.

spectral density N_0 W/Hz, and $d(t)$ is a binary-valued (± 1) random pulse train defined by the rate $1/T$ binary data sequence $\{d_i\}$ and the rectangular pulse shape, $p(t)$, as

$$d(t) = \sum_{i=-\infty}^{\infty} d_i p(t - iT), \quad p(t) = \begin{cases} 1; & 0 \leq t \leq T \\ 0; & \text{otherwise} \end{cases} \quad (2)$$

For an observation interval of L bits [we assume without loss of generality the interval $(0, LT)$], the conditional pdf of the received signal (observation) given the unknown phase and the particular data sequence, d_i , transmitted in that interval is easily shown to be

$$p(r(t)|\theta, d_i(t)) = C_0 \exp\left(\frac{2\sqrt{2S}}{N_0} \int_0^{LT} r(t) d_i(t) \sin(\omega_c t + \theta) dt\right) \triangleq q_i(\theta) \quad (3)$$

where $d_i(t)$ is the transmitted waveform corresponding to the transmitted sequence in accordance with Eq. (2) and C_0 is a constant of proportionality. To proceed further, we must now choose between AL and ML approaches.

III. Closed Loops Motivated by the AL Approach

A. Structures

Suppose that we are interested in estimating only the carrier phase, θ . Then, as previously mentioned, the appropriate approach is to average $p(r(t)|\theta, d_i(t))$ over all possible (2^L) and equally likely data sequences yielding the conditional pdf $p(r(t)|\theta) \triangleq q_{AL}(\theta)$. One AL open-loop phase estimate (herein referred to as "AL open-loop estimator no. 1") is obtained by finding the value of θ that maximizes $q_{AL}(\theta)$, namely (see Fig. 1: $\hat{\theta}_{AL} \triangleq \max_{\theta}^{-1} q_{AL}(\theta)$, unpartitioned observation)

$$\hat{\theta}_{AL1} \triangleq \max_{\theta}^{-1} \sum_{i=1}^{2^L} \exp\left(\frac{2\sqrt{2S}}{N_0} \int_0^{LT} r(t) d_i \sin(\omega_c t + \theta) dt\right) \quad (4)$$

where the inverse maximum notation " $\max^{-1} f(\theta)$ " denotes the value of θ that maximizes $f(\theta)$. Alternately, breaking up the integration over the entire observation into a sum of integrals on each bit interval and recognizing that the data bits are independent, identically distributed (iid) binary random variables, then $p(r(t)|\theta)$ can be expressed as a product of hyperbolic cosine functions. A second AL open-loop phase estimate (herein referred to as "AL open-loop estimator no. 2") is obtained by finding the value of θ that maximizes this product form of $q_{AL}(\theta)$, which corresponds to partitioning the observation into its individual bit intervals. The result is (see Fig. 1: $\hat{\theta}_{AL} \triangleq \max_{\theta}^{-1} q_{AL}(\theta)$, partitioned observation)

$$\hat{\theta}_{AL2} \triangleq \max_{\theta}^{-1} \prod_{k=0}^{L-1} \cosh\left(\frac{2\sqrt{2S}}{N_0} \int_{kT}^{(k+1)T} r(t) \sin(\omega_c t + \theta) dt\right) \quad (5)$$

It is important to emphasize here (and we shall repeat this emphasis later on in the closed-loop discussion) that *partitioning or not partitioning the observation interval has no effect on the value of the optimum estimator nor on its performance*. That is, optimum open-loop $\hat{\theta}_{AL1}$ and $\hat{\theta}_{AL2}$ are mathematically identical. The difference between the two lies solely in their implementation and likewise the difference in the closed-loop implementations motivated by these estimates, as we shall see shortly.

Finally, one could obtain an AL open-loop estimator by maximizing *any* monotonic function of $q_{AL}(\theta)$, for example $\ln q_{AL}(\theta)$. The reason for choosing the natural logarithm as the monotonic function is to simplify the mathematics, i.e., to convert the L -fold product in Eq. (5) to an L -fold sum. Thus, the third AL open-loop phase estimate (herein referred to as "AL open-loop estimator no. 3") is obtained by finding the value of θ that maximizes $\ln q_{AL}(\theta)$ with $q_{AL}(\theta)$ in its partitioned form. The result is (see Fig. 1: $\hat{\theta}_{AL} \triangleq \max_{\theta}^{-1} \ln q_{AL}(\theta)$, partitioned observation)

$$\hat{\theta}_{AL_3} \triangleq \max_{\theta}^{-1} \sum_{k=0}^{L-1} \ln \cosh \left(\frac{2\sqrt{2S}}{N_0} \int_{kT}^{(k+1)T} r(t) \sin(\omega_c t + \theta) dt \right) \quad (6)$$

Block diagram implementations of AL open-loop estimator no. 1 [Eq. (4)] and AL open-loop estimator no. 3 [Eq. (6)] are illustrated in Fig. 2, no. 3 being the form most commonly found in discussions of open-loop maximum a posteriori (MAP) carrier phase estimation. In drawing these implementations, we have quantized the unknown phase into Q values, and thus the maximization over the continuous phase parameter θ in Eqs. (4) and (6) is approximated by maximization over a Q -quantized version of this parameter.

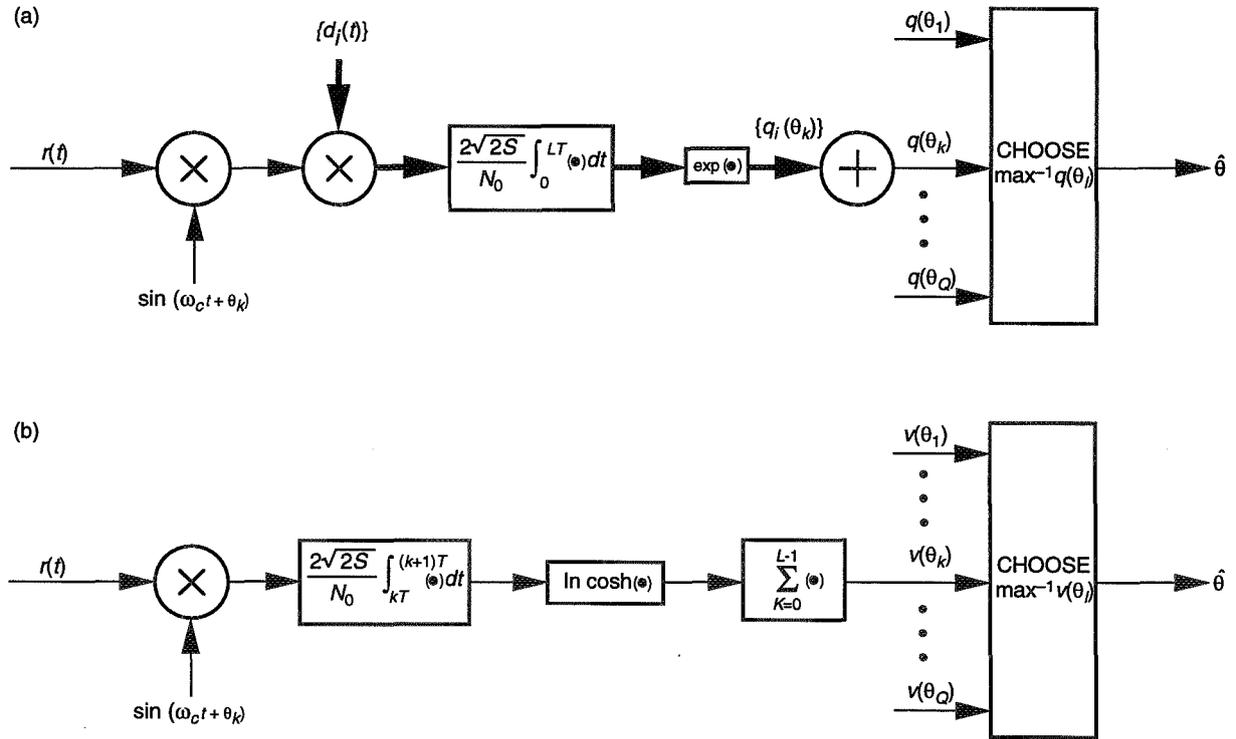


Fig. 2. Implementation of two AL open-loop phase estimators: (a) AL open-loop estimator no. 1—observation unpartitioned (quantized parallel implementation) and (b) AL open-loop estimator no. 3—observation partitioned (quantized parallel implementation).

Conceptually, a fourth optimum AL open-loop estimator, $\hat{\theta}_{AL_4}$, could be obtained by maximizing $\ln q_{AL}(\theta)$ with $q_{AL}(\theta)$ in its unpartitioned form. However, in view of the above discussion, $\hat{\theta}_{AL_3}$ and $\hat{\theta}_{AL_4}$ would be mathematically identical and, since $\hat{\theta}_{AL_4}$ appears to have no implementation advantage, we do not pursue it here.

Closed-loop phase synchronization structures⁴ based on the four AL open-loop estimators are obtained by choosing as error signals, e , the functions respectively given by $dq_{AL}(\theta)/d\theta$ and $d \ln q_{AL}(\theta)/d\theta$ where $q_{AL}(\theta)$ and $\ln q_{AL}(\theta)$ each takes on its unpartitioned or partitioned form. For simplicity of notation, we shall refer to these four closed-loop structures as AL closed-loop nos. 1, 2, 3 and 4. The implementations corresponding to AL closed-loop no. 1 and AL closed-loop no. 3 (the two simplest implementations of the four) are illustrated in Figs. 3(a) and (b), the latter being what is commonly called an "I-Q MAP estimation loop" [8,9]. The special cases of Fig. 3(b), wherein the hyperbolic tangent nonlinearity is approximated by linear and hard limiter devices, corresponding respectively to low and high signal-to-noise ratio (SNR) conditions, are commonly called the "I-Q Costas loop" [4] and "I-Q polarity-type Costas loop" [5].

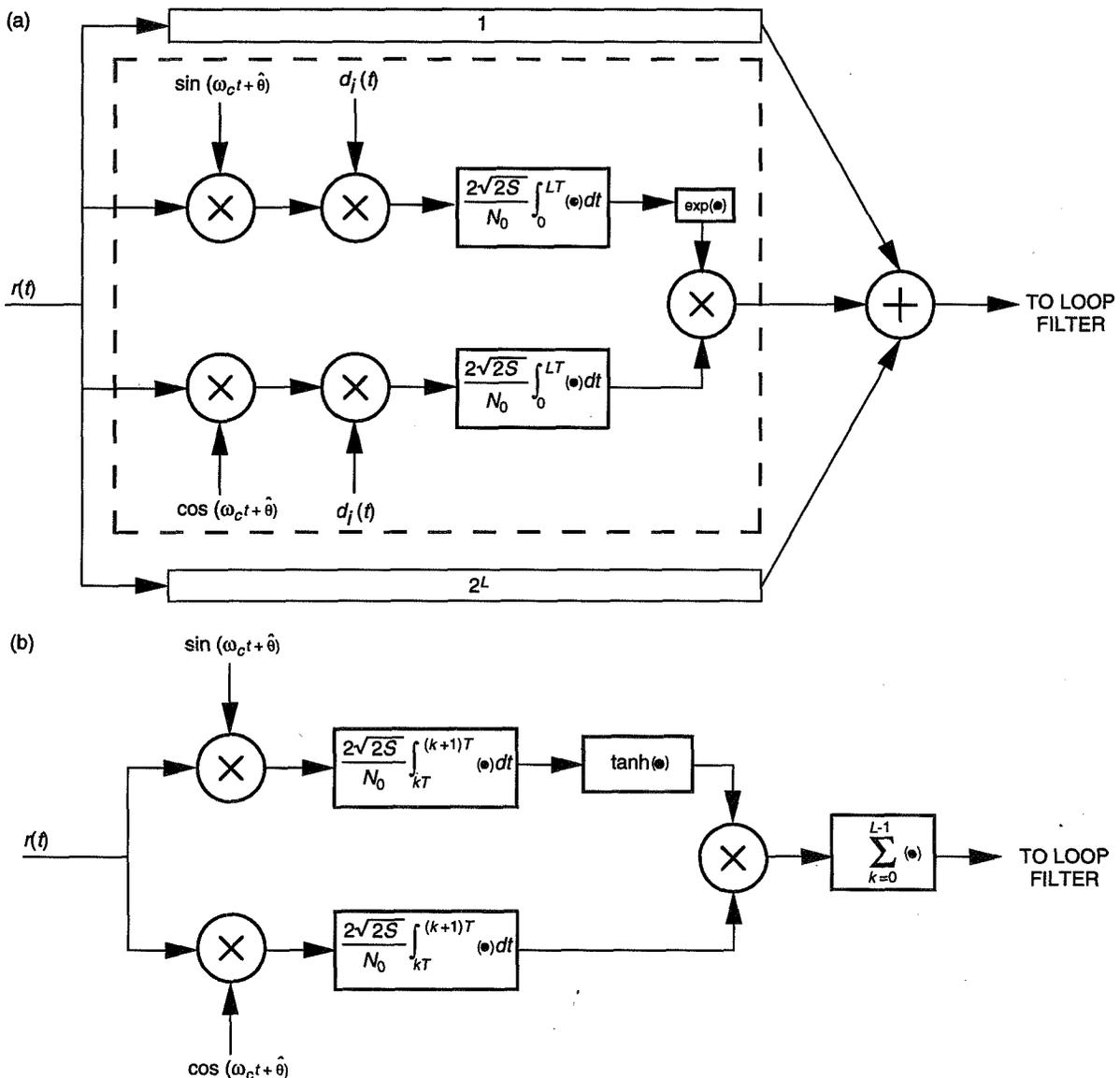


Fig. 3. Implementation of two AL closed loops: (a) AL closed loop no. 1—observation unpartitioned and (b) AL closed loop no. 3—observation partitioned.

⁴For ease of illustration, we show only the portion of the closed loop that generates the loop error signal, which in the actual implementation becomes the input to the loop filter.

Before proceeding, it is important to reemphasize that because of the monotonicity of the logarithm function, the AL open-loop phase estimates $\hat{\theta}_{AL3}$ and $\hat{\theta}_{AL4}$ are mathematically identical to $\hat{\theta}_{AL1}$ and $\hat{\theta}_{AL2}$ and thus yield identical performance. However, the equivalent statement is *not necessarily* true when considering the performances of the closed loops motivated by these four different AL formulations. More specifically, the closed loops motivated by $\hat{\theta}_{AL3}$ and $\hat{\theta}_{AL4}$ do not necessarily yield the same performance as those motivated by $\hat{\theta}_{AL1}$ and $\hat{\theta}_{AL2}$. The reason for this stems from the fact that the closed-loop performance (when properly normalized) is proportional to the derivative of $q_{AL}(\theta)$ (or $\ln q_{AL}(\theta)$ as appropriate) in the neighborhood of its maximum, which in general is different for $q_{AL}(\theta)$ and $\ln q_{AL}(\theta)$. However, we hasten to add that since partitioning does not change the functions $q_{AL}(\theta)$ or $\ln q_{AL}(\theta)$ themselves, the closed loops derived from either the partitioned or unpartitioned forms of the likelihood (or log likelihood) function should yield identical performance, i.e., AL closed-loop no. 1 and AL closed-loop no. 2 will have identical performance, as will AL closed-loop no. 3 and AL closed-loop no. 4.

B. Performance

In assessing the performance of one closed-loop scheme versus another, one must be careful to normalize the loop parameters to allow a fair basis of comparison. In this article, the comparison will be made on the basis of mean-squared phase error, σ_ϕ^2 , for a fixed-loop bandwidth, B_L .⁵ This is the typical measure of performance used to describe a closed-loop phase synchronization structure when it is operating in its tracking mode.

An analysis of the closed-loop performance of AL closed-loop no. 1 [Fig. 3(a)] results in an expression for the mean-squared phase error given by⁶

$$\sigma_\phi^2 = \frac{1}{\rho} \left[\frac{L \sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_{ij} \exp \{2R_d(L + D_i + D_j + D_{ij})\}}{\left[\sum_{m=0}^L \binom{L}{m} (L - 2m) \exp \{R_d(3L - 4m)\} \right]^2} \right] \triangleq \frac{1}{\rho S_L} \quad (7)$$

where

$$\rho = \frac{S}{N_0 B_L}, \quad R_d = \frac{ST}{N_0} \quad (8)$$

and

$$D_i = \sum_{k=0}^{L-1} d_k d_{ik}, \quad \sum_{k=0}^{L-1} d_{ik} d_{jk} \quad (9)$$

with

$$\mathbf{d} \triangleq (d_0, d_1, \dots, d_{L-1}) = \text{transmitted data sequence}$$

$$\mathbf{d}_i \triangleq (d_{i0}, d_{i1}, \dots, d_{i,L-1}) = \textit{i} \text{th data sequence; } i = 1, 2, \dots, 2^L \quad (10)$$

⁵ It is important at this point to emphasize that B_L , being proportional to the total loop gain, includes the slope of the loop S-curve at the origin as one of its factors. Since, in general, this slope is different for the various loops being investigated, it is absolutely essential to include this normalization (as we have done) in the definition of B_L when comparing the performance of these loops.

⁶ All of the performance results given in this article will be based upon the so-called "linear theory" [3], which assumes that the loop operates in a region of high loop SNR.

In Eq. (9), D_i represents the correlation of the i th data sequence with the transmitted sequence, and D_{ij} represents the correlation between the i th and the j th data sequences. Some properties of D_i and D_{ij} that are particularly useful in obtaining many of the results that follow are summarized as

$$\begin{aligned}
\sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_{ij} &= 0 \\
\sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_i D_{ij} &= \sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_j D_{ij} = 0 \\
\sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_{ij}^2 &= 2L \sum_{m=0}^L \binom{L}{m} (L-2m)^2 = 2^{2L} L \\
\sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_i D_j D_{ij} &= 2^{2L} L
\end{aligned} \tag{11}$$

The factor \mathcal{S}_L represents the loss of the *effective* loop SNR, $\rho' \triangleq \sigma_\phi^{-2}$, relative to the loop SNR, ρ , of a phase-locked loop (PLL). For certain configurations, as we shall see, this loss is synonymous with what is commonly referred to as “squaring loss” [4,10].

At first glance, it might appear that, for given values of ρ , R_d , and the observation length, L , the mean-squared phase error would be a function of the particular sequence chosen as the transmitted sequence. It is easy to show that indeed this is not the case, i.e., σ_ϕ^2 is independent of the sequence selected for \mathbf{d} .⁷ To see this, consider a sequence $\mathbf{d}_l \triangleq (d_{l0}, d_{l1}, \dots, d_{l,L-1}) \neq \mathbf{d}$ and rewrite D_i and D_{ij} as

$$\begin{aligned}
D_i &= \sum_{k=0}^{L-1} d_k \underbrace{d_{lk} d_{lk}}_{=1} d_{ik} = \sum_{k=0}^{L-1} d'_k d'_{ik} \\
D_{ij} &= \sum_{k=0}^{L-1} d_{ik} \underbrace{d_{lk} d_{lk}}_{=1} d_{jk} = \sum_{k=0}^{L-1} d'_{ik} d'_{jk}
\end{aligned} \tag{12}$$

where $d'_k = d_k d_{lk}$ represents the k th element of some other possible transmitted sequence $\mathbf{d}' \triangleq (d'_{i0}, d'_{i1}, \dots, d'_{i,L-1})$ and $d'_{ik} = d_{ik} d_{lk}$, $d'_{jk} = d_{jk} d_{lk}$ are the k th elements of two other possible sequences $\mathbf{d}'_i \triangleq (d'_{i0}, d'_{i1}, \dots, d'_{i,L-1})$ and $\mathbf{d}'_j \triangleq (d'_{j0}, d'_{j1}, \dots, d'_{j,L-1})$, respectively. Since, in general, $\mathbf{d}' \neq \mathbf{d}$ and since the summations on i and j in Eq. (7) range over all possible (2^L) sequences, then substitution of Eq. (9) into Eq. (7) shows that σ_ϕ^2 evaluated for a transmitted sequence equal to \mathbf{d}' is identical to that evaluated for a transmitted sequence equal to \mathbf{d} .

Special cases of Eq. (7) corresponding to $L = 1, 2$, and 3 are given below:

⁷ For convenience in the evaluation of Eq. (7), we may choose the all-1's sequence for \mathbf{d} , in which case D_i simplifies to $\sum_{k=0}^{L-1} d_{ik}$, which takes on values of $L - 2m$, $m = 0, 1, 2, \dots, L$.

$$\sigma_{\phi}^2 = \frac{1}{\rho} \left[\frac{e^{8R_d} - 1}{(e^{3R_d} - e^{-R_d})^2} \right]; \quad L = 1$$

$$\sigma_{\phi}^2 = \frac{1}{\rho} \left[\frac{e^{16R_d} + 2e^{8R_d} - 3}{(e^{6R_d} - e^{-2R_d})^2} \right]; \quad L = 2$$

$$\sigma_{\phi}^2 = \frac{1}{\rho} \left[\frac{e^{24R_d} + 5e^{16R_d} + 3e^{8R_d} - 9}{(e^{9R_d} + e^{5R_d} - e^{R_d} - e^{-3R_d})^2} \right]; \quad L = 3 \quad (13)$$

Figure 4 is a plot of \mathcal{S}_L (in dB) versus R_d (in dB) corresponding to the three cases in Eq. (13). We observe that the performance of AL closed-loop no. 1 as implemented in Fig. 3(a) is clearly a function of the observation length of the corresponding open-loop estimator that motivated the structure.

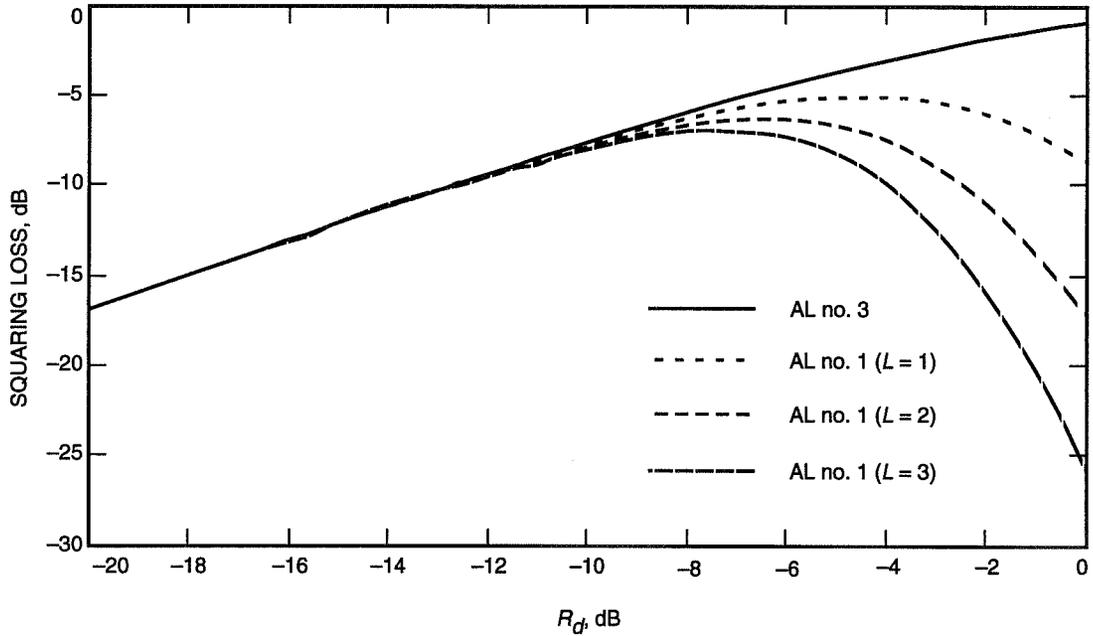


Fig. 4. Squaring-loss performance of AL closed-loop no. 1 with observation length L as a parameter. I&D weighting coefficients as determined by MAP estimation theory.

For large R_d , it is straightforward to show that σ_{ϕ}^2 has the asymptotic behavior

$$\sigma_{\phi}^2 \cong \frac{1}{\rho} e^{2LR_d} \rightarrow \mathcal{S}_L \cong e^{-2LR_d} \quad (14)$$

For small R_d , σ_{ϕ}^2 has the asymptotic form

$$\sigma_{\phi}^2 \cong \frac{1}{\rho} \left(\frac{1}{2R_d} \right) \rightarrow \mathcal{S}_L \cong 2R_d \quad (15)$$

which is independent of L .

Looking at Eq. (14) and Fig. 4, one gets the impression (and rightfully so) that the mean-squared phase error of AL closed-loop no. 1 becomes unbounded as $R_d \rightarrow \infty$. This singular behavior can be traced to the fact that the $2\sqrt{2S}/N_0$ weighting coefficient of the two integrate-and-dump (I&D) circuits in the closed loop of Fig. 3(a) becomes unbounded as $R_d \rightarrow \infty$ ($N_0 \rightarrow 0$). Suppose instead that we were to replace this coefficient by an arbitrary constant, say K_0 . From the standpoint of open-loop estimation of θ , AL open-loop estimator no. 1 of Eq. (4) with $2\sqrt{2S}/N_0$ now replaced by K_0 would remain unchanged. That is, *the choice of the weighting constant preceding the L -bit integration has no effect on the open-loop estimate*. On the other hand, the choice of this weighting coefficient for the closed-loop scheme has a very definite bearing on its performance. In particular, with $2\sqrt{2S}/N_0$ replaced by K_0 in Fig. 3(a), the mean-squared phase error, previously given by Eq. (7), now becomes

$$\sigma_\phi^2 = \frac{1}{\rho} \left[\frac{L \sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_{ij} \exp \left\{ K(D_i + D_j) + K^2 \left(\frac{L}{2R_d} \right) \left(1 + \frac{D_{ij}}{L} \right) \right\}}{\left[\sum_{m=0}^L \binom{L}{m} (L-2m) \exp \left\{ K(L-2m) + K^2 \left(\frac{L}{4R_d} \right) \right\} \right]^2} \right] \triangleq \frac{1}{\rho \mathcal{S}_L} \quad (16)$$

where we have further normalized the weighting coefficient as $K \triangleq (\sqrt{S/2}) K_0 T$. Note that if we set $K_0 = 2\sqrt{2S}/N_0$ as before, then $K = 2R_d$ and Eq. (16) reduces to Eq. (7).

From Eq. (16), we see that as long as K_0 (or equivalently K) is finite (which would be the case in a practical implementation of the AL closed-loop scheme), the large SNR asymptotic behavior of AL closed-loop no. 1 now becomes

$$\lim_{R_d \rightarrow \infty} \sigma_\phi^2 = \lim_{N_0 \rightarrow 0} \frac{N_0 B_L}{S} \left[\frac{L \sum_{i=1}^{2^L} \sum_{j=1}^{2^L} D_{ij} \exp \{K(D_i + D_j)\}}{\left[\sum_{m=0}^L \binom{L}{m} (L-2m) \exp \{K(L-2m)\} \right]^2} \right] = 0 \quad (17)$$

which is what one would expect. What is interesting is that, for any value of R_d , the value of K that minimizes Eq. (16), which, from the standpoint of *closed-loop* performance as measured by mean-squared phase error, would be considered optimum, is $K \rightarrow 0$, *independent of R_d* . In fact, if one takes the limit of Eq. (16) as $K \rightarrow 0$ [this must be done carefully using the properties in Eq. (11)], the following result is obtained:

$$\lim_{K \rightarrow 0} \sigma_\phi^2 \triangleq (\sigma_\phi^2)_{min} = \frac{1}{\rho} \left[1 + \frac{1}{2R_d} \right] \rightarrow (\mathcal{S}_L)_{max} = \frac{1}{1 + (1/2R_d)} = \frac{2R_d}{1 + 2R_d} \quad (18)$$

Interestingly enough, the result in Eq. (18), *which is now independent of L* , is also characteristic of the performance of the I-Q Costas loop [4], which is obtained as a low SNR approximation to AL closed-loop no. 3. It is important to understand that the optimum closed-loop performance of Eq. (18) is a consequence of optimizing the weight (gain) K for each value of L . If instead of doing this, one were to fix the gain K for all values of L (as suggested by the MAP estimation approach), the closed-loop performance (as measured by σ_ϕ^2 with fixed-loop bandwidth) is suboptimum and indeed depends once again on L . One final note is that the small SNR behavior of Eq. (18) is identical to that of Eq. (15), the reason being that the value of $K = 2R_d$ used in arriving at Eq. (15) approaches the optimum value ($K = 0$) as $R_d \rightarrow 0$.

As previously stated, the performance of AL closed-loop no. 2 is identical to that of AL closed-loop no. 1, and thus no further discussion is necessary. The performance of AL closed-loop no. 3 (and also

AL closed-loop no. 4) has been obtained previously [8]. In particular, the mean-squared phase error performance of this loop is given by

$$\sigma_\phi^2 = \frac{1}{\rho} \left[\frac{\overline{\tanh^2\{2R_d - \sqrt{2R_d}X\}}}{\left[\overline{\tanh\{2R_d - \sqrt{2R_d}X\}}\right]^2} \right] \triangleq \frac{1}{\rho\mathcal{S}_L} \quad (19)$$

where X is a zero-mean, unit-variance Gaussian random variable, and the over bar denotes statistical averaging over X . A plot of \mathcal{S}_L versus R_d is superimposed on the curves of Fig. 4. We first note that the performance as given by Eq. (18) is independent of L . Furthermore, a comparison of the squaring loss as determined from Eq. (18) with that calculated from Eq. (17) reveals that the performance of AL closed-loop no. 3 is superior to that of AL closed-loop no. 1 with optimized gain for all values of R_d (see Fig. 3 of [8]). As mentioned previously, if the hyperbolic tangent nonlinearity in Fig. 3(b) is approximated by a linear device (i.e., $\tanh x \cong x$), then the two loops have the same performance.

What is particularly interesting for AL closed-loop no. 3 is that even though the performance in Eq. (19) is computed assuming a weighting coefficient in front of the I&Ds in Fig. 3(b) equal to $2\sqrt{2S}/N_0$, the behavior of this loop is not singular in the limit as $R_d \rightarrow \infty$. Furthermore, it is natural to ask whether the above weighting coefficient is indeed optimum in the sense of minimizing σ_ϕ^2 . To answer this question, we proceed as we did for AL closed-loop no. 1, namely, we replace the weighting coefficient $2\sqrt{2S}/N_0$ by an arbitrary constant, say K_0 , and proceed to optimize the performance with respect to the choice of this gain.⁸ Making this replacement produces a mean-squared phase error, analogous to Eq. (19), given by

$$\sigma_\phi^2 = \frac{1}{\rho} \left[\frac{\overline{\tanh^2\{K[2R_d - \sqrt{2R_d}X]\}}}{\left[\overline{\tanh\{K[2R_d - \sqrt{2R_d}X]\}}\right]^2} \right] \triangleq \frac{1}{\rho\mathcal{S}_L} \quad (20)$$

where, as before, we have further normalized the weighting coefficient as $K \triangleq K_0 N_0 / 2\sqrt{2S}$. Maximizing the squaring loss factor \mathcal{S}_L (i.e., minimizing σ_ϕ^2) in Eq. (20) results in $K = 1$ ($K_0 = 2\sqrt{2S}/N_0$) for all values of R_d . Thus, for AL closed-loop no. 3, *the optimum gain from the standpoint of closed-loop performance is precisely that dictated by the open-loop MAP estimation of θ* , and the best performance is that described by Eq. (19).

We conclude our discussion of AL closed loops by pointing out that, in view of the superiority of Eq. (19) over Eq. (18), AL closed-loop no. 3 outperforms AL closed-loop no. 1 for all values of R_d .

IV. Closed Loops Motivated by the ML Approach

A. Structures

The ML approach to estimating the carrier phase, θ , is to maximize (rather than to average) $p(r(t)|d_i(t), \theta)$ over all possible (2^L) and equally likely data sequences. Analogous to AL open-loop estimator no. 1, "ML open-loop estimator no. 1" is defined by (see Fig. 1: $\hat{\theta}_{ML} \triangleq \max_\theta^{-1} q_{ML}(\theta)$, unpartitioned observation)

⁸ Again we note that this replacement does not affect the open-loop estimation of θ using Eq. (6).

$$\hat{\theta}_{ML_1} = \max_{\theta}^{-1} q_i(\theta)$$

$$q_i(\theta) \triangleq \exp \left(\max_{\{d_i(t)\}} \frac{2\sqrt{2S}}{N_0} \int_0^{LT} r(t) d_i(t) \sin(\omega_c t + \theta) dt \right) \quad (21)$$

where \hat{i} is the particular value of i corresponding to the data waveform $d_i(t)$ that achieves the maximization. Alternately, by breaking up the integration over the entire observation into a sum of integrals on each bit interval (the partitioned form of the observation) and recognizing that the data bits are iid binary random variables, then Eq. (21) evaluates to (see Fig. 1: $\hat{\theta}_{ML} \triangleq \max_{\theta}^{-1} q_{ML}(\theta)$, partitioned observation)

$$\hat{\theta}_{ML_2} = \max_{\theta}^{-1} \prod_{k=0}^{L-1} \exp \left(\left| \frac{2\sqrt{2S}}{N_0} \int_{kT}^{(k+1)T} r(t) \sin(\omega_c t + \theta) dt \right| \right) \quad (22)$$

This estimator is analogous to Eq. (5) and is called “ML open-loop estimator no. 2.” Next, we obtain ML open-loop estimates by maximizing the natural logarithm of $q_i(\theta)$. Using the product form of $q_i(\theta)$ as in Eq. (22), one obtains (see Fig. 1: $\hat{\theta}_{ML} \triangleq \max_{\theta}^{-1} \ln q_{ML}(\theta)$, partitioned observation)

$$\hat{\theta}_{ML_3} = \max_{\theta}^{-1} \sum_{k=0}^{L-1} \left| \frac{2\sqrt{2S}}{N_0} \int_{kT}^{(k+1)T} r(t) \sin(\omega_c t + \theta) dt \right| \quad (23)$$

which is analogous to Eq. (6) and therefore called the “ML open-loop estimator no. 3.” Finally, we consider a fourth ML open-loop estimator, which is based on maximizing the natural logarithm of $q_i(\theta)$ in its unpartitioned form of Eq. (15). This leads to “ML open-loop estimator no. 4,” which is defined by (see Fig. 1: $\hat{\theta}_{ML} \triangleq \max_{\theta}^{-1} \ln q_{ML}(\theta)$, unpartitioned observation)

$$\hat{\theta}_{ML_4} = \max_{\theta}^{-1} \frac{2\sqrt{2S}}{N_0} \int_0^{LT} r(t) d_i(t) \sin(\omega_c t + \theta) dt \quad (24)$$

Block diagram implementations of ML open-loop estimator no. 1 [Eq. (21)] and ML open-loop estimator no. 3 [Eq. (23)] are illustrated in Fig. 5 as representative of the four possibilities. In drawing these implementations, we have again quantized the unknown phase into Q values and, thus, the maximization over the continuous phase parameter θ in Eqs. (21) and (23) is approximated by maximization over a Q -quantized version of this parameter.

As was true for the AL case, it is important to emphasize that the four ML open-loop phase estimates as described by Eqs. (21) through (24) are identical. However, we shall again see that this same statement is *not* true when considering the performances of the closed loops motivated by these four different ML formulations.

Closed-loop phase synchronization structures based on the four ML open-loop estimators are obtained as analogies of their AL counterparts, choosing as error signals, e , the derivatives with respect to θ of the functions being maximized in Eqs. (21) through (24), respectively. Analogous to the terminology used for the AL case, we shall refer to these four closed-loop structures as ML closed-loop nos. 1, 2, 3,

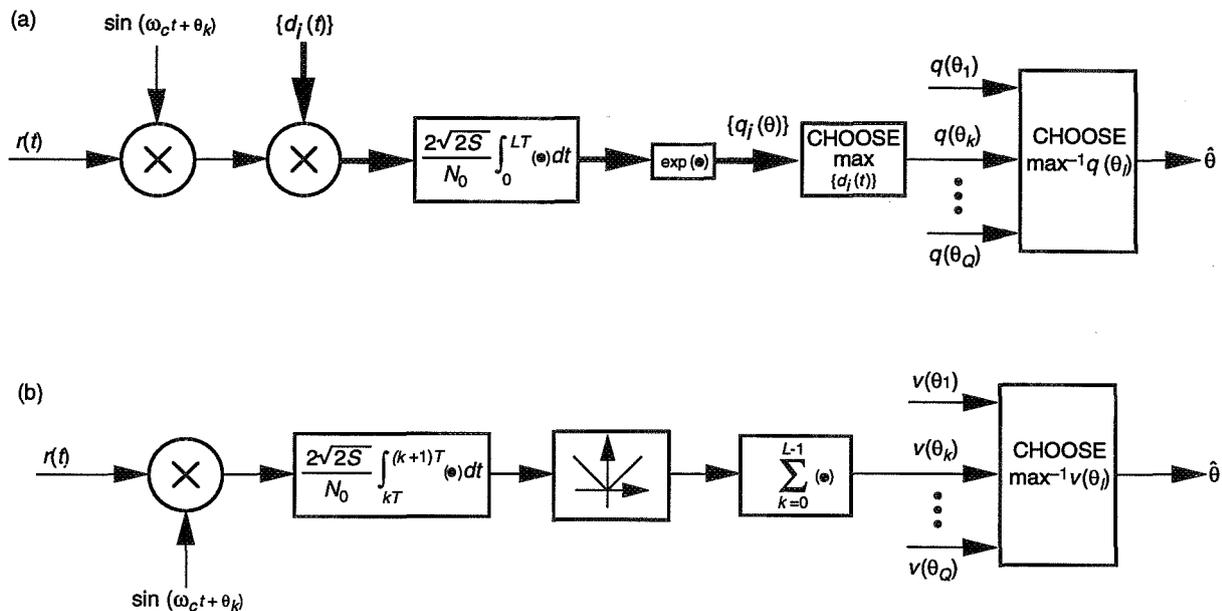


Fig. 5. Implementation of two ML open-loop phase estimators: (a) ML open-loop estimator no. 1—observation unpartitioned (quantized parallel implementation) and (b) ML open-loop estimator no. 3—observation partitioned (quantized parallel implementation).

and 4. An implementation of ML closed-loop no. 1 is illustrated in Fig. 6(a). We also show here in Fig. 6(b) an implementation of ML closed-loop no. 1 (or ML closed-loop no. 2) for the special case of $L = 1$ since, as we shall see shortly, this particular of L yields the best performance. It is worthy of note that ML closed-loop no. 3 is identical in form to the I-Q polarity-type Costas loop [5], as can be seen in Fig. 6(c). (Note that the L -fold accumulator that precedes the loop filter can be omitted since it can be absorbed into the loop filter itself by renormalizing its bandwidth.) We recall that, in the AL case, the I-Q polarity-type Costas loop is obtained only as a high SNR approximation to closed-loop no. 3.

B. Performance

An analysis of the closed-loop performance of Fig. 6(a) results in an expression for the mean-squared phase error given by (see the Appendix for the derivation)

$$\sigma_{\phi}^2 = \frac{1}{\rho} e^{LK^2/2R_d} \frac{[(1 - p_{2+}(0)) e^{2K} + p_{2-}(0) e^{-2K}]^L}{\left\{ [(1 - p_+(0)) e^K - p_-(0) e^{-K}] [(1 - p_+(0)) e^K + p_-(0) e^{-K}]^{L-1} \right\}^2}$$

$$\triangleq \frac{1}{\rho \mathcal{S}_L} \quad (25)$$

where

$$p_{\pm}(\phi) \triangleq \frac{1}{2} \operatorname{erfc} \left(\sqrt{R_d} \cos \phi \pm \frac{K}{2\sqrt{R_d}} \right)$$

$$p_{2\pm}(\phi) \triangleq p_{\pm}(\phi)|_{K \rightarrow 2K} = \frac{1}{2} \operatorname{erfc} \left(\sqrt{R_d} \cos \phi \pm \frac{K}{\sqrt{R_d}} \right) \quad (26)$$

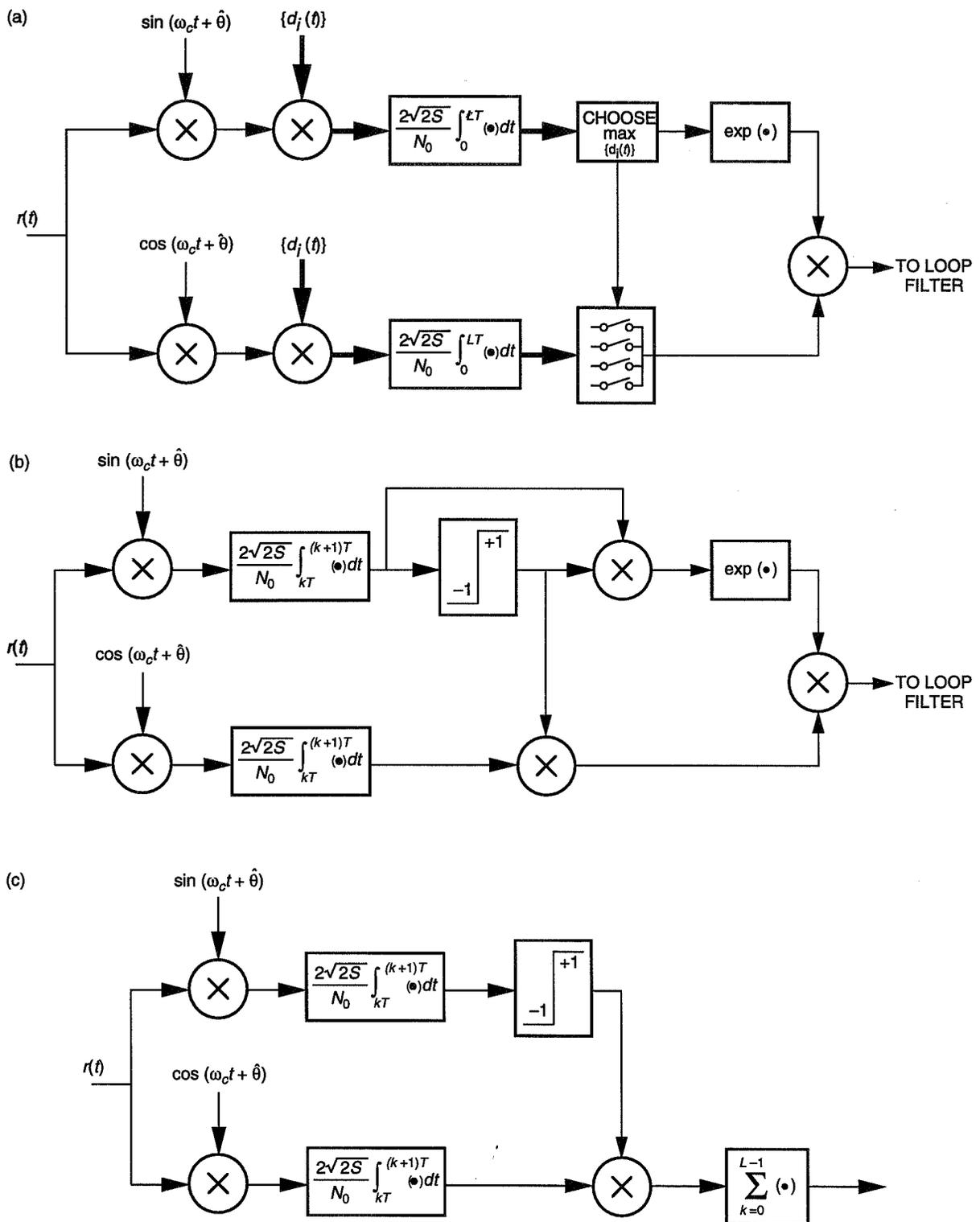


Fig. 6. Implementation of ML closed-loops: (a) ML closed-loop no. 1—observation unpartitioned, (b) ML closed-loop no. 2 ($L = 1$), and (c) ML closed-loop no. 3—observation partitioned.

As we did for the analogous AL closed loop [see Fig. 3(a)], we have avoided the singular behavior of the mean-squared phase error as $R_d \rightarrow \infty$ by replacing the $2\sqrt{2S}/N_0$ coefficient in front of the I&Ds in Fig. 6(a) by an arbitrary constant, say K_0 , that remains finite as $N_0 \rightarrow 0$ and further normalized the weighting coefficient as $K \triangleq (\sqrt{S/2})K_0T$. As long as K_0 (or equivalently K) is finite (which would be the case in a practical implementation of the ML closed-loop scheme), the large SNR asymptotic behavior of ML closed-loop no. 1 is

$$\lim_{R_d \rightarrow \infty} \sigma_\phi^2 = \lim_{N_0 \rightarrow 0} \frac{N_0 B_L}{S} = 0 \quad (27)$$

as one would expect. What is indeed interesting is that, *unlike* the AL case, the value of K that minimizes Eq. (25), which from the standpoint of *closed-loop* performance as measured by mean-squared phase error would be considered optimum, is *not* $K \rightarrow 0$. In fact, for each value of R_d and L , there exists an optimum value of K that unfortunately cannot be determined in closed form. Nevertheless, the optimum values of K can be found numerically as a function of R_d by maximizing S_L as determined from Eq. (25) for each value of L . The results are illustrated in Fig. 7. The corresponding values of $(S_L)_{max}$ are plotted versus R_d in dB in Fig. 8 for the same values of L as those in Fig. 7. Results obtained from a computer simulation of Fig. 6(b) agree with these analytically obtained numerical results for $(S_L)_{max}$ within 0.1 dB at $R_d = -6$ dB.

From Fig. 8, we observe that the performance of ML closed-loop no. 1 becomes worse with increasing L , i.e., $L = 1$ gives the best performance. Thus, the special case of the implementation in Fig. 6(a) corresponding to $L = 1$, i.e., Fig. 6(b), is the configuration of most interest. Also in the limit as $L \rightarrow \infty$, the optimum value of K approaches 0 independent of R_d . The corresponding value of S_L is determined by noting that for $K \rightarrow 0$ we have $p_+(0) = p_-(0) = p_{2+}(0) = p_{2-}(0) \triangleq p = 1/2 \operatorname{erfc} \sqrt{R_d}$. Then from Eq. (25), we get

$$\lim_{K \rightarrow 0} \sigma_\phi^2 = (\sigma_\phi^2)_0 = \frac{1}{\rho} (1 - 2p)^{-2} = \frac{1}{\rho} \left(\operatorname{erfc}^2 \sqrt{R_d} \right)^{-1} \rightarrow (S_L)_0 = \operatorname{erfc}^2 \sqrt{R_d} \quad (28)$$

which also *is independent of the observation length* L . Since the optimum value of K is always greater than 0 (see Fig. 7), Eq. (28) also serves as a lower bound on the squaring-loss performance of ML closed-loop no. 1. Other reasons for including this limiting squaring-loss behavior in Fig. 8 will become apparent shortly when we consider the other ML closed-loop configurations.

As in the AL case, the performance of ML closed-loop no. 2 is identical to ML closed-loop no. 1 and needs no further discussion. Moving on to ML closed-loop no. 3, we previously identified this as being identical in form to the I-Q polarity-type Costas loop. Hence, its performance is independent of L and is given by Eq. (28). Similarly, the performance of ML closed-loop no. 4 is also independent of L and given by Eq. (28). Thus, we see that of the four ML closed loops, ML closed-loops nos. 1 and 2 are superior to ML closed-loops nos. 3 and 4, which have performances that are identical and equal to those of the former in the worst case ($L \rightarrow \infty$).

When the performance of the best ML closed-loop scheme (i.e., nos. 1 or 2) is compared with that of the best AL closed-loop scheme (i.e., nos. 3 or 4), we find that the latter, e.g., the I-Q MAP estimation loop, is superior to the former for all values of R_d . This comparison is illustrated in Fig. 9, where the squaring-loss performance of the two schemes is plotted versus R_d .

C. Loop S-Curves

It is of interest to examine the S-curve behavior of ML closed-loop no. 1 and compare it with that of ML closed-loop no. 3 and AL closed-loop no. 3. The equation describing the loop S-curve, $\eta(\phi)$, of ML

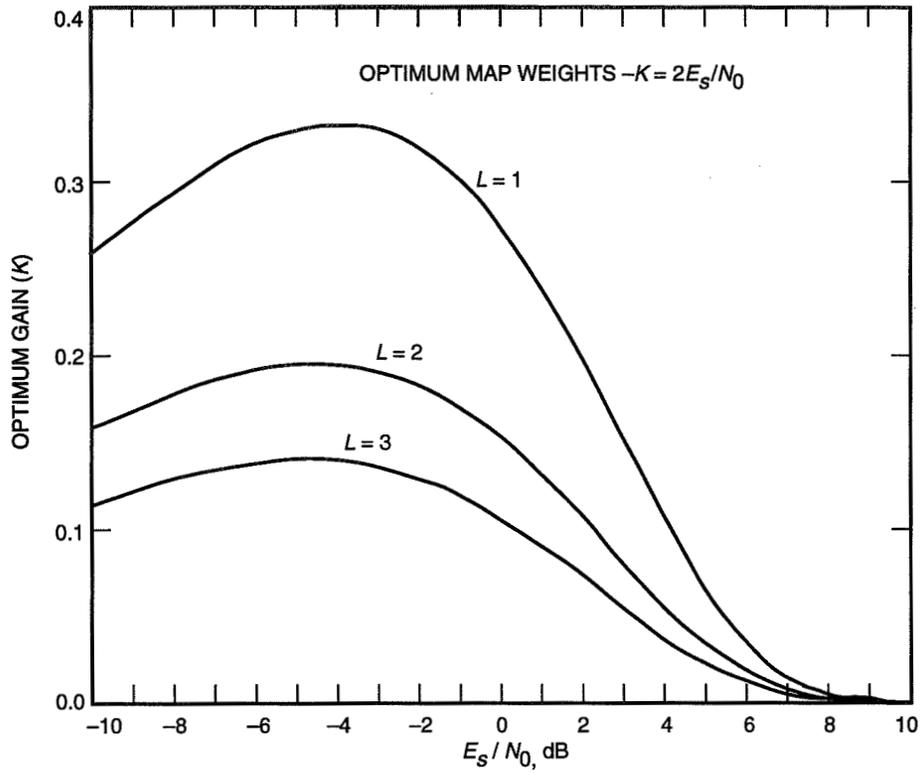


Fig. 7. Optimum weights (normalized) versus symbol SNR.

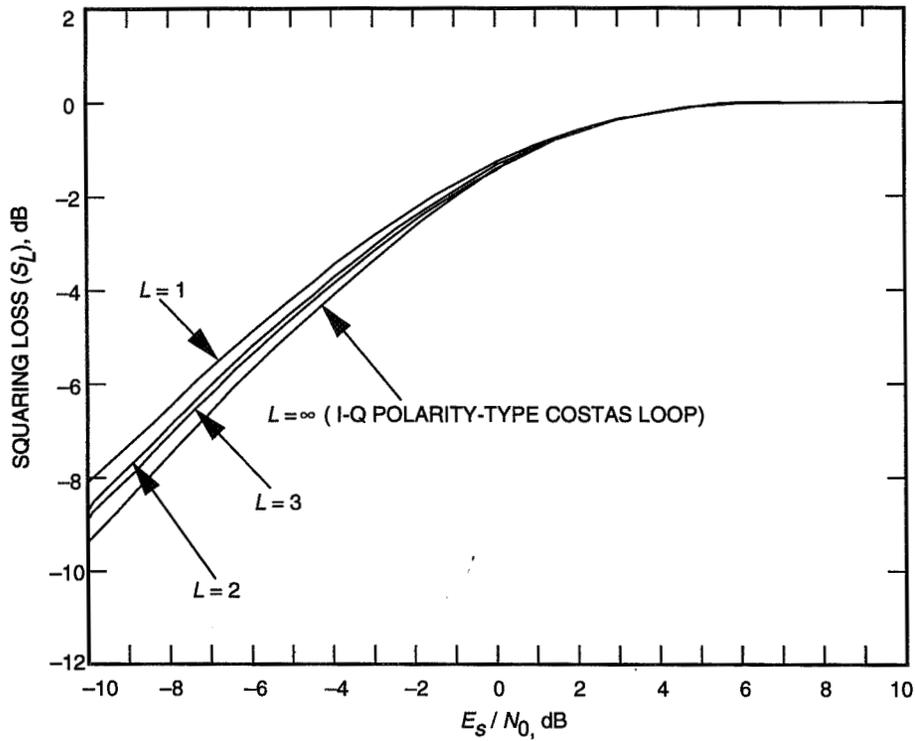


Fig. 8. Squaring loss versus symbol SNR for ML closed-loop no. 1 or no. 2.

closed-loop no. 1 is derived in the Appendix as Eq. (A-9) with the special case of $L = 1$ (already shown to yield the best tracking performance) given by Eq. (A-10). Figure 10 illustrates plots of $\eta(\phi)$ versus ϕ over one cycle of π rad for $R_d = -5, 2,$ and 5 dB, respectively, where in each case, K has been chosen equal to the optimum value as determined from Fig. 7. In the limit of small and large R_d , the S-curve approaches the following functional forms:

$$\eta(\phi) \propto \begin{cases} \sin 2\phi, & \text{small } R_d \\ \sin \phi \times \text{sgn}(\cos \phi), & \text{large } R_d \end{cases} \quad (29)$$

These limiting forms are identical to the same limiting behavior of the S-curves corresponding to ML closed-loop no. 3—the I-Q polarity-type Costas loop, and AL closed-loop no. 3—the I-Q MAP estimation loop.

V. Conclusions

Motivated by the theory of MAP carrier phase estimation, we have developed a number of closed-loop structures suitably derived from ML and AL functions. Several of these structures reduce to previously known closed-loop carrier phase synchronizers while others appear to be new. One of the new structures derived from ML considerations gives improved performance over the I-Q polarity-type Costas loop, which is also derived from these very same considerations. Of all the loops considered, however, the I-Q MAP estimation loop, which is derived from average log-likelihood considerations, is the best overall from a performance standpoint. We leave the reader with the thought that the structures proposed in this article are not exhaustive of the ways that closed-loop phase synchronizers can be derived from open-loop MAP estimation theory. Rather, they are given here primarily to indicate the variety of different closed-loop schemes that can be constructed simply from likelihood and log-likelihood functions.

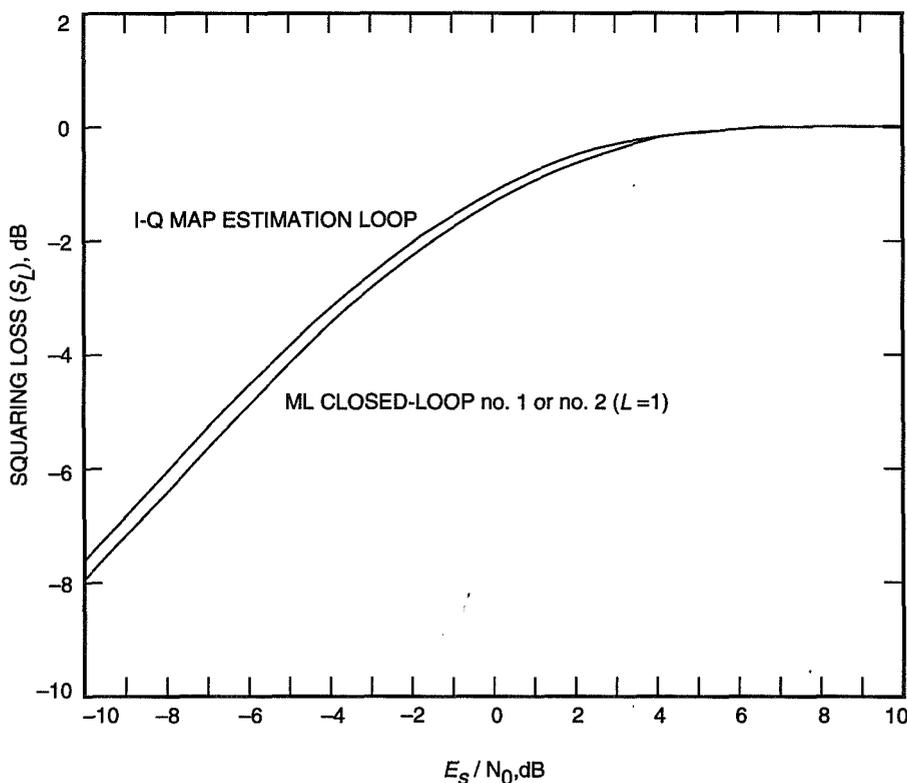


Fig. 9. A comparison of the squaring performances of ML closed-loop no. 1 or no. 2 and the I-Q MAP estimation loop (AL closed loop no. 3).

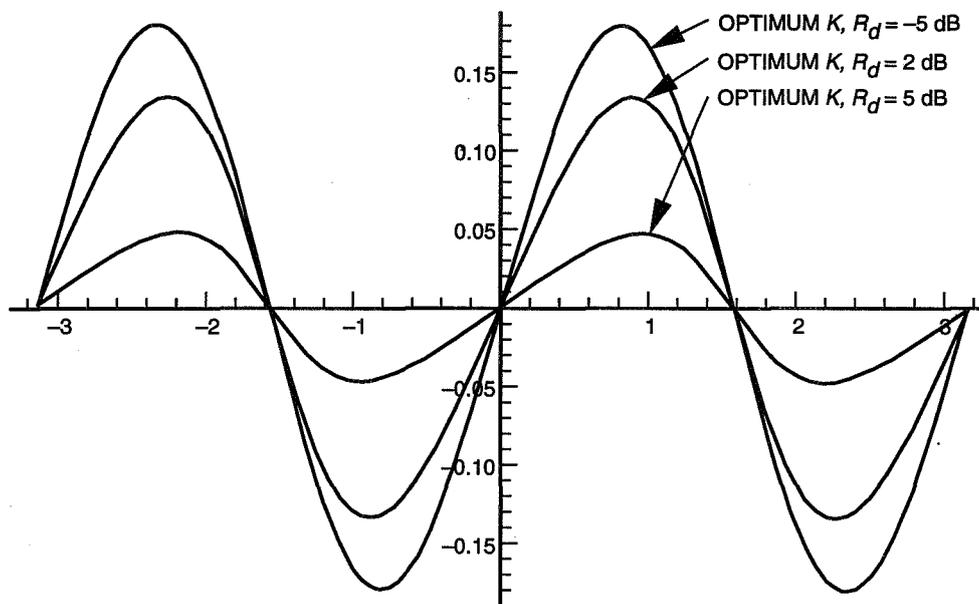


Fig. 10. Loop S-curves for ML closed-loop no. 1 or no. 2 ($L = 1$).

References

- [1] C. W. Helstrom, *Statistical Theory of Signal Detection*, New York: Pergamon Press, 1960.
- [2] H. L. Van Trees, *Detection Estimation, and Modulation Theory*, Part I, New York: John Wiley & Sons, 1968.
- [3] J. J. Stiffler, *Theory of Synchronous Communications*, Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1971.
- [4] W. C. Lindsey and M. K. Simon, "Optimum Design and Performance of Suppressed Carrier Receivers With Costas Loop Tracking," *IEEE Transactions on Communications*, vol. COM-25, no. 2, pp. 215-227, February 1977.
- [5] M. K. Simon, "Tracking Performance of Costas Loops With Hard Limited In-Phase Channel," *IEEE Transactions on Communications*, vol. COM-26, no. 4, pp. 420-432, April 1978.
- [6] S. Dolinar, "A New Code for Galileo," *The Telecommunications and Data Acquisition Progress Report 42-93*, vol. January-March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 83-96, May 15, 1988.
- [7] A. Mileant and S. Hinedi, "Overview of Arraying Techniques for Deep Space Communications," *IEEE Transactions on Communications*, vol. COM-42, nos. 2, 3, and 4, pp. 1856-1865, February, March, and April 1994.
- [8] M. K. Simon, "On the Optimality of the MAP Estimation Loop for Tracking BPSK and QPSK Signals," *IEEE Transactions on Communications*, vol. COM-27, no. 1, pp. 158-165, January 1979.

- [9] M. K. Simon, "Optimum Receiver Structures for Phase-Multiplexed Modulations," *IEEE Transactions on Communications*, vol. COM-26, no. 6, pp. 865–872, June 1978.
- [10] M. K. Simon, "On the Calculation of Squaring Loss in Costas Loops With Arbitrary Arm Filters," *IEEE Transactions on Communications*, vol. COM-26, no. 1, pp. 179–184, January 1978.

Appendix

Derivation of the Closed-Loop Tracking Performance of ML Closed-Loop No. 1

Consider the closed loop in Fig. 5(a), whose error signal, $e(t)$, at time $t = LT$ is characterized by

$$e = \exp \left(K_0 \int_0^{LT} r(t) d_{\hat{i}}(t) \sin(\omega_c t + \hat{\theta}) dt \right) \times K_0 \int_0^{LT} r(t) d_{\hat{i}}(t) \cos(\omega_c t + \hat{\theta}) dt \quad (\text{A-1})$$

Substituting $r(t)$ of Eq. (1) into Eq. (A-1) results in

$$e = \exp \left\{ K_0 \sqrt{\frac{S}{2}} \left(\int_0^{LT} d(t) d_{\hat{i}}(t) dt \right) \cos \phi + K_0 \int_0^{LT} n(t) d_{\hat{i}}(t) \sin(\omega_c t + \hat{\theta}) dt \right\} \\ \times \left[K_0 \sqrt{\frac{S}{2}} \left(\int_0^{LT} d(t) d_{\hat{i}}(t) dt \right) \sin \phi + K_0 \int_0^{LT} n(t) d_{\hat{i}}(t) \cos(\omega_c t + \hat{\theta}) dt \right] \quad (\text{A-2})$$

In view of the rectangular phase shape assumed in Eq. (2) for the transmitted data waveform, $d(t)$, Eq. (A-2) can be written in the discrete form

$$e = \exp \left\{ K_0 \sum_{k=0}^{L-1} d_{\hat{i}k} \left(\sqrt{\frac{S}{2}} d_k \cos \phi + n_{sk} \right) \right\} \times \left[K_0 \sum_{k=0}^{L-1} d_{\hat{i}k} \left(\sqrt{\frac{S}{2}} d_k \sin \phi + n_{ck} \right) \right] \quad (\text{A-3})$$

where

$$n_{sk} \triangleq \int_{kT}^{(k+1)T} n(t) \sin(\omega_c t + \hat{\theta}) dt; \quad n_{ck} \triangleq \int_{kT}^{(k+1)T} n(t) \cos(\omega_c t + \hat{\theta}) dt \quad (\text{A-4})$$

are zero mean iid Gaussian random variables with variance $\sigma_{n_{ck}}^2 = \sigma_{n_{sk}}^2 = N_0 T / 4$ and $d_{\hat{i}k} \triangleq \text{sgn} \left(\sqrt{S/2} d_k \cos \phi + n_{sk} \right)$. Introducing the further normalization $K = K_0 T \sqrt{S/2}$ (note that when $K_0 = 2\sqrt{2S}/N_0$, i.e., the gain suggested by the open-loop MAP estimation theory, then $K = 2R_d$) and

normalizing n_{sk} and n_{ck} to unit variance Gaussian random variables, N_{sk} and N_{ck} , respectively, Eq. (A-4) becomes

$$e = \exp \left\{ K \sum_{k=0}^{L-1} d_{ik} \left(d_k \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\} \times \left[K \sum_{k=0}^{L-1} d_{ik} \left(d_k \sin \phi + \frac{1}{\sqrt{2R_d}} N_{ck} \right) \right] \quad (\text{A-5})$$

with $d_{ik} \triangleq \text{sgn} \left(d_k \cos \phi + (1/\sqrt{2R_d}) N_{sk} \right)$.

Let $\eta(\phi)$ denote the signal component (mean) of the error sample e . Then, because of the independence of the N_{sk} 's and N_{ck} 's, we have

$$\begin{aligned} \eta(\phi) &= K \sin \phi \left(\overline{\sum_{k=0}^{L-1} d_{ik} d_k \exp \left\{ K d_{ik} \left(d_k \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\}}^{N_{sk}} \right) \\ &\quad \times \prod_{\substack{l=0 \\ l \neq k}}^{L-1} \overline{\exp \left\{ K d_{il} \left(d_l \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sl} \right) \right\}}^{N_{sl}} \end{aligned} \quad (\text{A-6})$$

where the over bar denotes statistical averaging. It is straightforward to show that the statistical averages required in Eq. (A-6) are independent of the data bits. That is,

$$\overline{d_{ik} d_k \exp \left\{ K d_{ik} \left(d_k \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\}}^{N_{sk}}$$

is independent of whether $d_k = 1$ or $d_k = -1$ and

$$\overline{\exp \left\{ K d_{il} \left(d_l \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sl} \right) \right\}}^{N_{sl}}$$

is independent of whether $d_l = 1$ or $d_l = -1$. Performing these statistical averages gives the closed form results

$$\overline{d_{ik} d_k \exp \left\{ K d_{ik} \left(d_k \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\}}^{N_{sk}} = e^{K^2/4R_d} [(1 - p_+(\phi)) e^{K \cos \phi} - p_-(\phi) e^{-K \cos \phi}] \quad (\text{A-7a})$$

$$\overline{\exp \left\{ K d_{il} \left(d_l \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sl} \right) \right\}}^{N_{sl}} = e^{K^2/4R_d} [(1 - p_+(\phi)) e^{K \cos \phi} + p_-(\phi) e^{-K \cos \phi}] \quad (\text{A-7b})$$

where

$$p_{\pm}(\phi) = \frac{1}{2} \operatorname{erfc} \left(\sqrt{R_d} \cos \phi \pm \frac{K}{2\sqrt{R_d}} \right) \quad (\text{A-8})$$

Finally, since Eq. (A-7a) is independent of k and Eq. (A-7b) is independent of l , then substituting these results into Eq. (A-6), we get

$$\begin{aligned} \eta(\phi) &= (K \sin \phi) L e^{LK^2/4R_d} [(1 - p_+(\phi)) e^{K \cos \phi} - p_-(\phi) e^{-K \cos \phi}] \\ &\quad \times [(1 - p_+(\phi)) e^{K \cos \phi} + p_-(\phi) e^{-K \cos \phi}]^{L-1} \end{aligned} \quad (\text{A-9})$$

which represents the S-curve of the loop. For $L = 1$, Eq. (A-9) simplifies to

$$\eta(\phi) = (K \sin \phi) e^{K^2/4R_d} [(1 - p_+(\phi)) e^{K \cos \phi} - p_-(\phi) e^{-K \cos \phi}] \quad (\text{A-10})$$

which, using the definition of p_{\pm} , is periodic in ϕ with period π .

The slope of the S-curve at $\phi = 0$ is needed for computing the closed-loop mean-squared phase error performance. Differentiating Eq. (A-10) with respect to ϕ and evaluating the result at $\phi = 0$ gives

$$\begin{aligned} K_{\eta} \triangleq \frac{d\eta(\phi)}{d\phi} \Big|_{\phi=0} &= K L e^{LK^2/4R_d} [(1 - p_+(0)) e^{K \cos \phi} - p_-(0) e^{-K \cos \phi}] \\ &\quad \times [(1 - p_+(0)) e^{K \cos \phi} + p_-(0) e^{-K \cos \phi}]^{L-1} \end{aligned} \quad (\text{A-11})$$

The noise component of e evaluated at $\phi = 0$ is

$$N = \exp \left\{ K \sum_{k=0}^{L-1} d_{ik} \left(d_k \cos \phi + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\} \times \left[K \sum_{k=0}^{L-1} d_{ik} \frac{1}{\sqrt{2R_d}} N_{ck} \right] \quad (\text{A-12})$$

Which is zero mean and has variance

$$\sigma_N^2 = \exp \left\{ 2K \sum_{k=0}^{L-1} d_{ik} \left(d_k + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\} \times \left[K^2 \left(\sum_{k=0}^{L-1} d_{ik} \frac{1}{\sqrt{2R_d}} N_{ck} \right)^2 \right] \quad (\text{A-13})$$

Averaging first over the N_{ck} 's, we get

$$\sigma_N^2 = \frac{K^2}{2R_d} \left(\sum_{k=0}^{L-1} \exp \left\{ 2K d_{ik} \left(d_k + \frac{1}{\sqrt{2R_d}} N_{sk} \right) \right\}^{N_{sk}} \right) \times \prod_{\substack{l=0 \\ l \neq k}}^{L-1} \exp \left\{ 2K \left(d_l + \frac{1}{\sqrt{2R_d}} N_{sl} \right) \right\}^{N_{sl}} \quad (\text{A-14})$$

Using Eq. (A-7) to evaluate the averages over the N_{sk} 's, we get

$$\sigma_N^2 = \frac{K^2 L}{2R_d} e^{LK^2/R_d} [(1 - p_{2+}(0)) e^{2K} + p_{2-}(0) e^{-2K}]^L \quad (\text{A-15})$$

where

$$p_{2\pm}(\phi) \triangleq p_{\pm}(\phi)|_{K \rightarrow 2K} = \frac{1}{2} \operatorname{erfc} \left(\sqrt{R_d} \cos \phi \pm \frac{K}{\sqrt{R_d}} \right) \quad (\text{A-16})$$

Since $e(t)$ is a piecewise constant (over intervals of length LT) random process with independent increments, its statistical autocorrelation function is triangular and given by

$$R_e(\tau) \triangleq \langle E \{e(t)e(t+\tau)\} \rangle = \begin{cases} \sigma_N^2, & |\tau| \leq LT \\ 0, & \text{otherwise} \end{cases} \quad (\text{A-17})$$

Where $\langle \bullet \rangle$ denotes time averaging, which is necessary because of the cyclostationarity of $e(t)$. As is customary in analyses of this type, we assume a narrow-band loop, i.e., a loop bandwidth $B_L \ll 1/T$. Then, $e(t)$ is approximated as a delta-correlated process with effective power spectral density:

$$\frac{N'_0}{2} \triangleq \int_{-\infty}^{\infty} R_e(\tau) d\tau = LT \sigma_N^2 \quad (\text{A-18})$$

Finally, the mean-squared phase error for the closed loop is

$$\sigma_\phi^2 = \frac{N'_0 B_L}{K_\eta^2} \quad (\text{A-19})$$

which, with substitution of Eqs. (A-11) and (A-18), results in Eq. (26) of the main text.

110870

33272

p. 24

358425

On the Application of Under-Decimated Filter Banks

Y.-P. Lin and P. P. Vaidyanathan

California Institute of Technology

Maximally decimated filter banks have been extensively studied in the past. A filter bank is said to be under-decimated if the number of channels is more than the decimation ratio in the subbands. A maximally decimated filter bank is well known for its application in subband coding. Another application of maximally decimated filter banks is in block filtering. Convolution through block filtering has the advantages that parallelism is increased and data are processed at a lower rate. However, the computational complexity is comparable to that of direct convolution. More recently, another type of filter bank convolver has been developed. In this scheme, the convolution is performed in the subbands. Quantization and bit allocation of subband signals are based on signal variance, as in subband coding. Consequently, for a fixed rate, the result of convolution is more accurate than is direct convolution. This type of filter bank convolver also enjoys the advantages of block filtering, parallelism, and a lower working rate. Nevertheless, like block filtering, there is no computational saving.

In this article, under-decimated systems are introduced to solve the problem. The new system is decimated only by half the number of channels. Two types of filter banks can be used in the under-decimated system: the discrete Fourier transform (DFT) filter banks and the cosine modulated filter banks. They are well known for their low complexity. In both cases, the system is approximately alias free, and the overall response is equivalent to a tunable multilevel filter. Properties of the DFT filter banks and the cosine modulated filter banks can be exploited to simultaneously achieve parallelism, computational saving, and a lower working rate. Furthermore, for both systems, the implementation cost of the analysis or synthesis bank is comparable to that of one prototype filter plus some low-complexity modulation matrices. The individual analysis and synthesis filters have complex coefficients in the DFT filter banks but have real coefficients in the cosine modulated filter banks.

I. Introduction

The M channel maximally decimated filter bank shown in Fig. 1 has been studied extensively in [1-8]. A filter bank is said to be under-decimated if the number of channels is more than the decimation ratio in the subbands. When the system in Fig. 1 is alias free, it is a linear time invariant system with transfer function $T(z)$, as indicated in Fig. 1. In the following discussion, $T(z)$ will be called the distortion function, or the overall response.

A maximally decimated filter bank is well known for its application in subband coding. Another

application of maximally decimated filter banks is in block filtering [9]. Convolution through block filtering has the advantages that parallelism is increased and data are processed at a lower rate. However, the computational complexity is comparable to that in direct convolution. In [10], filter banks are used to map long convolutions into smaller ones in the subbands. Computations are then performed in parallel at a lower rate.

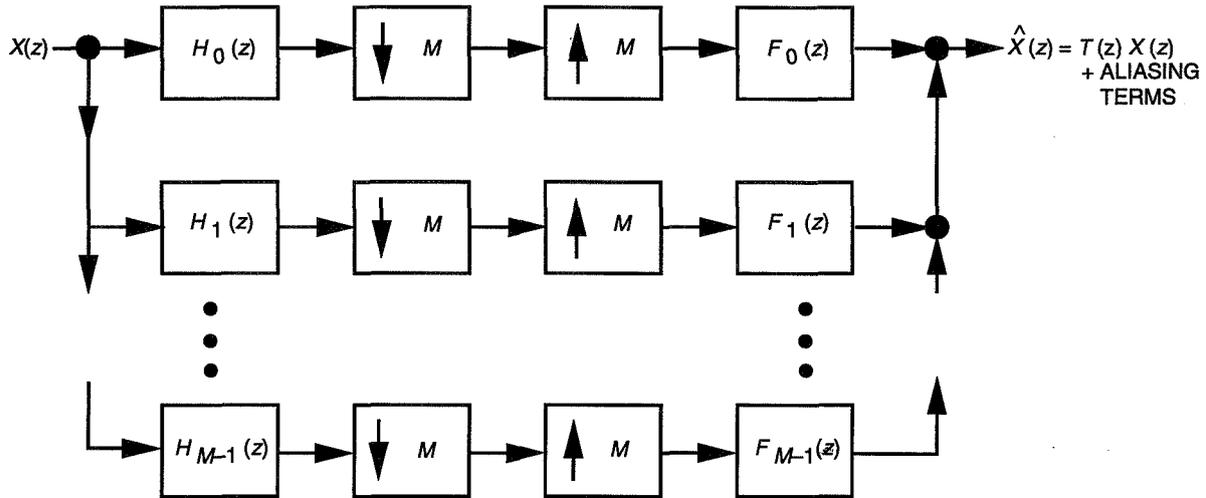


Fig. 1. The M -channel maximally decimated filter bank.

More recently [8], another type of filter bank convolver has been developed. In this scheme, the convolution is performed in the subbands. Quantization and bit allocation of subband signals are based on signal variance as in subband coding. Consequently, for a fixed rate, the result of convolution is more accurate than is direct convolution. This type of filter bank convolver also enjoys the advantages of block filtering, parallelism, and a lower working rate. Nevertheless, like block filtering, there is no computational saving.

In this article, under-decimated systems are introduced to solve the problem. Figure 2 shows the setup of the under-decimated system; it has $2M$ channels but is decimated only by M . Two types of filter banks can be used in the under-decimated system: the discrete Fourier transform (DFT) filter banks and the cosine modulated filter banks. They are recognized for their low complexity. In both cases, the system is approximately alias free and the overall response $T(z)$ is equivalent to a tunable multilevel filter. Properties of the DFT filter banks and the cosine modulated filter banks can be exploited to simultaneously achieve parallelism, computational saving, and a lower working rate. Furthermore, for both systems, the implementation cost of the analysis bank or the synthesis bank is comparable to that of one prototype filter plus some low complexity modulation matrices. The individual analysis and synthesis filters have complex coefficients in the DFT filter banks but have real coefficients in the cosine modulated filter banks.

This work is organized as follows: Section II is devoted to the construction of the new $2M$ -channel under-decimated DFT filter bank. Suppression of alias error due to decimation in the subbands is elaborated, and implementation and complexity of this DFT filter bank are discussed therein. The new $2M$ -channel cosine modulated filter bank is discussed in a similar manner in Section III. In the cosine modulated system, filters are real-coefficient and hence have positive and negative spectral occupancy. As a result, alias error in this case is more complicated than in DFT filter banks and requires careful treatment. Also in Section III, we study different types of spectral configuration for the under-decimated cosine modulated filter banks. Design examples are given in Section IV.

In this article, the following conventions are used:

- (1) Bold faced lower-case letters are used to represent vectors, and bold faced upper-case letters are used to represent matrices.
- (2) The notations \mathbf{A}^T , \mathbf{A}^* , and \mathbf{A}^\dagger represent the transpose, conjugate, and transpose conjugate of \mathbf{A} , respectively. The “tilde” notation is defined as follows: $\tilde{\mathbf{A}}(z) = \mathbf{A}^\dagger(1/z^*)$.
- (3) Matrix \mathbf{I}_k denotes a $k \times k$ identity matrix, and \mathbf{J}_k denotes a $k \times k$ reversal matrix with

$$\mathbf{J}_k = \begin{pmatrix} 0 & \dots & 0 & 1 \\ 0 & \dots & 1 & 0 \\ \vdots & & \vdots & \vdots \\ 1 & \dots & 0 & 0 \end{pmatrix}$$

- (4) The delay chain $\mathbf{e}_k(z)$ is the vector $\mathbf{e}_k(z) = [1 \ z^{-1} \dots \ z^{-(k-1)}]^T$.
- (5) The unit pulse, denoted as $\delta(n)$, is defined according to

$$\delta(n) = \begin{cases} 1 & n = 0, \\ 0 & \text{otherwise} \end{cases}$$
- (6) The value of the function $[x]$ is the largest integer less or equal to x .
- (7) The $2M \times 2M$ DFT matrix, \mathbf{W} , is defined such that $[\mathbf{W}]_{mn} = W^{mn}$. The quantity W is given by $W = e^{-j\pi/M}$, where $j = \sqrt{-1}$.
- (8) A filter $H(z)$ is called a Nyquist (M) filter if its impulse response $h(n)$ satisfies $h(Mn) = c\delta(n)$, for some constant c .

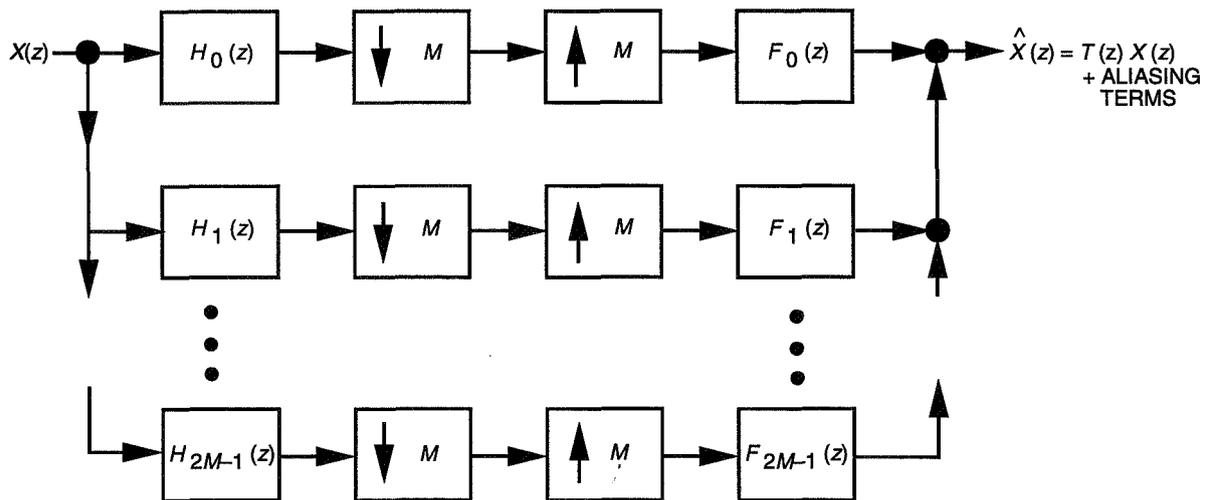


Fig. 2. A $2M$ -channel under-decimated filter bank.

II. DFT Filter Banks and Their Application in Tunable Multilevel Filtering

The system in Fig. 2 is called a DFT filter bank if the analysis filters are shifted versions of the same prototype. The same holds for the synthesis bank. The prototype of the analysis bank and the prototype

of the synthesis bank need not be the same. To be more specific, let $P_0(z)$ be the prototype filter of the analysis bank and $Q_0(z)$ be the prototype filter of the synthesis bank. The filters $P_k(z)$ and $Q_k(z)$, $k = 1, 2, \dots, 2M - 1$, are, respectively, the shifted versions of $P_0(z)$ and $Q_0(z)$.

$$P_k(z) = P_0(zW^k), \quad Q_k(z) = Q_0(zW^k), \quad k = -M, -M + 1, \dots, M - 1$$

Notice that unit circle $P_k(z)$ is only a shift of $P_0(z)$ by $k\pi/M$, since $P_k(e^{j\omega}) = P_0(e^{j(\omega - k\pi/M)})$. Figure 3 shows this relationship. The analysis filters and synthesis filters of the DFT filter bank have the following form:

$$H_k(z) = a_k P_0(zW^k), \quad \text{and} \quad F_k(z) = a_k^* Q_0(zW^k), \quad W = e^{-j\pi/M} \quad (1)$$

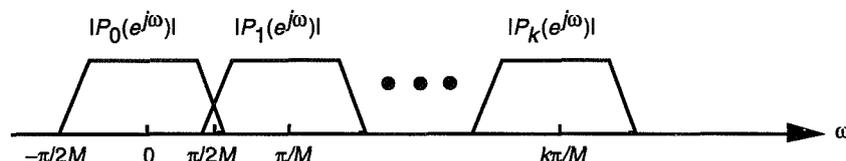


Fig. 3. Magnitude responses of $P_k(z)$.

(The definition of a DFT filter bank here is slightly different from the conventional DFT filter banks [7].) It follows that $H_k(e^{j\omega})$ is just a shift of $P_0(e^{j\omega})$ by $k\pi/M$ except for a scalar. The same holds for the synthesis filters.

We now show that with proper design of the two prototypes, this DFT filter is approximately alias free and the overall response is equivalent to a tunable multilevel filter. Moreover, the overall response can be a real-coefficient linear-phase filter as desired. Efficient implementation of the DFT filter bank will also be discussed.

A. Suppression of Aliasing Error

Consider the under-decimated system in Fig. 2, a $2M$ -channel filter bank with decimation ratio M . The suppression of aliasing error due to downsampling in the subbands can be explained pictorially. Take the first subband as an example. Because of decimation followed by expansion, there will be $M - 1$ image copies of $H_0(z)$, as shown in Fig. 4. We can see from Fig. 4 that these image copies will be suppressed if both $H_0(z)$ and $F_0(z)$ have stop-band edges less than π/M . When the spectral supports of $F_0(z)$ and the image copy of $H_0(z)$ do not overlap, the aliasing error will be suppressed to the level of the stop-band attenuation of $H_0(z)$ or $F_0(z)$, which is equivalent to the stop-band attenuation of $P_0(z)$ or $Q_0(z)$. In the other subbands, the same reasoning for aliasing suppression applies.

We now present the mathematical counterpart of the above discussion. The output $\hat{X}(z)$ is related to the input $X(z)$ by

$$\hat{X}(z) = \sum_{i=0}^{M-1} A_i(z) X(zW^{2i}) \quad (2)$$

The alias transfer function, $A_i(z)$, is defined as

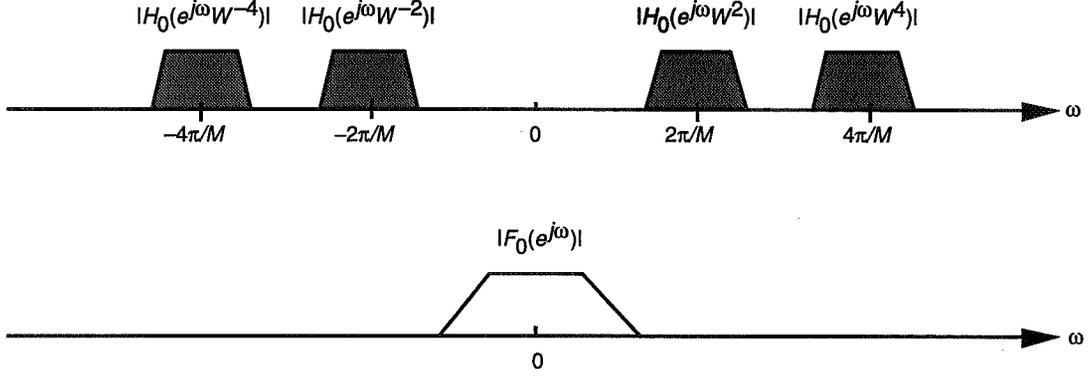


Fig. 4. Image copies of $H_0(z)$ due to decimation followed by expansion and the spectral support of $F_0(z)$.

$$A_i(z) = \frac{1}{M} \sum_{k=0}^{2M-1} H_k(zW^{2i})F_k(z) \quad (3)$$

The system in Fig. 2 is alias free if $A_i(z) = 0$ for $i = 1, 2, \dots, M-1$. With analysis filters and synthesis filters chosen as in Eq. (1), $A_i(z)$ can be written as

$$A_i(z) = \frac{1}{M} \sum_{k=0}^{2M-1} |a_k|^2 P_0(zW^{2i+k})Q_0(zW^k) \quad (4)$$

Assume the nonadjacent bands of $P_0(z)$ and $Q_0(z)$ do not overlap, i.e.,

$$P_0(e^{j\omega}W^{2i})Q_0(e^{j\omega}) \approx 0, \quad i = 1, \dots, M-1 \quad (5)$$

This assumption is reasonable if $P_0(z)$ and $Q_0(z)$ have stop-band edges less than π/M and large enough stop-band attenuation. Equation (5) gives us

$$H_k(zW^{2i})F_k(z) \approx 0, \quad k = 0, 1, \dots, 2M-1, \quad i = 1, 2, \dots, M-1$$

which implies $A_i(z) \approx 0, i \neq 0$. We conclude that the DFT filter bank is almost alias free. Also notice that the degree of alias suppression improves with the stop-band attenuation of the two prototypes.

B. The Overall Response of the DFT Filter Bank

For a $2M$ -channel system decimated by M as shown in Fig. 2, the distortion function $T(z)$ can be expressed as [7]

$$T(z) = \frac{1}{M} \sum_{k=0}^{2M-1} H_k(z)F_k(z) \quad (6)$$

Let $R_0(z) = P_0(z)Q_0(z)$. Substitute the expression of $H_k(z)$ and $F_k(z)$ in Eq. (1); then

$$T(e^{j\omega}) = \frac{1}{M} \sum_{k=0}^{2M-1} |a_k|^2 R_0(e^{j(\omega-k\pi/M)}) \quad (7)$$

When $R_0(z)$ is a Nyquist ($2M$) filter, it can be shown the addition of $|a_k|^2 R_0(e^{j(\omega-k\pi/M)})$ in Eq. (7) will not result in any bumps or dips in the response of $T(z)$ because of the Nyquist property of $R_0(z)$. The definition of a Nyquist filter is given in Section I. A detailed explanation can be found in [7].

With Eq. (7), we can plot a typical magnitude response of $T(z)$, as in Fig. 5, which shows that the overall response is equivalent to a multilevel filter. Since the value of a_k can be chosen freely, $T(z)$ is actually a tunable multilevel filter. In particular, we can get lowpass filters with stop-band edges adjustable in integer multiples of π/M .

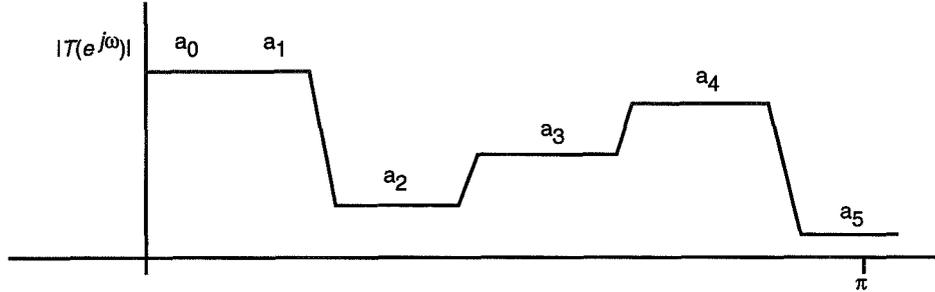


Fig. 5. A typical magnitude response of $T(z)$, a multilevel filter.

Remarks:

- (1) If $R_0(z)$ is a real-coefficient filter and we choose $a_k = a_{2M-k}$ to be some real number, $k = 1, 2, \dots, M$, it can be verified that the resulting $T(z)$ is also real-coefficient.
- (2) Let $R_0(z)$ be linear phase with order N_r , a multiple of M . In this case, $R_0(zW^k)$ is linear phase. By Eq. (7), this implies that $T(z)$ has linear phase.

Summarizing, we have shown that if $R_0(z)$ is Nyquist ($2M$) and Eq. (5) is valid, the DFT filter in Fig. 2 is nearly alias free and the overall response is equivalent to a tunable multilevel filter.

C. Implementation of the DFT Filter Bank

There exists an efficient implementation for the DFT filter bank. To see this, express the prototype $P_0(z)$ in polyphase representation, i.e.,

$$P_0(z) = \sum_{i=0}^{2M-1} E_i(z^{2M})z^{-i} \tag{8}$$

where $E_i(z)$ is the i th type 1 polyphase component of $P_0(z)$ [7]. The analysis filters can be rewritten as

$$H_k(z) = a_k P_0(zW^k) = a_k \sum_{i=0}^{2M-1} E_i(z^{2M})W^{-ki}z^{-i}, \quad k = 0, 1, \dots, 2M-1 \tag{9}$$

Let

$$\mathbf{h}(z) = [H_0(z) H_1(z) \dots H_{2M-1}(z)]^T \tag{10}$$

The matrix representation of Eq. (9) is

$$\mathbf{h}(z) = \begin{pmatrix} a_0 & 0 & \cdots & 0 \\ 0 & a_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{2M-1} \end{pmatrix} \mathbf{W}^* \begin{pmatrix} E_0(z^{2M}) & 0 & \cdots & 0 \\ 0 & E_1(z^{2M}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & E_{2M-1}(z^{2M}) \end{pmatrix} \mathbf{e}_{2M}(z) \quad (11)$$

Observing Eq. (11), we can draw the polyphase implementation of the analysis bank as in Fig. 6. The implementation cost is that of the prototype filter $P_0(z)$ plus a DFT matrix. The same holds for the synthesis bank. The computational complexity of the analysis bank is comparable to that of the analysis prototype filter plus one DFT matrix. Notice that all the computations involved in the filter bank are performed after the M -fold decimators; lower rate and lower complexity are achieved at the same time.

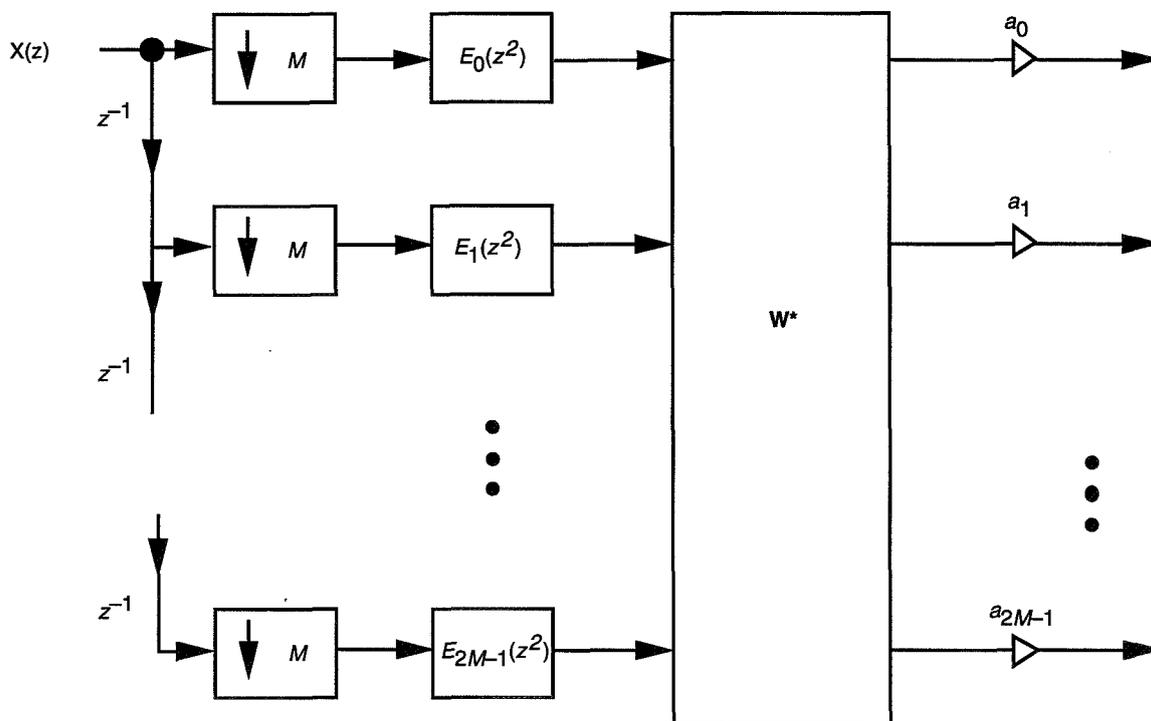


Fig. 6. Efficient implementation of the analysis bank of the $2M$ -channel DFT filter bank. The DFT matrix, \mathbf{W} , is of size $2M \times 2M$.

III. Cosine Modulated Filter Banks and Application in Tunable Multilevel Filtering

In the DFT filter bank described in the previous section, the analysis and synthesis filters have complex coefficients. If it is desirable for the individual filters to have real coefficients, then we can use the new under-decimated cosine modulated filter bank to be discussed in this section.

The system in Fig. 2 is said to be a cosine modulated filter bank if all analysis and synthesis filters are generated by cosine or sine modulation of one or two prototype filters. In this section, we introduce two classes of new under-decimated cosine modulated filter banks. The systems are nearly alias free. Aliasing error decreases as the stop-band attenuation of the prototype increases. Individual analysis and synthesis filters have real coefficients. We can design the prototypes so that the overall response of the

filter bank is a linear-phase multilevel filter. Furthermore, there exists efficient implementation of this cosine modulated filter bank. The implementation cost of the analysis bank is that of the prototype filter plus two DCT matrices (Appendix). The complexity of an $M \times M$ DCT matrix is only of the order $M \log(M)$ [11]. The same holds true for the synthesis bank. Two types of filter stacking can be applied in the new under-decimated cosine modulated filter bank. In our discussion, a cosine modulated filter bank will be identified as type 1 or type 2 according to the stacking of its analysis and synthesis filters.

A. Type 1 Cosine Modulated Filter Bank

1. Construction of the Type 1 Cosine Modulated Filter Bank. In the cosine modulated filter bank, all analysis and synthesis filters have real coefficients. Each filter has positive and negative spectral occupancy as opposed to single-sided spectral occupancy in the DFT filter bank. This incurs a problem that we do not have in the DFT filter bank. Details and a proposed solution of this new problem will be given below.

Let $P_0(z)$ and $Q_0(z)$ be respectively the prototype filters of the analysis bank and the synthesis bank. The definitions of $P_k(z)$ and $Q_k(z)$ are as in Section II. To get real-coefficient analysis and synthesis filters from the prototypes, we can combine $P_k(z)$ and $P_{-k}(z)$:

$$H_k(z) = a_k P_k(z) + a_k^* P_{-k}(z), \quad F_k(z) = b_k Q_k(z) + b_k^* Q_{-k}(z), \quad k = 1, 2, \dots, M-1$$

Since $P_0(z)$ and $P_M(z)$ are already real filters, we can directly choose

$$H_k(z) = 2a_k P_k(z), \quad F_k(z) = 2b_k Q_k(z), \quad k = 0 \text{ or } M$$

Figure 7 shows the spectral supports of the analysis filters for the type 1 cosine modulated filter bank. The stacking of the spectral supports of the synthesis filters is similar.

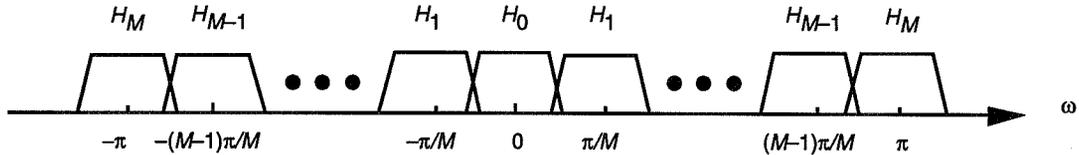


Fig. 7. Normalized magnitude responses of the analysis filters of the type 1 cosine modulated filter bank.

Aliasing error due to M -fold decimation in the 0th and the M th subband can be suppressed on the synthesis side as we did in the DFT filter bank. The situation in the other subbands is different because now $H_k(z)$ and $F_k(z)$, $k = 1, 2, \dots, M-1$, are bandpass filters. Referring to Fig. 8, decimation by M followed by expansion by M in the subbands will cause one image copy of $P_k(z)$ to overlap completely with $Q_{-k}(z)$, $k = 1, 2, \dots, M-1$. This type of aliasing error cannot be suppressed in the synthesis bank.

Our solution to this problem is to introduce a second subsystem that has exactly the same aliasing error to cancel the existing one. Let the second subsystem have analysis filters $H'_k(z)$ and synthesis filters $F'_k(z)$, $k = 1, 2, \dots, M-1$. To create the same aliasing error, the filters of the second subsystem are required to have similar stacking as that of the first subsystem. In particular,

$$H'_k(z) = a'_k P_k(z) + a'^*_k P_{-k}(z), \quad F'_k(z) = b'_k Q_k(z) + b'^*_k Q_{-k}(z), \quad k = 1, 2, \dots, M-1$$

The configuration of the analysis filters in the second subsystem is shown in Fig. 9. Notice that the spectral occupancy of $H_0(z)$ and $H_M(z)$ are not covered in the second subsystem.

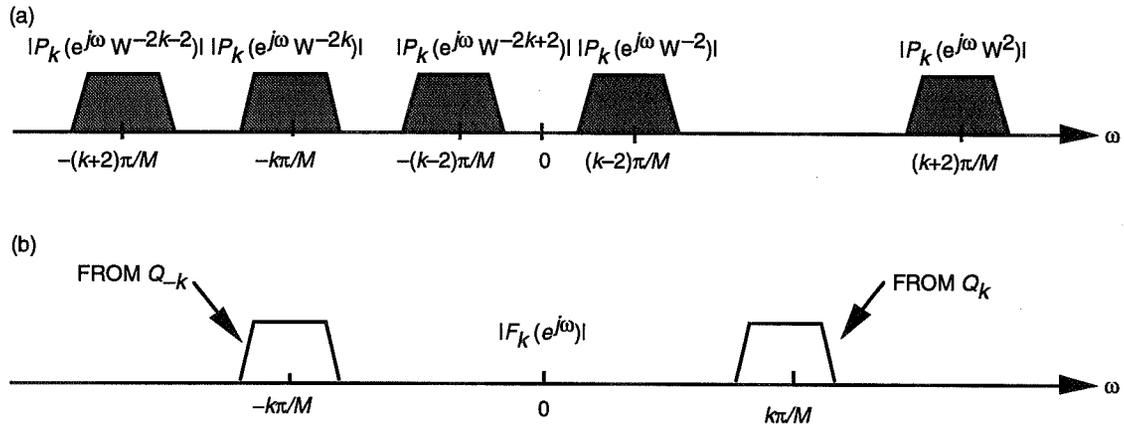


Fig. 8. Type 1 cosine modulated filter bank: (a) image copies of $P_k(z)$ due to decimation followed by expansion and (b) the spectral support of $F_k(z)$.

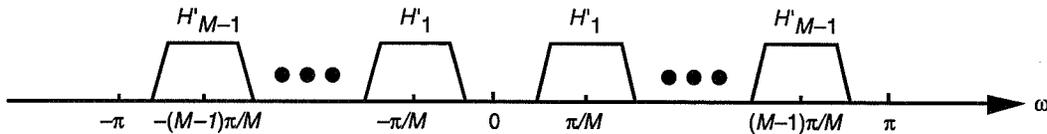


Fig. 9. Normalized magnitude responses of the analysis filters of the second subsystem in the type 1 cosine modulated filter bank.

The setup of the new system is now complete and is shown in Fig. 10. It is a connection of two subsystems. The first subsystem has $M + 1$ channels, and the second subsystem has $M - 1$ channels. The whole system is under-decimated; it has $2M$ channels but is decimated only by M . The analysis and synthesis filters can be summarized as follows:

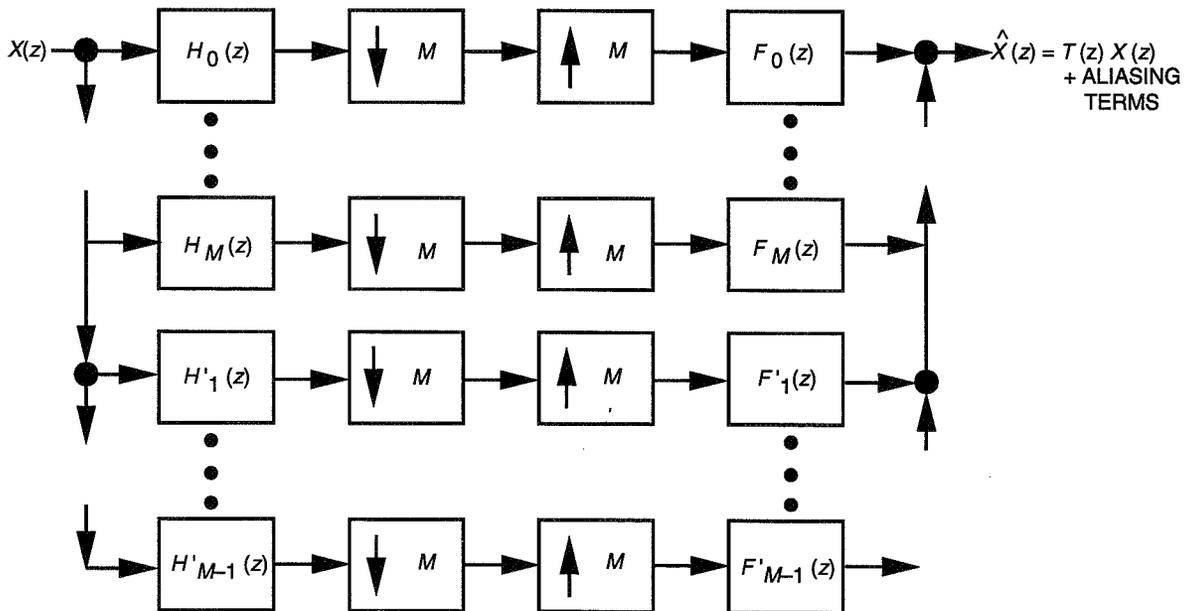


Fig. 10. The setup for the type 1 under-decimated cosine modulated filter bank.

$$\begin{aligned}
H_k(z) &= 2a_k P_k(z), & k &= 0, M \\
H_k(z) &= a_k P_k(z) + a_k^* P_{-k}(z), & k &= 1, 2, \dots, M-1 \\
H'_k(z) &= a'_k P_k(z) + a'^*_k P_{-k}(z), & k &= 1, 2, \dots, M-1 \\
F_k(z) &= 2b_k Q_k(z), & k &= 0, M \\
F_k(z) &= b_k Q_k(z) + b_k^* Q_{-k}(z), & k &= 1, 2, \dots, M-1 \\
F'_k(z) &= b'_k Q_k(z) + b'^*_k Q_{-k}(z), & k &= 1, 2, \dots, M-1
\end{aligned} \tag{12}$$

The values of a_k , a'_k , b_k , and b'_k will be determined later.

In the following we show that with proper design of the prototypes and appropriate choices of a_k , a'_k , b_k , and b'_k , this filter bank is almost alias free. The overall response of the type 1 cosine modulated filter bank can be designed to be a linear-phase tunable multilevel filter.

2. Cancellation and Suppression of Aliasing Error. As we mentioned in the construction of filters, the aliasing error in the 0th and the M th subbands will be *suppressed* in the synthesis bank. Only the subbands with bandpass filters require alias *cancellation*. The physical picture is as follows:

Consider the k th subband, $1 \leq k \leq M-1$. Due to decimation followed by expansion, $P_k(z)$ has $M-1$ image copies and $P_{-k}(z)$ also has $M-1$ image copies. The image copies of $P_k(z)$ will be suppressed by $Q_k(z)$ provided that both $P_0(z)$ and $Q_0(z)$ have stop-band edges less than π/M and large enough stop-band attenuation. Of the $M-1$ image copies of $P_k(z)$, $M-2$ of them are in the stop band of $Q_{-k}(z)$ and, hence, will be suppressed by $Q_{-k}(z)$, as depicted in Fig. 8. However, one of the image copies of $P_k(z)$ will fall on top of the spectral support of $Q_{-k}(z)$. Only this image copy requires cancellation. In the k th subband of the second subsystem, the same aliasing occurs. It can be shown that the aliasing error of the second subsystem cancels that of the first subsystem if the values of a_k , b_k , a'_k , and b'_k are chosen properly. Mathematical proof of this claim is as follows.

With filters constructed as in Eq. (12) and the expression of alias transfer functions in Eq. (3), we have

$$A_i(z) = \frac{1}{M} \left(A_i^{(1)}(z) + A_i^{(2)}(z) + A_i^{(3)}(z) + A_i^{(4)}(z) \right) \tag{13}$$

where

$$A_i^{(1)}(z) = (a_0 b_0 + a_0^* b_0) P_0(z W^{2i}) Q_0(z) + \sum_{k=1}^{M-1} (a_k b_k + a_k^* b_k^*) P_k(z W^{2i}) Q_k(z)$$

$$A_i^{(2)}(z) = \sum_{k=1}^{M-1} (a_k^* b_k + a_k'^* b_k') P_{-k}(z W^{2i}) Q_k(z)$$

$$A_i^{(3)}(z) = \sum_{k=1}^{M-1} (a_k b_k^* + a_k' b_k'^*) P_k(z W^{2i}) Q_{-k}(z)$$

$$A_i^{(4)}(z) = (a_0^* b_0^* + a_0 b_0^*) P_0(zW^{2i}) Q_0(z) + \sum_{k=1}^{M-1} (a_k^* b_k^* + a_k' b_k'^*) P_{-k}(zW^{2i}) Q_{-k}(z)$$

Assume $P_0(z)$ and $Q_0(z)$ satisfy Eq. (5). It follows that $A_i^{(1)}(z) \approx 0$ and $A_i^{(4)}(z) \approx 0$. Let $a_k, k = 0, \dots, M$ be real and choose

$$\begin{aligned} b_k &= a_k, & k &= 0, \dots, M \\ a_k' &= -ja_k, & b_k' &= ja_k, & k &= 1, \dots, M-1 \end{aligned} \quad (14)$$

If a_k, b_k, a_k' , and b_k' are chosen as above, it can be verified that $a_k^* b_k + a_k' b_k' = 0$, which implies $A_i^{(2)}(z) = A_i^{(3)}(z) = 0$. So the condition $A_i(z) \approx 0$ is ensured, provided that Eq. (5) is valid and a_k, b_k, a_k' , and b_k' are chosen according to Eq. (14).

With Eq. (14), we can write down the time domain description of the analysis and synthesis filters for the type 1 cosine modulated filter bank. Let $p_0(n)$ be the impulse response of $P_0(z)$ and $q_0(n)$ be the impulse response of $Q_0(z)$.

$$\begin{aligned} h_k(n) &= 2a_k p_0(n) \cos(kn\pi/M), & k &= 0, 1, \dots, M \\ h_k'(z) &= 2a_k p_0(n) \sin(kn\pi/M), & k &= 1, 2, \dots, M-1 \\ f_k(z) &= 2a_k q_0(n) \cos(kn\pi/M), & k &= 0, 1, \dots, M \\ f_k'(n) &= 2a_k q_0(n) \sin(kn\pi/M), & k &= 1, 2, \dots, M-1 \end{aligned} \quad (15)$$

From the expression in Eq. (15), we can see that each individual filter is a sine or cosine modulation of the prototype filters.

3. The Magnitude Response of the Overall Response $T(z)$. Using Eqs. (12) and (6), we get

$$T(e^{j\omega}) = \frac{2}{M} \sum_{k=0}^M |a_k|^2 \left(R_0(e^{j(\omega+k\pi/M)}) + R_0(e^{j(\omega-k\pi/M)}) \right) \quad (16)$$

The above expression for the overall response is similar to that in the case of DFT filter bank, Eq. (7). If $R_0(z)$ is a Nyquist ($2M$) filter, this is a tunable multilevel filter bank like in a DFT filter bank.

4. The Phase of the Overall Response $T(z)$. The overall response $T(z)$ has linear phase provided that $R_0(z)$ is linear phase and N_r , the order of $R_0(z)$, is an even multiple of M . The reason is given below. The linear phase property of $R_0(z)$ entitles us to write

$$R_0(e^{j\omega}) = e^{-j\omega N_r/2} R(\omega) \quad (17)$$

where $R(\omega)$ is a real-valued function. Substitute Eq. (17) into Eq. (16) and we get

$$T(e^{j\omega}) = \frac{4}{M} e^{-j\omega N_r/2} \sum_{k=0}^M |a_k|^2 \left(e^{jk\pi N_r/2M} R\left(\frac{\omega - k\pi}{M}\right) + e^{-jk\pi N_r/2M} R\left(\frac{\omega + k\pi}{M}\right) \right)$$

which has linear phase when N_r is a multiple of $2M$.

Notice that if $Q_0(z)$ is the time reversed version of $P_0(z)$, i.e., $Q_0(z) = z^{-N_p} \widetilde{P}_0(z)$, then $F_k(z)$ and $F'_k(z)$ are the time-reversed version of $H_k(z)$ and $H'_k(z)$, respectively. With this choice of $Q_0(z)$, the overall response is

$$T(z) = \frac{z^{-N_p}}{M} \left(\sum_{k=0}^M H_k(z) \widetilde{H}_k(z) + \sum_{k=1}^{M-1} H'_k(z) \widetilde{H}'_k(z) \right)$$

In this case, the overall response has linear phase regardless of the order of $R_0(z)$.

5. The Phase of the Individual Analysis and Synthesis Filters. Let the prototype filter $P_0(z)$ be linear phase. If the center of symmetry of $p_0(n)$ coincides with that of the sine or cosine functions that modulate $p_0(n)$, then the resulting analysis filters also have linear phase. In the case of the type 1 cosine modulated filter bank, the condition can be further reduced. It can be verified that if N_p , the order of $P_0(z)$, is a multiple of M , every analysis filter has linear phase. The same holds for the synthesis bank.

When N_p is an even multiple of M , all the analysis filters in the second subsystem, $H'_k(z)$, $k = 1, 2, \dots, M-1$, are found to have a null at zero frequency and π . We can also verify that as N_p is an odd multiple of M , the analysis filters $H_k(e^{j\omega})$, $k = 1, 2, \dots, M-1$, have zeros at zero frequency and π . We conclude that some $M-1$ of the $2M$ analysis filters have zeros at zero frequency and π if the analysis filters have linear phase. More on this property will be addressed when we present a similar result for the type 2 cosine modulated filter bank.

Summarizing, we have shown that the filter bank in Fig. 10 is equivalent to a linear-phase tunable multilevel filter if the following two conditions hold: (1) The nonadjacent bands of $P_0(z)$ and $Q_0(z)$ do not overlap, [Eq. (5)], and (2) $R_0(z)$ is linear phase and close to a Nyquist ($2M$) filter. The implementation cost of the analysis bank, as will be shown in the Appendix, is the prototype filter $P_0(z)$ plus two DCT matrices. Complexity follows the buildup of the architecture; the computational cost of the analysis bank is that of $P_0(z)$ plus two DCT matrices working at an M -fold decimated rate.

B. Type 2 Cosine Modulated Filter Bank

In Figs. 7 and 9, we show the configuration of the analysis filter for the type 1 under-decimated cosine modulated filter bank. In the type 2 cosine modulated under-decimated filter bank, a different stacking of filters is applied. We show the new setup in Fig. 11. The filter bank can still be conceived as a connection of two subsystems, both with M channels. The spectral supports of the second set of analysis filters are exactly the same as the spectral supports of the first set of analysis filters (Fig. 12). The same holds for the synthesis bank. To be more specific, let $P_0(z)$ and $Q_0(z)$ be the two prototypes as before. Define

$$\widehat{P}_k(z) = P_0(zW^{k+0.5}) \quad \text{and} \quad \widehat{Q}_k(z) = Q_0(zW^{k+0.5}), \quad W = e^{-j\pi/M}$$

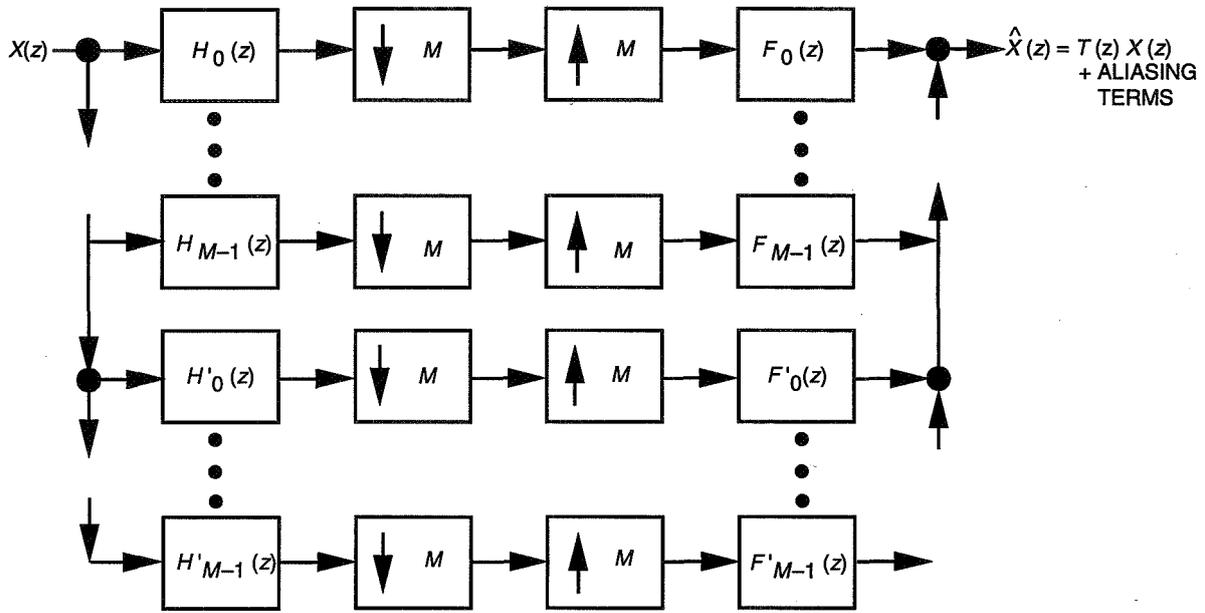


Fig. 11. The setup for the type 2 under-decimated cosine modulated filter bank.

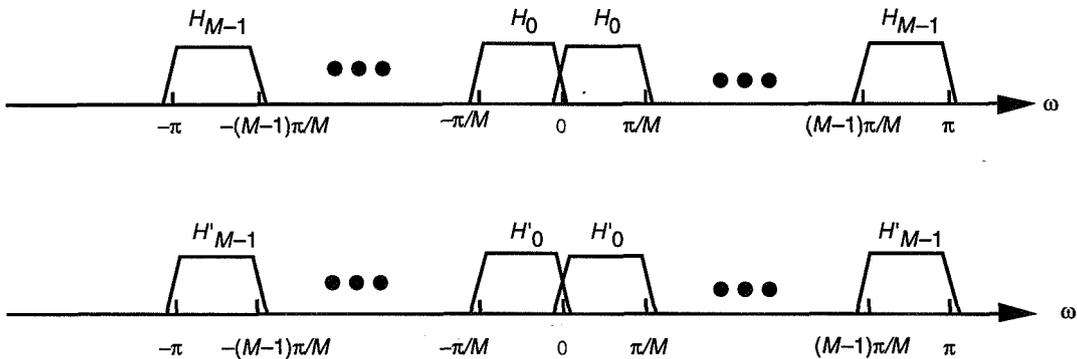


Fig. 12. Normalized magnitude responses of the analysis filters of the type 2 cosine modulated filter bank.

Spectral support of $\hat{P}_k(z)$ is shown in Fig. 13. It is similar for $\hat{Q}_k(z)$. Notice the difference between the definition of $\hat{P}_k(z)$ and the definition of $P_k(z)$ in the type 1 cosine modulated filter bank, Fig. 3; if the type 1 system and the type 2 system have the same analysis prototype, then there is on the unit circle $P'_k(z)$ a shift of $P_k(z)$ by $\pi/2M$, and this comes from the extra $W^{0.5}$ in the preceding equation.

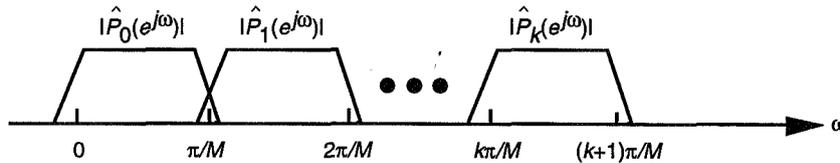


Fig. 13. Magnitude responses of $\hat{P}_k(z)$.

We choose the analysis and synthesis filters, similar to the construction of the type 1 system, as follows:

$$\begin{aligned}
H_k(z) &= a_k \widehat{P}_k(z) + a_k^* \widehat{P}_{-k}(z), & k = 0, 1, \dots, M-1 \\
H'_k(z) &= -j \left(a_k \widehat{P}_k(z) - a_k^* \widehat{P}_{-k}(z) \right), & k = 0, 1, \dots, M-1 \\
F_k(z) &= b_k \widehat{Q}_k(z) + b_k^* \widehat{Q}_{-k}(z), & k = 0, 1, \dots, M-1 \\
F'_k(z) &= j \left(b_k \widehat{Q}_k(z) - b_k^* \widehat{Q}_{-k}(z) \right), & k = 0, 1, \dots, M-1
\end{aligned} \tag{18}$$

In this case, we found that the overall response is still equivalent to a tunable multilevel filter, and the argument of alias cancellation continues to hold after minor adjustments. More details are given below. The implementation cost and the computational complexity are the same as for the type 1 system. This can be proved in a manner similar to that used for the type 1 system in the Appendix.

1. Cancellation and Suppression of Aliasing Error. The cancellation and suppression of aliasing error is very similar to that of the type 1 cosine modulated filter bank. The image copies of $\widehat{P}_k(z)$, Fig. 14(a), can be suppressed by $\widehat{Q}_k(z)$, provided that $P_0(z)$ and $Q_0(z)$ have large enough stop-band attenuation and their nonadjacent bands do not overlap. We can see from Fig. 14(a) that two image copies of $\widehat{P}_k(z)$ overlap with $\widehat{Q}_{-k}(z)$. Figure 14(b) shows the resulting aliasing that cannot be suppressed by the synthesis filters. As in the type 1 case, the second subsystem contributes another alias term, which can be used to cancel the alias from the first subsystem. This can be verified by directly substituting the analysis and synthesis filter in Eq. (18) into the definition of alias transfer functions in Eq. (3).

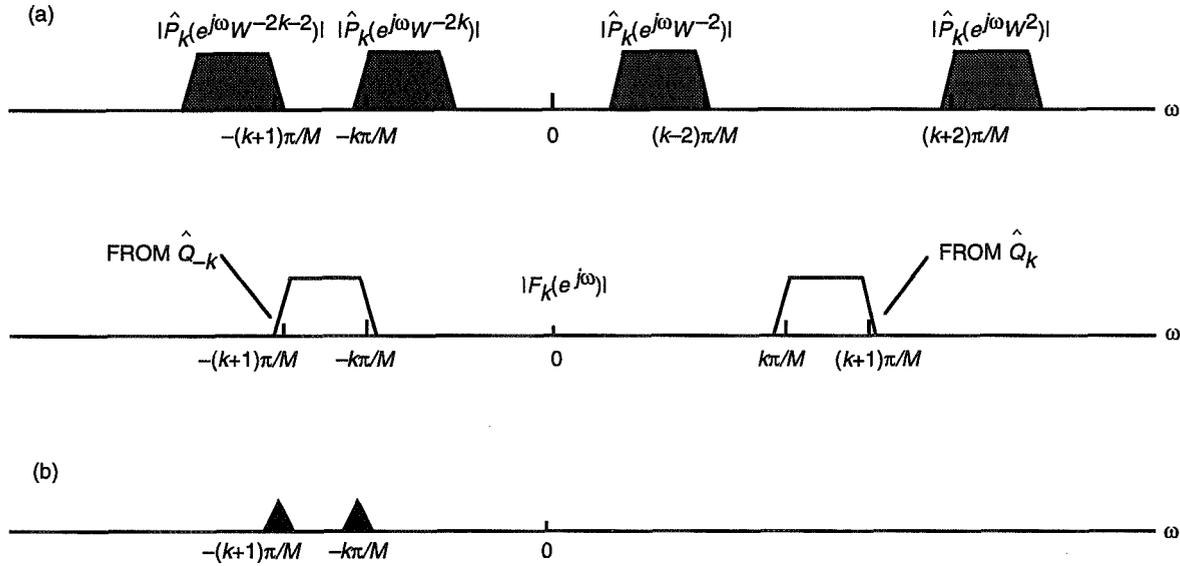


Fig. 14. Type 2 cosine modulated filter bank: (a) image copies of $\widehat{P}_k(z)$ due to decimation followed by expansion and the spectral support of $F_k(z)$ and (b) residual alias component that cannot be suppressed by $\widehat{Q}_k(z)$.

2. The Magnitude Response of the Overall Response $T(z)$. The overall response $T(z)$ in the type 2 cosine modulated filter bank can still be expressed in a form similar to that in the type 1 case. Let $R_0(z)$ be defined as before, that is, $R_0(z) = P_0(z)Q_0(z)$. With analysis and synthesis filters chosen as in Eq. (18), the overall response is

$$T(z) = \frac{2}{M} \sum_{k=-M}^{M-1} |a_k|^2 R_0 \left(e^{j(\omega - \pi(k+0.5)/M)} \right)$$

As in the DFT filter bank, the magnitude response of $T(z)$ will resemble that of a multilevel filter if $R_0(z)$ is a Nyquist ($2M$) filter or close to a Nyquist ($2M$) filter. Again, we can tune the values of a_k as desired; the overall response is actually a tunable multilevel filter.

3. The Phase Response of the Overall Response $T(z)$. In the type 1 cosine modulated filter bank, the phase of $T(z)$ is linear if $R_0(z)$ is linear phase and N_r , the order of $R_0(z)$, is an even multiple of M . It is still true in the type 2 system. We can verify this by following the same steps as in the type 1 case.

4. The Phase Responses of the Individual Filters. Each analysis filter in the type 1 cosine modulated filter bank has linear phase if $P_0(z)$ has linear phase and N_p , the order of $P_0(z)$, is an even multiple of M . The same holds for the synthesis bank. This is still true in the type 2 case. However, the type 2 system is found to have the following two additional properties when the individual filters are linear-phase:

- (i) The filters $H'_0(z)$ and $F'_0(z)$ have zeros at zero frequency, i.e., $H'_0(e^{j0}) = F'_0(e^{j0}) = 0$. The highpass filters $H'_{M-1}(z)$ and $F'_{M-1}(z)$ have zeros at π .
- (ii) Either $H_0(z)$ or $F_0(z)$ has a bump at zero frequency and either $H_{M-1}(z)$ or $F_{M-1}(z)$ has a bump at π .

Property (i): Let $P_0(z)$ be linear phase and a_0 be real. Referring to the construction of filters in Eq. (18), the impulse responses of $H_k(z)$ and $H'_k(z)$ are, respectively, $h_k(n)$ and $h'_k(n)$:

$$h_k(n) = 2a_0p_0(n) \cos \left(\pi(k+0.5)\frac{n}{M} \right)$$

$$h'_k(n) = 2a_0p_0(n) \sin \left(\pi(k+0.5)\frac{n}{M} \right)$$

Their time reversed versions are

$$h_k(N_p - n) = 2a_0p_0(n) \cos \left(\pi(k+0.5)\frac{n}{M} - \pi(k+0.5)\frac{N_r}{M} \right)$$

$$h'_k(N_p - n) = -2a_0p_0(n) \sin \left(\pi(k+0.5)\frac{n}{M} - \pi(k+0.5)\frac{N_r}{M} \right)$$

For the above two filters to have linear phase, N_r is necessarily a multiple of $2M$. If N_r is an even multiple of $2M$, then $h_k(n) = h_k(N_p - n)$ and $h'_k(n) = -h'_k(N_p - n)$. When N_r is an odd multiple of $2M$, $H_k(z)$ becomes odd symmetric and $H'_k(z)$ becomes even symmetric. Without loss generality, we can assume that N_r is an even multiple of $2M$.

Notice that an even N_p , $h'_k(n) = -h'_k(N_p - n)$ implies that $H'_k(e^{j0}) = H'_k(e^{j\pi}) = 0$ [7]. In particular, the lowpass filter $H'_0(z)$ has a notch at zero frequency, and the highpass filter $H'_{M-1}(z)$ has a notch at π .

The other filters, $H'_k(z)$, $k = 1, 2, \dots, M - 2$, are all bandpass filters; zeros at zero frequency do not have a significant effect on their shapes. The same holds true for the synthesis bank.

In the type 1 cosine modulated filter bank, we obtained similar results. In that case, when the analysis filters have linear phase, some $M - 1$ of them have zeros at zero frequency and π . But those are all bandpass filters; zeros at zero frequency and π are not of particular importance.

Property (ii): For the flatness of the passband of $T(z)$, $R_0(z)$ is required to be a Nyquist ($2M$) filter or close to a Nyquist ($2M$) filter, as we have discussed previously. The Nyquist property of $R_0(z)$ ensures that

$$\sum_{k=-M}^{M-1} R(e^{j\omega} W^{k+0.5}) = c$$

for some constant c . Without loss of generality, we can consider $c = 1$. At zero frequency, only the term $k = -1$ and the term $k = 0$ contribute significantly, so

$$R_0(e^{j\pi/2M}) + R_0(e^{-j\pi/2M}) \approx 1 \quad (19)$$

Since $R_0(z)$ is a real filter, we have

$$|R_0(e^{j\pi/2M})| = |R_0(e^{-j\pi/2M})| \quad (20)$$

Combining Eqs. (19) and (20) leads to $|R_0(e^{j\pi/2M})| \geq 0.5$, i.e., $|P_0(e^{j\pi/2M}) Q_0(e^{j\pi/2M})| \geq 0.5$. This in turn implies $|P_0(e^{j\pi/2M})| \geq \sqrt{0.5}$ or $|Q_0(e^{j\pi/2M})| \geq \sqrt{0.5}$. Suppose $|P_0(e^{j\pi/2M})| \geq \sqrt{0.5}$. Moreover, by the linear phase property of $P_0(z)$, we can write $P_0(e^{j\omega}) = e^{-j\omega N_p/2} P(\omega)$, where $P(\omega)$ is a real-valued function. With this expression, $H_0(e^{j0})$ becomes

$$H_0(e^{j0}) = e^{-j\omega N_p} \left(e^{jN_p\pi/4M} P \frac{-\pi}{2M} + e^{-jN_p\pi/4M} P \frac{\pi}{2M} \right)$$

From the discussion of property (i), we know N_p is an even multiple of $2M$. This gives us

$$|H_0(e^{j0})| = 2|P \frac{\pi}{2M}| \geq \sqrt{2}$$

This means that $H_0(z)$ has a bump of about 3 dB at zero frequency. If we assume $|Q_0(e^{j\pi/2M})| \geq \sqrt{0.5}$, then $F_0(z)$ will have a bump of about 3 dB at zero frequency. The filters $H_{M-1}(z)$ and $F_{M-1}(z)$ are shifts of $H_0(z)$ and $F_0(z)$ by π ; if either $H_0(z)$ or $F_0(z)$ has a bump at zero frequency, then either $H_{M-1}(z)$ or $F_{M-1}(z)$ has a bump at zero frequency.

In the above derivation, we assume a_k is a real number. The readers will find that for complex a_k , properties (i) and (ii) continue to hold after minor adjustments of the above argument.

It is noteworthy that the null of $H'_0(z)$ and $F'_0(z)$ at zero frequency and the bump of $H_0(z)$ or $F_0(z)$ at zero frequency do not affect the overall response. The overall response is a summation of responses of the first subsystem and the second subsystem. The bump from the first subsystem compensates for the null of the second subsystem to ensure the flatness of the overall response $T(z)$.

IV. Design Procedures

We now discuss four techniques for designing $P_0(z)$ and $Q_0(z)$. The resulting $R_0(z) = P_0(z)Q_0(z)$ is close to a Nyquist ($2M$) filter and, at the same time, the nonadjacent bands of $P_0(z)$ and $Q_0(z)$ do not overlap, i.e., $P_0(z)$ and $Q_0(z)$ satisfy Eq. (5). The prototypes $P_0(z)$ and $Q_0(z)$ designed by these methods can be applied to both the DFT filter banks and the two classes of cosine modulated filter banks.

A. Eigenfilter Design

The basic idea of this method is the following. Let $P_0(z)$ be a lowpass filter with a stop-band edge less than π/M . (Any lowpass filter design technique can be used for designing $P_0(z)$.) First we find the condition for $Q_0(z)$ such that the product $R_0(z) = P_0(z)Q_0(z)$ is Nyquist ($2M$). Then we use an eigenfilter design and incorporate this condition to design $Q_0(z)$ [7]. The condition is derived below.

Let N_r , N_p , and N_q be, respectively, the orders of $R_0(z)$, $P_0(z)$, and $Q_0(z)$. Let the impulse responses of $R_0(z)$, $P_0(z)$, and $Q_0(z)$ be $r_0(n)$, $p_0(n)$ and $q_0(n)$, respectively. Then $r_0(n)$ is the convolution of $p_0(n)$ and $q_0(n)$, i.e.,

$$r_0(n) = \sum_{m=0}^{N_q} p_0(n-m)q_0(m) \quad (21)$$

Given $p_0(n)$, we want to constrain $q_0(n)$ so that $r_0(n)$ is a Nyquist ($2M$) filter. In this case, $q_0(n)$ must satisfy

$$\sum_{m=0}^{N_q} p_0(2Mn-m)q_0(m) = c\delta(n-n_0) \quad (22)$$

for some n_0 and c . Constant c is a scalar factor; we can consider $c = 1$. The equivalent of Eq. (22) in matrix form is

$$\mathbf{r} = \mathbf{P}\mathbf{q} \quad (23)$$

where \mathbf{r} is a $(\lfloor(N_p + N_q)/2M\rfloor + 1) \times 1$ column vector with the n_0 th entry equal to 1 and all others equal to 0, \mathbf{P} is a $(\lfloor(N_p + N_q)/2M\rfloor + 1) \times (N_q + 1)$ matrix with

$$[\mathbf{P}]_{mn} = \begin{cases} p_0(2Mm-n+1), & \text{if } 2Mm-n+1 \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

and $\mathbf{q} = [q_0(0) \ q_0(1) \ \cdots \ q_0(N_q)]^T$.

The condition in Eq. (23) can be easily incorporated in the eigenfilter approach [7].

B. A Shortcut Design

Let $P_0(z)$ be a Nyquist ($2M$) lowpass filter with zero phase. Let $Q_0(z)$ also be zero phase with its passband covering $P_0(z)$ as depicted in Fig. 15, i.e., $Q_0(e^{j\omega}) \approx 1$ in the passband and the transition band of $P_0(e^{j\omega})$. Because $R_0(z)$ differs from $P_0(z)$ only in the stop band where the magnitude of $P_0(z)$ is small, we have

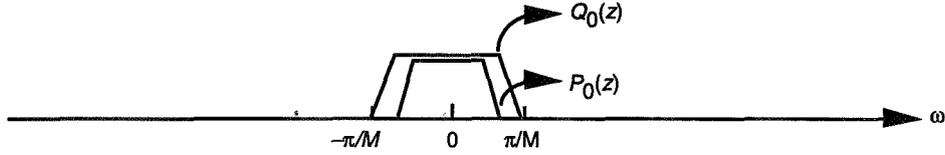


Fig. 15. The passband of $Q_0(z)$ covers the passband and transition band of $P_0(z)$.

$$R_0(z) \approx P_0(z) \quad (24)$$

In this case, $R_0(z)$ will still be close to a Nyquist ($2M$) lowpass filter.

In the above discussion, $P_0(z)$ and $Q_0(z)$ are zero phase and, hence, noncausal. Some delays can be added to make the filter bank causal, since both prototypes are FIR. In the two following methods, the synthesis prototype $Q_0(z)$ is chosen as the time-reversed version of $P_0(z)$, i.e., $Q_0(z) = z^{-N_p} \tilde{P}_0(z)$, where N_p is the order of $P_0(z)$.

C. Nonlinear Optimization of $P_0(z)$

The filter coefficients of $P_0(z)$ can be directly optimized to minimize the stop-band energy of $P_0(z)$ subject to the Nyquist condition in Eq. (22). The stop-band energy of $P_0(z)$ is

$$\phi_s = \int_{(\pi/2M)+\epsilon}^{\pi} |P_0(e^{j\omega})|^2 d\omega$$

But with $Q_0(z) = z^{-N_p} P_0(z)$, the right-hand side of Eq. (22) becomes a quadratic form of $p_0(n)$. To minimize ϕ_s , we need a nonlinear optimization package that can incorporate nonlinear constraints. The detail of this design technique is documented in [12].

Instead of optimizing the coefficients of $P_0(z)$ with nonlinear constraints, we can use a objective function to reflect how close $R_0(z)$ is to a Nyquist ($2M$) filter. A possible objective function is

$$\phi_p = \int_0^{\pi/M} \left(|P_0(e^{j\omega})|^2 + |P_0(e^{j(\omega - (\pi/M))})|^2 - 1 \right)^2 d\omega$$

Combining both ϕ_s and ϕ_p , let

$$\phi = \alpha \phi_p + (1 - \alpha) \phi_s$$

where $0 < \alpha < 1$. The new objective function ϕ can be minimized using the usual nonlinear optimization programs (e.g., [13]) without any constraints.

D. Kaiser Window Design

In this approach, $P_0(z)$ is a lowpass filter with an even order obtained through Kaiser window design. That is,

$$p_0(n) = \frac{\sin(\omega_c n)}{\pi n} w(n)$$

where $w(n)$ is a Kaiser window [7]. After we choose the stop-band attenuation and the width of the transition band, the length of the window can be estimated by a formula developed by Kaiser. In this case, the window is completely determined [7]. The cutoff frequency ω_c of the ideal filter is the only free parameter left for the design of $P_0(z)$.

According to Eq. (22), we choose a simple objective function, ϕ_{Kaiser} , to reflect the closeness of $R_0(z)$ to a Nyquist ($2M$) filter:

$$\phi_{Kaiser} = \max_n |r_0(2Mn) - \delta(n - n_0)|$$

Experiments show that ϕ_{Kaiser} is a convex function of ω_c ; we can adjust the parameter ω_c to find the best $P_0(z)$, which yields the smallest ϕ_{Kaiser} . Design examples will be given in the next section.

E. Comparison of the Four Design Methods

Of all the four approaches stated above, the shortcut design and the Kaiser window design are probably the easiest. In the shortcut design approach, we only need to design two lowpass filters with certain specifications described above. In the Kaiser window design, only the parameter ω_c needs to be optimized, yet no nonlinear optimization is involved. Also, the value of the objective function ϕ_{Kaiser} at a given ω_c can be computed easily.

V. Design Example

We now present one design example of the under-decimated system. The type 1 cosine modulated filter bank is used in this example. The Kaiser window approach is adopted for the design of the prototype filter.

Example 1: Tunable Multilevel Filter. The system has 20 channels. In this case, $M = 10$. The analysis bank prototype filter $P_0(z)$ is linear phase with order $N_p = 120$, stop-band attenuation 85 dB, passband edge $\omega_p = 0.04\pi$, and stop-band edge $\omega_s = 0.099\pi$. The synthesis bank prototype $Q_0(z)$ is chosen as the time-reversed version of $P_0(z)$. As elaborated in Section III, the resulting overall response will have linear phase. Figure 16(a) shows the magnitude response of $P_0(z)$. The normalized magnitude responses of the analysis filters are shown in Fig. 16(b) and (c). The synthesis filters are time-reversed versions of the analysis filters; the magnitude responses of the synthesis filters are the same as those of analysis filters.

After designing the prototype filters, the values of a_k can be changed freely to obtain the desired overall response, $T(z)$. For instance, we set $a_0 = a_1 = 1$, $a_2 = a_3 = a_4 = 0$, $a_5 = a_6 = a_7 = 0.7$, and $a_8 = a_9 = a_{10} = 0.3$. The magnitude response of the resulting $T(z)$ is plotted in Fig. 16(b). Since $T(z)$ has linear phase, we did not show the phase response. The corresponding dB plot of Fig. 16(d) is shown in Fig. 16(e).

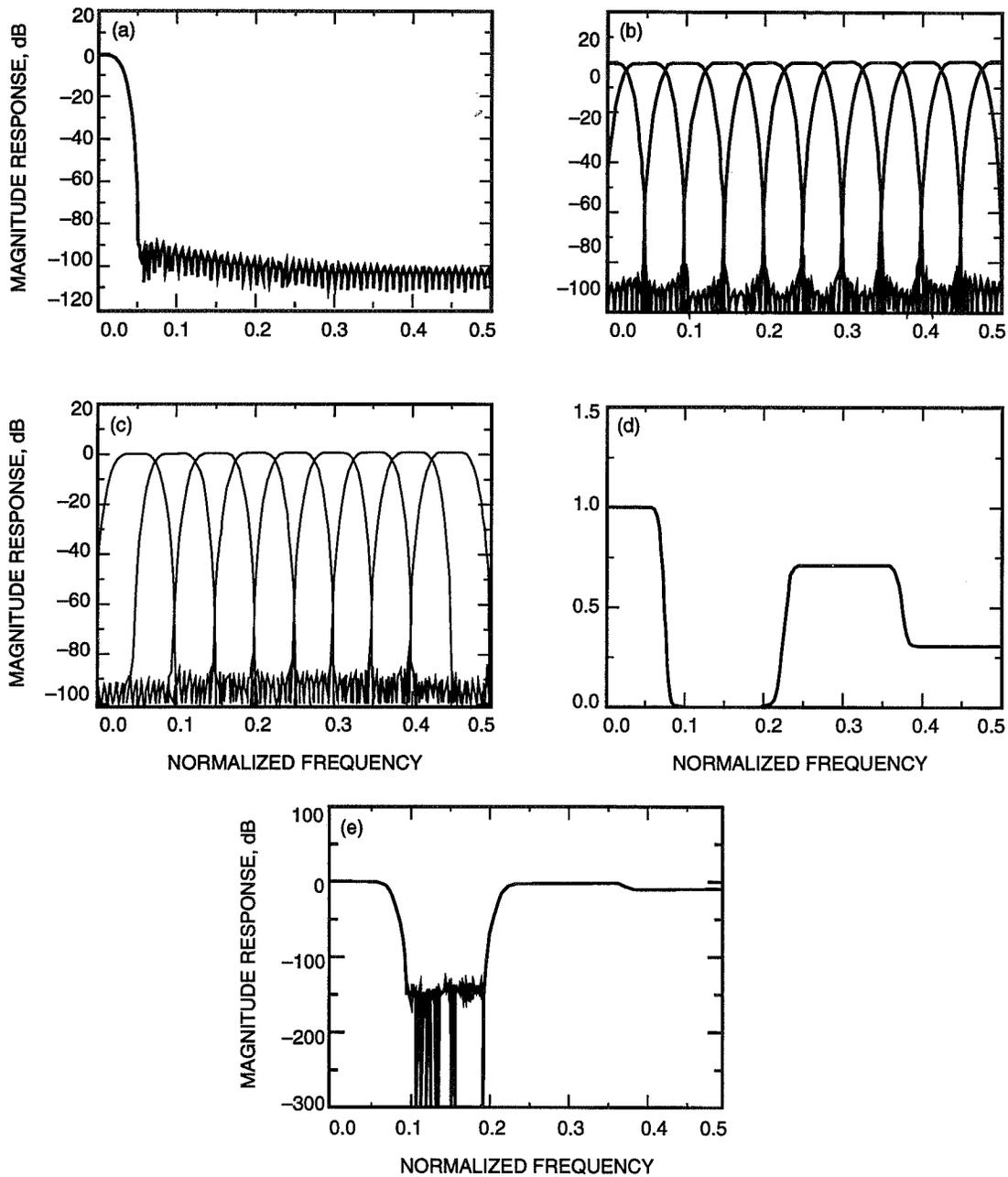


Fig. 16. Example: (a) the magnitude response of the prototype filter, $P_0(z)$, (b) the normalized magnitude responses of the analysis filters in the first subsystem, (c) the normalized magnitude responses of the analysis filters in the second subsystem, (d) the magnitude response of the overall response $T(z)$, and (e) the magnitude response of the overall response $T(z)$ in a dB plot.

References

- [1] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, New Jersey: Prentice Hall, 1983.
- [2] M. J. T. Smith and III. T. P. Barnwell, "A New Filter Bank Theory for Time-Frequency Representation," *IEEE Trans. on Acoustic, Speech and Signal Processing*, vol. ASSP-35, pp. 314-327, March 1987.

- [3] M. Vetterli, "A Theory of Multirate Filter Banks," *IEEE Trans. on Acoustic, Speech and Signal Processing*, vol. ASSP-35, pp. 356–372, March 1987.
- [4] P. P. Vaidyanathan, "Quadrature Mirror Filter Banks, M-Band Extensions and Perfect Reconstruction Techniques," *IEEE ASSP Magazine*, vol. 4, pp. 420, June 1987.
- [5] H. J. Nussbaumer, "Pseudo QMF Filter Bank," *IBM Tech. Disclosure Bulletin*, vol. 24, pp. 3081–3087, November 1981.
- [6] J. H. Rothweiler, "Polyphase Quadrature Filters, a New Subband Coding Technique," *Proc. of the IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, pp. 1980–1983, April 1983.
- [7] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, New Jersey: Prentice Hall, 1993.
- [8] P. P. Vaidyanathan, "Orthonormal and Biorthonormal Filter Banks as Convolver, and Convolutional Coding Gain," *IEEE Tran. on Signal Processing*, vol. SP-41, pp. 2110–2130, June 1993.
- [9] C. S. Burrus. "Block Implementation of Digital Filters," *IEEE Trans. on Circuit Theory*, vol. CT-18, pp. 697–701, November 1971.
- [10] M. Vetterli, "Running FIR and IIR Filtering Using Multirate Filter Banks," *IEEE Trans. on Acoustic, Speech and Signal Processing*, vol. ASSP-36, pp. 730–738, May 1988.
- [11] P. Yip and K. R. Rao, "Fast Discrete Transforms," *Handbook of Digital Signal Processing*, edited by D. F. Elliott, San Diego, California: Academic Press, 1987.
- [12] T. Nguyen., "Near-Perfect-Reconstruction Pseudo-QMF Banks," *IEEE Tran. on Signal Processing*, vol. SP 41, pp. 2110–2130, December 1993.
- [13] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes*, Cambridge: Cambridge University Press, 1989.
- [14] Y.-P. Lin and P. P. Vaidyanathan, "Linear Phase Cosine Modulated Maximally Decimated Filter Banks With Perfect Reconstruction," Technical Report, California Institute of Technology, Pasadena, California, November 1993.

Appendix

Implementation of the Type 1 Cosine Modulated Filter Bank

In this appendix, we prove that the implementation cost of the analysis bank of the type 1 cosine modulated filter bank is that of the analysis prototype filter plus two DCT matrices.

Let

$$P_0(z) = \sum_{n=0}^{2M-1} G_n(z^{2M}) z^{-n}$$

where $G_n(z)$ is the n th type 1 polyphase component of $P_0(z)$. Then

$$P_k(z) = \sum_{n=0}^{2M-1} G_n(z^{2M}) z^{-n} W^{-kn} \quad (\text{A-1})$$

Rewriting analysis filters in Eq. (13) in terms of the polyphase components of $P_0(z)$ with a_k, b_k, a'_k , and b'_k as in Eq. (15), we obtain

$$H_k(z) = 2 \sum_{n=0}^{2M-1} a_k G_n(z^{2M}) z^{-n} \cos\left(\frac{\pi}{M} kn\right), \quad k = 0, 1, \dots, M$$

$$H'_k(z) = 2 \sum_{n=0}^{2M-1} a_k G_n(z^{2M}) z^{-n} \sin\left(\frac{\pi}{M} kn\right), \quad k = 1, 2, \dots, M-1 \quad (\text{A-2})$$

Define a $2M$ -component vector $\mathbf{h}(z)$ given by

$$\mathbf{h}(z) = \begin{pmatrix} H_0(z) \\ \vdots \\ H_M(z) \\ H'_1(z) \\ \vdots \\ H'_{M-1}(z) \end{pmatrix}$$

Using Eq. (A-2), $\mathbf{h}(z)$ can be written as

$$\mathbf{h}(z) = 2 \begin{pmatrix} \mathbf{A}_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 \end{pmatrix} \begin{pmatrix} \mathbf{C} & \Lambda_0 \mathbf{C} \\ \mathbf{S} & \Lambda_1 \mathbf{S} \end{pmatrix} \begin{pmatrix} \mathbf{g}_0(z^{2M}) & \mathbf{0} \\ \mathbf{0} & \mathbf{g}_1(z^{2M}) \end{pmatrix} \begin{pmatrix} \mathbf{e}(z) \\ z^{-M} \mathbf{e}(z) \end{pmatrix} \quad (\text{A-3})$$

where $\mathbf{g}_i(z)$, Λ_i , and \mathbf{A}_i are diagonal matrices with

$$[\mathbf{g}_0(z)]_{kk} = G_k(z), \quad [\mathbf{g}_1(z)]_{kk} = G_{k+M}(z), \quad k = 0, 1, \dots, M-1 \quad (\text{A-4})$$

$$[\Lambda_0]_{kk} = (-1)^k, \quad k = 0, 1, \dots, M$$

$$[\Lambda_1]_{kk} = (-1)^k, \quad k = 1, 2, \dots, M-1 \quad (\text{A-5})$$

$$[\mathbf{A}_0]_{kk} = a_k, \quad k = 0, 1, \dots, M$$

$$[\mathbf{A}_1]_{kk} = a_k, \quad k = 1, 2, \dots, M-1 \quad (\text{A-6})$$

and \mathbf{C} and \mathbf{S} are $(M+1) \times M$ and $(M-1) \times M$ matrices with

$$[\mathbf{C}]_{mn} = \cos\left(\frac{\pi}{M}mn\right), \quad m = 0, \dots, M, \quad n = 0, \dots, M-1$$

$$[\mathbf{S}]_{mn} = \sin\left(\frac{\pi}{M}mn\right), \quad m = 1, \dots, M-1, \quad n = 0, \dots, M-1 \quad (\text{A-7})$$

Define two $M \times 2M$ matrices

$$\mathbf{T}_0 = (\mathbf{C} \quad \Lambda_0 \mathbf{C}) \quad \text{and} \quad \mathbf{T}_1 = (\mathbf{S} \quad \Lambda_1 \mathbf{S})$$

From Eq. (A-3), we can draw Fig. A-1, a schematic implementation of the $2M$ channel cosine modulated system. The input to \mathbf{T}_0 and \mathbf{T}_1 , $\mathbf{a}(n)$, can be partitioned into two $M \times 1$ vectors:

$$\mathbf{a}(n) = \begin{pmatrix} \mathbf{a}_0(n) \\ \mathbf{a}_1(n) \end{pmatrix}$$

Their dependence on n will be dropped for convenience. As indicated in Fig. A-1, \mathbf{d}_0 and \mathbf{d}_1 are the outputs of \mathbf{T}_0 and \mathbf{T}_1 , respectively:

$$\mathbf{d}_0 = \mathbf{T}_0 \mathbf{a} \quad \text{and} \quad \mathbf{d}_1 = \mathbf{T}_1 \mathbf{a}$$

From the definitions of \mathbf{T}_0 and \mathbf{T}_1 , we know

$$\mathbf{d}_0 = \mathbf{C}\mathbf{a}_0 + \Lambda_0 \mathbf{C}\mathbf{a}_1 \quad \text{and} \quad \mathbf{d}_1 = \mathbf{S}\mathbf{a}_0 + \Lambda_1 \mathbf{S}\mathbf{a}_1 \quad (\text{A-8})$$

In [14], it is pointed out that \mathbf{C} and \mathbf{S} have the following properties:

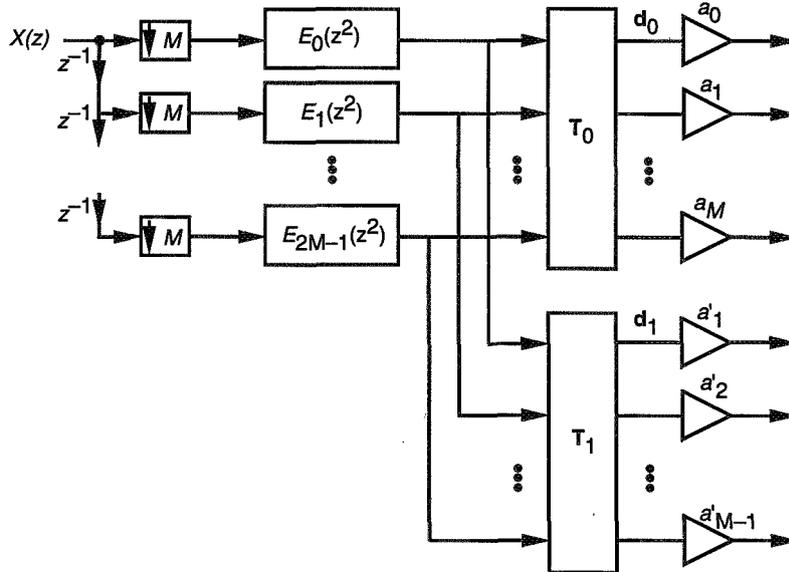


Fig. A-1. A schematic for the implementation of the analysis bank of the type 1 under-decimated cosine modulated $2M$ -channel filter bank.

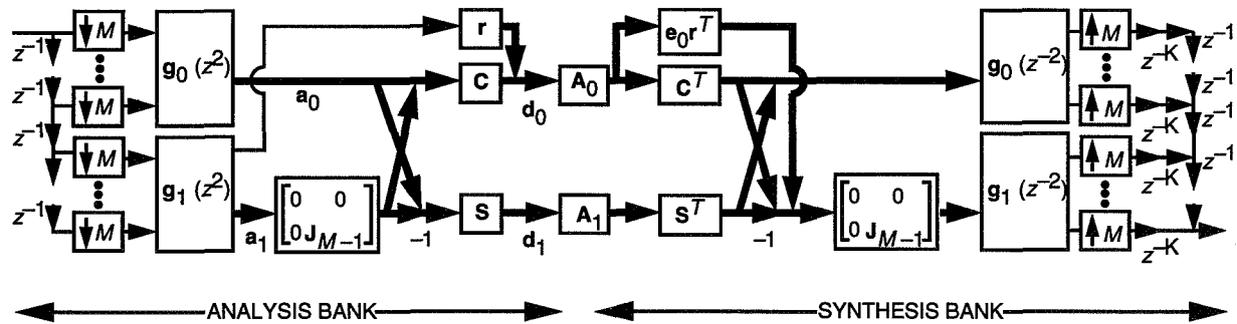
$$\Lambda_0 \mathbf{C} = \mathbf{C} \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{M-1} \end{pmatrix} + \mathbf{r} (1 \ 0 \ \dots \ 0) \quad \text{and} \quad \Lambda_1 \mathbf{S} = -\mathbf{S} \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{M-1} \end{pmatrix}$$

Using these relations, Eq. (A-8) becomes

$$\mathbf{d}_0 = \mathbf{C} \left(\mathbf{a}_0 + \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{M-1} \end{pmatrix} \mathbf{a}_1 \right) + [\mathbf{a}_1]_0 \mathbf{r}, \quad \mathbf{d}_1 = \mathbf{S} \left(\mathbf{a}_0 - \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_{M-1} \end{pmatrix} \mathbf{a}_1 \right) \quad (\text{A-9})$$

where $[\mathbf{a}_1]_0$ is the first element of \mathbf{a}_1 , and $\mathbf{r} = [1 \ -1 \ \dots \ (-1)^{M-1}]^T$. Equation (A-9) allows us to have a more clear idea of the implementation of \mathbf{T}_0 and \mathbf{T}_1 . The implementation of the synthesis bank is similar to the above.

Now we can draw a more detailed and complete picture for the implementation of the $2M$ channel filter bank. For simplicity, we can choose the synthesis filters to be the time-reversed versions of the corresponding analysis filters without the scalars a_k . The efficient implementation of the filter bank is shown in Fig. A-2. From Fig. A-2, we can see that the major complexity of the analysis bank is the cost of the analysis prototype plus matrix \mathbf{C} and matrix \mathbf{S} . Matrices \mathbf{C} and \mathbf{S} can be implemented by fast algorithms for DCT and DST matrices [11]. The process is similar for the synthesis bank.



SIZE OF MATRICES: $g_0(z^2), g_1(z^2) M \times M$
 $\mathbf{C} (M+1) \times M$
 $\mathbf{S} (M-1) \times M$
 $\mathbf{A}_0, \mathbf{A}_1$ SQUARE MATRICES

Fig. A-2. Efficient implementation of the type 1 under-decimated cosine modulated $2M$ -channel filter bank, where $k = N - 2M + 1$ and $\mathbf{e}_0 = [1 \ 0 \ \dots \ 0]^T$.

110871
33273
358427

A Simplified Integer Cosine Transform and Its Application in Image Compression

M. Costa and K. Tong
Communications Systems Research Section

A simplified version of the integer cosine transform (ICT) is described. For practical reasons, the transform is considered jointly with the quantization of its coefficients. It differs from conventional ICT algorithms in that the combined factors for normalization and quantization are approximated by powers of two. In conventional algorithms, the normalization/quantization stage typically requires as many integer divisions as the number of transform coefficients. By restricting the factors to powers of two, these divisions can be performed by variable shifts in the binary representation of the coefficients, with speed and cost advantages to the hardware implementation of the algorithm. The error introduced by the factor approximations is compensated for in the inverse ICT operation, executed with floating point precision. The simplified ICT algorithm has potential applications in image-compression systems with disparate cost and speed requirements in the encoder and decoder ends. For example, in deep space image telemetry, the image processors on board the spacecraft could take advantage of the simplified, faster encoding operation, which would be adjusted on the ground, with high-precision arithmetic. A dual application is found in compressed video broadcasting. Here, a fast, high-performance processor at the transmitter would precompensate for the factor approximations in the inverse ICT operation, to be performed in real time, at a large number of low-cost receivers.

I. Introduction

The integer cosine transform (ICT) [1,2] is an approximation of the discrete cosine transform. It can be implemented exclusively with integer arithmetic, with associated advantages in cost and speed for hardware implementations. One particular version of the ICT has been chosen for the compression algorithm of the image telemetry to be transmitted from Jupiter by the Galileo spacecraft [3].

As part of an image-compression system, the role of the ICT is to decorrelate the picture elements of image blocks, typically of size 8×8 , for subsequent quantization and entropy encoding. In this article, a simplified version of the ICT is investigated. The simplification involves approximating integer divisions by right shifts in the binary representations of the numbers and allows for faster and simpler hardware realizations. This version of the ICT is an orthogonal transform and is only approximately normalized. The departure from exact normalization is of little consequence since it can be compensated for in the inverse ICT operation, performed with real arithmetic.

In the next section, we review the calculation of the ICT and its interface with the quantization stage in an image compression system. In Section III, we look at two examples of quantization arrays. Then, in Section IV, we discuss the application of the modified algorithm to the compression of a set of planetary images. Concluding remarks are given in Section V.

II. The ICT Operation in Image Coding

We initially consider the ICT of a one-dimensional data vector \mathbf{X} of size 8. We choose the ICT version adopted for the Galileo image telemetry compression. This ICT is implemented by premultiplying \mathbf{X} by the orthogonal matrix C , given by

$$C = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 3 & 2 & 1 & -1 & -2 & -3 & -5 \\ 3 & 1 & -1 & -3 & -3 & -1 & 1 & 3 \\ 3 & -1 & -5 & -2 & 2 & 5 & 1 & 3 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 2 & -5 & 1 & 3 & -3 & -1 & 5 & -2 \\ 1 & -3 & 3 & -1 & -1 & 3 & -3 & 1 \\ 1 & -2 & 3 & -5 & 5 & -3 & 2 & -1 \end{pmatrix} \quad (1)$$

One attractive characteristic of the ICT is that the absolute values of its coefficients are equal to powers of two, or powers of two plus one. Thus, the matrix multiplication can be efficiently executed by shifts and additions.

Denoting the ICT vector by \mathbf{Y} , we write $\mathbf{Y} = C\mathbf{X}$. As C is an orthogonal matrix, we have $CC^T = D$, where D is diagonal. Now, let Δ denote the inverse of D . Clearly, we can factor the identity matrix I as $I = \sqrt{\Delta}CC^T\sqrt{\Delta}$. Thus, we can identify $M = \sqrt{\Delta}C$ as an orthonormal matrix, so that $M^T = C^T\sqrt{\Delta} = M^{-1}$. The matrix M represents the normalized ICT.

In source-compression applications, the ICT coefficients of image blocks are typically individually quantized, and the entire block is then entropy encoded. The quantization can be implemented by rounded integer divisions of the transformed coefficients by quantization factors. The rounded integer division of a coefficient a by a factor q is given¹ by $\lfloor (a/q) + 0.5 \rfloor$, where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x . Since normalization and quantization involve integer divisions, it makes sense to combine both steps into a single operation. In the sequel, we extend the ICT to two-dimensional arrays and review the details of the combined normalization and quantization operations.

First we note that

$$C^{-1} = C^T \Delta \quad (2)$$

To verify this equation, we write $CC^{-1} = I$. Substitute $C = \sqrt{\Delta^{-1}}M$ to obtain $\sqrt{\Delta^{-1}}MC^{-1} = I$. Premultiplying this equation by $M^T\sqrt{\Delta}$ gives the desired result.

To recover the data vector \mathbf{X} from the transform \mathbf{Y} , we premultiply \mathbf{Y} by C^{-1} . Thus, we have

$$\mathbf{X} = C^{-1}\mathbf{Y} = C^{-1}C\mathbf{X} = (C^T\sqrt{\Delta})(\sqrt{\Delta}C)\mathbf{X} \quad (3)$$

¹ In integer arithmetic calculations, the alternative expression $\lfloor (a + \lfloor q/2 \rfloor) / q \rfloor$ is often used.

This equation highlights $\sqrt{\Delta}C$ and $C^T\sqrt{\Delta}$ as matrix representations of the direct normalized ICT and its inverse transform, respectively.

The extension to two-dimensional blocks is straightforward. The two-dimensional ICT is a separable transform that can be computed by successively applying one-dimensional ICTs to the columns of the data block and then to the rows of the intermediate result. In matrix form, this is represented by premultiplying by the transform matrix and postmultiplying by its transpose. Letting X and Y denote the two-dimensional data block and its transform, we have

$$Y = \sqrt{\Delta} CXC^T\sqrt{\Delta} \quad (4)$$

and

$$X = C^T\sqrt{\Delta} Y\sqrt{\Delta}C \quad (5)$$

The above equations can be simplified by noting that CXC^T and Y are premultiplied and postmultiplied by the diagonal matrix $\sqrt{\Delta}$. Given two diagonal matrices, D_1 and D_2 , and an arbitrary matrix, A , the product D_1AD_2 can be more efficiently calculated by term-by-term multiplying all the elements of A by the corresponding elements of $D_1\mathbf{1}D_2$, where $\mathbf{1}$ is a matrix of the same size as A , composed only of 1's. This equivalent procedure reduces the number of multiplications by one-half. Denoting the term-by-term multiplication by the operator symbol $\#$, we write

$$Y = (CXC^T) \# N \quad (6)$$

and

$$X = C^T(Y \# N) C \quad (7)$$

where the normalization matrix is defined as

$$N = \sqrt{\Delta} \mathbf{1} \sqrt{\Delta} \quad (8)$$

It is clear in these equations that the original data block X can be exactly reconstructed from the transform block Y . In practice, though, the transform coefficients are quantized and reproduced with finite precision at the user end. As we have mentioned, the quantization procedure can be implemented by rounded term-by-term integer division of the transform coefficients by an array of quantization factors. This operation can be represented by the term-by-term multiplication of the transform matrix Y by H , the array of inverses of the quantization factors, it being understood that the results are rounded to the nearest integer.

Denoting the quantized transform by Y^* , we have

$$Y^* = Y (\#) H = (CXC^T) \# N (\#) H \quad (9)$$

where we use $(\#)$ to represent the nearest integer operation.

As it is easily verifiable, the above # operations can be permuted, and we can combine the normalization and quantization arrays into a single array $Q \triangleq N \# H$ so that

$$Y^* = (CXC^T) (\#) Q \quad (10)$$

The reconstructed image block X^* can be obtained by

$$X^* = C^T(Y^* \# Q^*) C \quad (11)$$

where Q^* , the counterpart of the array Q , satisfies the equation

$$Q^* \# Q = \Delta \mathbf{1} \Delta \quad (12)$$

We are now ready to consider the simplified ICT implementation. Suppose the optimized normalization/quantization matrix Q for a particular class of images has been obtained. Typically, the entries of this matrix are approximated by integer inverses. Let Q_a denote an alternative approximation of Q , with entries composed exclusively of negative powers of two. For the approximation rule, we select the nearest power of two, breaking ties in favor of the smaller factors. Substituting Q_a for Q , the quantization procedure can be implemented with simple binary shifts in the representations of the transformed coefficients, followed by addition of 0.5, and truncation.²

Performing the quantization by binary shifts may cut the execution time by more than 50 percent as compared with the time required by the operation with integer divisions. The exact time savings depends on the numbers involved and on the processor architecture. In general, binary shifts are executed in a fraction of the time required for integer divisions, which are usually executed by repeated shifts and subtractions.

To compensate for the approximation in Q , the reconstruction is performed with the matrix Q_a^* obtained from

$$Q_a^* \# Q_a = \Delta \mathbf{1} \Delta \quad (13)$$

III. Examples

To illustrate, we examine two examples with commonly used quantization templates.

A. Uniform Quantization

In the first example, we consider a uniform quantization template. This is one of the quantization templates to be used in the Galileo image-compression system. In this case, $H = \mathbf{1}$, so that Q and Q^* are equal to $N = \sqrt{\Delta} \mathbf{1} \sqrt{\Delta}$. In practice, Q is only an integer approximation of N , so that Q^* and Q are slightly different.

Recalling the expression for C , we have

² To avoid the addition of a noninteger, the alternative quantization rule given in Footnote 1 can be used.

$$D = CC^T = \begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 78 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 40 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 78 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 78 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 40 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 78 \end{pmatrix} \quad (14)$$

$$D \mathbf{1} D = \begin{pmatrix} 64 & 624 & 320 & 624 & 64 & 624 & 320 & 624 \\ 624 & 6084 & 3120 & 6084 & 624 & 6084 & 3120 & 6084 \\ 320 & 3120 & 1600 & 3120 & 320 & 3120 & 1600 & 3120 \\ 624 & 6084 & 3120 & 6084 & 624 & 6084 & 3120 & 6084 \\ 64 & 624 & 320 & 624 & 64 & 624 & 320 & 624 \\ 624 & 6084 & 3120 & 6084 & 624 & 6084 & 3120 & 6084 \\ 320 & 3120 & 1600 & 3120 & 320 & 3120 & 1600 & 3120 \\ 624 & 6084 & 3120 & 6084 & 624 & 6084 & 3120 & 6084 \end{pmatrix} \quad (15)$$

and

$$\sqrt{D} \mathbf{1} \sqrt{D} \cong \begin{pmatrix} 8 & 25 & 18 & 25 & 8 & 25 & 18 & 25 \\ 25 & 78 & 56 & 78 & 25 & 78 & 56 & 78 \\ 18 & 56 & 40 & 56 & 18 & 56 & 40 & 56 \\ 25 & 78 & 56 & 78 & 25 & 78 & 56 & 78 \\ 8 & 25 & 18 & 25 & 8 & 25 & 18 & 25 \\ 25 & 78 & 56 & 78 & 25 & 78 & 56 & 78 \\ 18 & 56 & 40 & 56 & 18 & 56 & 40 & 56 \\ 25 & 78 & 56 & 78 & 25 & 78 & 56 & 78 \end{pmatrix} \quad (16)$$

Therefore, the normalization/quantization template is given by the matrix

$$Q = \begin{pmatrix} 8^{-1} & 25^{-1} & 18^{-1} & 25^{-1} & 8^{-1} & 25^{-1} & 18^{-1} & 25^{-1} \\ 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} & 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} \\ 18^{-1} & 56^{-1} & 40^{-1} & 56^{-1} & 18^{-1} & 56^{-1} & 40^{-1} & 56^{-1} \\ 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} & 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} \\ 8^{-1} & 25^{-1} & 18^{-1} & 25^{-1} & 8^{-1} & 25^{-1} & 18^{-1} & 25^{-1} \\ 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} & 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} \\ 18^{-1} & 56^{-1} & 40^{-1} & 56^{-1} & 18^{-1} & 56^{-1} & 40^{-1} & 56^{-1} \\ 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} & 25^{-1} & 78^{-1} & 56^{-1} & 78^{-1} \end{pmatrix} \quad (17)$$

and the reconstruction template Q^* is expressed as

$$Q^* = \begin{pmatrix} 1/8 & 25/624 & 18/320 & 25/624 & 1/8 & 25/624 & 18/320 & 25/624 \\ 25/624 & 1/78 & 56/3120 & 1/78 & 25/624 & 1/78 & 56/3120 & 1/78 \\ 18/320 & 56/3120 & 1/40 & 56/3120 & 18/320 & 56/3120 & 1/40 & 56/3120 \\ 25/624 & 1/78 & 56/3120 & 1/78 & 25/624 & 1/78 & 56/3120 & 1/78 \\ 1/8 & 25/624 & 18/320 & 25/624 & 1/8 & 25/624 & 18/320 & 25/624 \\ 25/624 & 1/78 & 56/3120 & 1/78 & 25/624 & 1/78 & 56/3120 & 1/78 \\ 18/320 & 56/3120 & 1/40 & 56/3120 & 18/320 & 56/3120 & 1/40 & 56/3120 \\ 25/624 & 1/78 & 56/3120 & 1/78 & 25/624 & 1/78 & 56/3120 & 1/78 \end{pmatrix}$$

Our proposed simplified ICT algorithm uses a power-of-two approximation for the elements of Q , so that Q is replaced by Q_a , given by

$$Q_a = \begin{pmatrix} 8^{-1} & 32^{-1} & 16^{-1} & 32^{-1} & 8^{-1} & 32^{-1} & 16^{-1} & 32^{-1} \\ 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} \\ 16^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 16^{-1} & 64^{-1} & 32^{-1} & 64^{-1} \\ 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} \\ 8^{-1} & 32^{-1} & 16^{-1} & 32^{-1} & 8^{-1} & 32^{-1} & 16^{-1} & 32^{-1} \\ 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} \\ 16^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 16^{-1} & 64^{-1} & 32^{-1} & 64^{-1} \\ 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} & 32^{-1} & 64^{-1} & 64^{-1} & 64^{-1} \end{pmatrix} \quad (18)$$

and the reconstruction template Q_a^* is expressed by

$$Q_a^* = \begin{pmatrix} 1/8 & 32/624 & 16/320 & 32/624 & 1/8 & 32/624 & 16/320 & 32/624 \\ 32/624 & 64/6084 & 64/3120 & 64/6084 & 32/624 & 64/6084 & 64/3120 & 64/6084 \\ 16/320 & 64/3120 & 32/1600 & 64/3120 & 16/320 & 64/3120 & 32/1600 & 64/3120 \\ 32/624 & 64/6084 & 64/3120 & 64/6084 & 32/624 & 64/6084 & 64/3120 & 64/6084 \\ 1/8 & 32/624 & 16/320 & 32/624 & 1/8 & 32/624 & 16/320 & 32/624 \\ 32/624 & 64/6084 & 64/3120 & 64/6084 & 32/624 & 64/6084 & 64/3120 & 64/6084 \\ 16/320 & 64/3120 & 32/1600 & 64/3120 & 16/320 & 64/3120 & 32/1600 & 64/3120 \\ 32/624 & 64/6084 & 64/3120 & 64/6084 & 32/624 & 64/6084 & 64/3120 & 64/6084 \end{pmatrix}$$

B. JPEG Quantization

The Joint Photographic Experts Group (JPEG) has put forth an image-compression standard based on the discrete cosine transform. A typical quantization array for this standard is given by [4]

$$J = \begin{pmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 60 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{pmatrix} \quad (19)$$

Combining this array with the ICT normalization gives the matrix

$$Q = \begin{pmatrix} 128^{-1} & 275^{-1} & 179^{-1} & 400^{-1} & 192^{-1} & 999^{-1} & 912^{-1} & 1524^{-1} \\ 300^{-1} & 936^{-1} & 782^{-1} & 1482^{-1} & 649^{-1} & 4524^{-1} & 3351^{-1} & 4290^{-1} \\ 250^{-1} & 726^{-1} & 640^{-1} & 1341^{-1} & 716^{-1} & 3184^{-1} & 2760^{-1} & 3128^{-1} \\ 350^{-1} & 1326^{-1} & 1229^{-1} & 2262^{-1} & 1274^{-1} & 6786^{-1} & 3351^{-1} & 4836^{-1} \\ 144^{-1} & 550^{-1} & 662^{-1} & 1399^{-1} & 544^{-1} & 2723^{-1} & 1843^{-1} & 1923^{-1} \\ 600^{-1} & 2730^{-1} & 3072^{-1} & 4992^{-1} & 2023^{-1} & 8112^{-1} & 6312^{-1} & 7176^{-1} \\ 877^{-1} & 3575^{-1} & 3120^{-1} & 4860^{-1} & 1843^{-1} & 6759^{-1} & 4800^{-1} & 5642^{-1} \\ 1799^{-1} & 7176^{-1} & 5306^{-1} & 7644^{-1} & 2798^{-1} & 7800^{-1} & 5753^{-1} & 7722^{-1} \end{pmatrix}$$

and its counterpart

$$Q^* = \begin{pmatrix}
128/64 & 275/624 & 179/320 & 400/624 & \dots \\
300/624 & 936/6084 & 782/3120 & 1482/6084 & \dots \\
250/320 & 726/3120 & 640/1600 & 1341/3120 & \dots \\
350/624 & 1326/6084 & 1229/3120 & 2262/6084 & \dots \\
144/64 & 550/624 & 662/320 & 1399/624 & \dots \\
600/624 & 2730/6084 & 3072/3120 & 4992/6084 & \dots \\
877/320 & 3575/3120 & 3120/1600 & 4860/3120 & \dots \\
1799/624 & 7176/6084 & 5306/3120 & 7644/6084 & \dots \\
\dots & 192/64 & 999/624 & 912/320 & 1524/624 \\
\dots & 649/624 & 4524/6084 & 3351/3120 & 4290/6084 \\
\dots & 716/320 & 3184/3120 & 2760/1600 & 3128/3120 \\
\dots & 1274/624 & 6786/6084 & 3351/3120 & 4836/6084 \\
\dots & 544/64 & 2723/624 & 1843/320 & 1923/624 \\
\dots & 2023/624 & 8112/6084 & 6312/3120 & 7176/6084 \\
\dots & 1843/320 & 6759/3120 & 4800/1600 & 5642/3120 \\
\dots & 2798/624 & 7800/6084 & 5753/3120 & 7722/6084
\end{pmatrix} \tag{20}$$

Therefore, a possible power-of-two approximation for Q is given by

$$Q_a = \begin{pmatrix}
128^{-1} & 256^{-1} & 128^{-1} & 512^{-1} & 128^{-1} & 1024^{-1} & 1024^{-1} & 1024^{-1} \\
256^{-1} & 1024^{-1} & 512^{-1} & 1024^{-1} & 512^{-1} & 4096^{-1} & 4096^{-1} & 4096^{-1} \\
256^{-1} & 512^{-1} & 512^{-1} & 1024^{-1} & 512^{-1} & 2048^{-1} & 2048^{-1} & 2048^{-1} \\
256^{-1} & 1024^{-1} & 1024^{-1} & 2048^{-1} & 1024^{-1} & 4096^{-1} & 4096^{-1} & 4096^{-1} \\
128^{-1} & 512^{-1} & 512^{-1} & 1024^{-1} & 512^{-1} & 2048^{-1} & 2048^{-1} & 2048^{-1} \\
512^{-1} & 2048^{-1} & 2048^{-1} & 4096^{-1} & 2048^{-1} & 8192^{-1} & 4096^{-1} & 8192^{-1} \\
1024^{-1} & 4096^{-1} & 2048^{-1} & 4096^{-1} & 2048^{-1} & 8192^{-1} & 4096^{-1} & 4096^{-1} \\
2048^{-1} & 8192^{-1} & 4096^{-1} & 8192^{-1} & 2048^{-1} & 8192^{-1} & 4096^{-1} & 8192^{-1}
\end{pmatrix}$$

and the corresponding reconstruction array is

$$Q_a^* = \begin{pmatrix}
128/64 & 256/624 & 128/320 & 512/624 & \dots \\
256/624 & 1024/6084 & 512/3120 & 1024/6084 & \dots \\
256/320 & 512/3120 & 512/1600 & 1024/3120 & \dots \\
256/624 & 1024/6084 & 1024/3120 & 2048/6084 & \dots \\
128/64 & 512/624 & 512/320 & 1024/624 & \dots \\
512/624 & 2048/6084 & 2048/3120 & 4096/6084 & \dots \\
1024/320 & 4096/3120 & 2048/1600 & 4096/3120 & \dots \\
2048/624 & 8192/6084 & 4096/3120 & 8192/6084 & \dots \\
\dots & 128/64 & 1024/624 & 1024/320 & 1024/624 \\
\dots & 512/624 & 4096/6084 & 4096/3120 & 4096/6084 \\
\dots & 512/320 & 2048/3120 & 2048/1600 & 2048/3120 \\
\dots & 1024/624 & 4096/6084 & 4096/3120 & 4096/6084 \\
\dots & 512/64 & 2048/624 & 2048/320 & 2048/624 \\
\dots & 2048/624 & 8192/6084 & 4096/3120 & 8192/6084 \\
\dots & 2048/320 & 8192/3120 & 4096/1600 & 4096/3120 \\
\dots & 2048/624 & 8192/6084 & 4096/3120 & 8192/6084
\end{pmatrix} \tag{21}$$

It is interesting to note that the power-of-two approximation can also be considered in a dual application, where the encoder is capable of performing floating point arithmetic at high speed, and the compressed images need to be recovered at a large number of low-cost receivers. In this situation, which could be encountered, for example, in broadcast video applications, it would be beneficial to replace the reconstruction array Q^* given in Eq. (20) by a power-of-two approximation. This approximation would have to be precompensated by a quantization array composed of real numbers, such that the term-by-term product of the two arrays produces $\Delta 1 \Delta$ [cf. Eq. (13)].

IV. Practical Results

The simplified algorithm was compared with the conventional ICT algorithm in the compression of three 800×800 , 8-bit typical planetary images. The comparisons were made in terms of the rate-distortion trade-off, expressed in plots of the peak signal-to-noise ratio (PSNR) as a function of the achieved compression ratio (CR). The PSNR of 8-bit images is defined by $10 \log_{10}(255^2/MSE)$, where MSE denotes the mean square error of the reconstructed image. The CR is defined as the ratio between the number of bits in the original image and the number of bits in the compressed image.

Figure 1 shows the rate-distortion performance obtained for image "ant9," with the standard ICT and the simplified algorithms. The original is a noisy, hard-to-compress image, with a large number of streaks caused by α -particle effects on the CCD sensor array.

The quantization array (H) is uniform, and the set of 7 points corresponding to each algorithm was obtained by weighting the combined normalization/quantization array given in Eqs. (16) and (17) by 1, 2, 4, 8, 16, 32, and 64. The weights were chosen as powers of two, since the use of arbitrary weights would obviously interfere with the practicality of the simplified algorithm. A smoother variation of the rate-distortion points can be obtained by progressively weighting pieces of the quantization array. For example, the bottom-right half of the array, which corresponds to higher frequency terms, could be weighted first, before the adjustment is extended to the rest of the array. Clearly, a large number of weighting strategies are possible and their relative performances will depend on the image being compressed.

From Fig. 1, we notice that the two algorithms are quite comparable in performance. The maximum separation between the interpolated curves, occurring at the higher compression ratios, is approximately one-third of a dB in PSNR (approximately a 7-percent variation in MSE). Figure 2 presents a similar plot for image "r4," a somewhat "busy" image with little sensor noise. Figure 3 shows results with "saturn1," an easier image to compress, reaching a compression ratio of 60 at the reasonable PSNR of 34 dB. In both cases, the observed performances obtained with uniform quantization are very similar, the maximum separation between the interpolated curves amounting to about one-quarter of a dB, or a 6-percent variation in MSE.

To exemplify the potential savings in computation time, we consider the processor of Galileo's attitude and articulation control subsystem (AACS) [5], to be used also for onboard image compression. The times required for an addition and an integer division are 5 and 13.25 μsec , respectively. Thus, normalization and quantization of a conventional 8×8 ICT block requires $64 \times 18.25 \mu\text{sec} = 1.168 \text{ msec}$ (one addition and one division per transform coefficient, according to the formula in Footnote 1). Alternatively, a fixed-length binary shift operation takes from 2.5 to 6.5 μsec , depending on the length. Estimating an average duration of 4.5 μsec for binary shifts, the time required for normalization/quantization of an 8×8 ICT block is $64 \times 9.5 \mu\text{sec} = 608 \mu\text{sec}$ with the simplified implementation. Thus, a savings of 560 μsec per ICT block can be realized. Nevertheless, two factors contribute to making the proposed scheme not applicable to Galileo. First, the AACS processor does not take full advantage of the binary shift implementation. There are no instructions for variable-length binary shifts, so additional time is required for setting the length. Second, a significant fraction (74 percent) of the average time needed for compressing images with the AACS processor is spent on the entropy encoding stage (a combination of run length and Huffman code), due to the intense memory accessing required.

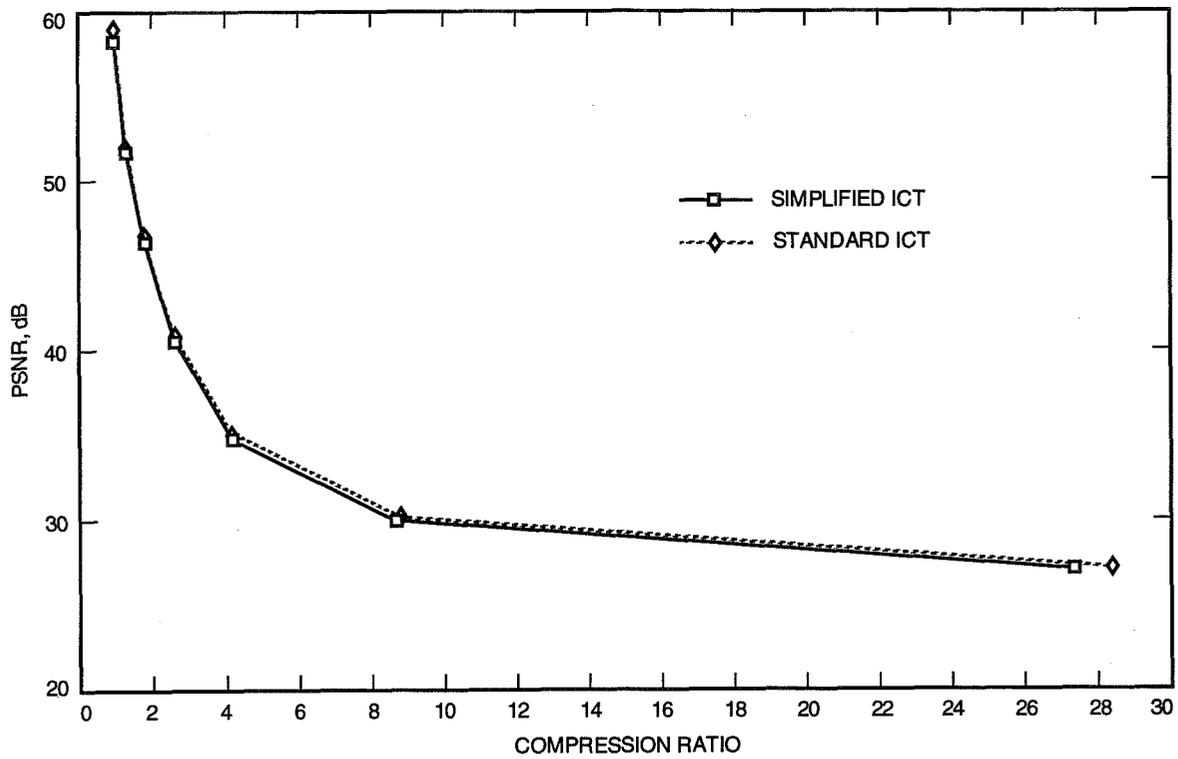


Fig. 1. PSNR as a function of compression ratio for image "ant9," with uniform quantization.

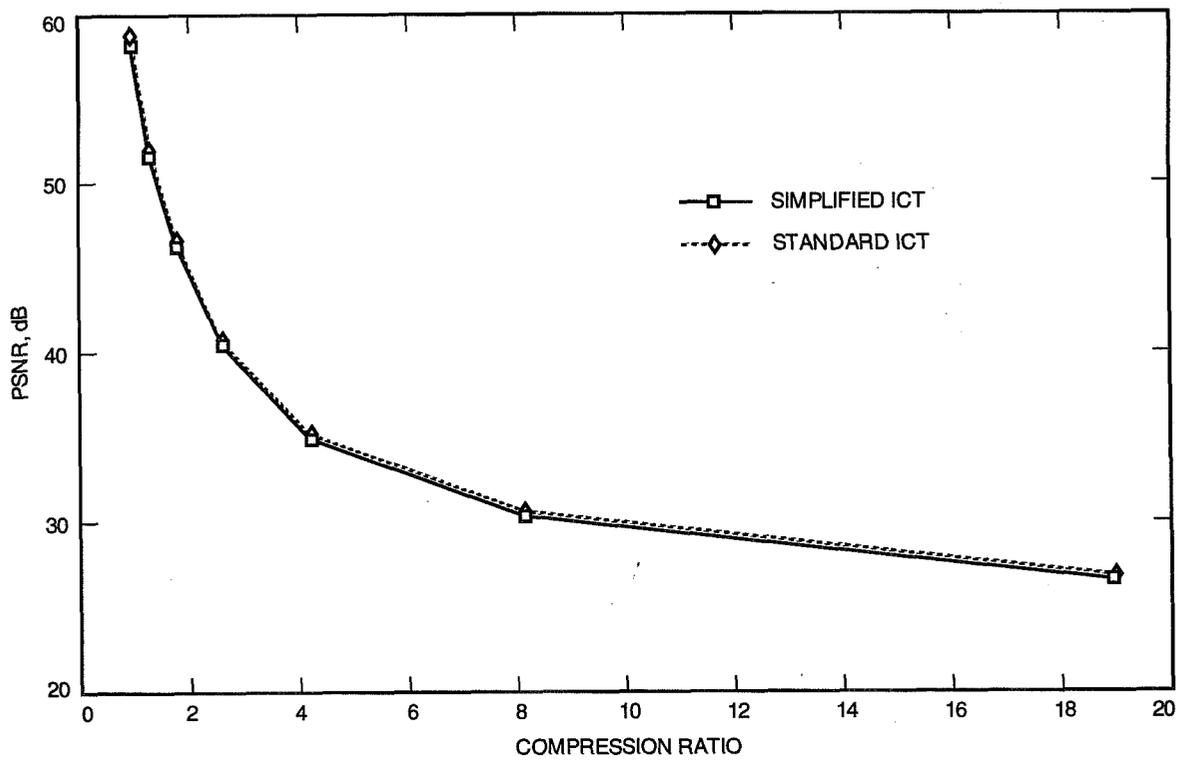


Fig. 2. PSNR as a function of compression ratio for image "r4," with uniform quantization.

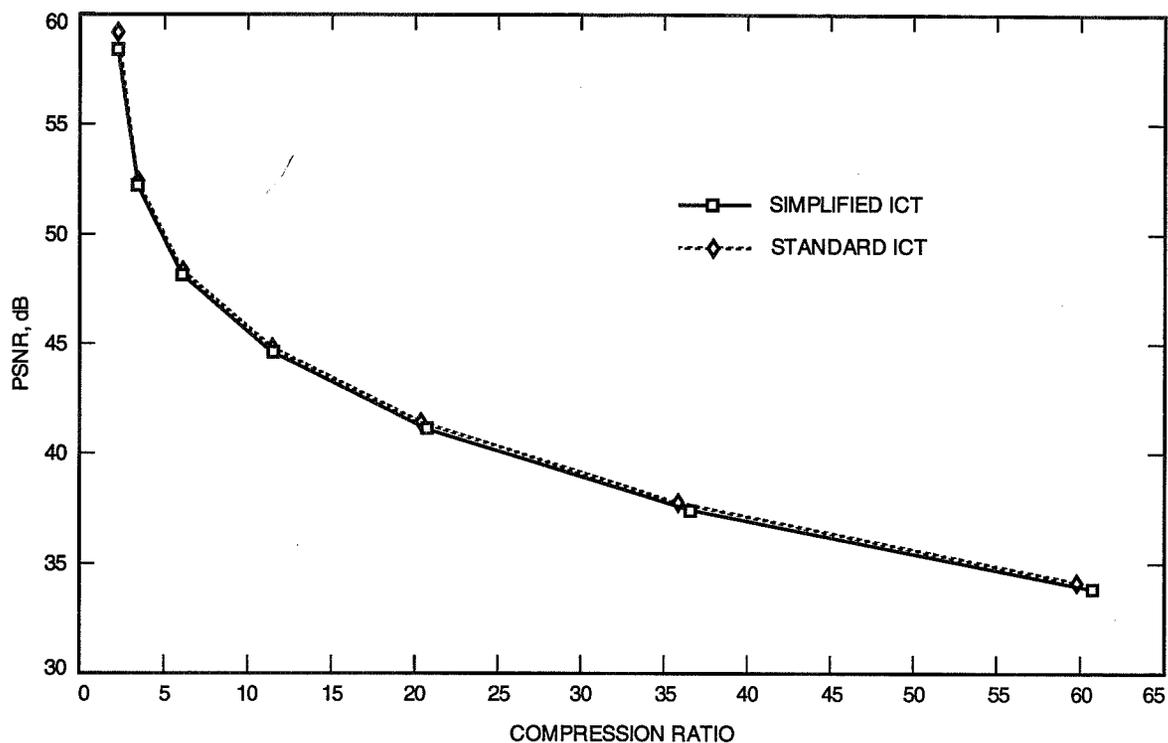


Fig. 3. PSNR as a function of compression ratio for image "saturn1," with uniform quantization.

V. Conclusions

This article described a simplification of the integer cosine transform. The simplified transform is orthogonal and approximately normalized. This transform is considered in the context of an image compression system in which normalization and quantization operations are performed jointly, exclusively by means of additions and binary shifts. The departure from a perfectly normalized transform is compensated for in the inverse transform operation, performed with real arithmetic. Alternatively, the power-of-two approximation can be used at the inverse transform operation, having been preconditioned by the proper real arithmetic during the direct transform calculation. The algorithm is useful in applications with disparate speed and complexity constraints at the points where the direct and inverse transforms are calculated. The performance of the simplified transform is compared with that of a conventional ICT in the compressing of a set of three planetary images. The distortion-rate characteristics of the two algorithms are found to be very similar, differing only by a fraction of a dB in PSNR. The advantage of using the simplified transform is the reduction in time required for normalization and quantization. This reduction is generally processor-dependent and can be significant in special-purpose hardware implementations.

References

- [1] W.-K. Cham, "Development of Integer Cosine Transforms by the Principle of Dyadic Symmetry," *IEE Proceedings*, vol. 136, pt. I, no. 4, pp. 276-282, August 1989.
- [2] K.-M. Cheung, F. Pollara, and M. Shahshahani, "Integer Cosine Transform for Image Compression," *The Telecommunications and Data Acquisition Progress Report 42-105*, vol. January-March 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 45-53, May 15, 1991.
- [3] K.-M. Cheung and K. Tong, "Proposed Data Compression Schemes for the Galileo S-Band Contingency Mission," *1993 Space and Earth Science Data Compression Workshop*, NASA Conference Publication 3191, Snowbird, Utah, pp. 99-109, April 2, 1993.
- [4] W. Pennebaker and J. Mitchell, *JPEG Still Image Compression Standard*, New York: Van Nostrand Reinhold, 1993.
- [5] *HAL/S ATAC 16M Reference Documents*, Document 52-054614, Applied Technologies Advanced Computer.

58-61
33274

110872

R-20

TDA Progress Report 42-119

November 15, 1994

358430

Adaptive Line Enhancers for Fast Acquisition

H.-G. Yeh

Spacecraft Telecommunications Equipment Section

T. M. Nguyen

Communications Systems Research Section

Three adaptive line enhancer (ALE) algorithms and architectures—namely, conventional ALE, ALE with double filtering, and ALE with coherent accumulation—are investigated for fast carrier acquisition in the time domain. The advantages of these algorithms are their simplicity, flexibility, robustness, and applicability to general situations including the Earth-to-space uplink carrier acquisition and tracking of the spacecraft. In the acquisition mode, these algorithms act as bandpass filters; hence, the carrier-to-noise ratio (CNR) is improved for fast acquisition. In the tracking mode, these algorithms simply act as lowpass filters to improve signal-to-noise ratio; hence, better tracking performance is obtained. It is not necessary to have a priori knowledge of the received signal parameters, such as CNR, Doppler, and carrier sweeping rate. The implementation of these algorithms is in the time domain (as opposed to the frequency domain, such as the fast Fourier transform (FFT)). The carrier frequency estimation can be updated in real time at each time sample (as opposed to the batch processing of the FFT). The carrier frequency to be acquired can be time varying, and the noise can be non-Gaussian, nonstationary, and colored.

I. Introduction

The present spacecraft transponder acquires and tracks the carrier signal by using a phase-locked loop (PLL). Because the frequency sweeping technique is employed in the acquisition process, the time that it takes for the PLL to acquire the uplink carrier is relatively long. The sweeping rate is set to about 40 Hz/sec when the spacecraft receiver signal-to-noise ratio (SNR) equals -151 dBm for a deep-space mission. To sweep ± 10 kHz from the best locked frequency, 17 min are required to complete the acquisition process. Therefore, there is a need for fast-acquisition technique development with application to the transponder. In general, a fast-acquisition technique as shown in Fig. 1 will be very useful for deep-space missions, especially in emergencies. First, the receiver is in the acquisition process. Second, when the uplink carrier is acquired and indicated by the lock detector, the switch is then shifted to the tracking position and the tracking process takes over immediately. Although devised to support the space mission, the fast-acquisition technique proposed in this article is also applicable to other types of receivers, including fixed-ground and mobile receivers.

The problem of estimating certain parameters of a sinusoidal signal in the presence of noise is of general interest and has received considerable attention in the literature [1]. Examples may be found in vibration measurements, Doppler radar returns, geophysical processing, and communication systems [2].

Many techniques, such as the fast Fourier transform (FFT) [3] and adaptive least-squares algorithms [4], have been proposed in the literature to solve such problems. These methods provide excellent results but may require excessively long observation times because of batch processing.

Recently, time-domain spectral estimation techniques based on an adaptive line enhancer (ALE) system have been introduced [5-9]. The ALE system is depicted in Fig. 2. The system, which was introduced by

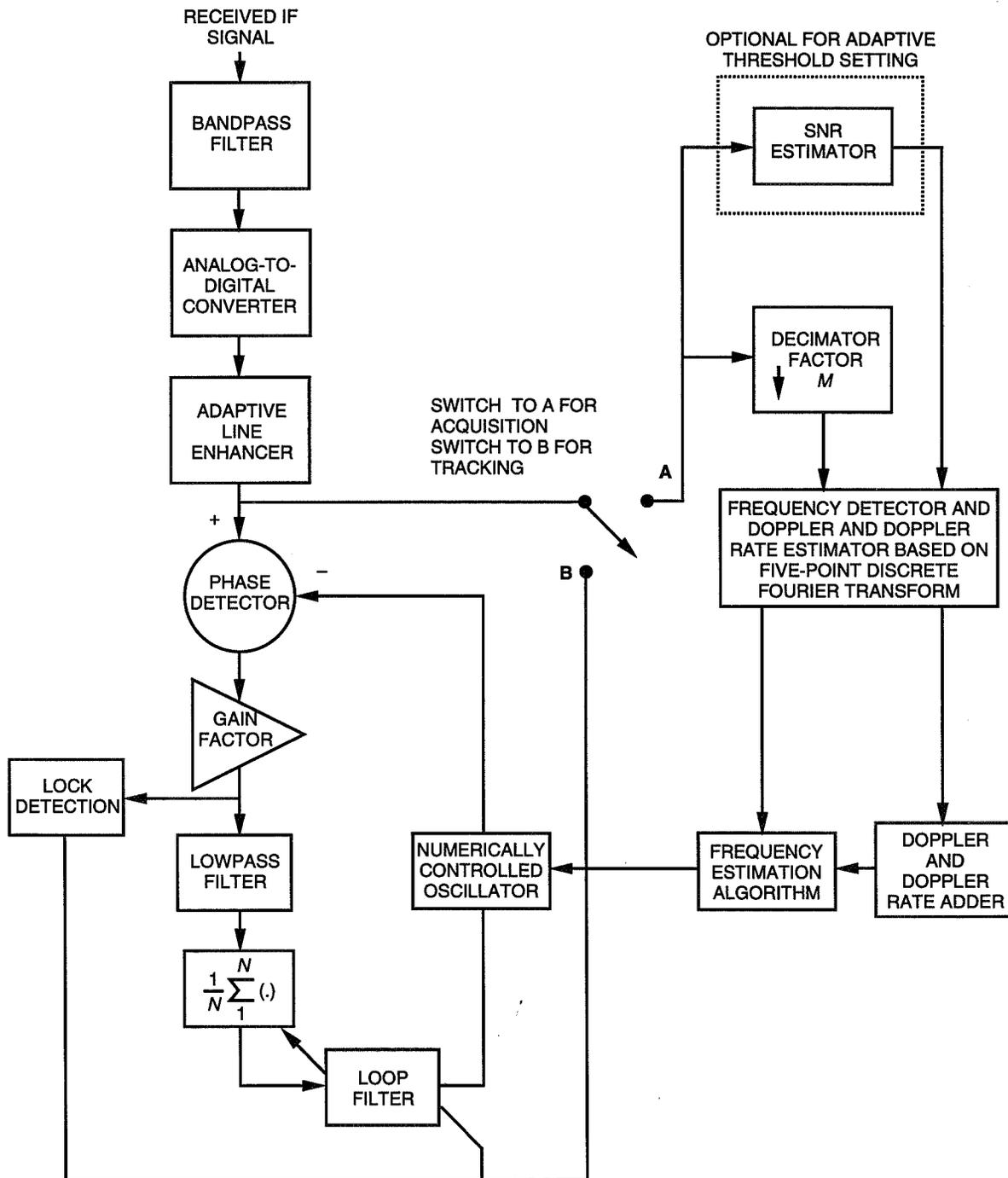


Fig. 1. Using ALE in the digital receiver for both acquisition and tracking.

Widrow [5], uses the measured signal as the desired response and a delayed version of itself as input. The principle is that the delay should decorrelate the noise between the primary and reference inputs while leaving the narrowband carrier signal correlated. When functioning ideally, the adaptive filter output is an enhanced version of the carrier components with a higher carrier-to-noise ratio (CNR). Both CNR and SNR are used in this article, and they are interchangeable.

The adaptive filter depicted in Fig. 2 is a time-varying system and the weight vector is updated, based on the least-mean-squares (LMS) algorithm. The LMS algorithm is based on the method of steepest descent [5]. Many applications have been developed using the LMS algorithm; the fast measurement of digital instantaneous frequency in [6] and [12] are examples. In addition, it is well known that the LMS-type algorithms are more robust to sudden variations in the environment than the FFT algorithms.

Three ALE algorithms and architectures for fast acquisition are presented in this article. The analysis of the general properties of an ALE is given in Section II. It contains discussion of an optimal adaptive filter, optimal gain, and a steady-state frequency response. Section III introduces two modified ALE algorithms: the ALE with double filtering (ALEDF) and the ALE with coherent accumulation (ALECA). Implementation is considered in Section IV. Simulations for acquiring fixed-frequency-signal acquisition are provided in Section V. Performance comparison of these adaptive line enhancers is also discussed in Section V. A discussion and the conclusion are given in Section VI.

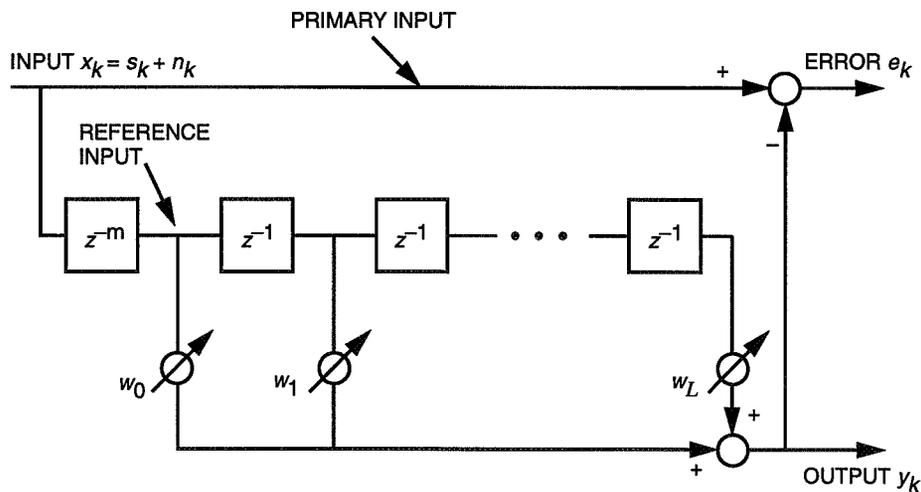


Fig. 2. The structure of the conventional adaptive line enhancer.

II. Analysis of the Optimal ALE

It is shown in [10] that, in general, the real LMS algorithm is not equivalent to the real part of the complex LMS algorithm. However, when the algorithm is configured as an ALE with sinusoidal inputs, the mean weight behavior of the real algorithm is identical to the real part of the mean weight of the complex algorithm. There are simplicities in the analysis of the complex model that do not exist in the analysis of the real model. Consequently, the analysis will be performed by using complex models in Sections II and III. Both the system gain and the steady-state frequency response of the ALE are provided in this section.

A. The Optimal Adaptive Filter

The input signal of Fig. 2 is x_k , which contains carrier component s_k and white noise component n_k with power σ_n^2 . The signal at the primary input is defined as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{s}_k + \mathbf{n}_k \\ &= a e^{j(\omega_o k T + \phi)} + n_k \end{aligned} \quad (1)$$

where a denotes the signal amplitude, ϕ denotes the signal phase, T is the sampling period, and ω_o is the carrier frequency.

In vector/matrix form,

$$\begin{aligned} \mathbf{x}_k &= \mathbf{s}_k + \mathbf{n}_k \\ &= a e^{j(\omega_o k T + \phi)} \mathbf{q} + \mathbf{n}_k \end{aligned} \quad (2)$$

where

$$\mathbf{x}_k = [x_k \ x_{k-1} \ \cdots \ x_{k-L}]^T \quad (3)$$

$$\mathbf{q} = [1 \ e^{-j\omega_o T} \ e^{-j2\omega_o T} \ \cdots \ e^{-jL\omega_o T}]^T \quad (4a)$$

$$\mathbf{n}_k = [n_k \ n_{k-1} \ \cdots \ n_{k-L}]^T \quad (4b)$$

are vectors of length $(L + 1)$. The input signal vector to the adaptive filter is \mathbf{x}_{k-m} , where m is the delay unit. The delay unit m chosen must be of sufficient length to cause the broadband noise components in the filter (reference) input to become uncorrelated from those in the primary input. The carrier signal components, because of their periodic nature, will remain correlated with each other. The adaptive filter output is

$$y_k = \mathbf{w}_k^H \mathbf{x}_{k-m} \quad (5)$$

where $\mathbf{w}_k = [w_0 \ w_1 \ \cdots \ w_L]^T$ and H denotes the conjugate transpose. The error sequence is defined as

$$e_k = x_k - y_k \quad (6)$$

The weight vector is updated as follows:

$$\mathbf{w}_{k+1} = \mathbf{w}_k + 2\mu e_k \mathbf{x}_{k-m}^* \quad (7)$$

where $*$ denotes the conjugate operation and μ is the step size of the adaptation.

The convergence of the weight vector is assured by [5]

$$0 < \mu < \frac{1}{(L+1)(\text{carrier} + \text{noise power})} \quad (8)$$

where $L + 1$ is the number of taps of the adaptive filter. The optimal weight vector \mathbf{w}_{opt} , called the Wiener weight vector, is found in [5] as

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{p} \quad (9)$$

where

$$\mathbf{R} = E[\mathbf{x}_{k-m}\mathbf{x}_{k-m}^H] = \mathbf{R}_s + \mathbf{R}_N = a^2\mathbf{q}\mathbf{q}^H + \sigma_n^2\mathbf{I} \quad (10)$$

\mathbf{R}_s = autocorrelation matrix of the carrier

\mathbf{R}_n = autocorrelation matrix of the noise with power σ_n^2

$$\mathbf{p} = E[\mathbf{x}_{k-m}\mathbf{x}_k^*] = a^2e^{-j\omega_0 mT}\mathbf{q} \quad (11)$$

By applying the matrix inversion lemma, also called the “ABCD lemma,”

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1}[\mathbf{I} - \mathbf{B}(\mathbf{DA}^{-1}\mathbf{B} + \mathbf{C}^{-1})^{-1}\mathbf{DA}^{-1}] \quad (12)$$

The \mathbf{R}^{-1} is obtained as follows:

$$\mathbf{R}^{-1} = (\sigma_n^2\mathbf{I} + a^2\mathbf{q}\mathbf{q}^H)^{-1} = \frac{1}{\sigma_n^2} \left[\mathbf{I} - \frac{a^2}{\sigma_n^2\mathbf{I} + (L+1)a^2}\mathbf{q}\mathbf{q}^H \right] \quad (13)$$

where

$$\mathbf{A} = \sigma_n^2\mathbf{I} \quad (14)$$

$$\mathbf{B} = a^2\mathbf{q} \quad (15)$$

$$\mathbf{C} = \mathbf{I} \quad (16)$$

$$\mathbf{D} = \mathbf{q}^H \quad (17)$$

The optimal weight is then obtained as follows:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{p} = \beta z_0^{-m}\mathbf{q} \quad (18)$$

where

$$\beta = \frac{a^2}{\sigma_n^2 + (L+1)a^2} \quad (19)$$

$$z_0 = e^{j\omega_0 T} \quad (20)$$

When the adaptive filter converges to its steady state, the weight vector fluctuates around its optimal solution.

B. The Optimal Coherent Processing Gain of the ALE

The optimal linear solution for selecting the weight vector of an ALE is similar to that for the so-called “matched filter.” For a carrier at frequency ω_o embedded in white noise, the matched filter response is a sampled sinusoidal signal whose frequency is ω_o . The matched filter produces the peak SNR at each sample, but does not preserve the carrier signal waveform at the output, especially when the input signal has time-varying parameters. The matched filter solution does provide the best SNR gain obtainable by linear processing. However, the solution can be constructed only by giving prior knowledge of the frequency ω_o . On the other hand, the ALE output y_k preserves the carrier signal waveform as shown in Eq. (21). Furthermore, it is not necessary to have a priori knowledge of the received signal parameters, such as carrier SNR, Doppler, and carrier sweeping rate. For example, the carrier frequency sweeping rate depends on the uplink carrier signal level for a deep-space mission. The uplink carrier frequency ω_o sweeping rate is set to about 544 and 40 Hz/sec around the best-lock frequency when the carrier signal level is equal to -110 and -151 dBm, respectively. Therefore, the ALE method is designed to approximate the optimal SNR gain obtained by the matched filter solution for this problem.

The optimal steady-state carrier component at the ALE system output is

$$y_{sk} = \mathbf{w}_{opt}^H \mathbf{x}_{k-m} = g a e^{j(\omega_o k T + \phi)} = g s_k \quad (21)$$

where the optimal coherent processing gain is

$$g = \frac{(L+1)a^2}{(L+1)a^2 + \sigma_n^2} = (L+1)\beta \quad (22)$$

Clearly, the carrier component at the ALE output has the same phase (or delayed by $2\pi n$, n being an integer) as the input signal and is multiplied by the real processing gain factor g . The ALE system total power output at steady state is

$$E[y_k y_k^*] = E\left\{[\mathbf{w}_{opt}^H \mathbf{x}_{k-m}][\mathbf{w}_{opt}^H \mathbf{x}_{k-m}]^H\right\} = \mathbf{w}_{opt}^H \mathbf{R} \mathbf{w}_{opt} \quad (23)$$

Substituting Eqs. (10) and (18) into Eq. (23), the ALE system total power output becomes

$$E[y_k y_k^*] = g^2 \left[a^2 + \frac{\sigma_n^2}{L+1} \right] \quad (24)$$

This ALE system output CNR is then obtained as follows:

$$\text{CNR}_{out} = \frac{\text{output carrier power}}{\text{output noise power}} = \frac{E[y_s^2(k)]}{E[y_n^2(k)]} = \frac{(L+1)a^2}{\sigma_n^2} \quad (25)$$

The ALE system input CNR power ratio is

$$\text{CNR}_{input} = \frac{\text{input carrier power}}{\text{input noise power}} = \frac{a^2}{\sigma_n^2} \quad (26)$$

Therefore, the ALE optimal steady-state CNR gain is

$$\mathbf{G}_{\text{ALE}} = \frac{\text{CNR}_{\text{output}}}{\text{CNR}_{\text{input}}} = L + 1 \quad (27)$$

Equation (27) shows that the ALE optimal CNR gain is proportional to the length of the adaptive filter.

C. The Steady-State Frequency Response of the Optimal ALE

From Fig. 2 and using the z -transform, one can derive the transfer function of the ALE. It is obtained as follows:

$$\mathbf{H}(z) = \frac{\mathbf{Y}(z)}{\mathbf{X}(z)} = z^{-m} \mathbf{W}(z) \quad (28)$$

where

$$\mathbf{W}(z) = \mathbf{Z} \{ \mathbf{w}_{\text{opt}}^*(k) \} = \beta z_0^m \sum_{i=0}^L z_0^i z^{-i} = \beta z_0^m \frac{1 - (z_0 z^{-1})^{(L+1)}}{1 - z_0 z^{-1}} \quad (29)$$

Therefore, $\mathbf{H}(z)$ is

$$\mathbf{H}(z) = \beta z^{-m} z_0^m \frac{1 - (z_0 z^{-1})^{(L+1)}}{1 - z_0 z^{-1}} \quad (30)$$

Consequently, the optimal steady-state frequency response of the ALE is

$$\begin{aligned} \mathbf{H}(\omega) &= \beta e^{j(\omega_o - \omega)mT} \frac{1 - e^{j(\omega_o - \omega)T(L+1)}}{1 - e^{j(\omega_o - \omega)T}} \\ &= \beta e^{j(\omega_o - \omega)(m+L/2)T} \frac{\sin [((L+1)/2)(\omega_o - \omega)T]}{\sin [(1/2)(\omega_o - \omega)T]} \end{aligned} \quad (31)$$

At $\omega = \omega_o$, the optimal frequency response (the peak value of the transfer function) becomes

$$\mathbf{H}(\omega_o) = \beta(L+1) = g \quad (32)$$

Equation (32) shows a real constant gain g at the frequency ω_o , which is the acquired or tracked carrier frequency.

Steady-state magnitude responses of the optimal ALE with filter lengths of 16- and 32-taps are shown in Fig. 3, where the sampling frequency is equal to $10\omega_o$. It is observed that the magnitude response of the 32-tap ALE has much sharper cutoff frequency and twice the frequency resolution of the 16-tap ALE. It also shows that the ALE acts as a bandpass filter and the center frequency of the ALE is adapted to track the frequency of the input carrier signal. Consequently, when the ω_o of the input carrier is changed (either increased or decreased), the center frequency of the bandpass filter will follow. Furthermore, the

steady-state response of the ALE at frequency ω_o is related to the input CNR and the length of the weight as depicted in Eq. (33):

$$20 \log |\mathbf{H}(\omega_o)| = 20 \log(g) = 20 \log \frac{(L+1)\text{CNR}_{input}}{1 + (L+1)\text{CNR}_{input}} \text{ dB} \quad (33)$$

Equation (33) is plotted and shown in Fig. 4. It shows that, at high input CNR, the adaptive filter has a gain close to 1. This means that the carrier will pass through the adaptive filter (filter weighting function is about 1) nearly 100 percent; the error sequence obtained after the subtraction is minimized in the least-mean-square sense. At low input carrier SNR ($\text{SNR} < -20$ dB), the adaptive filter gain is close to 0. This indicates that when the noise component is much stronger than the carrier component, the best filter gain obtained should be small so that the error sequence will not increase in the least-square sense.

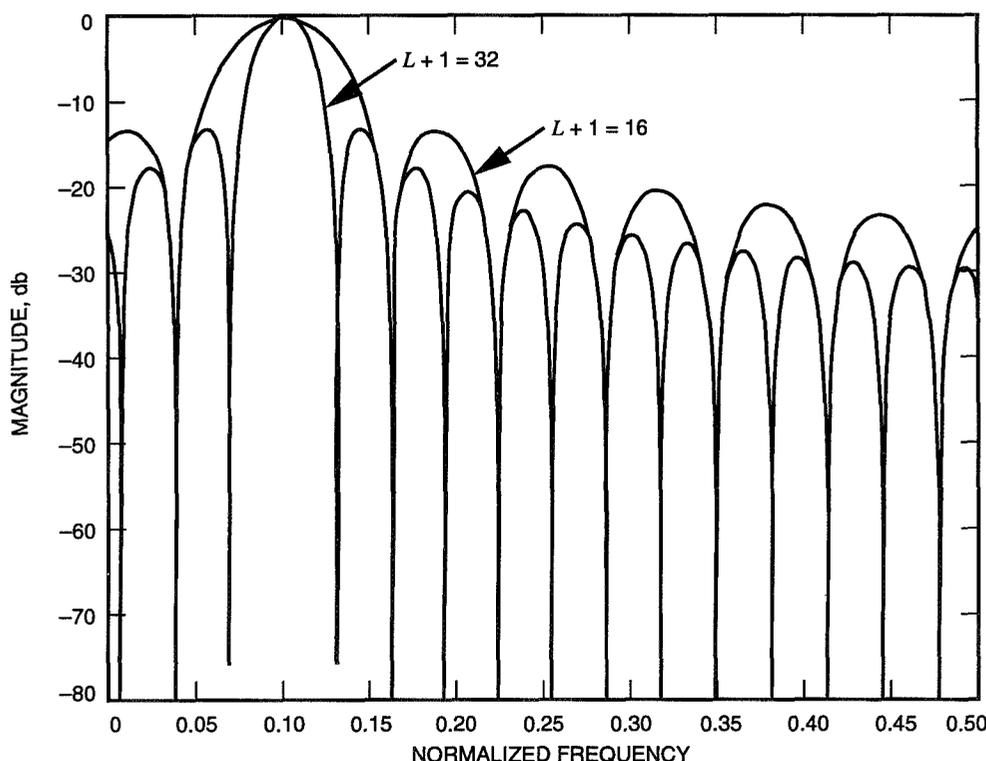


Fig. 3. The magnitude response of the optimal ALE with filter lengths of 16 and 32.

III. The ALEDF and ALECA Algorithms

In general, the ALE's CNR gain can be doubled by doubling the length based on Eq. (27), which also doubles the frequency resolution. However, this means that the total number of operations required will be doubled. Two ALEs, modified to improve the ALE system performance at the given frequency resolution and without increasing the computational load too much, are presented in this section.

A. The Adaptive Line Enhancer With Double Filtering

The first modified ALE is the so-called ALE with double filtering (ALEDF) and is shown in Fig. 5. A finite-impulse response (FIR) filter is cascaded with the ALE as the second stage. Coefficients of this FIR

filter are a real-time copy of those of the adaptive filter in the ALE. By filtering the received signal using two identical filters, the overall system gain is squared in linear scale or doubled in dB scale. Consequently, the overall processing gain at the desired signal of the ALEDF is equal to g^2 . However, the processing time of this ALEDF is $2(L + 1)T$, where T is the sampling period. Note that the processing time of an ALE with $2(L + 1)$ taps is the same as that of the ALEDF with $(L + 1)$ taps per filter. The magnitude responses of both the ALEDF and ALE with 16 taps per filter are given in Fig. 6. Both magnitude responses have the same frequency resolution. Due to double filtering, the magnitude response of the ALEDF has a sharper cutoff frequency and much lower sidelobes compared with those of the ALE, as shown in Fig. 6. However, the magnitude response of the ALEDF of Fig. 6 has half-frequency resolution and much lower sidelobes compared with those of the ALE with 32 taps, as shown in Fig. 3. This is due to the fact that the ALEDF has two 16-tap filters in cascade while the ALE has 32-tap filters.

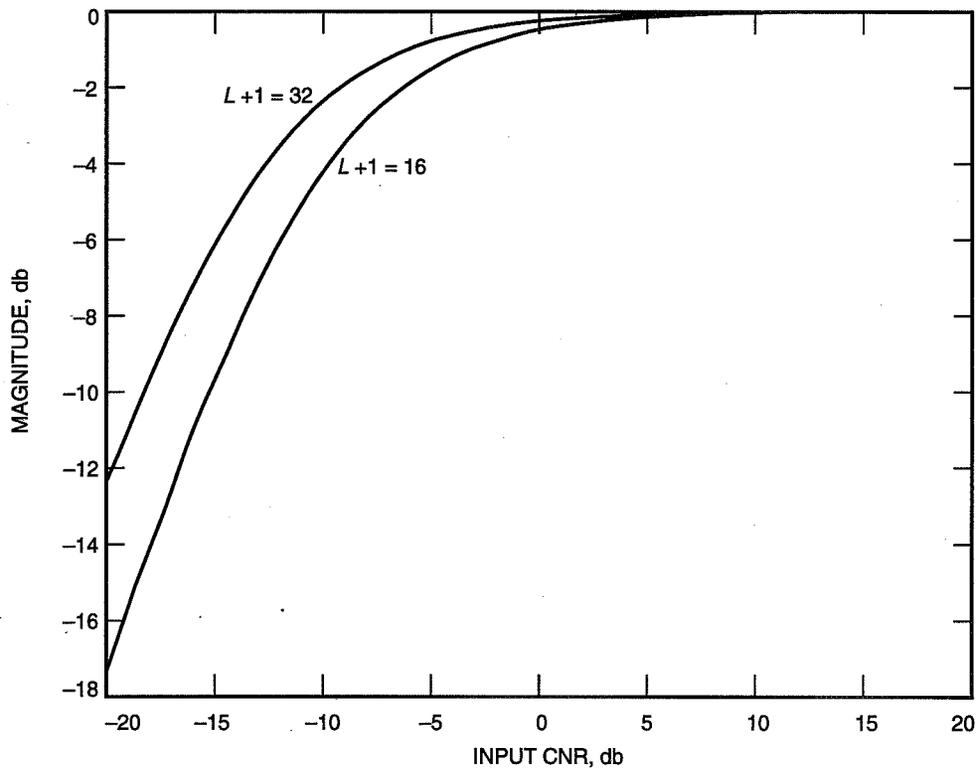


Fig. 4. The coherent processing gain magnitude response of the optimal ALE at the tracked frequency ω_0 versus the input CNR.

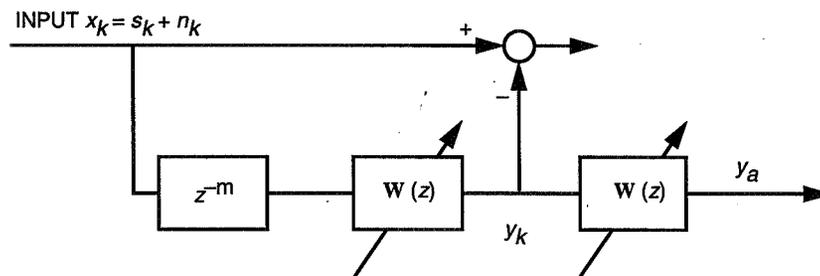


Fig. 5. The structure of the adaptive line enhancer with double filtering (ALEDF).

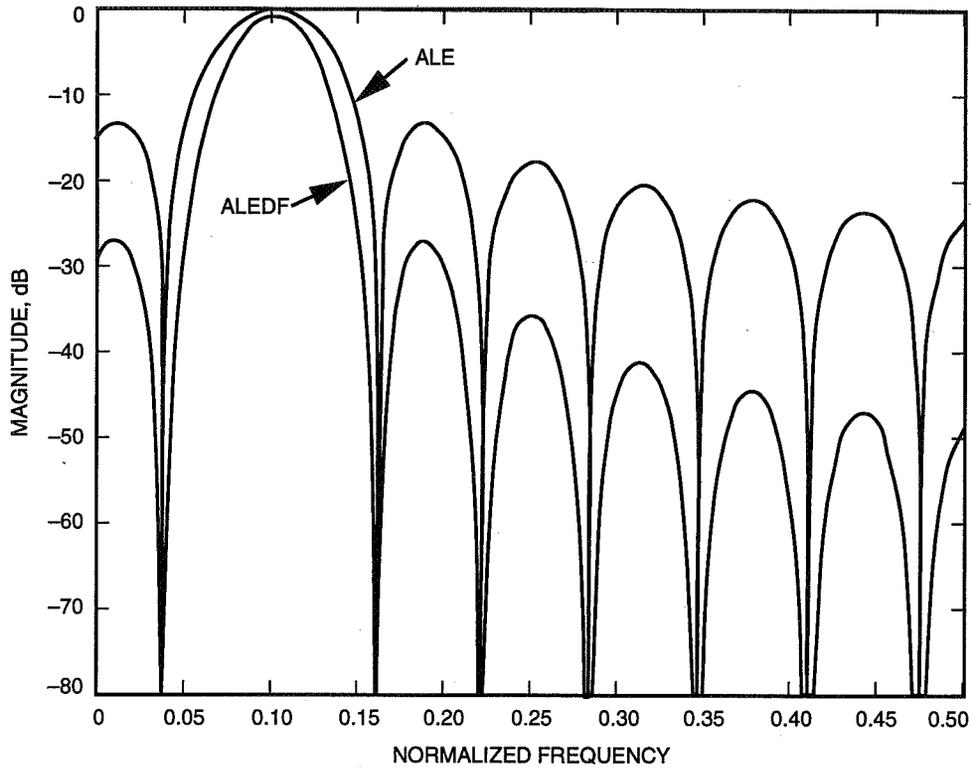


Fig. 6. The optimal magnitude response of both the ALEDF and ALE; filter length = 16.

B. The Adaptive Line Enhancer With Coherent Accumulation

The second modified architecture is the so-called ALE with coherent accumulation (ALECA); it is depicted in Fig. 7. The ALECA was first introduced in [11]. Figure 7 shows that an ALE output is cascaded with a closed feedback loop. This loop contains real-time copied filter coefficients from the adaptive filter of the ALE, an m -delay unit, and a multiplication parameter c . The output of the ALE is the input of the second stage. It has been shown in Eq. (21) that the carrier component of the ALE output y_k has the same phase as the input carrier component. By applying the same processing in the feedback loop (i.e., the same delay units and same filter as that of the reference input line in the ALE), the carrier component of y_b will have the same phase as that of y_k . Therefore, it is an ALE with coherent accumulation to produce the final output y_a .

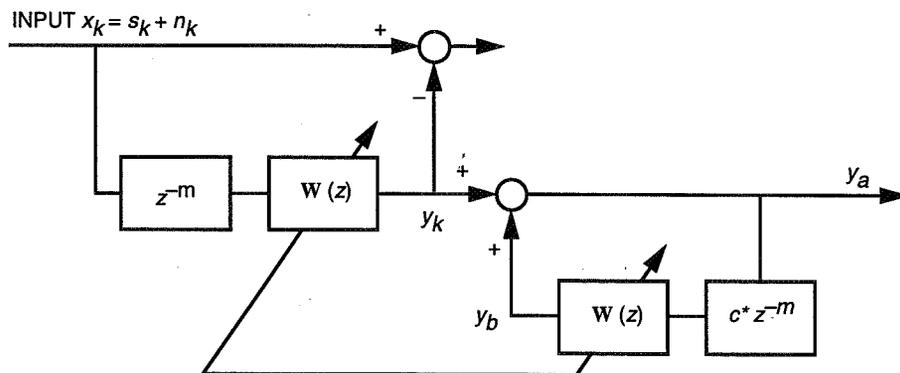


Fig. 7. The structure of the adaptive line enhancer with coherent accumulation (ALECA).

This ALECA architecture uses a recursive loop. Consequently, the stability of the system is of concern. To ensure the stability of the ALECA system, the feedback parameter c must be less than 1 and greater than or equal to 0 (Appendix). The ALECA becomes a conventional ALE when $c = 0$. The ALECA system performance is analyzed and given in the following sections.

1. The Steady-State Frequency Response of the Optimal ALECA. From Fig. 7, the overall transfer function is

$$\mathbf{H}_a(z) = \frac{\mathbf{Y}_a(z)}{\mathbf{X}(z)} = \frac{\mathbf{H}(z)}{1 - c\mathbf{H}(z)} \quad (34)$$

Therefore, the steady-state frequency response to the input s_k is

$$\mathbf{H}_a(\omega) = \frac{\beta e^{j(\omega_o - \omega)(m+L/2)T} \frac{\sin[((L+1)/2)(\omega_o - \omega)T]}{\sin[(1/2)(\omega_o - \omega)T]}}{1 - c\beta e^{j(\omega_o - \omega)(m+L/2)T} \frac{\sin[((L+1)/2)(\omega_o - \omega)T]}{\sin[(1/2)(\omega_o - \omega)T]}} \quad (35)$$

At $\omega = \omega_o$, the frequency response becomes

$$\mathbf{H}_a(\omega_o) = \frac{g}{1 - cg} \quad (36)$$

Note that g is less than or equal to 1. When c approaches $1/g$, $\mathbf{H}_a(\omega_o)$ becomes infinite. Consequently, the magnitude response at $\omega = \omega_o$ versus input CNR is obtained and given in Eq. (37):

$$20 \log |\mathbf{H}_a(\omega_o)| = 20 \log \left| \frac{g}{1 - cg} \right| = 20 \log \left| \frac{(L+1)\text{CNR}_{input}/(1 + (L+1)\text{CNR}_{input})}{1 - c[(L+1)\text{CNR}_{input}/(1 + (L+1)\text{CNR}_{input})]} \right| \text{ dB} \quad (37)$$

2. ALECA System Performance. Figure 8 shows the magnitude response of the optimal ALECA at $\omega = \omega_o$ versus input CNR based on Eq. (37). Figure 9 shows the magnitude response of the optimal ALECA versus the frequency based on Eq. (35) with several different c values. When parameter c equals zero, the ALECA becomes the conventional ALE. When parameter c is close to 1, the peak value of the ALECA transfer function is significantly higher than that of the conventional ALE, which also leads to some improvement in CNR gain due to the smaller effective noise bandwidth.

IV. Implementation Considerations

Starting with this section, only the real (not complex) case is considered. The total number of operations required is $2L + 2$ for both multiplications and accumulations of an ALE system using Eqs. (15) and (17) with a filter length of $L + 1$. However, the computation of the ALEDF with $L + 1$ taps would require $3L + 3$ multiplications and $3L + 2$ additions. This is because Eq. (17) is employed once while Eq. (15) is used twice. The total number of multiplications and additions required of an ALE with $2(L + 1)$ taps is $4L + 4$ and $4L + 4$, respectively. However, this $(2L + 2)$ -tap ALE has twice the frequency resolution and CNR gain of the $2(L + 1)$ -tap ALE. Clearly, 25 percent of the multiplications and additions are saved in the ALEDF with $L + 1$ taps in comparison with $2(L + 1)$ -tap ALE. The total number of multiplications and accumulations required of the ALECA with the filter length of $L + 1$ is $3L + 4$ and $3L + 3$, respectively. Table 1 shows the computational load comparison between several ALE architectures.

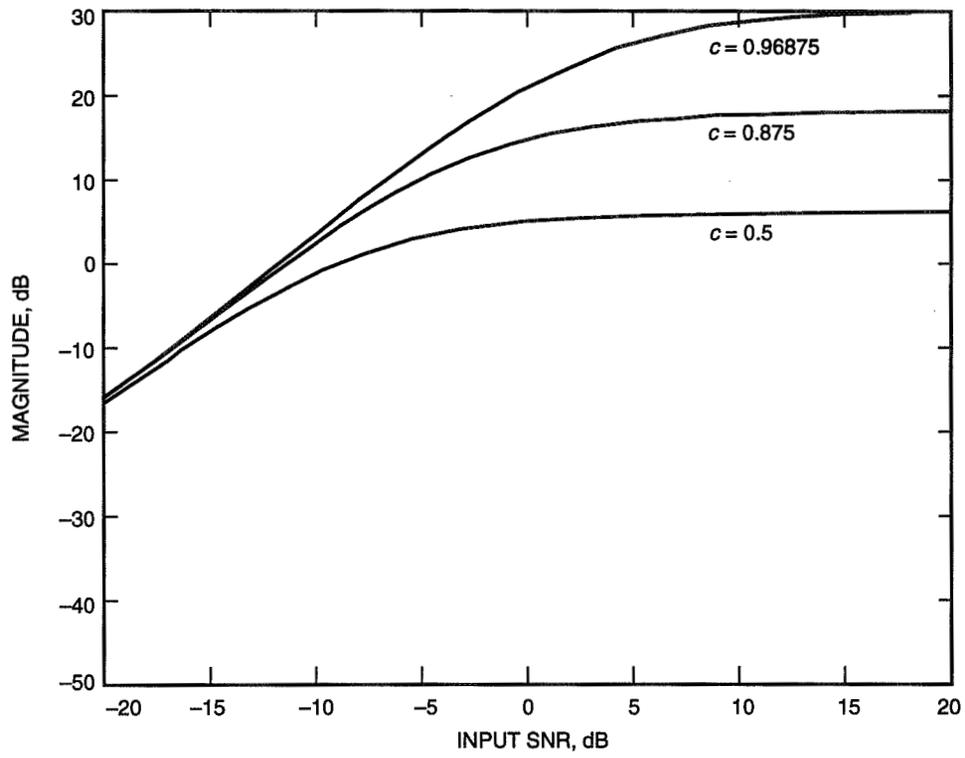


Fig. 8. The magnitude response of the optimal ALECA at $\omega = \omega_0$ versus the input CNR.

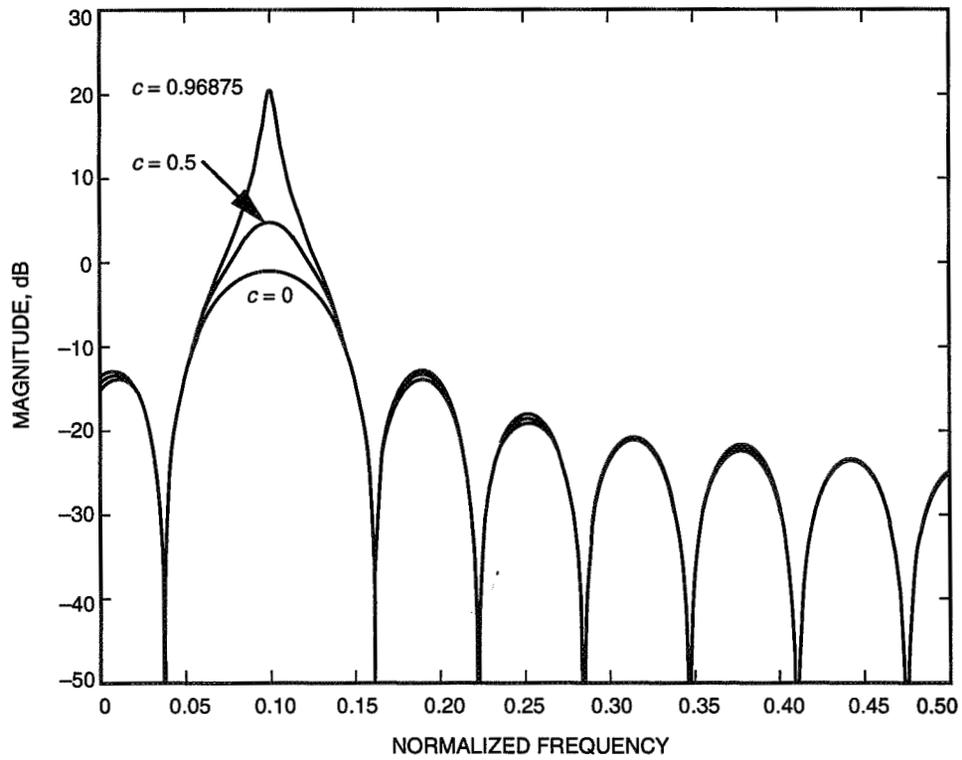


Fig. 9. The magnitude response of the optimal ALECA versus the frequency.

Table 1. The computational load comparison between several ALE architectures.

Number of operations	ALE (filter length: $L + 1$)	ALEDF (length: $L + 1/\text{filter}$)	ALECA (length: $L + 1/\text{filter}$)
Multiplications	$2L + 2$	$3L + 3$	$3L + 4$
Additions	$2L + 2$	$3L + 2$	$3L + 3$

V. Simulations

The tracking performances of the 16-tap ALE, ALEDF, and ALECA are studied via simulations with the same frequency resolution for comparison. The carrier is a sinusoidal signal with a fixed frequency, and the sampling rate is 256 times the carrier frequency. Figure 10 shows the input signal in the time domain with $\text{SNR} = 0$ dB. Figures 11, 12, and 13 present the carrier estimation obtained by the ALE, ALEDF, and ALECA, respectively. The total number of weights is 16 per filter, the number of delays is chosen as 1, and the parameter c is equal to 0.96875. The step size chosen is 0.000625. Visual examination indicates that the ALECA output provides the best estimated carrier signal, the ALEDF output is second, and the conventional ALE is last. To show the frequency response of the filtered sequence, the sampling frequency is changed to eight times the carrier frequency for better resolution. All other conditions are the same as the time-domain simulation.

Figures 14 through 17 show the frequency-domain response of the input sequence, the ALE output data, the ALEDF output data, and the ALECA output data, respectively. The location of the simulated carrier frequency is indicated on Figs. 14 through 16 by vertical arrows on the x-axis. The desired signal gain of the ALECA is greater than that of either the ALE or the ALEDF.

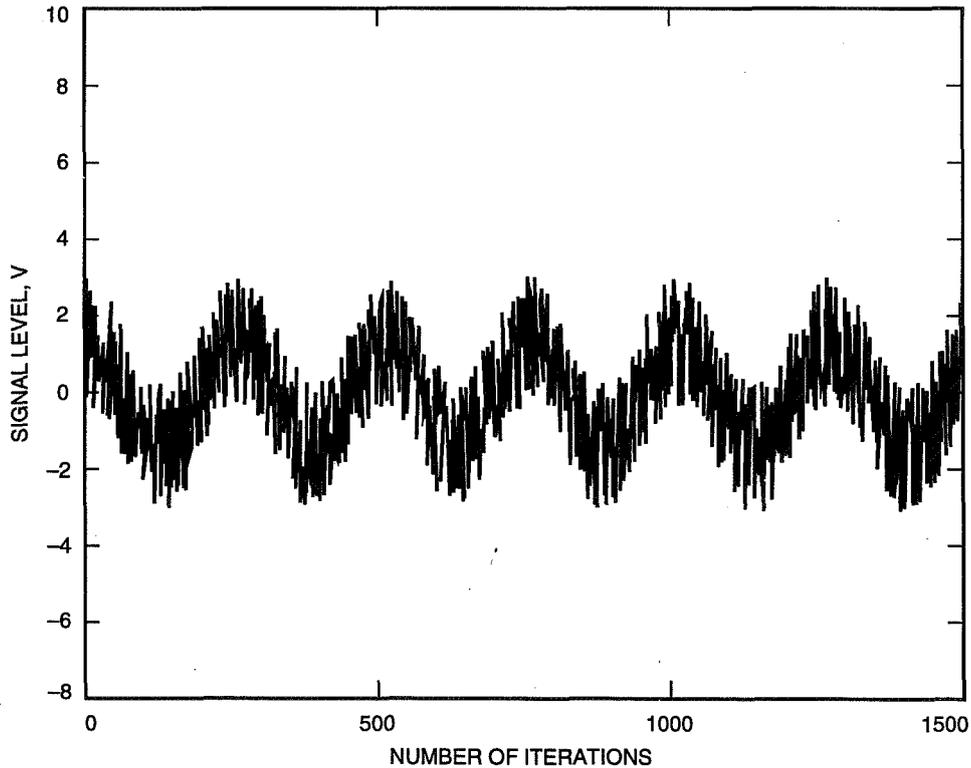


Fig. 10. The ALE input signal in time domain with $\text{SNR} = 0$ dB.

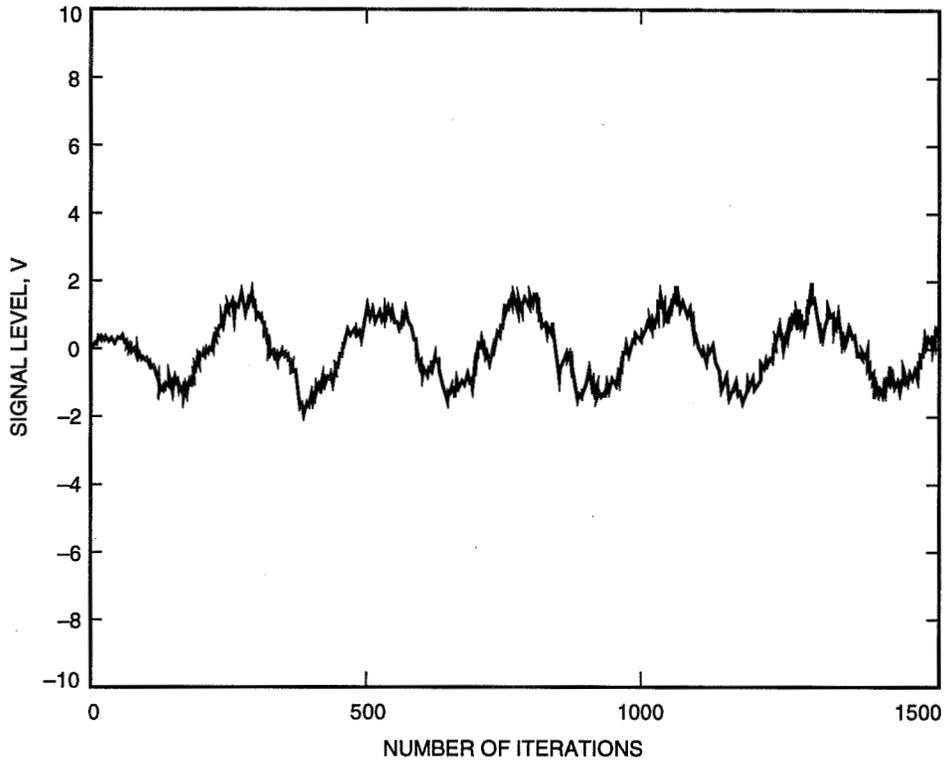


Fig. 11. The 16-tap ALE output: an estimated carrier signal.

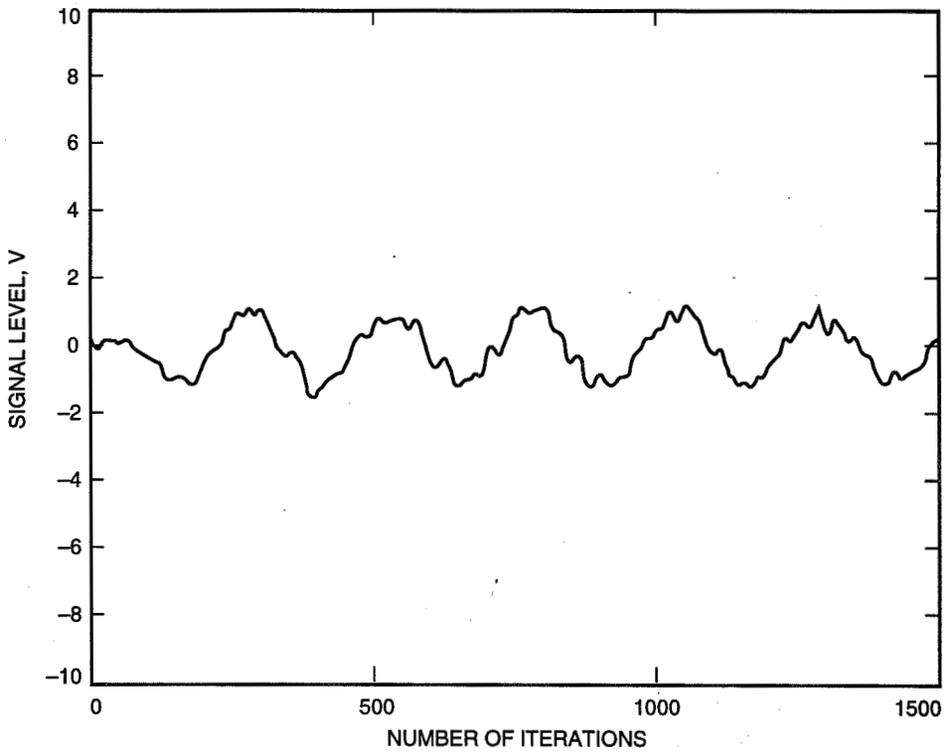


Fig. 12. The 16-tap ALEDF output: an estimated carrier signal.

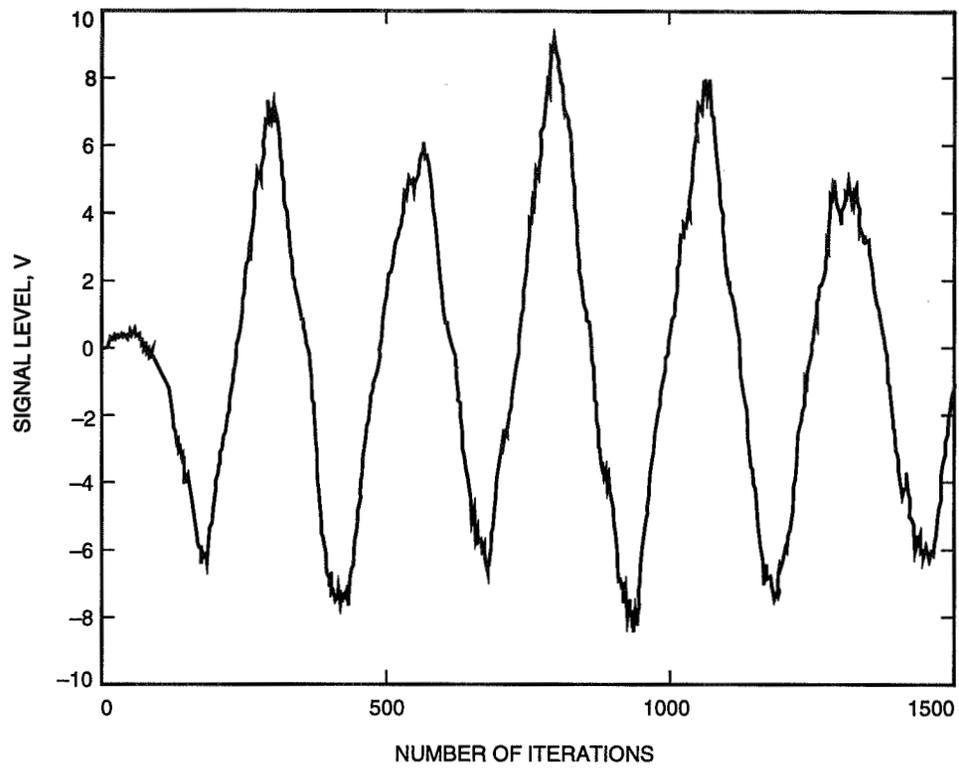


Fig. 13. The 16-tap ALECA output: an estimated carrier signal.

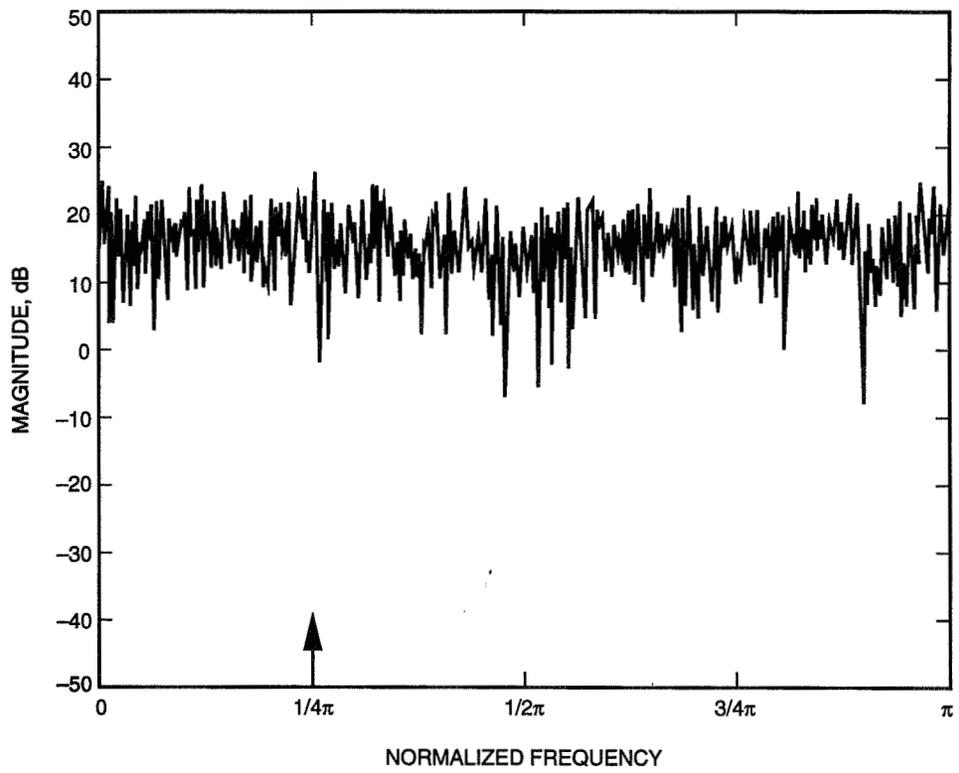


Fig. 14. Magnitude of the input data to the ALECA.

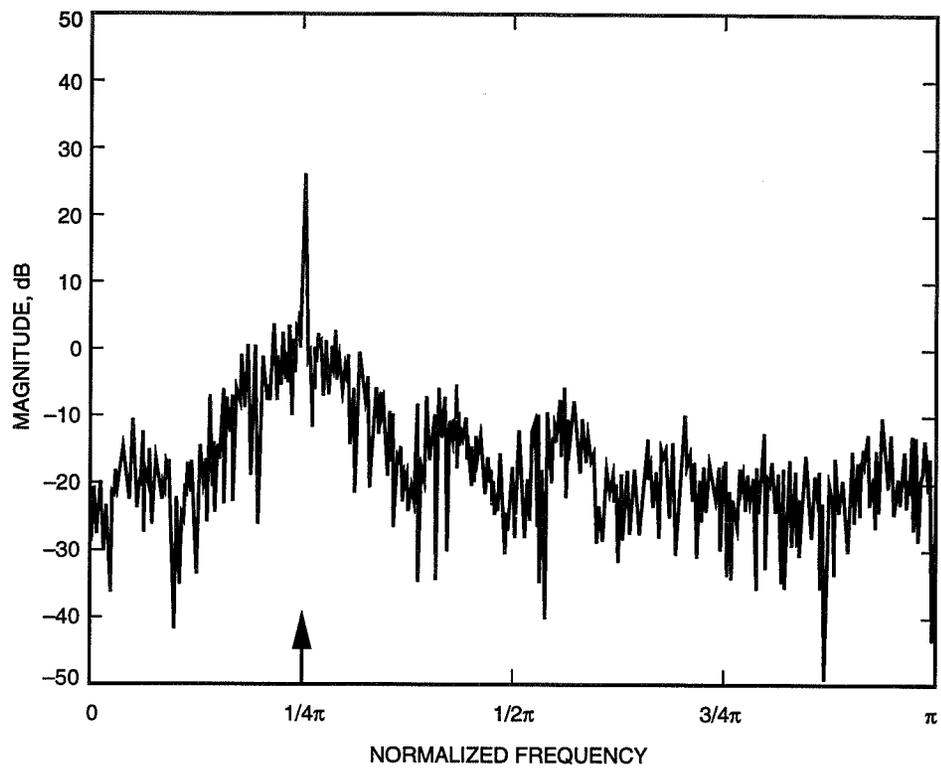


Fig. 15. Magnitude of the ALE output data.

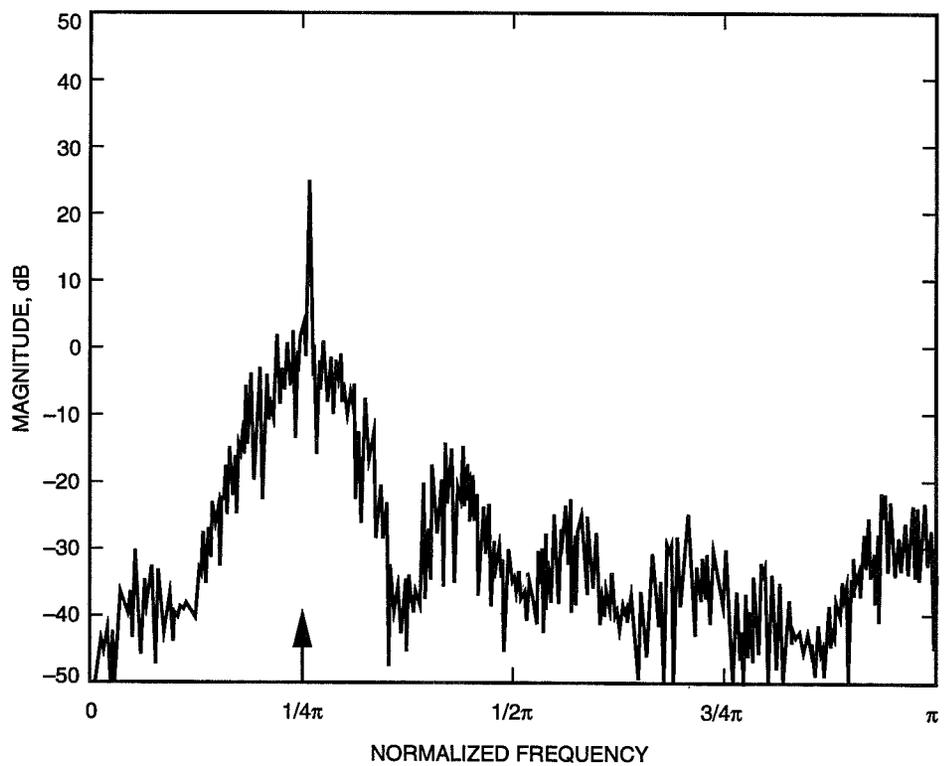


Fig. 16. Magnitude of the ALEDF output data.

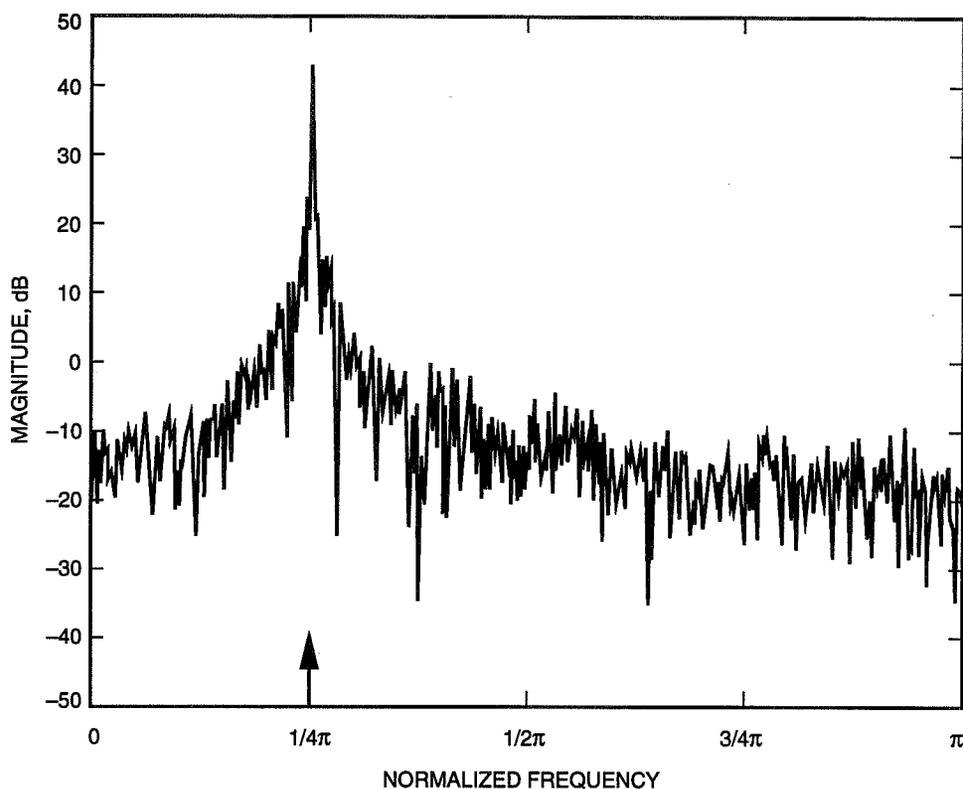


Fig. 17. Magnitude of the ALECA output data.

Furthermore, the ALECA output data provide a very sharp frequency response. These observations all agree with our analysis in Section III.B and as shown in Figs. 6 and 9 for the ALEDF and ALECA, respectively. Theoretically, the SNR of a system using the 16-tap ALE is improved by 12 dB over the same system without the ALE. The performance of both the ALEDF and the ALECA should be better than that of the 16-tap ALE by at most 3 dB, as indicated by Figs. 6 and 9. However, because the weights of these systems slowly fluctuate even in the steady state, the true SNRs at the output are always slightly less than their theoretical values. Note that if the filter length of the ALE is doubled, its theoretical SNR gain becomes 15 dB. However, both the processing time and frequency resolution are doubled.

VI. Discussion and Conclusion

In this article, we presented three ALEs for fast acquisition in the time domain. Conventional ALE, ALEDF, and ALECA systems are introduced. The theoretical performances of these algorithms are presented with computer simulations that support their validity. A fair conclusion is made about both the frequency resolution and processing time.

A. Frequency Resolution

To keep the same frequency resolution, all ALEs have the same number of taps per filter. The positive features of the ALEDF over this conventional ALE are a sharper cutoff frequency in passband and much lower sidelobes in stop band. The advantages of the ALECA over the ALE are a very narrow passband and nearly the same sidelobes in stop band, and a signal gain that can be adjusted by parameter c . The maximum additional CNR gain of either the ALEDF or the ALECA over this conventional ALE is 3 dB due to the second filtering architecture. However, both the ALEDF and the ALECA require 50 percent more computational operations than does the conventional ALE.

B. Processing Time

The filter length of this conventional ALE is twice that of either the ALEDF or the ALECA. Consequently, its frequency resolution is twice that of other ALEs. The advantages of the ALEDF and ALECA over this ALE are a 25-percent saving in multiplications and additions and a 50-percent saving in the memory location required for the weight vector. The CNR gains of both the ALEDF and the ALECA are less than or close to that of this ALE.

These algorithms can be easily implemented via a digital signal processor or application-specific integrated circuits (ASICs) at a sampling frequency of around 100 kHz to acquire the uplink carrier without sweeping the uplink frequency. Furthermore, these algorithms can be easily integrated with either a conventional voltage-controlled oscillator (VCO) in a closed-loop acquisition/tracking architecture, as the present deep-space transponder is, or with a numerically controlled oscillator (NCO) in an open-loop ALE-DFT scheme for acquiring and a closed-loop scheme for tracking the carrier signal.¹

Acknowledgments

The authors wish to acknowledge the valuable support of Rick Crist and Arthur W. Kermode, and special thanks to Selahattin Kayalar, Sami M. Hinedi, and V. A. Vilnrotter for their careful review, correction, and discussions.

References

- [1] S. Acquirre, D. H. Brown, and W. J. Hurd, "Phase Lock Acquisition for Sampled Data PLLs Using the Sweep Technique," *The Telecommunications and Data Acquisition Progress Report 42-86*, vol. April-June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 95-102, August 15, 1986.
- [2] B. Shah and S. Hinedi, "A Comparison of Open-Loop Frequency Acquisition Techniques for Suppressed Carrier Biphase Signals," *The Telecommunications and Data Acquisition Progress Report 42-103*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 170-188, November 15, 1990.
- [3] M. Aung, W. J. Hurd, C. M. Buu, J. B. Berner, S. A. Stephens, and J. M. Gevargiz, "Block V Receiver Fast Acquisition Algorithm for the Galileo S-band Mission," *The Telecommunications and Data Acquisition Progress Report 42-118*, vol. April-June 1994, Jet Propulsion Laboratory, Pasadena, California, pp. 83-114, August 15, 1994.
- [4] R. Kumar, "Fast Frequency Acquisition via Adaptive Least Squares Algorithm," *The Telecommunications and Data Acquisition Progress Report 42-85*, vol. January-March 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 53-61, May 15, 1986.

¹ T. M. Nguyen and H. G. Yeh, "A New Carrier Frequency Acquisition Technique for Future Digital Transponders," in preparation for publication.

- [5] B. Widrow, J. R. Tlover, J. M. Cool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, E. Dong, Jr., R. C. Goodlin, "Adaptive Noise Canceling Principles and Applications," *Proceedings of the IEEE*, vol. 63, pp. 1692–1716, December 1975.
- [6] L. Griffiths, "Rapid Measurement of Digital Instantaneous Frequency," *IEEE Trans. Acoustic, Speech, Signal Processing*, vol. ASSP-23, no. 2, pp. 207–222, April 1975.
- [7] J. R. Zeidler and D. M. Chabries, "An Analysis of the LMS Adaptive Filter Used as a Spectral Line Enhancer," Tech. Note 1476, Naval Undersea Center, Alexandria, Virginia, February 1975.
- [8] W. S. Burdick, "Detection of Narrow-Band Signals Using Time-Domain Adaptive Filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-14, pp. 578–591, July 1978.
- [9] J. T. Richard, J. R. Zeidler, M. J. Dentino, and M. Shensa, "A Performance Analysis of Adaptive Line Enhancer-Augmented Spectral Detectors," *IEEE Trans. Acoustic, Speech, Signal Processing*, vol. ASSP-29, pp. 694–701, June 1981.
- [10] N. J. Bershad, "On the Real and Complex Least Mean Square Adaptive Filter Algorithms," *Proceedings of the IEEE*, vol. 69, no. 4, pp. 469–470, April 1981.
- [11] B. T. Ho, "An Improvement of Adaptive Line Enhancer by Means of Coherent Accumulation Algorithm," *Chinese Journal of Acoustics*, vol. 9, no. 4, 1991.
- [12] D. W. Tufts, L. J. Griffiths, B. Widrow, J. Glover, J. McCool, and J. Treichler, "Adaptive Line Enhancement and Spectrum Analysis," *Proceedings of the IEEE, Letter*, pp. 169–173, January 1977.

Appendix

Stability of the ALECA

From the transfer function $\mathbf{H}_a(z)$, the pole of the ALECA can be found from the characteristic equation:

$$1 - z_0 z^{-1} - c\beta(z_0 z^{-1})^m (1 - (z_0 z^{-1})^{L+1}) = 0 \quad (\text{A-1})$$

Let $z_0 z^{-1} = x$; Eq. (A-1) becomes

$$1 - x - c\beta x^m (1 - x^{L+1}) = 0 \quad (\text{A-2})$$

or

$$\frac{1}{c\beta x^m} = \frac{1 - x^{L+1}}{1 - x} \quad (\text{A-3})$$

Let $x = 1 - \Delta x$ and consider the pole closed to the unit circle (i.e., $\Delta x \approx 0$); make the Taylor's series expansion of x in Eq. (A-3) and take only the first-order term as the approximation:

$$\frac{1}{c\beta(1 - m\Delta x)} = \frac{1 - (1 - (L+1)\Delta x)}{1 - (1 - \Delta x)} \quad (\text{A-4})$$

Solving for Δx ,

$$\Delta x = \frac{(L+1)\beta c - 1}{(L+1)\beta c m} < 0 \quad \text{if } 0 \leq c < 1 \quad (\text{A-5})$$

This implies

$$\text{pole} = z = \frac{z_0}{1 - \Delta x} \quad (\text{A-6})$$

Since Δx is less than 0 by choosing $0 \leq c < 1$ and the magnitude of z_0 is equal to 1, the pole is always located inside a unit circle. Consequently, the ALECA is always stable.

110873

358433

Antenna Noise Temperatures of the 34-Meter Beam-Waveguide Antenna With Horns of Different Gains Installed at F1

T. Y. Otoshi, P. R. Lee, and M. M. Franco
Ground Antennas and Facilities Engineering Section

This article presents a set of theoretical and measured zenith-antenna noise temperatures at 8.45 GHz for the DSS-13 34-m beam-waveguide antenna when horns of different gains are installed at F1. The methodology for calculations is shown in detail. The major differences between calculated and measured values are attributed to changes in subreflector support leg scattering when illuminated by the various horns.

I. Introduction

A study program was initiated to find causes of higher than expected noise temperatures measured at three focal points in a beam-waveguide (BWG) antenna system. The program consists of developing theoretical models for making predictions and obtaining experimental data for model verifications. Theoretical methods are being developed to predict noise temperatures at focal points F1, F2, and F3. Focal point F1 is the Cassegrain focal point near the main reflector vertex. An intermediate focal point F2 lies above the azimuth track, and focal point F3 is the final BWG focal point, located in a subterranean pedestal room. Degradations caused by the BWG system mirrors and shrouds were experimentally determined from comparisons made of values measured at the different focal points [1].

The purpose of this article is to present theoretical and measured zenith-antenna noise temperature values at 8.45 GHz for the DSS-13 34-m BWG antenna with different horns installed at F1. The horns are corrugated, each horn having the same 6.25-deg semiflare angle but of different lengths and aperture diameters.

II. Analytical Procedure and Results

The calculation methodology differs from others used previously in that (1) the noise temperature generated by the horn spillover (into the sky regions) between the edges of the subreflector and main reflector is taken into account, (2) subreflector and main reflector efficiencies (to be defined) are applied as necessary in the calculations of noise temperature contributions, and (3) brightness temperatures are computed for each theta for horn, subreflector, and main reflector radiation patterns for the weather conditions that prevailed at the time of the experiments.

A. Radiation Patterns and Spillover Power Ratios

Table 1 is a summary of the power ratios of the subreflector forward spillover, the main reflector spillover towards ground, and the spillover into the BWG opening on the main reflector. These power ratios were obtained by first calculating horn patterns at the frequency of 8.45 GHz using the circular waveguide mode matching program CWG.F [2]. The term “horn gain” used throughout this article is the directivity of the horn calculated from these patterns. The radiation pattern from the subreflector for each incident horn pattern was then calculated for eight phi-plane cuts with the physical optics program POSUB.F. Next, the spherical wave azimuthal expansion program AZEXP.F was used to integrate the total power in the radiation pattern. The total power was then normalized and subtracted from the horn input power to obtain the subreflector spillover power ratio.

Main reflector spillover towards ground was obtained by using a radiation pattern from the subreflector calculated at a near-field radius of 1780.5 cm, the distance to the outer edge of the main reflector (see Fig. 1). This pattern was used as input into the programs AZEXP.F and EFFIC.F to calculate the fraction of the power in the radiation pattern incident between the edge of the main reflector surface and the horizon. Spillover into the BWG opening of 243.8-cm diameter was calculated similarly, using a radiation pattern from the subreflector calculated at a near-field radius of 1191.3 cm, the distance to the edge of the BWG opening. In addition, Table 1 shows the spillover into the total BWG opening of the 304.8-cm diameter, which includes the Cassegrain cone mounting ring. However, since this mounting ring reflects power to the sky, only the spillover into the actual opening (determined by the 243.8-cm diameter of the BWG shroud walls) will be considered for noise temperature analyses.

For purposes of determining noise temperature contributions due to radiation of reflected power from the main reflector into the far field, antenna patterns and beam efficiencies were obtained for the 34-m-diameter main reflector when illuminated by the subreflector patterns for the various horn-gain cases. Figure 2 shows the far-field radiation patterns for the 34-m BWG antenna for the 29.7-dBi and the 22.5-dBi horn-gain cases. The patterns for the other intermediate horn-gain cases fall somewhere between the patterns for these largest and smallest aperture diameters. As an example, see Fig. 3. These patterns were obtained by using the calculated horn patterns as input to the program POJB.F, which is a physical optics program developed by the Ground Antennas and Facilities Engineering Section. The shapes of the patterns are approximately $J_1(u)/u$, where u is related to theta [3].

Table 2 shows the far-field gains of the BWG antenna at 8.45 GHz for the various horn-gain cases. The gain includes only the losses from the main reflector spillover and the subreflector spillover, and the illumination loss at the main reflector aperture. The actual gain would be lower due to any resistive losses in the horn, subreflector, and main reflector surfaces. Beam efficiencies are tabulated for theta at 0.5 and 1.0 deg for the various horn-gain cases.

B. Antenna Noise Temperature Contributions

For calculations of antenna noise temperature contributions, it is important that power ratios be multiplied by the subreflector and main reflector efficiencies defined in Appendix A. Use of the correct power ratios is a necessary step that must be taken to preserve the conservation of power principle in noise temperature calculations [4]. In Appendix A, it is shown that the requirement is met because the sum of all fractional powers involved in the overall noise temperature calculations is equal to unity.

Table 3 shows a summary of the calculated zenith antenna noise temperature contributions for the 34-m BWG antenna for the various horn-gain cases. Supporting data as well as the calculation methodologies used to obtain these final values may be found in Appendix A.

To compute horn contributions, the brightness temperatures and horn beam efficiencies were first computed from TYO61M4A, which is a FORTRAN program similar to that described in [5]. Then noise temperature contributions from the horn patterns in the sky region (between subreflector and main

Table 1. Predicted F1 spillover amounts, 8.45 GHz, physical optics analysis.^a

Horn diameter, cm (in.)	Horn gain, dBi	p_{S1} (subreflector spillover)	p_{S2} (main reflector spill to ground)	p_{S3} (spill into 304.8-cm BWG opening)	p_{S3} (spill into 243.8-cm BWG opening)
48.346 (19.034)	29.71	0.0294	0.0022	0.0039	0.0023
40.754 (16.045)	28.70	0.0503	0.0030	0.0031	0.0018
31.458 (12.385)	26.90	0.0827	0.0036	0.0020	0.0012
24.948 (9.822)	25.12	0.1465	0.0082	0.0014	0.00086
17.976 (7.077)	22.52	0.3437	0.0149	0.0010	0.00057

^a See Appendix A for further definitions of p_S symbols.

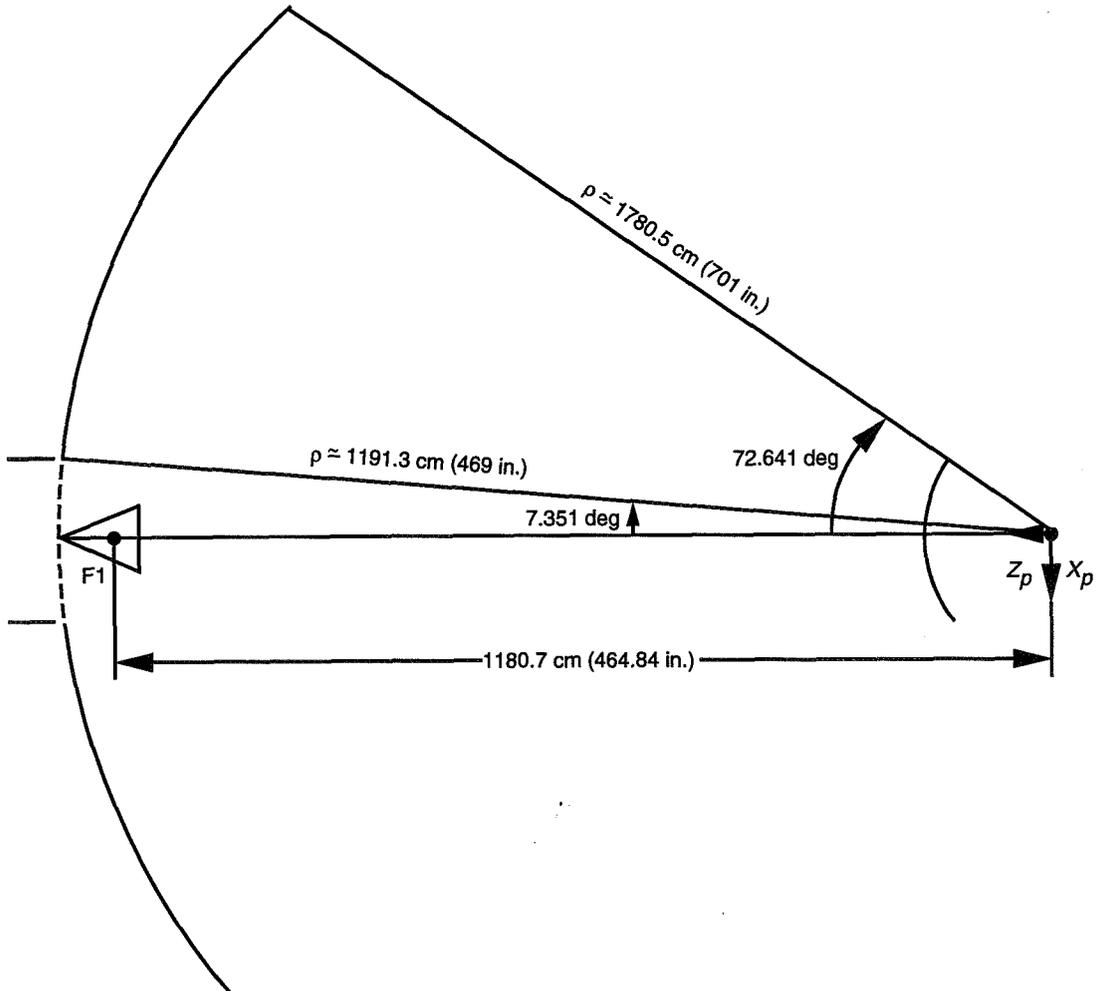


Fig. 1. Geometry for calculating the main reflector and BWG opening contribution using the POSUB.F program.

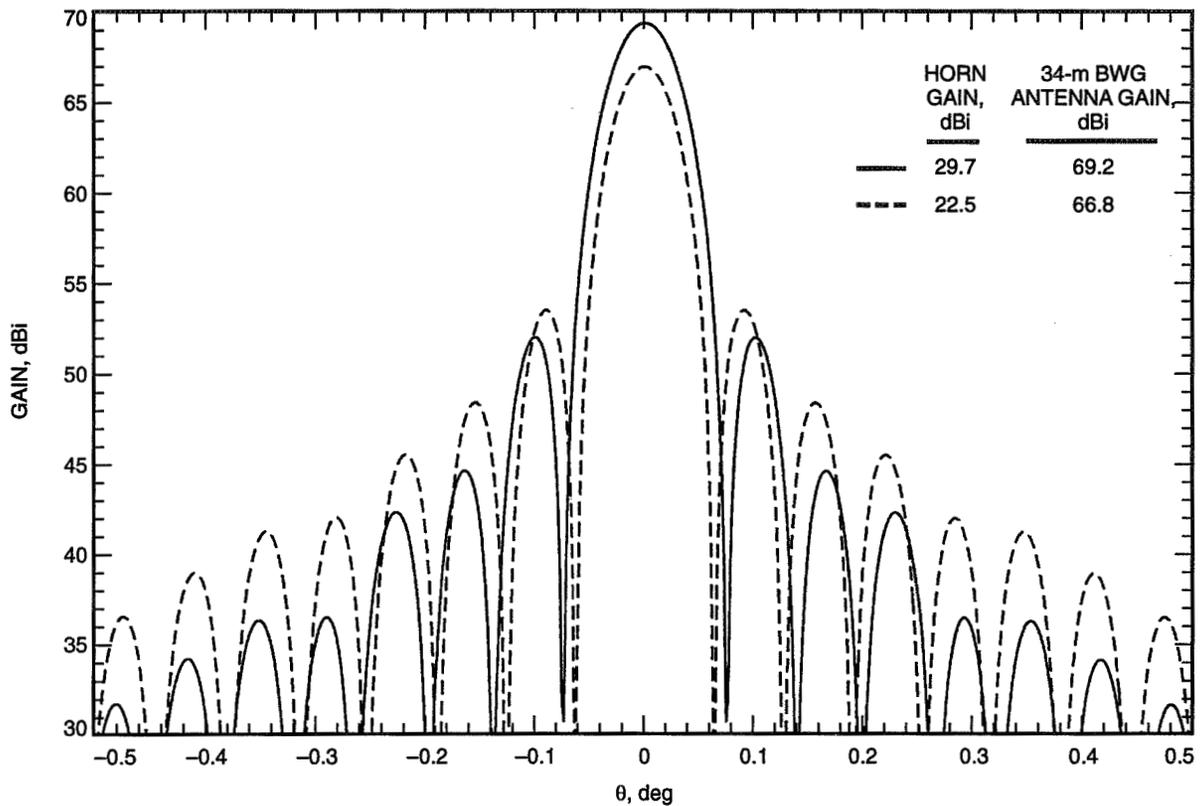


Fig. 2. The 34-m BWG antenna far-field patterns and gain at 8.45 GHz for the largest and smallest aperture horn at F1.

Table 2. Far-field data for the 34-m-diameter antenna at 8.450 GHz for various horns at F1.

Horn gain, dBi	34-m-diameter antenna			
	Gain, dBi	First null, deg	Beam efficiency	
			At 0.5 deg	At 1.0 deg
29.71	69.21	0.075	0.9909	0.9982
28.70	69.10	0.073	0.9921 ^a	0.9988 ^a
26.90	68.83	0.069	0.9838	0.9973
25.12	68.28	0.066	0.9704	0.9939
22.52	66.80	0.063	0.9567	0.9907

^aThe efficiencies at 0.5 deg and 1.0 deg are slightly larger for the 28.7-dBi horn than those for the 29.7-dBi horn. This is attributed to slight differences in the main-lobe beamwidths and the close-in sidelobe patterns. These differences affect the power contained inside the patterns up to the arbitrary 0.5-deg and 1.0-deg angles selected.

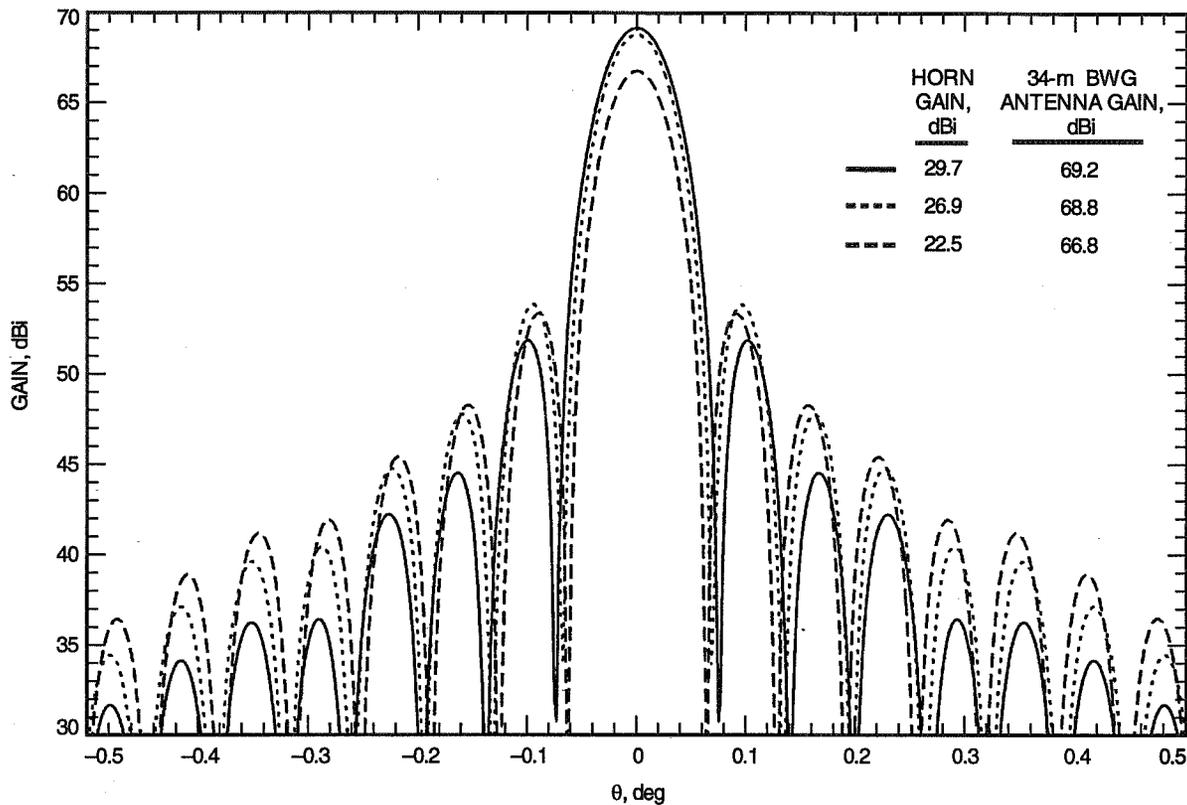


Fig. 3. The patterns of Fig. 2 in addition to the pattern for an intermediate-sized horn aperture.

Table 3. Summary of antenna noise temperature contributions at F1.^a

Horn gain, dBi	ΔT_{A1} (main reflector to zenith), K	ΔT_{A2} (subreflector to sky/ground), K	ΔT_{A3} (subreflector to BWG hole), K	ΔT_{A4} (horn to sky, 8.7 to 68.2 deg), K	ΔT_{A5} (horn to other spills), K	$\sum_{i=1}^5$ ΔT_{Ai} , K	Noise temperature difference, K
29.7	4.370	0.455	0.657	0.121	0.018	5.621	—
28.7	4.263	0.595	0.510	0.194	0.046	5.608	-0.01
26.9	4.125	0.706	0.329	0.307	0.094	5.561	-0.06
25.1	3.826	1.471	0.210	0.564	0.140	6.211	0.59
22.5	2.919	2.059	0.120	1.396	0.232	6.726	1.11

^a See Appendix A for details.

reflector edges) were determined from the program output. A sample case output of this program is included in Appendix B.

Next, as discussed in Section II.A, from the spillover power ratios computed for the region outside the outer edge of the main reflector, the noise temperature contribution was determined by first determining an effective brightness temperature of the ground and low-horizon sky region. Due to the variability of this region with changes in weather and ambient temperature as well as zenith angle, the brightness

temperatures were calculated for the applicable conditions that existed during the corresponding measurement periods. The effective brightness temperatures were obtained through the use of a BASIC program named SREFLNT.BAS. A sample case output may be seen in Appendix B. Effective brightness temperatures for this ground and low-horizon sky region may be seen in Table A-3. It is of interest that these effective brightness temperatures vary from 210 K to 217 K, as contrasted with the 240-K value that has been used at JPL in past years for these types of calculations for Cassegrain antennas.

The BWG opening noise contribution is computed from the power ratio values and equations given in Appendix A. Since the Cassegrain mounting ring surface will reflect power into the sky close to zenith, only the 243.8-cm-diameter portion of the BWG shroud opening should contribute significantly to the antenna noise temperature at F1. Noise temperature calculations are shown in Table A-4. The brightness temperature of the shroud opening was assumed to be equal to the ambient temperature of the antenna at the times that system noise temperature measurements were being performed for corresponding horn configurations.

To determine the contribution due to sky absorption of the power radiated by the main reflector, it was necessary to compute beam efficiencies as a function of zenith angle for the various horn-gain cases. These values as tabulated in Table 2 show that, for all horn-gain cases, 99 percent of the power radiated by the main reflector will be contained inside the main beam between 0 deg and 1 deg of zenith θ angles. Since sky brightness temperature is nearly constant (within 1.5 percent) from 0 to 10 deg of zenith angle, it can be assumed that, for these theoretical cases, 100 percent of the power radiated from the 34-m BWG main reflector surface will be absorbed by sky that has an effective brightness temperature equal to that at the zenith angle of 0 deg. It should be pointed out that the brightness temperature of the zenith sky, as calculated from program TY061M4A, includes the cosmic background and atmospheric loss contributions (see Appendix B).

Table 3 shows most of the contributions to antenna noise temperatures. If the total absolute antenna noise temperatures need to be computed, then knowledge is also required of such additional terms as noise due to tripod scattering of the far-field plane wave illuminating the entire antenna, main reflector panel leakage, main reflector and subreflector resistive losses, and other terms presented in [6]. It will be shown later in this article that this residual term is approximately equal to 3.0 K.

III. Experimental Work

Operating noise temperature measurements were sequentially made at 8.450 GHz with the phase centers of the 29.7-, 28.7-, 26.9-, 25.1-, and 22.5-dBi horns aligned with the geometric focal point F1. Figure 4 shows the experimental test setup of the 29.7-dBi horn at F1.

The experimental data are presented in Tables 4 and 5. In Table 4,

$$\Delta T_{op1} = L_{wg}[T_{op} - (T_{op})_{ref}] \quad (1)$$

where $(T_{op})_{ref}$ is the normalized operating noise temperature of the 29.7-dBi horn, and L_{wg} is the waveguide loss from the horn apertures to the high-electron-mobility-transistor (HEMT) input. It was assumed that L_{wg} measured for the 22.5-dBi horn [1] applied to all horn-gain cases.

To experimentally determine only the noise temperature contribution from spillover into the 243.8-cm-diameter BWG opening at F1 below the feedhorn, a ground plane was installed below the feedhorns. Figure 5 shows the installation of a Thermax¹ ground plane below the 29.7-dBi horn. The

¹ A trademark of the Celotex Corporation, Tampa, Florida.

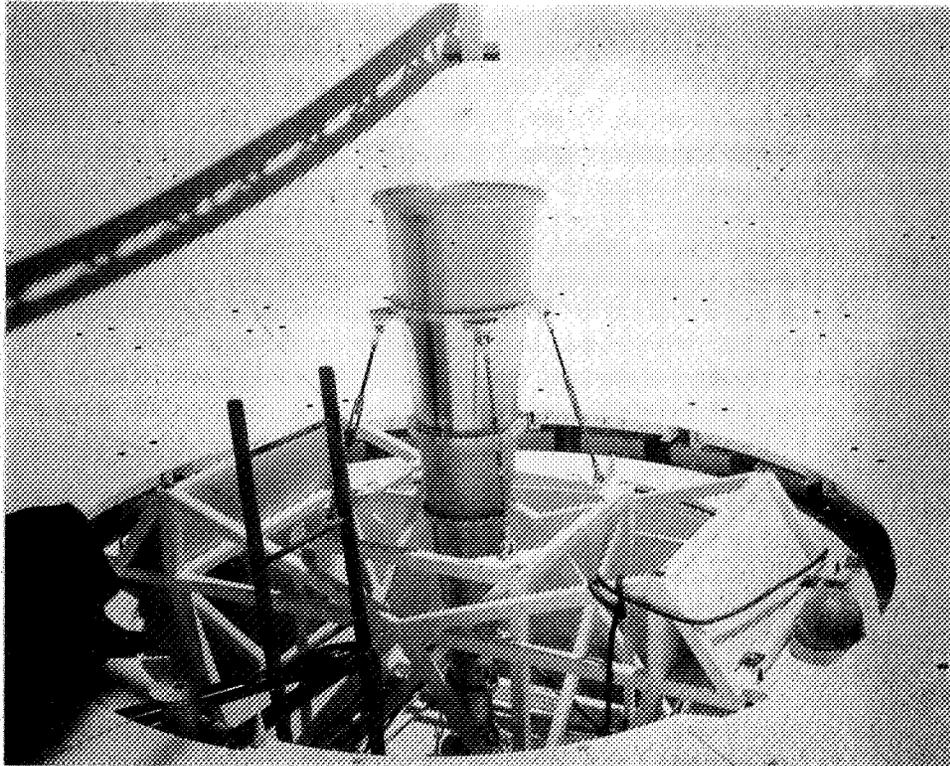


Fig. 4. Partial view of the X-band 29-dBi horn test package and mounting structure installed at the Cassegrain focal point F1.

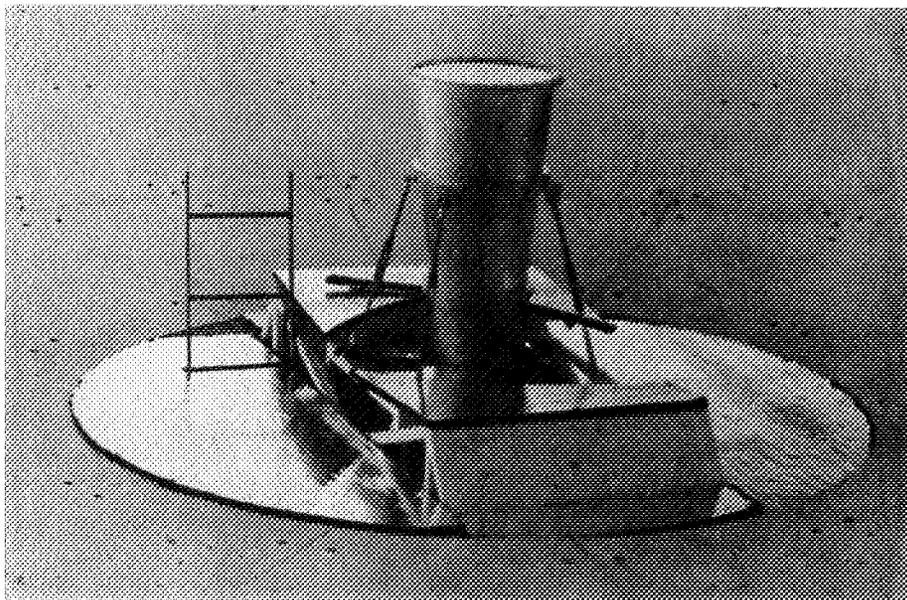


Fig. 5. The installation shown in Fig. 4 with the addition of a Thermax™ ground plane that covers the 243.8-cm-diameter BWG opening below the feedhorn.

Table 4. DSS-13 measured and normalized noise temperature (T_{op}) without a Thermax™ conducting plane below the feedhorn at F1.^a

Horn gain, dBi	Horn phase center, ^b cm	Relative humidity, percent	Temperature, deg C		Barometric pressure, mb	T_{op} , K		Measured ΔT_{op1} , ^e K
			Load	Outdoor		Measured ^c	Normalized ^d	
29.7	116.8	5.00	28.56	25.47	896.88	27.08	27.07	0.00
28.7	82.6	4.42	28.43	26.90	895.80	27.13	27.14	0.05
26.9	44.5	5.01	28.80	25.62	895.59	27.12	27.12	0.04
25.1	25.4	5.17	29.60	26.67	895.30	28.07	28.04	1.01
22.5	7.9	4.92	29.97	26.79	894.99	29.03	29.00	1.98

^a Frequency = 8.45 GHz.

^b Phase center location is measured down from the horn aperture.

^c Data were taken with the research and development (R&D) X-band test package described in [1]. $L_{wg} = 1.0163$ (0.07 dB).

^d Normalized T_{op} corrected for standard atmosphere and waveguide ambient temperature [1].

^e See Eq. (1).

Table 5. DSS-13 measured and normalized noise temperature (T_{op}) with a Thermax™ conducting plane below the feedhorn.^a

Horn gain, dBi	Horn phase center, ^b cm	Relative humidity, percent	Temperature, deg C		Barometric pressure, mb	T_{op} , K		Measured ΔT_{op2} , ^e K
			Load	Outdoor		Measured ^c	Normalized ^d	
29.7	116.8	5.35	27.95	24.87	896.99	26.67	26.67	0.41
28.7	82.6	4.33	28.28	26.39	895.59	26.74	26.76	0.39
26.9	44.5	5.17	28.73	25.81	895.59	26.83	26.82	0.31
25.1	25.4	5.04	29.76	26.59	895.29	27.80	27.78	0.26
22.5	7.9	4.93	29.92	27.07	894.99	28.74	28.71	0.30

^a Frequency = 8.45 GHz.

^b Phase center location is measured down from the horn aperture.

^c Data were taken with the R&D X-band test package described in [1]. $L_{wg} = 1.0163$ (0.07 dB).

^d Normalized T_{op} corrected for standard atmosphere and waveguide ambient temperature [1].

^e See Eq. (2).

last column of Table 5 shows the change in T_{op} values measured with and without the Thermax™ ground plane. The values were computed from

$$\Delta T_{op2} = L_{wg} [(T_{op})_{w/o} - (T_{op})_{with}] \quad (2)$$

This difference provides a measured value of the noise contribution due to spillover into the 243.8-cm-diameter BWG opening at F1, and the difference is compared to the theoretical ΔT_{A3} values shown in Table 3.

IV. Discussion of Results

The zenith T_{op} value of the 34-m BWG antenna with the X-band test package and a particular feedhorn installed at F1 can be expressed as

$$(T_{op})_{F1} = L_{wg}^{-1}T_{A,F1} + T_{wg} + T_{hemt} + T_{fu} \quad (3)$$

where $T_{A,F1}$ is the 34-m BWG antenna noise temperature at F1 and the other symbols are defined in Appendix A. Algebraic manipulation of Eq. (3) gives

$$T_{A,F1} = (T_{op})_{F1} - L_{wg}(T_{wg} + T_{hemt} + T_{fu}) \quad (4)$$

From [1], $L_{wg} = 1.0163$, $T_{wg} = 4.69$ K, $T_{hemt} = 13.0$ K, and $T_{fu} = 0.4$ K, and from Table 4, the measured $(T_{op})_{F1}$ for the 29.7-dBi horn is shown to be 27.08 K. Substitution of these values in Eq. (4) gives

$$T_{A,F1} = 8.70 \text{ K}$$

which includes the cosmic background and atmospheric loss contributions. It is also valid to state that

$$T_{A,F1} = \sum_{i=1}^5 \Delta T_{Ai} + T_u \quad (5)$$

which gives

$$T_u = T_{A,F1} - \sum_{i=1}^5 \Delta T_{Ai} \quad (6)$$

where T_u is a residual (currently unknown) contribution due to strut scattering, leakage through gaps in the main reflector surface, holes in the main reflector surface panels, and subreflector and main reflector resistive losses. Some of the values for these contributions can be obtained from [6].

Substitution of $T_{A,F1} = 8.70$ K and $\sum_{i=1}^5 \Delta T_{Ai} = 5.62$ K (from Table 3) in Eq. (6) gives

$$T_u = 3.08 \text{ K}$$

Another method of getting an approximate value for T_u is from the relationship

$$(T_u)_{approx} = (T_{op})_{F1} - (T_{op})_{ground} \quad (7)$$

where $(T_{op})_{ground}$ is the operating noise temperature of the portable front-end X-band test package (with horn) on the ground [1]. The value on the right-hand side of Eq. (7) was shown in [1] to be 3.2 K for the 29.7-dBi horn configuration.

For interest, differential antenna noise temperatures (relative to the 29.7-dBi horn) are presented in Table 6. Values for this table were derived from values given in Tables 3 and 4. The differential values

shown in the last column of Table 6 are attributed primarily to residual tripod (strut) contributions not taken into account correctly by the theoretical method. Tripod contributions are difficult to calculate, but an estimate of the changes can be obtained by first accounting for all known or best-estimate contributions (excluding tripod contributions) and then subtracting the calculated total from the measured value. Note that more noise contribution from the tripod occurs for the horns of smaller gain. As shown in Table 7, horns of smaller gain illuminate more of the tripod legs away from the tripod-subreflector connection points and thus have higher tripod-scatter contributions.

Table 6. Relative differences between antenna temperatures of the 29.7-dBi horn and other horns installed at F1.^a

Horn gain, dBi	ΔT_A , ^b K		
	Measured ^c	Calculated ^d	Difference
29.7	—	—	—
28.7	0.05	-0.01	0.06
26.9	0.04	-0.06	0.10
25.1	1.01	0.59	0.42
22.5	1.98	1.11	0.87

^a Frequency = 8.45 GHz.

^b $\Delta T_A = [(T_A)_{horn} - (T_A)_{29.7 \text{ dBi horn}}] \approx \Delta T_{opt}$.

^c From Table 4.

^d From Table 3.

Table 7. Comparison of fractional powers in scatter regions near the subreflection edge.

Horn gain, dBi	θ_i , deg	Fraction of total power contained in θ_i region
29.7	8.7 to 11	0.0173
	11 to 17	0.0084
28.7	8.7 to 11	0.0291
	11 to 17	0.0121
26.9	8.7 to 11	0.0357
	11 to 17	0.0282
25.1	8.7 to 11	0.0747
	11 to 17	0.0429
22.5	8.7 to 11	0.1533
	11 to 17	0.1293

V. Conclusions

A methodology has been presented for calculating antenna noise temperature contributions of a BWG antenna for various horns placed at F1. A reference antenna noise temperature is based on the 29.7-dBi horn because this is the equivalent horn for which the BWG antenna was designed. As the

horn gains become smaller towards 22.5 dBi, a larger discrepancy in noise temperature, of about 0.87 K, occurs. This discrepancy is attributed to larger unknown noise contributions from tripod scattering as the horn gain becomes smaller.

References

- [1] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "Portable Microwave Test Packages for Beam-Waveguide Antenna Performance Evaluations," *IEEE Trans. on Microwave Theory and Techniques, Special Issue: Microwaves in Space*, vol. MTT-40, no. 6, pp. 1286–1293, June 1992.
- [2] D. Hoppe, "Scattering Matrix Program for Circular Waveguide Junctions," *Cosmic Software Catalog*, 1987 edition, NASA-CR-179669, NTO-17245, NASA's Computer Software Management and Information Center, Athens, Georgia, 1987.
- [3] S. Silver, ed., *Microwave Antenna Theory and Design*, Radiation Laboratory Series, New York: McGraw-Hill Book Company, Inc., p. 194, 1949.
- [4] W. W. Mumford and E. H. Scheibe, *Noise Performance Factors in Communication Systems*, Dedham, Massachusetts: Horizon House-Microwave, Inc., pp. 16–17 and 20, 1968.
- [5] T. Y. Otoshi and C. T. Stelzried, "Antenna Temperature Analysis," *Space Program Summary No. 37-36*, vol. IV, Jet Propulsion Laboratory, Pasadena, California, pp. 262–267, December 31, 1965.
- [6] D. A. Bathker, W. Veruttipong, T. Y. Otoshi, and P. W. Cramer, Jr., "Beam-Waveguide Antenna Performance Predictions with Comparisons to Experimental Results," *IEEE Trans. on Microwave Theory and Techniques, Special Issue: Microwaves in Space*, vol. MTT-40, no. 6, pp. 1274–1285, June 1992.

Appendix A

Analysis of Antenna Noise Temperature Contributions at F1

I. Nomenclature (See Fig. A-1)

- α_{H1} Fraction of the total horn-radiated power that is captured by the subreflector. This term is equal to the subreflector efficiency.
- α_{H2} Fraction of the total horn-radiated power that is absorbed by the sky region between the edges of the subreflector and main reflector.
- α_{H3} Fraction of the total horn-radiated power that goes into cross-polarization and becomes absorbed primarily by the sky region, between the subreflector edge and the 50-deg elevation angle.
- α_{A1} Fraction of the total horn-radiated power that becomes absorbed by the zenith sky after reflections from the subreflector and main reflector.
- α_{A2} Fraction of the total horn-radiated power that becomes absorbed by ground and low-horizon sky region outside the main reflector edge after reflection from the subreflector.
- α_{A3} Fraction of the total horn-radiated power that becomes absorbed by the environment inside the BWG opening on the main reflector.
- η_{SR} Subreflector efficiency.
- η_{MR} Main reflector efficiency.
- η_B Beam efficiency.
- L_{wg} Loss factor for the waveguide between the horn aperture and the input flange of the HEMT, power ratio.
- p_{S1} Fraction of the total horn-radiated power not captured by the subreflector.
- p_{S2} Fraction of the subreflector reflected power that radiates and becomes absorbed by the region outside the main reflector. Normalization is done with respect to the subreflector power pattern and not with respect to the horn-radiated power.
- p_{S3} Fraction of the subreflector reflected power that radiates to the BWG opening on the main reflector. Normalization is done with respect to the subreflector power pattern and not with respect to the horn-radiated power.
- $(T_b)_{A1}$ Brightness temperature of the zenith sky, K.
- $(\overline{T_b})_{A2}$ Effective brightness temperature of the nonoptical path environment inside the BWG hole opening, K. (Assume ambient temperature environment.)
- $(\overline{T_b})_{A3}$ Effective brightness temperature of the ground and low-horizon sky region for spillover past the main reflector edge, K.
- $(\overline{T_b})_{H2}$ Effective brightness temperature for the sky region between the subreflector edge and main reflector edge, K.
- $(\overline{T_b})_{H3}$ Effective brightness temperature of the cross-polarization environment, K. (Assume $(\overline{T_b})_{H3}$ to be sky at a 50-deg elevation angle.)
- T_{fu} Effective noise temperature of the follow-up receiver, K.
- T_{hemt} Effective noise temperature of the HEMT as defined at the input flange of the HEMT, K.

- T_{op} Operating noise temperature, K.
- T_p Physical temperature of absorber (approximately 300 K).
- $T_{sky,zen}$ Brightness temperature of zenith sky, including cosmic background and atmospheric contributions, K.
- T_u Unknown residual contribution not taken into account by the method presented in this article, K.
- T_{wg} Noise temperature due to waveguide loss between the horn aperture and the input flange of the HEMT, K.

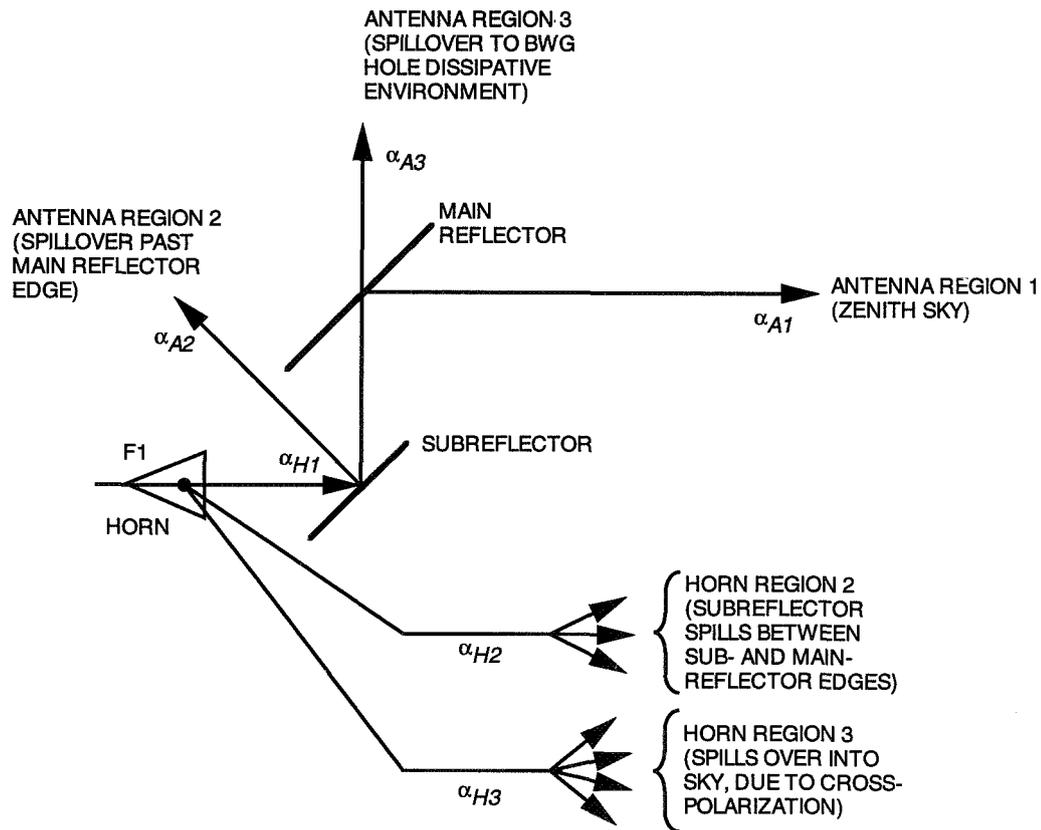


Fig. A-1. Brightness temperature regions contributing to antenna noise temperatures for the horn at $F1$.

II. Antenna Noise Temperature Contribution Relationships

The following assumptions are made:

- (1) The p_{S1} , p_{S2} , and p_{S3} data were obtained from programs discussed in Section II and presented in Table 1.
- (2) The α_{H2} data were obtained from Program TYO61M4A, as discussed in Section II.
- (3) Normalization of power ratios was done according to the definitions given in Section A-I.

In Fig. A-1, the following relationships hold:

$$\alpha_{H1} = 1 - p_{S1} = \eta_{SR} \quad (\text{A-1})$$

$$\alpha_{H3} = p_{S1} - \alpha_{H2} = 1 - \alpha_{H1} - \alpha_{H2} \quad (\text{A-2})$$

$$\eta_{MR} = 1 - p_{S2} - p_{S3} \quad (\text{A-3})$$

$$\alpha_{A1} = \eta_{SR}\eta_{MR} \quad (\text{A-4})$$

$$\alpha_{A2} = p_{S2}\eta_{SR} = p_{S2}(1 - p_{S1}) \quad (\text{A-5})$$

$$\alpha_{A3} = p_{S3}\eta_{SR} = p_{S3}(1 - p_{S1}) \quad (\text{A-6})$$

Note that the following conservation of power relationships hold:

$$\alpha_{H1} + \alpha_{H2} + \alpha_{H3} = 1 \quad (\text{A-7})$$

and since

$$\alpha_{H1} = \alpha_{A1} + \alpha_{A2} + \alpha_{A3} \quad (\text{A-8})$$

then

$$\alpha_{A1} + \alpha_{A2} + \alpha_{A3} + \alpha_{H2} + \alpha_{H3} = 1 \quad (\text{A-9})$$

The antenna noise temperature contributions are

$$\Delta T_{A1} = \alpha_{A1} T_{sky,zen} \quad (\text{A-10})$$

$$\Delta T_{A2} = \alpha_{A2} (\overline{T_b})_{A2} \quad (\text{A-11})$$

$$\Delta T_{A3} = \alpha_{A3} (\overline{T_b})_{A3} = \alpha_{A3} T_p \quad (\text{A-12})$$

$$\Delta T_{A4} = \alpha_{H2} (\overline{T_b})_{H2} \quad (\text{A-13})$$

$$\Delta T_{A5} = \alpha_{H3} (\overline{T_b})_{H3} \quad (\text{A-14})$$

The above relations given by Eqs. (A-10) through (A-14) neglect the fact that theoretical patterns do not consider scattering from the tripod subreflector supports that raise the far-out side-lobe levels. Neglecting the tripod-subreflector scattering, the theoretical total antenna noise temperature of the 34-m-diameter antenna at F1 is

$$T_A = \sum_{i=1}^5 \Delta T_{Ai} + T_u \quad (\text{A-15})$$

where T_u is other unknown contributions from panel leakage, resistive subreflector and main reflector losses, and tripod-scattering effects. As discussed in Section IV, an approximate value of T_u can be obtained from calculated values using the methodology given in [1] or from the measured value of

$$(T_u)_{approx} = (T_{op})_{F1} - (T_{op})_{ground} \quad (A-16)$$

In a future more rigorous analysis, some of the fractional powers associated with T_u might be calculable and then taken into account by including a new term, α_{A4} , in the power ratio relationships given in Eqs. (A-8) and (A-9).

III. Calculated Contributions to Antenna Temperature at 8.450 GHz

Tables A-1 through A-6 are tabulations of calculated antenna temperature contributions for the regions shown in Fig. A-1.

Table A-1. Tabulation of fractional powers and efficiencies at 8.45 GHz.^a

Horn gain, dBi	$\alpha_{H1} = \eta_{SR}$	α_{H2}^b	α_{H3}	η_{MR}^c	α_{A1}	α_{A2}	α_{A3}
29.7	0.9706	0.0264	0.0030	0.9955	0.9662	0.0021	0.0022
28.7	0.9497	0.0427	0.0076	0.9952	0.9451	0.0028	0.0017
26.9	0.9173	0.0671	0.0156	0.9952	0.9129	0.0033	0.0011
25.1	0.8535	0.1231	0.0234	0.9909	0.8457	0.0070	0.0007
22.5	0.6563	0.3051	0.0386	0.9845	0.6461	0.0098	0.0004

^a See Eqs. (A-1) through (A-6) and Table 1 p_S values for calculation of values presented in this table.

^b The α_{H2} values were obtained from Program TYO61M4A. See Table A-5.

^c Use data for the 243.8-cm-diameter BWG opening.

Table A-2. Antenna region 1 contributions (shown in Fig. A-1).

Horn gain, dBi	Average temperature, deg C	Relative humidity, percent	$T_{pg},^a$ K	$(T_b)_{A1},^b$ K	α_{A1} (see Table A-1)	ΔT_{A1} , K (see Eq. (A-10))
29.7	25.47	5.00	298.6	4.523	0.9662	4.370
28.7	26.90	4.42	300.1	4.511	0.9451	4.263
26.9	25.62	5.01	298.8	4.519	0.9129	4.125
25.1	26.67	5.17	299.8	4.524	0.8457	3.826
22.5	26.79	4.92	300.0	4.518	0.6461	2.919

^a Ground level air temperature.

^b $(T_b)_{A1} = T_{sky}$ looking at zenith, K. Obtain values from TYO61M4A.

Table A-3. Contributions from antenna region 2 (shown in Fig. A-1).

Horn gain, dBi	α_{A2} (see Table A-1)	$(\overline{T}_b)_{A2},^a$ K	$\Delta T_{A2},^b$ K
29.7	0.0021	216.7	0.455
28.7	0.0028	212.5	0.595
26.9	0.0033	213.9	0.706
25.1	0.0070	210.1	1.471
22.5	0.0098	210.1	2.059

^a $(\overline{T}_b)_{A2}$ is derived by dividing ΔT_{A2} by α_{A2} .

^b ΔT_{A2} is obtained from Program SREFLNT.BAS. See sample output in Appendix B.

Table A-4. Contributions from antenna region 3 (shown in Fig. A-1).^a

Horn gain, dBi	α_{A3} (see Table A-1)	T_p , K	Calculated ΔT_{A3} , K	Measured $\Delta T_{A3},^b$ K	Difference, K
29.7	0.0022	298.6	0.657	0.41	0.25
28.7	0.0017	300.1	0.510	0.39	0.12
26.9	0.0011	298.8	0.329	0.31	0.02
25.1	0.0007	299.8	0.210	0.26	-0.05
22.5	0.0004	300.0	0.120	0.30	-0.18

^a Based on the 243.8-cm-diameter BWG opening.

^b See Table 5.

Table A-5. Contributions from horn region 2 (shown in Fig. A-1).^a

Horn gain, dBi	8.7 deg		68.2 deg		$\Delta T_{A4},^c$ K	α_{H2}^d
	$\sum T_{Ai},^b$ K	η_B^b	$\sum T_{Ai},^b$ K	η_B^b		
29.7	4.4081	0.9736	4.5288	0.99998	0.1207	0.0264
28.7	4.3238	0.9573	4.5182	0.99997	0.1944	0.0427
26.9	4.2223	0.9329	4.5289	0.99995	0.3066	0.0671
25.1	3.9745	0.8768	4.5384	0.99997	0.5639	0.1231
22.5	3.1458	0.6946	4.5419	0.9997	1.3961	0.3051

^a Horn radiation between subreflector edge (8.7 deg) and main reflector edge (68.2 deg).

^b The values for $\sum T_{Ai}$ and η_B were obtained from the TYO61M4A program. See Appendix B for a sample case printout.

^c $\Delta T_{A4} = \sum T_{Ai}$ at 68.2 deg - $\sum T_{Ai}$ at 8.7 deg.

^d $\alpha_{H2} = \eta_B$ at 68.2 deg - η_B at 8.7 deg.

**Table A-6. Contributions for horn region 3
spillover losses (see Fig. A-1).**

Horn gain, dBi	α_{H3} (see Table A-1)	$(\overline{T_b})_{H3}$, K	ΔT_{A5} , K
29.7	0.0030	6.0	0.018
28.7	0.0076	6.0	0.046
26.9	0.0156	6.0	0.094
25.1	0.0234	6.0	0.140
22.5	0.0386	6.0	0.232

Appendix B

Sample Printouts of Computer Programs Used to Compute Antenna Temperature Contributions

I. Program TYO61M4A.FOR (Only partial printout shown)

PROGRAM TYO61M4A.FOR
BY T. OTOSHI
LATEST REVISION: JUNE 30, 1993

PROGRAM TO COMPUTE ANTENNA GAIN AND ANTENNA
TEMPERATURE FROM PATTERN DATA
AND BRIGHTNESS TEMPERATURE DATA
THAT IS EITHER READ IN OR GENERATED BY PROGRAM

DATE OF THIS RUN: 3-15-94

OUTPUT FILENAME: OUTDAT\29_8DBNC.OD1

ANT PATTERN FILENAME: PAT\19_034M.PAT

BRIGHTNESS TEMPERATURES ARE GENERATED BY THIS PROGRAM

BRIGHTNESS TEMP PARAMETER FILENAME: NBTP\19_034.BT1

29.7 DBI CORR HORN PATTERN AT 8.450 GHZ FROM P.LEE 5/26/93
X-BAND AT F1, 29.8 DB HORN, NO CELOTEX, 14SEP 1993, 8.45 GHZ

ANTENNA BEAM EFFICIENCY, TEMPERATURE, AND GAIN E AND H PLANE PATTERN CASE

I	THETA1 (DEG)	F1 (DB)	F2 (DB)	TBI (DEG K)	ETA1	SUM ETA1	TAI (DEG K)	SUMTAI (DEG K)	THETA1 (DEG)
					ZENITH VALUE				
1	.0	.0	.0	4.523	.000000	.000000	.000000	.000000	.5
2	1.0	-.4	-.3	4.524	.137819	.137819	.623431	.623431	1.5
3	2.0	-1.4	-1.4	4.524	.215778	.353597	.976212	1.599643	2.5
4	3.0	-3.2	-3.2	4.526	.214926	.568524	.972756	2.572399	3.5
5	4.0	-5.7	-5.6	4.528	.162605	.731129	.736249	3.308648	4.5
6	5.0	-8.7	-8.6	4.531	.102488	.833617	.464366	3.773014	5.5
7	6.0	-11.7	-11.6	4.534	.061269	.894886	.277796	4.050810	6.5
8	7.0	-14.2	-14.2	4.538	.039816	.934702	.180700	4.231510	7.5
9	8.0	-16.5	-16.4	4.543	.027227	.961929	.123685	4.355195	8.5
8.7 deg									
10	9.0	-19.2	-18.9	4.548	.016613	.978541	.075560	4.430755	9.5
11	10.0	-22.8	-22.3	4.554	.008287	.986829	.037739	4.486494	10.5
12	11.0	-26.3	-26.0	4.561	.003995	.990823	.018220	4.486713	11.5
13	12.0	-28.1	-28.0	4.568	.002782	.993605	.012709	4.499422	12.5
14	13.0	-29.4	-29.1	4.576	.002284	.995890	.010454	4.509876	13.5
15	14.0	-32.0	-31.3	4.585	.001413	.997303	.006478	4.516354	14.5
16	15.0	-36.1	-35.2	4.594	.000609	.997912	.002797	4.519151	15.5
17	16.0	-38.1	-38.4	4.604	.000354	.998265	.001629	4.520780	16.5
18	17.0	-37.6	-38.0	4.615	.000415	.998681	.001918	4.522697	17.5
19	18.0	-38.8	-38.3	4.627	.000372	.999053	.001719	4.524417	18.5
20	19.0	-42.5	-41.1	4.639	.000186	.999239	.000864	4.525280	19.5

21	20.0	-46.7	-45.8	4.652	.000070	.999309	.000328	4.525609	20.5
22	21.0	-44.8	-45.8	4.667	.000091	.999400	.000426	4.526034	21.5
23	22.0	-44.0	-44.1	4.681	.000126	.999527	.000592	4.526627	22.5
24	23.0	-46.4	-45.0	4.697	.000091	.999618	.000427	4.527054	23.5
25	24.0	-51.9	-49.1	4.714	.000033	.999650	.000154	4.527207	24.5
26	25.0	-51.7	-53.3	4.732	.000020	.999671	.000097	4.527304	25.5
27	26.0	-48.6	-50.1	4.750	.000044	.999715	.000208	4.527513	26.5
28	27.0	-48.9	-48.8	4.770	.000051	.999765	.000242	4.527754	27.5
29	28.0	-52.6	-50.6	4.790	.000028	.999794	.000136	4.527891	28.5
30	29.0	-58.9	-55.9	4.812	.000008	.999802	.000039	4.527929	29.5
31	30.0	-54.9	-57.9	4.834	.000010	.999812	.000050	4.527979	30.5
32	31.0	-52.3	-53.7	4.859	.000022	.999835	.000109	4.528088	31.5
33	32.0	-53.2	-52.8	4.884	.000023	.999858	.000112	4.528200	32.5
34	33.0	-57.4	-55.0	4.911	.000012	.999869	.000057	4.528257	33.5
35	34.0	-61.9	-61.1	4.938	.000003	.999873	.000017	4.528275	34.5
36	35.0	-57.2	-62.1	4.968	.000006	.999879	.000031	4.528305	35.5
37	36.0	-55.1	-57.1	4.998	.000013	.999892	.000064	4.528368	36.5
38	37.0	-55.9	-55.9	5.031	.000013	.999905	.000067	4.528435	37.5
39	38.0	-60.0	-57.5	5.064	.000007	.999912	.000037	4.528472	38.5
40	39.0	-68.2	-62.6	5.101	.000002	.999914	.000010	4.528482	39.5
41	40.0	-62.6	-68.1	5.137	.000002	.999916	.000010	4.528492	40.5
42	41.0	-58.9	-61.8	5.178	.000005	.999921	.000028	4.528520	41.5
43	42.0	-58.7	-59.2	5.218	.000007	.999929	.000038	4.528559	42.5
44	43.0	-61.2	-59.4	5.263	.000006	.999934	.000029	4.528588	43.5
45	44.0	-67.2	-62.4	5.308	.000002	.999937	.000012	4.528600	44.5
46	45.0	-67.4	-69.3	5.357	.000001	.999938	.000005	4.528605	45.5
47	46.0	-62.1	-69.5	5.407	.000002	.999940	.000012	4.528617	46.5
48	47.0	-60.0	-63.6	5.462	.000004	.999944	.000025	4.528642	47.5
49	48.0	-60.1	-61.6	5.517	.000005	.999950	.000029	4.528671	48.5
50	49.0	-62.2	-61.8	5.578	.000004	.999954	.000023	4.528694	49.5
51	50.0	-66.5	-64.1	5.640	.000002	.999956	.000011	4.528706	50.5
52	51.0	-69.5	-69.2	5.709	.000001	.999956	.000004	4.528710	51.5
53	52.0	-65.6	-75.8	5.777	.000001	.999958	.000006	4.528716	52.5
54	53.0	-62.8	-69.1	5.854	.000002	.999960	.000013	4.528728	53.5
55	54.0	-61.9	-65.5	5.931	.000003	.999963	.000019	4.528748	54.5
56	55.0	-62.3	-64.4	6.018	.000003	.999966	.000020	4.528768	55.5
57	56.0	-64.2	-64.9	6.106	.000002	.999969	.000015	4.528783	56.5
58	57.0	-67.6	-67.2	6.204	.000001	.999970	.000008	4.528791	57.5
59	58.0	-73.2	-71.9	6.303	.000000	.999970	.000003	4.528973	58.5
60	59.0	-75.2	-80.9	6.416	.000000	.999971	.000001	4.528794	59.5
61	60.0	-70.9	-74.5	6.529	.000000	.999971	.000003	4.528797	60.5
62	61.0	-68.6	-69.4	6.658	.000001	.999972	.000006	4.528803	61.5
63	62.0	-67.9	-67.1	6.788	.000001	.999973	.000009	4.528812	62.5
64	63.0	-68.5	-66.2	6.938	.000001	.999975	.000010	4.528822	63.5
65	64.0	-70.3	-66.4	7.089	.000001	.999976	.000009	4.528831	64.5
66	65.0	-73.3	-67.5	7.265	.000001	.999977	.000006	4.528837	65.5
67	66.0	-76.3	-69.6	7.442	.000001	.999977	.000004	4.528841	66.5
68	67.0	-75.0	-73.2	7.651	.000000	.999978	.000002	4.528843	67.5
69	68.0	-71.7	-79.9	7.860	.000000	.999978	.000002	4.528846	68.5
68.2 deg									
70	69.0	-69.3	-89.3	8.111	.000000	.999979	.000004	4.528850	69.5
71	70.0	-67.7	-77.9	8.363	.000001	.999979	.000006	4.528856	70.5
72	71.0	-66.8	-73.4	8.671	.000001	.999980	.000009	4.528865	71.5
73	72.0	-66.3	-71.0	8.979	.000001	.999982	.000011	4.528876	72.5
74	73.0	-66.2	-69.7	9.363	.000001	.999983	.000013	4.528890	73.5
75	74.0	-66.3	-69.0	9.748	.000001	.999985	.000014	4.528904	74.5

II. Program SREFLNT1.BAS
(Only partial printout shown)

PROGRAM ID: SREFLNT1.BAS
 WRITTEN BY TOM OTOSHI LAST REV: MARCH 14, 1994

DATE OF THIS RUN: 3-22-93

CASEID THIS RUN: 29.8_DBI HORN NO CELOTEX 14SEP93 MEAS AT F1

FINAL REDUCED DATA

DBEFF = DELTA BEAM EFFICIENCY AT ANGLE THETA
 TBI = INTERPOLATED BRIGHTNESS TEMP
 DTA = DELTA ANTENNA TEMPERATURE
 CUMTA = CUMULATIVE ANTENNA TEMP

THETA (DEG)	DBEFF	TBI (K)	DTA (K)	CUMTA (K)
72.8205	0.000400	226.922	0.0908	0.0908
73.2500	0.000400	224.986	0.0900	0.1808
73.7500	0.000300	222.650	0.0688	0.2476
74.2500	0.000200	220.080	0.0440	0.2916
74.7500	0.000100	217.277	0.0217	0.3133
75.2500	0.000200	214.475	0.0429	0.3562
75.7500	0.000000	211.672	0.0000	0.3562
76.2500	0.000100	208.588	0.0209	0.3771
76.7500	0.000100	205.223	0.0205	0.3976
77.2500	0.000000	201.857	0.0000	0.3976
77.7500	0.000100	198.492	0.0199	0.4175
78.2500	0.000000	194.792	0.0000	0.4175
78.7500	0.000000	190.756	0.0000	0.4175
79.2500	0.000100	186.719	0.0187	0.4361
79.7500	0.000000	182.683	0.0000	0.4361
80.2500	0.000000	178.256	0.0000	0.4361
80.7500	0.000000	173.439	0.0000	0.4361
81.2500	0.000000	168.621	0.0000	0.4361
81.7500	0.000100	163.804	0.0164	0.4525
82.2500	0.000000	158.564	0.0000	0.4525
82.7500	0.000000	152.903	0.0000	0.4525
83.2500	0.000000	147.241	0.0000	0.4525
83.7500	0.000000	141.580	0.0000	0.4525
84.2500	0.000000	135.637	0.0000	0.4525
84.7500	0.000000	129.412	0.0000	0.4525
85.2500	0.000100	123.096	0.0123	0.4648
85.7500	0.000000	116.688	0.0000	0.4648
86.2500	0.000000	110.398	0.0000	0.4648
86.7500	0.000000	104.225	0.0000	0.4648
87.2500	0.000000	98.671	0.0000	0.4648
87.7500	0.000000	93.735	0.0000	0.4648
88.2500	0.000000	90.554	0.0000	0.4648
88.7500	0.000000	89.128	0.0000	0.4648

89.2500	0.000000	91.331	0.0000	0.4648
89.7500	0.000000	97.162	0.0000	0.4648
90.2500	0.000000	91.778	0.0000	0.4648
90.7500	0.000000	75.179	0.0000	0.4648
91.2500	0.000000	61.955	0.0000	0.4648
91.7500	0.000000	52.109	0.0000	0.4648
92.2500	0.000000	44.342	0.0000	0.4648
92.7500	0.000000	38.655	0.0039	0.4687
93.2500	0.000000	34.057	0.0000	0.4687
93.7500	0.000000	30.547	0.0000	0.4687
94.2500	0.000000	27.626	0.0000	0.4687
94.7500	0.000000	25.294	0.0000	0.4687
95.2500	0.000000	23.305	0.0000	0.4687
95.7500	0.000000	21.659	0.0000	0.4687
96.2500	0.000000	20.298	0.0000	0.4687
96.7500	0.000000	19.222	0.0000	0.4687
97.2500	0.000000	18.145	0.0000	0.4687
97.7500	0.000000	17.069	0.0000	0.4687
98.2500	0.000000	16.197	0.0000	0.4687
98.7500	0.000000	15.529	0.0000	0.4687
99.2500	0.000000	14.860	0.0000	0.4687
99.7500	0.000000	14.192	0.0000	0.4687

TOTAL COMPUTED ANTENNA TEMP = 0.469 K

SUM OF DELTA BEAM EFF FOR GROUND REGION ONLY = 0.00220

SUM OF DELTA BEAM EFF FOR GROUND + SKY REGION = 0.00230

SUBREFL EFF RELATIVE TO HORN INPUT = 0.970600

SUBREFL EFF \times ANT TEMP CONTRIBUTION = 0.455 K

110874

4033
33276

p. 20

358436

Low-Complexity Wavelet Filter Design for Image Compression

E. Majani

Imaging and Spectrometry Systems Section

Image compression algorithms based on the wavelet transform are an increasingly attractive and flexible alternative to other algorithms based on block orthogonal transforms. While the design of orthogonal wavelet filters has been studied in significant depth, the design of nonorthogonal wavelet filters, such as linear-phase (LP) filters, has not yet reached that point. Of particular interest are wavelet transforms with low complexity at the encoder. In this article, we present known and new parametrizations of two families of LP perfect reconstruction (PR) filters. The first family is that of all PR LP filters with finite impulse response (FIR), with equal complexity at the encoder and decoder. The second family is one of LP PR filters, which are FIR at the encoder and infinite impulse response (IIR) at the decoder, i.e., with controllable encoder complexity. These parametrizations are used to optimize the subband/wavelet transform coding gain, as defined for nonorthogonal wavelet transforms. Optimal LP wavelet filters are given for low levels of encoder complexity, as well as their corresponding integer approximations, to allow for applications limited to using integer arithmetic. These optimal LP filters yield larger coding gains than orthogonal filters with an equivalent complexity. The parametrizations described in this article can be used for the optimization of any other appropriate objective function.

I. Introduction

Tree-structured subband coding is an increasingly attractive and flexible alternative to other subband coding techniques based on block orthogonal transforms, which exhibit annoying blocky artifacts at low bit rates. The main building block of tree-structured subband coders is the two-channel subband coder. Given a number of levels of decomposition, L , the corresponding subband transform is a function of the analysis lowpass and highpass filters, $H_0(z)$ and $H_1(z)$, and the synthesis lowpass and highpass filters, $G_0(z)$ and $G_1(z)$ (see Fig. 1). Most filter banks of interest to subband coding are known as perfect reconstruction (PR) filter banks, and most filter banks with near-perfect reconstruction are approximations or truncations of PR filter banks.

Of particular interest to onboard image compression is the issue of computational complexity of the subband/wavelet transform; onboard resources are often limited, and so time and space complexity constraints are common. Also of particular interest are good subband transforms that are less complex than their inverse transforms, i.e., impose less complexity at the encoder, at the cost of more at the decoder. Since this feature cannot be obtained with the use of orthogonal transforms (which always have

equal complexity at the encoder and at the decoder), most of our design efforts will be turned toward nonorthogonal filter banks, more precisely linear-phase (LP) filter banks.

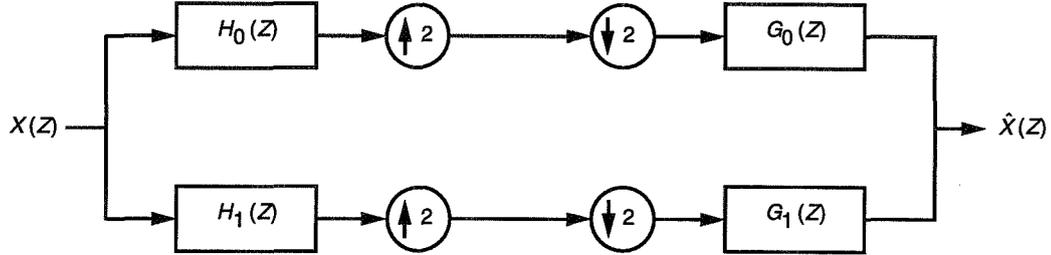


Fig. 1. Analysis and synthesis sections for a maximally decimated two-band filter bank.

In this article, we report on tree-structured subband coder (or wavelet filter) designs that satisfy complexity constraints at the encoder. First, we review the theory of PR filter banks and describe two families of solutions that are appropriate for our purposes. Then, we review definitions and known upper bounds on the subband coding gain of arbitrary linear transforms (not necessarily orthogonal). Finally, we derive solutions that maximize the subband coding gain, an incomplete, but informative measure of the performance of a subband transform-based coder, and give good integer approximations of these solutions. We conclude with the design of optimal boundary filters.

II. Two-Channel PR Filter Banks: A Review

Consider the two-channel filter bank presented in Fig. 1 (see [1] as a general reference).

To derive the equations satisfied by such a filter bank, an expression is needed for the z -transform of the reconstructed signal, $\hat{X}(z)$, in terms of the original signal, $X(z)$, and the analysis and synthesis filters, $H_k(z)$ and $G_k(z)$, respectively. The expressions for the $X_k(z)$ are given by

$$X_k(z) = H_k(z)X(z), \quad k = 0, 1 \quad (1)$$

After decimation, the z -transforms are given by

$$V_k(z) = \frac{1}{2} [X_k(z^{1/2}) + X_k(-z^{1/2})], \quad k = 0, 1 \quad (2)$$

The z -transform of the signals $y_k(n)$ are given by

$$\begin{aligned} Y_k(z) &= V_k(z^2), \quad k = 0, 1 \\ &= \frac{1}{2} [X_k(z) + X_k(-z)] \\ &= \frac{1}{2} [H_k(z)X(z) + H_k(-z)X(-z)] \end{aligned} \quad (3)$$

Finally, the reconstructed signal becomes

$$\hat{X}(z) = \frac{1}{2} [G_0(z)G_1(z)] \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix} \begin{bmatrix} X(z) \\ X(-z) \end{bmatrix} \quad (4)$$

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2} (H_0(z)G_0(z) + H_1(z)G_1(z)) (X(z)) \\ &\quad + \frac{1}{2} (H_0(-z)G_0(z) + H_1(-z)G_1(z)) (X(-z)) \end{aligned} \quad (5)$$

For a PR filter bank, we must have, by definition,

$$\hat{X}(z) = z^{-l} (X(z))$$

i.e., the reconstructed signal $\hat{X}(z)$ is a delayed copy of input signal $X(z)$, which yields the following two PR conditions:

$$H_0(z)G_0(z) + H_1(z)G_1(z) = 2z^{-l} \quad (6)$$

$$H_0(-z)G_0(z) + H_1(-z)G_1(z) = 0$$

These equations can be rewritten as a linear system of two equations where $G_0(z)$ and $G_1(z)$ are the variables and $H_0(z)$ and $H_1(z)$ are assumed to be known:

$$\begin{bmatrix} H_0(z) & H_1(z) \\ H_0(-z) & H_1(-z) \end{bmatrix} \begin{bmatrix} G_0(z) \\ G_1(z) \end{bmatrix} = \begin{bmatrix} 2z^{-l} \\ 0 \end{bmatrix} \quad (7)$$

Define the transfer function $Q(z)$ as

$$Q(z) = H_0(z)H_1(-z)$$

and define $\Delta(z)$ as

$$\Delta(z) = Q(z) - Q(-z)$$

Note that $\Delta(-z) = -\Delta(z)$. A well-known condition for a unique solution to exist is that the determinant of the matrix on the left side of Eq. (7) be nonzero. This yields the following condition on $Q(z)$:

$$\Delta(z) = Q(z) - Q(-z) \neq 0$$

which yields the following solution for $G_0(z)$ and $G_1(z)$:

$$G_0(z) = 2z^{-l} \frac{H_1(-z)}{\Delta(z)} \quad (8)$$

$$G_1(z) = -2z^{-l} \frac{H_0(-z)}{\Delta(z)}$$

For finite impulse response (FIR) solutions to Eq. (8), one must require that $\Delta(z)$ be a simple delay ($\Delta(z) = 2 \times z^{-l}$), and so the equations now become

$$\begin{aligned} G_0(z) &= H_1(-z) \\ G_1(z) &= -H_0(-z) \end{aligned} \tag{9}$$

If $\Delta(z)$ is not a simple delay, then the PR synthesis filters $G_0(z)$ and $G_1(z)$ are infinite impulse response (IIR), irrespective of the choice of $H_0(z)$ and $H_1(z)$, i.e., whether the analysis filters are FIR or not.

The main two properties sought in PR filter banks for image compression are orthogonality and phase linearity. The only FIR wavelet filter that satisfies both properties simultaneously is the Haar filter. The IIR “sinc” wavelet [see Eq. (13)] also satisfies both properties simultaneously.

We now describe two families of LP PR filter banks among which we will look for those filters with the largest coding gain under some complexity constraints.

III. Linear-Phase PR Filter Banks

There are only two different types of LP (symmetric) filters (whether they are FIR or IIR), depending on whether there are one or two “center” coefficients. Symmetric filters with one center coefficient (odd-length for FIR filters) are called whole-sample symmetric (WSS), while those with two center coefficients (even-length for FIR filters) are called half-sample symmetric (HSS).

Since we are interested in solutions with low complexity at the encoder, we will consider the following solutions to the PR equations:

- (1) Both the analysis and synthesis filters are FIR
- (2) The analysis filters are FIR, while the synthesis filters are IIR (in practice, optimized truncations of IIR solutions are used)

A. FIR/FIR Solutions

Given that Eq. (9) is satisfied, the PR equations [see Eq. (6)] become

$$\Delta(z) = Q(z) - Q(-z) = H_0(z)H_1(-z) - H_0(-z)H_1(z) = 2z^{-l} \tag{10}$$

where the sum of the lengths of $H_0(z)$ and $H_1(z)$ ($|h_0| + |h_1|$) can be shown to always be a multiple of 4 (see [7]), and $l = (|h_0| + |h_1|)/2 - 1$. Given a choice for $H_0(z)$, there are an infinite number of solutions for $H_1(z)$ satisfying Eq. (10). The parametrization of this infinite family varies with the nature of the symmetry of the prototype filter $H_0(z)$: either HSS or WSS filters.

1. WSS Solutions. The following theorem (as well as its proof) is from [7]; it parametrizes all filters $H_1(z)$ that are complementary to a prototype filter $H_0(z)$.

Theorem 1: *If the lengths $|h_0|$ and $|h_1|$ of the two complementary filters $H_0(z)$ and $H_1(z)$ are odd and satisfy*

$$|h_0| = |h_1| + 2$$

and if $H_0(-1) = 0$, then all the highpass analysis filters $H'_1(z)$ complementary to $H_0(z)$ are of the form

$$H'_1(z) = z^{-2m} H_1(z) + E(z^2)H_0(z)$$

where

$$E(z^2) = \sum_{i=1}^m \alpha_i \left(z^{-2(i-1)} + z^{-2(2m-i)} \right)$$

The length of $H'_1(z)$ is clearly $|h'_1| = |h_1| + 4m$. It is sometimes appropriate to impose zeros at π for lowpass filters ($H_0(e^{-j\pi}) = H_0(-1) = 0$) or zeros at dc frequency for highpass filters ($H_1(e^{j\pi}) = H_1(1) = 0$). Since $E(1) = 2 \sum_{i=1}^m \alpha_i$ is nonzero in general, $H_0(-1) = 0$ is the only way to ensure that $H'_1(-1) = H_1(-1) = \sqrt{2}$. Finally, the requirement that $H_1(1) = 0$ (which we will find important later) translates into a constraint on the coefficients α_i :

$$H'_1(1) = H_1(1) + E(1)H_0(1) = H_1(1) + 2H_0(1) \sum_{i=1}^m \alpha_i$$

Therefore, to ensure that $H'_1(z)$ has a zero at dc frequency, we must have

$$\sum_{i=1}^m \alpha_i = -\frac{H_1(1)}{2H_0(1)} \quad (11)$$

If $H_1(1) = 0$, then we must have $\sum_{i=1}^m \alpha_i = 0$.

FIR filter banks will be referred to as $|h_0|/|h_1|$ (remember that $|g_0| = |h_1|$ and $|g_1| = |h_0|$). In Table 1, we give an example of the construction of a 5/7 PR filter pair from a 5/3 PR filter pair ($h_0 = [-1, 3, 4, 3, -1]$ and $h_1 = [-1, 3, -1]$) using Theorem 1. The solution with a zero at dc frequency is found using Eq. (11) and yields

$$\alpha_1 = \frac{-1}{16}$$

and

$$h'_1 = \frac{[1, -3, -19, 42, -19, -3, 1]}{16}$$

Table 1. Generating a 5/7 filter from a 5/3 filter.

Filter	1	z^{-1}	z^{-2}	z^{-3}	z^{-4}	z^{-5}	z^{-6}
$z^{-2}H_1(z)$			-1	3	-1		
$H_0(z)$	-1	3	4	3	-1		
$z^{-2}H_0(z)$			-1	3	4	3	-1
$(1 + z^{-2})H_0(z)$	-1	3	3	6	3	3	-1
$H'_1(z)$	$-\alpha_1$	$3\alpha_1$	$-1 + 3\alpha_1$	$3 + 6\alpha_1$	$-1 + 3\alpha_1$	$3\alpha_1$	$-\alpha_1$

2. HSS Solutions. While the following theorem (the equivalent to Theorem 1, but for HSS solutions) is not given in [7], it can be obtained in the same way as for WSS filters. We now give it without proof.

Theorem 2: *If the lengths $|h_0|$ and $|h_1|$ of the two complementary filters $H_0(z)$ and $H_1(z)$ are even and equal, then all the synthesis filters $H'_1(z)$ complementary to $H_0(z)$ are of the form*

$$H'_1(z) = z^{-2m}H_1(z) + E(z^2)H_0(z)$$

where

$$E(z^2) = \sum_{i=1}^m \alpha_i \left(z^{-2(i-1)} - z^{-2(2m-i+1)} \right)$$

Note that for HSS filters, there is always a zero at π for lowpass filters and at zero frequency for highpass filters, and so the construction implied by Theorem 2 always yields $H'(1) = H_1(-1) = 0$.

In Table 2, we give an example of the construction of a 2/6 PR filter pair from a 2/2 PR filter pair ($h_0 = [1, 1]$ and $h_1 = [1, -1]$) using Theorem 2. If we choose $\alpha_1 = 1/8$, we obtain a filter which has the largest possible number of zeros at dc frequency for a 2/6 PR filter: $h'_1 = [1, 1, -8, 8, -1, -1]$.

Table 2. Generating a 2/6 filter from a 2/2 filter.

Filter	1	z^{-1}	z^{-2}	z^{-3}	z^{-4}	z^{-5}
$H_1(z)$			1	-1		
$H_0(z)$	1	1				
$z^{-4}H_0(z)$					1	1
$\alpha_1(1 - z^{-4})H_0(z)$	α_1	α_1			$-\alpha_1$	$-\alpha_1$
$H'_1(z)$	α_1	α_1	1	-1	$-\alpha_1$	$-\alpha_1$

B. FIR/IIR Solutions

Another interesting set of solutions to the PR equations involves IIR filters. We have seen previously that if $H_0(z)$ is given, there exist an infinite number of highpass filters $H_1(z)$ which are complementary to it. However, if $H_1(z)$ is chosen conveniently to satisfy

$$|H_1(e^{jw})| = |H_0(e^{-jw})| \quad (12)$$

then the solution for the synthesis filters is unique and is defined by Eq. (8). Again, HSS and WSS filters must be treated separately.

1. WSS Solutions. For WSS filters, a good choice, from a coding perspective, for $H_1(z)$ that satisfies Eq. (12) is

$$H_1(z) = -z^{-1}H_0(-z)$$

This yields

$$Q(z) = z^{-1}H_0(z)^2$$

and

$$\Delta(z) = z^{-1}(H_0(z)^2 + H_0(-z)^2)$$

Replacing in Eq. (8), we obtain the following IIR solutions for the synthesis filters:

$$G_0(z) = 2z^{-1} \frac{z^{-l}H_0(z)}{H_0(z)^2 + H_0(-z)^2}$$

and

$$G_1(z) = (-1)^l z (G_0(-z))$$

We now consider two examples that will prove useful later.

Example 1: If we choose $H_0(z) = (1 + z^{-1})^2$, then

$$Q(z) = z^{-1}(1 + z^{-1})^4$$

and

$$\Delta(z) = 2z^{-1}(1 + 6z^{-2} + z^{-4})$$

This yields the IIR solution for $G_0(z)$:

$$G_0(z) = \frac{(1 + z^{-1})^2}{1 + 6z^{-2} + z^{-4}}$$

Note that the product filter defined as $P(z) = H_0(z)G_0(z)$ in this last example is the Butterworth filter of order two (see [8] for further insights on this remark), i.e.,

$$P(z) = H_0(z)G_0(z) = \frac{(1 + z^{-1})^4}{1 + 6z^{-2} + z^{-4}}$$

Example 2: If we choose

$$H_0(z) = -1 + 7z^{-2} + 12z^{-3} + 7z^{-4} - z^{-6}$$

then $\Delta(z)$ becomes

$$\Delta(z) = 2z^{-1} (1 - 14z^{-2} + 35z^{-4} + 244z^{-6} + 35z^{-8} - 14z^{-10} + z^{-12})$$

The product filter $Q(z)$ becomes

$$Q(z) = \frac{(-1 + 7z^{-2} + 12z^{-3} + 7z^{-4} - z^{-6})(1 + z^{-1})^2}{1 - 14z^{-2} + 35z^{-4} + 244z^{-6} + 35z^{-8} - 14z^{-10} + z^{-12}}$$

2. HSS Solutions. For HSS filters, the choice we make is

$$H_1(z) = H_0(-z)$$

which yields

$$Q(z) = H_0(z)^2$$

and

$$\Delta(z) = H_0(z)^2 - H_0(-z)^2$$

The IIR solution $G_0(z)$ is, therefore,

$$G_0(z) = (-1)^{l+1} \frac{2z^{-l}H_0(-z)}{H_0(z)^2 - H_0(-z)^2}$$

with

$$G_1(z) = (-1)^{l+1}G_1(z)$$

We now consider an FIR inversion example that will prove useful later.

Example 3: If we choose

$$H_0(z) = -1 + 2z^{-1} + 9z^{-2} + 9z^{-3} + 2z^{-4} - z^{-5}$$

then $\Delta(z)$ becomes

$$\Delta(z) = z^{-1}(-8 + 36z^{-2} + 344z^{-4} + 36z^{-6} - 8z^{-8})$$

and we obtain the following IIR solution for the synthesis lowpass filter:

$$G_0(z) = \frac{-1 + 2z^{-1} + 9z^{-2} + 9z^{-3} + 2z^{-4} - z^{-5}}{-8 + 36z^{-2} + 344z^{-4} + 36z^{-6} - 8z^{-8}}$$

IV. The Subband Coding Gain for PR Filter Banks

There exists a measure of the coding performance of transform-based coding schemes known as the transform coding gain [4]. The coding gain is simply defined as the ratio of the reconstruction error variance of the transform coding scheme and the reconstruction error variance of a pulse code modulation (PCM) scheme (i.e., without a linear transformation to decorrelate the signal). If the coding gain is expressed in decibels (dB), then a coding gain of 0 for a particular transform means that no gain is achieved by applying the transform to the signal. The expression most widely known is one that assumes very fine scalar quantization of the transformed coefficients and that is independent of the quantization level under that assumption. The coding gain then becomes a function of the linear transform only, and the source correlation model. For example, the coding gain of the 8-point discrete cosine transform is 8.83 dB, using a first-order Markov source model with correlation $\rho = 0.95$. Other correlation models can also be used, as well as the actual correlation statistics of the signal to be coded. The coding gain is, therefore, the proper measure of coding performance when high-rate quantization is assumed (i.e., at low compression ratios). At low bit rates, the approximations that led to its expression are no longer valid, and so the coding performance of the given transform is no longer guaranteed. It can be said, however, that quite often transforms that have a large coding gain at high bit rates tend to perform well at lower bit rates, as can be observed from actual rate-distortion curves [2].

In [3], Katto and Yasuda derive an expression for the coding gain that is valid for nonunitary subband transforms (such as the biorthogonal wavelet transform). If M is the number of subbands, $h_k(i)$ ($k = 1, \dots, M$) are the coefficients of the k th analysis filter, $g_k(j)$ ($k = 1, \dots, M$) are the coefficients of the k th synthesis filter, and ρ is the correlation of the source modeled as a one-dimensional Markov source, their expression for the coding gain is

$$G_{SBC}(\rho) = \frac{1}{\prod_{k=1}^M (A_k B_k)^{\alpha_k}}$$

where

$$A_k = \sum_i \sum_j h_k(i) h_k(j) \rho^{|j-i|}$$

$$B_k = \sum_i g_k(i)^2$$

and α_k is the subsampling ratio for the k th filter (i.e., $\alpha_k = 1/M$ for a uniform subband decomposition). The coding gain for the two-dimensional case is simply twice the value in the one-dimensional case.

Note that, for tree-structured subband decompositions, G_{SBC} is a function of the following:

- (1) The subband decomposition, which can be entirely defined by a binary tree
- (2) The analysis lowpass and highpass filters, $H_0(z)$ and $H_1(z)$, and the synthesis lowpass and highpass filters, $G_0(z)$ and $G_1(z)$
- (3) The source correlation model

We will consider two types of subband decomposition: “logarithmic” decompositions, for which only the downsampled output of the analysis lowpass filter may be further decomposed, and “arbitrary” decompositions, for which there is no constraint on which downsampled output may be decomposed next.

A “uniform” decomposition up to level L would correspond to a binary tree of depth L . The L will always be referred to as the maximum depth of the tree. The analysis lowpass and highpass filters, $H_0(z)$ and $H_1(z)$, and the synthesis lowpass and highpass filters, $G_0(z)$ and $G_1(z)$, are required to satisfy the PR conditions (Fig. 1). The model used is a two-dimensional separable Markov model. Even though our results will be limited to a Markov model, the design methodology used can easily be extended to any other source correlation model.

First, we review the known upper bounds on the subband coding gain. Then we examine the limiting case $\rho \rightarrow 1$ and show that a necessary condition for the coding gain to be adequate as $\rho \rightarrow 1$ is for $H_1(z)$ to have a zero at dc frequency. Finally, we describe a useful normalization of the coding gain, which compares coding gain values to the known upper bounds for orthogonal wavelet filters.

A. Known Upper Bounds

Almost all the theoretical results on subband transform coding are about orthogonal subband transforms. We know that the largest coding gain achievable by any linear transformation for a Markov source with correlation ρ is that obtained by the Karhunen–Loeve transform (KLT) (an orthogonal transform) as the block size goes to $+\infty$:

$$G_{KLT}(\rho) = \frac{1}{1 - \rho^2}$$

In [5], de Queiroz and Malvar give an expression for the largest coding gain attainable by any orthogonal subband transform for depth- L logarithmic subband decompositions, which corresponds to the coding gain of the IIR sinc filter:

$$h_k = \frac{\sin(\pi k/2)}{\pi k/2}, \quad -\infty < k < +\infty \quad (13)$$

the frequency response of which is an ideal lowpass filter. They use it to show that orthogonal wavelet transforms are asymptotically suboptimal. We will refer to that expression as $G_{LT}(\rho, L)$, where L is the number of levels of decomposition of the transform.

The same derivation used in [5] can be used to obtain an expression for the largest coding gain attainable by any orthogonal subband transform for depth- L uniform subband decompositions; we note it as $G_{UT}(\rho, L)$. Naturally, in the limit as $L \rightarrow +\infty$, we have $G_{UT}(\rho, L) \rightarrow G_{KLT}(\rho)$. Therefore, the subband coding gain of any orthogonal subband transform of maximum depth L is upper bounded by $G_{LT}(\rho, L)$ for logarithmic decompositions and by $G_{UT}(\rho, L)$ for uniform decompositions.

B. Limiting Case: $\rho \rightarrow 1$

For values of ρ close to 1, only the expression for A_k changes when evaluating the coding gain $G_{SBC}(\rho)$. By writing $\rho = 1 - \epsilon$, and assuming $\epsilon \approx 0$, we obtain for A_k

$$A_k = \sum_i \sum_j h_k(i) h_k(j) (1 - \epsilon)^{|j-i|}$$

$$A_k \approx \sum_i \sum_j h_k(i) h_k(j) (1 - |j-i|\epsilon)$$

$$A_k \approx \left(\sum_i h_k(i) \right)^2 - \left(\sum_i \sum_j h_k(i) h_k(j) |j - i| \right) \epsilon$$

If $\sum_i h_k(i) = 0$, the A_k simplifies to

$$A_k \approx - \left(\sum_i \sum_j h_k(i) h_k(j) |j - i| \right) \epsilon$$

Otherwise, it simplifies to

$$A_k \approx \left(\sum_i h_k(i) \right)^2$$

Therefore, in the limit as $\epsilon \rightarrow 0$, the coding gain is significantly larger by imposing $\sum_i h_k(i) = 0$, i.e., $H_1(1) = 0$, corresponding to a zero at π .

C. Coding Gain Notations

In the following, only dB values of the normalized coding gain (NCG) will be given. The NCG is obtained by computing the dB value of the subband coding gain, and subtracting the dB value of $G_{LT}(\rho, L)$, so that the wide range of coding values can be displayed on the same graph. All NCG values will be given for values of ρ between 1/2 and 1. As mentioned earlier, logarithmic and arbitrary subband decompositions will be considered. For logarithmic decompositions, values of the NCG will be given for a maximum depth L equal to 8, while for arbitrary decompositions, we chose $L = 6$. This means that the NCG values will slightly vary for logarithmic and arbitrary decompositions, although they remain comparable, due to the limited changes in subband coding gain for $L = 6$ and $L = 8$.

In all NCG curves (Fig. 2), the highest solid curve (which we refer to as the reference UT curve) corresponds to the upper bound provided by the $G_{UT}(\rho, L)$, itself normalized by subtracting the dB value of $G_{LT}(\rho, L)$. The other solid curve (flat and equal to 0) is simply the reference curve for logarithmic decompositions (which we refer to as the reference LT curve), i.e., the dB value of $G_{LT}(\rho, L)$. The gap between the two solid curves is simply the coding gain gap mentioned in [5] between the optimal uniform and logarithmic subband decompositions for orthogonal filters.

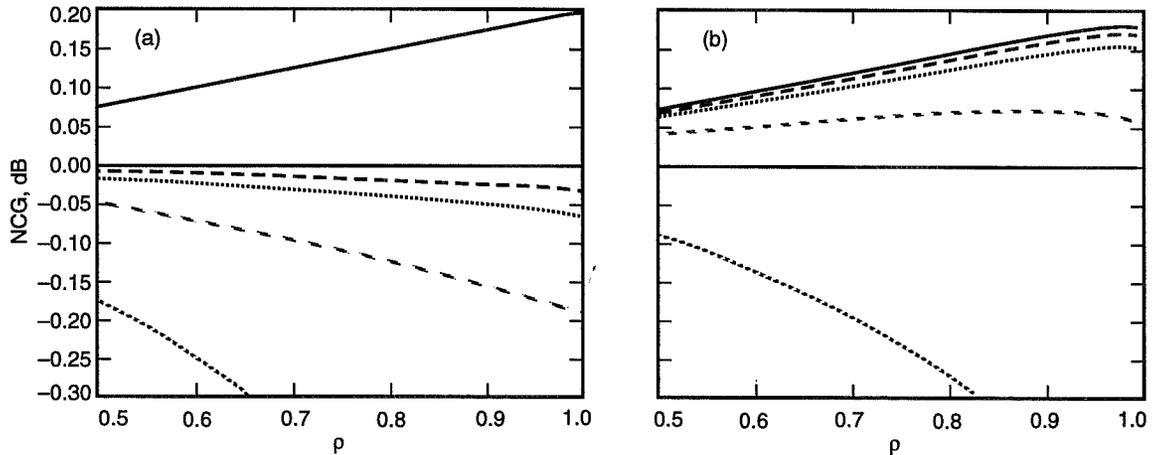


Fig. 2. Normalized coding gain of optimized n -tap orthogonal filters ($n = 4, 8, 14, 20$): (a) 8-level logarithmic decomposition orthogonal filter and (b) 6-level uniform decomposition orthogonal filter.

V. Optimal PR Filters

Given the expression for the subband coding gain, and a parametrization of HSS and WSS filters, we now look for filters with optimal coding gain performance as a function of average filter length. Then, we will look for integer approximations of these filters which yield similar coding gain performance. We now briefly review known results on orthogonal filters which maximize the subband coding gain [9].

A. Orthogonal Filters

The parametrization used here is the one described in [6], a parametrization of even n -tap orthogonal filters with one zero at π using $n/2 - 1$ free parameters. The optimization was performed for logarithmic and uniform decompositions for five levels of decomposition.

The NCG curves are provided in Fig. 2 for the usual values of $n = 4, 8, 14,$ and 20 , and were computed for 8 and 6 levels of decomposition, even though the optimization was done for $L = 5$. The corresponding coding gains for $L = 5$ and $\rho = 0.95$ can be found in Table 3. The filters obtained were very similar for both types of decomposition, and similar to those found in [9]. The filters obtained are local maxima of the coding gain, and are believed to be very close to the global maxima. An interesting characteristic of these filters is that they have a maximum number of roots on the unit circle.

Table 3. Coding gain performance of optimal FIR/FIR filters, logarithmic decompositions, $L = 5, \rho = 0.95$, for orthogonal filters.

Filter length	Coding gain, dB
2	8.24
4	9.29
6	9.62
8	9.74
14	9.85
∞	9.91

B. LP Biorthogonal Filters

Optimization of the subband coding gain is possible given an efficient parametrization of the various families of filters. The parametrization of FIR/FIR solutions is more complex due to the many combinations of filter lengths for a given average filter length. The parametrization of FIR/IIR solutions is simpler, since the parametrization of $H_0(z)$ is all that is required.

1. FIR/FIR Solutions. Given an even average filter length n , the possible filter length combinations $|h_0|/|h_1|$ are of the type $|h_0| = k$ and $|h_1| = n - k$, with $k = 1, \dots, n/2$. This means that the filters can be either HSS or WSS, resulting in different parametrizations.

For WSS filters, Theorem 1 describes how to generate all the filters $H_1'(z)$ of length $|h_0| - 2 + 4m$ complementary to a filter $H_0(z)$. In [7], it is shown that there exists a unique filter $H_1(z)$ complementary to $H_0(z)$ with $|h_1| = |h_0| - 2$. A useful consequence of that result is that there exists a unique filter $H_1(z)$ complementary to $H_0(z)$ with $|h_1| = |h_0| + 2$ with the added constraint that $H_1(1) = 0$. These unique solutions can easily be found by solving a system of linear equations corresponding to the PR conditions. Parametrizing WSS filters is, therefore, done in two steps: (1) parametrizing a filter $H_0(z)$ and calculating

its unique complementary filter $H_1(z)$ (with $|h_1| = |h_0| \pm 2$ depending on whether $H_1(1) = 0$ is required or not) and (2) parametrizing all the filters $H_1^i(z)$ of higher order and complementary to $H_0(z)$. Note that if the desired lengths $|h_0|$ and $|h_1|$ satisfy $|h_0| < |h_1|$, one simply can replace the analysis filters by the synthesis filters and note that $|g_1| = |h_0| < |h_1| = |g_0|$, allowing the suggested parametrization.

For HSS filters, Theorem 2 describes how to generate all the filters $H_1^i(z)$ of length $|h_0| + 4m$ complementary to a filter $H_0(z)$. The parametrization again takes place in two steps: (1) parametrizing a filter $H_0(z)$ and calculating its unique complementary filter $H_1(z)$ (with $|h_1| = |h_0|$) and (2) parametrizing all the filters $H_1^i(z)$ of higher order and complementary to $H_0(z)$. For the case of $|h_0| < |h_1|$, the same remark made above concerning WSS filters applies here.

All optimizations were carried out for average filter lengths of $n = 4, 6$, and 8 , with $\rho = 0.95$ and $L = 5$ levels of logarithmic decomposition. For $n = 14$, only the 17/11 filter length combination was examined. $H_1(1) = 0$ is imposed on all solutions. The optimal filter coefficients are given in Table 4 and the corresponding coding gain values in Table 5. In Table 5, note the higher coding gains obtained by LP filters, even higher than the upper bound for orthogonal filters (Table 3)!

Table 4. Optimal linear-phase FIR/FIR filter banks.

i	$(h_0)_{\pm i}$	$(h_1)_{\pm i}$	i	$(h_0)_i, (h_0)_{-1-i}$	$(h_1)_i, (h_1)_{-1-i}$
5/3 Filter			2/6 Filter		
0	1.02707904	0.70710678	1	0.70710678	0.70710678
1	0.38713452	-0.35355339	2		0.09733489
2	-0.19356726		3		-0.09733489
5/7 Filter			6/10 Filter		
0	0.95902785	0.75833803	1	0.79363797	0.60082030
1	0.36569130	-0.36322679	2	0.08023230	-0.14072753
2	-0.13809844	-0.02561563	3	-0.16676349	-0.07121045
3		0.00967340	4		-0.01194390
			5		0.02482550
9/7 Filter					
0	0.81096744	0.79365640			
1	0.39424588	-0.43412065			
2	-0.11475353	-0.04327481			
3	-0.02568087	0.08056725			
4	0.04781158				
17/11 Filter					
0	0.83851308	0.70235757			
1	0.45656233	-0.41589851			
2	-0.09573748	-0.02337038			
3	-0.11802962	0.09492166			
4	0.06386749	0.02574498			
5	0.01728699	-0.03257654			
6	-0.03776016				
7	-0.00625838				
8	0.00791907				

Table 5. Coding gain performance of optimal FIR/FIR filters, logarithmic decompositions, $L = 5$, $\rho = 0.95$, for LP filters.

Mean filter length	Coding gain	$ h_0 / h_1 $
2	8.24	2/2
4	9.60	5/3
6	9.71	5/7
8	9.88	9/7
14	9.96	17/11
∞	?	?

2. FIR/IIR Solutions. For FIR/IIR solutions to the PR equations, the parametrization is limited to that of the prototype filter $H_0(z)$. The corresponding inverse IIR filters are unique and can be approximated by long enough truncations of the IIR filters for coding gain calculations.

All optimizations were carried out for $L = 5$ levels of logarithmic decompositions and $\rho = 0.95$. $H_1(1) = 0$ (equivalently $H_0(-1) = 0$) is imposed on all solutions. Filter lengths considered for $H_0(z)$ (and therefore $H_1(z)$) were $n = 3, \dots, 11$. The optimal coding gains obtained are given in Table 6 as a function of filter length. The only filter lengths of interest from a coding gain perspective are clearly $n = 3, 6$, and 7 ; larger coding gains are attainable with FIR/FIR solutions for $n \geq 8$. The corresponding optimal analysis filter coefficients for $n = 3, 6$, and 7 are given in Table 7. Their IIR inverses can be computed using Eq. (8).

Table 6. Optimal coding gain as a function of filter length for FIR/IIR combinations.

Filter length	Coding gain, dB for $\rho = 0.95$
3	9.36
4	9.33
5	9.36
6	9.70
7	9.76
8	9.74
9	9.79
10	9.84
11	9.84

3. Integer Approximations of Optimal LP Filters. Because of the complexity constraints imposed by onboard processing, it is often useful to look for integer approximations of good filters. Rather than conduct an exhaustive search of all PR solutions with integer coefficients, we narrowed our search to a neighborhood of the optimal filters arrived at earlier. We also restricted the coding gain of the integer approximation to be within 0.05 dB of the coding gain of the filter it is approximating, a constraint that always leads to solutions of minimum integer ranges.

Good integer approximations to optimal FIR/FIR filters are given in Table 8. Again, $H_1(1) = 0$ is required of all integer solutions. A few others exist and are in the immediate neighborhood of those given.

Table 7. Optimal analysis FIR filters with inverse IIR filters.

Length 3	
i	$(h_0)_{\pm i}$
0	0.70710678
1	0.35355339
Length 6	
i	$(h_0)_i, (h_0)_{-1-i}$
1	0.64546192
2	0.13959559
3	-0.07795073
Length 7	
i	$(h_0)_{\pm i}$
0	0.75048476
1	0.42976974
2	-0.02168899
3	-0.07621635

Table 8. Integer approximations of optimal analysis FIR/FIR filters; the coding gain (CG) is given for $\rho = 0.95$ and $L = 5$.

$ h_0 / h_1 $	h_0	h_1	CG
2/6	[1, 1]	[1, 1, -8, 8, -1, -1]	9.59
5/3	[-1, 2, 6, 2, -1]	[-1, 2, -1]	9.59
6/6	[-1, -2, 32, 32, -2, -1]	[3, 6, -32, 32, -6, -3]	9.68
5/7	[-1, 3, 8, 3, -1]	[1, -3, -31, 66, -31, -3, 1]	9.70
9/7	[2, -1, -6, 19, 44, 19, -6, -1, 2]	[2, -1, -12, 22, -12, -1, 2]	9.86
6/10	[-2, 1, 10, 10, 1, -2]	[-2, 1, 6, 12, -57, 57, -12, -6, -1, 2]	9.87

Good integer approximations to the analysis FIR filters of FIR/IIR solutions are given in Table 9. Again, $H_1(1) = 0$ is required on all integer solutions. The integer range is, not surprisingly, lower than that of the integer FIR/FIR solutions, since the complexity (here the integer range) is minimized at the encoder, at the cost of more complexity at the decoder. The truncated IIR inverses of these three filters can be found in Table 10, while their closed-form expressions were derived earlier in Examples 1-3 (Section III.B). Actual implementation of these inverse filters should involve truncations of the IIR inverses, so as to minimize the average or maximum reconstruction error, for example.

Table 9. Integer approximations of optimal analysis FIR filters with inverse IIR filters; the CG is given for $\rho = 0.95$ and $L = 5$.

Filter length	Filter coefficients	Coding gain
3	[1, 2, 1]	9.36
6	[-1, 2, 9, 9, 2, -1]	9.69
7	[-1, 0, 7, 12, 7, 0, -1]	9.74

Table 10. IIR inverses of FIR filters with integer coefficients: coefficients with 8 decimal points only.

i	$h_0 = [1, 2, 1]$	$h_0 = [-1, 2, 9, 9, 2, -1]$	$h_0 = [-1, 0, 7, 12, 7, 0, -1]$
	$(g_0)_{\pm i}$	$(g_0)_i, (g_0)_{-1-i}$	$(g_0)_{\pm i}$
0	1.00000000	0.74953169	0.88621564
1	0.41421356	0.08132065	0.43634598
2	-0.17157288	-0.16264131	-0.14842319
3	-0.07106781	0.00882291	-0.11356570
4	0.02943725	0.03529163	0.07681856
5	0.01219331	0.00095724	0.04203769
6	-0.00505063	-0.00765795	-0.02465968
7	-0.00209204	0.00010386	-0.01521935
8	0.00086655	0.00166170	0.00908314
9	0.00035894	0.00001127	0.00537790
10	-0.00014868	-0.00036057	-0.00322211
11	-0.00006158	0.00000122	-0.00193014
12	0.00002551	0.00007824	0.00115210
13	0.00001057	0.00000013	0.00068826
14	-0.00000438	-0.00001698	-0.00041153
15	-0.00000181	0.00000001	-0.00024596
16	0.00000075	0.00000368	0.00014697
17	0.00000031	0.00000000	0.00008784
18	-0.00000013	-0.00000080	-0.00005250
19	-0.00000005	0.00000000	-0.00003138
20	0.00000002	0.00000017	0.00001875
21	0.00000001	0.00000000	0.00001121
22		-0.00000004	-0.00000670
23		0.00000000	-0.00000400
24		0.00000001	0.00000239
25			0.00000143
26			-0.00000085
27			-0.00000051
28			0.00000031
29			0.00000018
30			-0.00000011
31			-0.00000007
32			0.00000004
33			0.00000002
34			-0.00000001
35			-0.00000001

VI. Optimal Boundary Filters

The standard signal extension technique used for LP filters is known as the symmetric extension technique. We propose to choose the signal extension that will yield the largest coding gain and to compare the coding gain thus attained with that obtained by both the symmetric and circular extension techniques. Since there is an equivalence between signal extension and modification of the filters at the left and right boundaries of the signal, we will now talk exclusively about boundary filter design.

We will illustrate our method with the optimization of the boundary filters for the 5/3 filter:

$$H_0(z) = [-1, 2, 6, 2, -1]$$

and

$$H_1(z) = [-1, 2, -1]$$

Consider a signal of length 8. The following matrix A ,

$$A = \begin{pmatrix} a & b & c & & & & & \\ -2 & 4 & -2 & & & & & \\ -1 & 2 & 6 & 2 & -1 & & & \\ & & & \ddots & & & & \\ & & & & -2 & 4 & -2 & \\ & & & & d & e & f & g \\ & & & & & & h & i \end{pmatrix}$$

corresponds to the parametrization of the linear transform that is equivalent to the one-level decomposition of a signal of length 8, with the parameters $a, b, c, d, e, f, g, h,$ and i . The inverse matrix B should be of the type

$$B = 32 \times A^{-1} = \begin{pmatrix} ? & ? & & & & & & \\ ? & ? & 2 & -1 & & & & \\ ? & ? & 4 & 2 & & & & \\ ? & ? & 2 & 6 & & -1 & & \\ & & & 2 & \ddots & -2 & & \\ & & & -1 & & ? & ? & ? \\ & & & & & ? & ? & ? \\ & & & & & ? & ? & ? \end{pmatrix}$$

where the question marks indicate unspecified values that depend on the choice of the parameters of matrix A . A simple analysis shows that biorthogonality is preserved if the following equations are satisfied:

$$a + b + c = 8$$

$$b + 2c = 0$$

$$2d + e = 0$$

In Table 11, we give the coding gain that was computed for five levels of logarithmic decomposition, with $\rho = 0.95$, for the three types of signal extension and various signal lengths. The difference between symmetric and optimal extension becomes insignificant at large signal lengths, but is important for smaller ones, and so the use of optimal extensions is relevant, particularly in applications in which an image might be divided into smaller separate blocks for compression.

Table 11. Coding gain performance of finite length signals with 5/3 filter.

Signal length	Circular	Symmetric	Optimal
64	9.16	9.28	9.39
128	9.37	9.44	9.49
256	9.48	9.51	9.54
512	9.55	9.55	9.56
∞	9.59	9.59	9.59

VII. Conclusion

Two families of LP filter banks were examined for efficient implementation as part of an onboard image compression system. The first family (FIR/FIR solutions to the PR equations) exhibits the same complexity at the encoder and the decoder. The shortest filters of that family with the best coding gain performance were designed using a parametrization scheme specific to that family. Integer approximations for that family of filters with similar coding gain performance were also designed. The coding gain performance of this family of LP filters is better than that of orthogonal filters for logarithmic decompositions; for long enough filters, their coding gain is above the upper bound over all possible orthogonal filters. The second family (FIR/IIR solutions) exhibits an asymmetry between the computational complexity at the encoder and the decoder, allowing for less complexity at the encoder, at the cost of more complexity at the decoder. Filters that maximize the subband coding gain were designed, as well as integer approximations with similar coding gain performance. The range of the coefficients of the integer approximations was found to be less than that of counterparts from the first family of LP filters with equivalent performance, validating the approach consisting of designing filters with asymmetric characteristics at the encoder and at the decoder.

These summarized results point to the following conclusions when comparing orthogonal and biorthogonal wavelet filters:

- (1) For the same average filter length, biorthogonal wavelet filters yield larger coding gains than their orthogonal counterparts, when logarithmic decompositions are considered
- (2) Biorthogonal wavelet filters can be used in asymmetric applications, which require very low complexity at the encoder (or the decoder), unlike orthogonal filters

Further investigations that incorporate other design criteria, particularly at low bit rates such as minimal ringing around edges (usually obtained with short filters) and a visual evaluation of the distortion present in the reconstructed images, are required in wavelet filter design. This study illustrates the improved design flexibility of biorthogonal wavelet filters over orthogonal filters, which is likely to be confirmed when using objective functions other than the coding gain.

References

- [1] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Englewood Cliffs, New Jersey: Prentice Hall, 1993.
- [2] M. Lightstone and E. Majani, "Low Bit-Rate Design Considerations for Wavelet-Based Image Coding," *Proceedings of the SPIE, Visual Communications and Image Processing*, Chicago, Illinois, September 25–28, 1994.
- [3] J. Katto and Y. Yasuda, "Performance Evaluation of Subband Coding and Optimization of its Filter Coefficients," *SPIE Proceedings of Visual Communications and Image Processing*, vol. 1605, pp.95–106, November 1991.
- [4] H. S. Malvar, "Signal Processing with Lapped Transforms," Norwood, Massachusetts: Artech House, 1992.
- [5] R. de Queiroz and H. S. Malvar, "On the Asymptotic Performance of Hierarchical Transforms," *IEEE Transactions on Signal Processing*, vol. 40, no. 10, pp. 2620–2622, October 1992.
- [6] H. Zou and A. H. Tewfik, "Parametrization of Compactly Supported Orthonormal Wavelets," *IEEE Transactions on Signal Processing*, vol. 41, no. 3, pp. 1428–1431, March 1993.
- [7] M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design," *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2207–2232, September 1992.
- [8] C. Herley and M. Vetterli, "Wavelets and Recursive Filter Banks," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2536–2556, August 1993.
- [9] H. Caglar, Y. L. Liu, and A. N. Akansu, "Statistically Optimized PR-QMF Design," *Proceedings of the SPIE VCIP Conference '91*, vol. 1605, pp. 86–94, November 11–13, 1991.

110875

33277
p.9

358476

A Preliminary Optical Visibility Model

K. Cowles

Communications Systems Research Section

B. M. Levine

Optical Sciences and Applications Section

A model is being created to describe the effect of weather on optical communications links between space and ground sites. This article describes the process by which the model is developed and gives preliminary results for two sites. The results indicate nighttime attenuation of optical transmission at five wavelengths. It is representative of a sampling of nights at Table Mountain Observatory from January to June and Mount Lemmon Observatory from May and June. The results are designed to predict attenuation probabilities for optical communications links.

I. Introduction

Space-to-ground visibility statistics are being collected to develop a weather model for optical communications. The Autonomous Visibility Monitoring (AVM) task has deployed two remote observing stations to measure starlight transmission and to calculate attenuation probabilities from which we will create the model. The details of the observing stations have been discussed previously [1]. This article explains the process by which the data are turned into a visibility model for optical communications, and the preliminary data are interpreted. There are two remote observing sites and one test site. The AVM site at the Table Mountain Facility (TMF) has been officially operational since April 1, 1994, while the Mount Lemmon site began operations on May 16, 1994. The test site at the JPL mesa site is not currently operational. Data collected through June 30 are presented here. We will show the theory behind the model, analysis of the initial data, and calibration of those data.

II. Theory

The automated observatories conduct star observations at each site continuously, with pauses for data transfer back to JPL and closures due to weather. The operational station software algorithm directs the system to observe four stars per hour at each of five filters, saving the image and related information in a standard Flexible Image Transport System (FITS) format. Each observation contains information about what is contained in the image as well as conditions at the observatory, such as a weather report and status of the backup power system. The FITS files are compressed and then transferred by modem to JPL for analysis.

The basis of the atmospheric transmission measurement in the AVM system comes from radiation transfer theory, which states that the attenuation of light is logarithmically proportional to the amount of atmosphere it traverses [2]. This relationship is given by the equation

$$I = I_0 10^{-0.4m_e X} \quad (1)$$

where

- I = intensity measured on the ground
- I_0 = (absolute) intensity of the object above the atmosphere
- m_e = atmospheric loss factor
- X = amount of atmosphere traversed

The latter quantity is expressed in units of air mass. One air mass is the amount of atmosphere at zenith, and the amount of atmosphere traversed is roughly proportional to the inverse of the cosine of the zenith pointing angle [3], i.e.,

$$X \sim \frac{1}{\cos(\text{zenith angle})} \quad (2)$$

The atmospheric transmission is given by the ratio of received intensity to that above the atmosphere:

$$\text{transmission} = \frac{I}{I_0} \quad (3)$$

The transmission is related to the attenuation by the relation,

$$\text{attenuation (dB)} = -10 \log_{10}(\text{transmission}) = 4m_e X \quad (4)$$

From Eq. (1), there is a log linear relationship between the observed quantities of received intensity and the air mass. The unknowns are the absolute intensity at each wavelength above the atmosphere and the atmospheric loss. The absolute intensity, I_0 , is determined through calibration, and will be discussed later. If we assume I_0 is a known quantity, the atmospheric loss can be determined from a given observation.

Automatic algorithms locate the star within the frame of the charge-coupled device (CCD) and integrate the measured values of intensity. In practice, this is determined by summing all intensity values in the frame above a background threshold. All observations are normalized for known star magnitude and observation exposure time.

III. Data Analysis and Model Generation

The atmospheric transmission values over all five filters for each night of observation are saved in summary files along with other observational specific information. The transmission values are normalized by their zenith angle to yield a zenith attenuation according to the equation

$$a_{zen} = - \left(\frac{10}{X} \right) \log \text{transmission (dB)} \quad (5)$$

The AVM model is defined as a cumulative density function compiled from the ensemble of all zenith attenuation measurements at each filter for each site.

The AVM model based on current data is shown in Figs. 1 and 2. These curves are similar to the telecommunications interfaces for atmospheric and environmental effects data for X-band (8420 MHz), S-band (2295 MHz), and Ka-band (32 GHz).¹ Error bars are drawn corresponding to ± 2 standard deviations of the computed value for each cumulative probability [4].

Figure 1 shows the cumulative attenuation probability for the TMF at each of the five filters. While the TMF AVM telescope did not officially begin observations until April 1, some observations were taken during testing, beginning on January 29. The curves include nighttime observations only and exclude nights when the telescope was down for maintenance. The number of observations included for each curve is indicated.

Cloudy data and clear sky data are incorporated to indicate the probability of link attenuation less than the amounts given on the x-axis. For example, in Fig. 1(b) we see that for the TMF at 532 nm, attenuation was less than 2 dB for 42 percent of the nights and less than 8 dB for 69 percent of the nights during the months of February to June.

Data for Mount Lemmon at each filter is shown in Fig. 2. Comparison of the Mount Lemmon curves with the TMF curves shows that transmission is better at Mount Lemmon. This is expected, since the elevation of Mount Lemmon is 500 m higher than TMF. In Fig. 2(b), attenuation at Mount Lemmon at 532 nm is seen to be less than 2 dB for 62 percent of the nights and less than 8 dB for 73 percent of the nights during May and June.

Looking at the curves for each site, we see that the 532-nm-filter (filter no. 3) data and the V-filter (filter no. 6) data are similar, as are the 860-nm-filter (filter no. 2) and the I-filter (filter no. 4) data. The V, R (filter no. 5), and I filters are astronomical standards and have a much larger bandwidth. However, the V filter is centered near 532 nm and the I is centered near 860 nm, so we expect a similarity in these observations. The initial slope of the 860-nm and I curves is also much steeper than that of the 532-nm and V curves. This is expected because Earth's atmosphere is less transmissive at 532 nm (green) than at 860 nm (near infrared).

Similar curves will be generated for each quarter of the year. The data presented here are slightly misleading because they represent different periods of time. The Mount Lemmon data also have substantially fewer data points and, hence, a larger margin of error.

IV. Analysis and Calibration of I_0

Absolute intensity at each wavelength above the atmosphere can be solved by observing a set of intensity values taken from observations of stars normalized by their magnitudes and zenith pointing angles over a clear sky, and by performing a least-squares fit. From shot-noise considerations, the variations in the received intensity are proportional to the intensity itself, so a weighted least-squares fit is performed. The weighting is determined by the relationship between the intensity and measurements made on the CCD detector.

After the intensity of each observation is determined, a number of procedures examine the night's data for self-consistency and eliminate any outliers from the subsequent calibration routines. These procedures both perform the weighted least-squares fits and examine the resultant spread of the data from the fits.

¹S. C. Slobin, "TCI-40: Telecommunications Interfaces, Atmospheric and Environmental Effects," *Deep Space Network/Flight Project Interface Design Handbook, Volume I: Existing DSN Capabilities*, DSN Document 810-5, Rev. D (internal document), Jet Propulsion Laboratory, Pasadena, California, 1992.

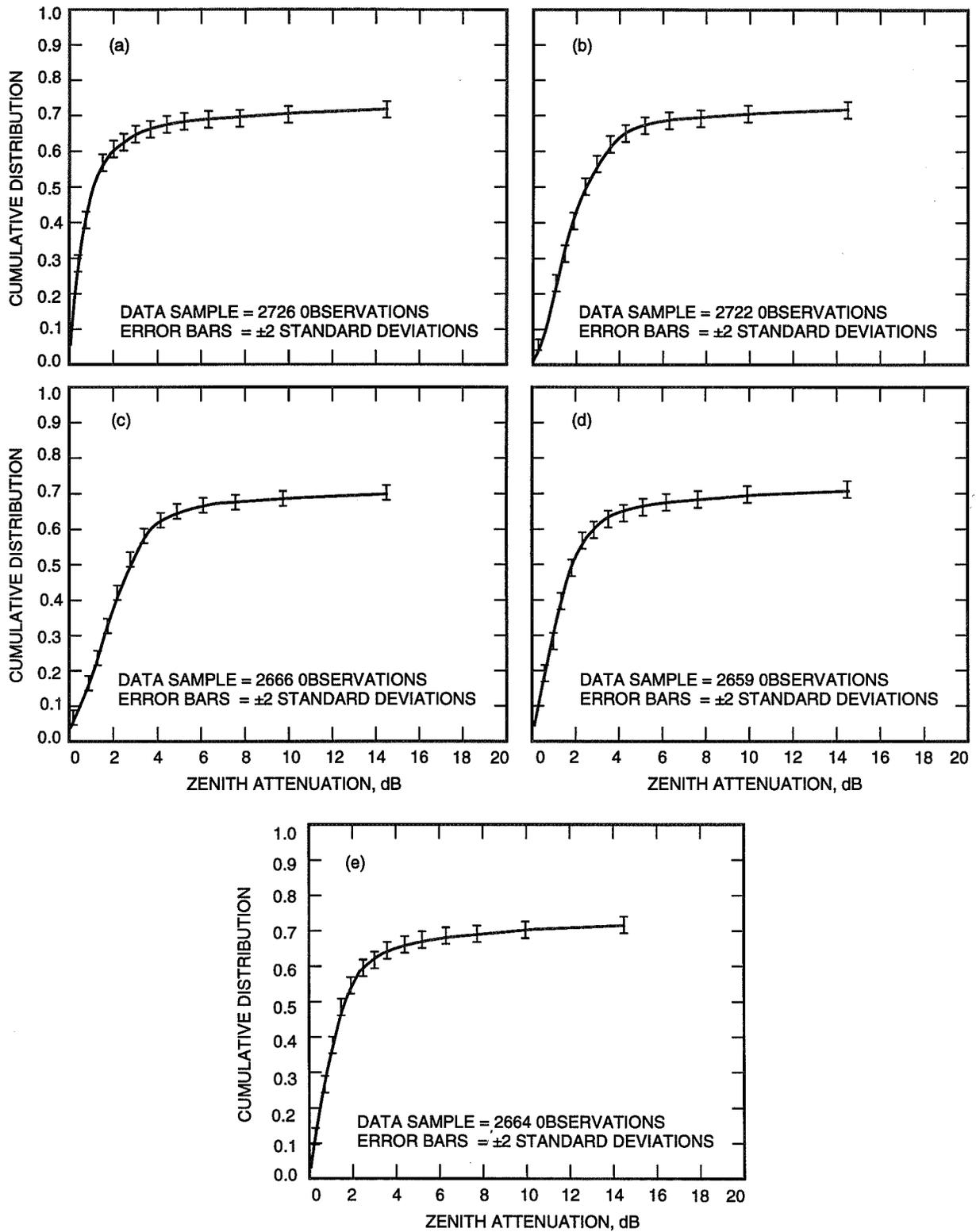


Fig. 1. TMF weather model cumulative attenuation probability from January 29 to June 30, 1994: (a) 860-nm filter; (b) 532-nm filter; (c) V (532-nm) filter; (d) R (680-nm) filter; and (e) I (860-nm) filter.

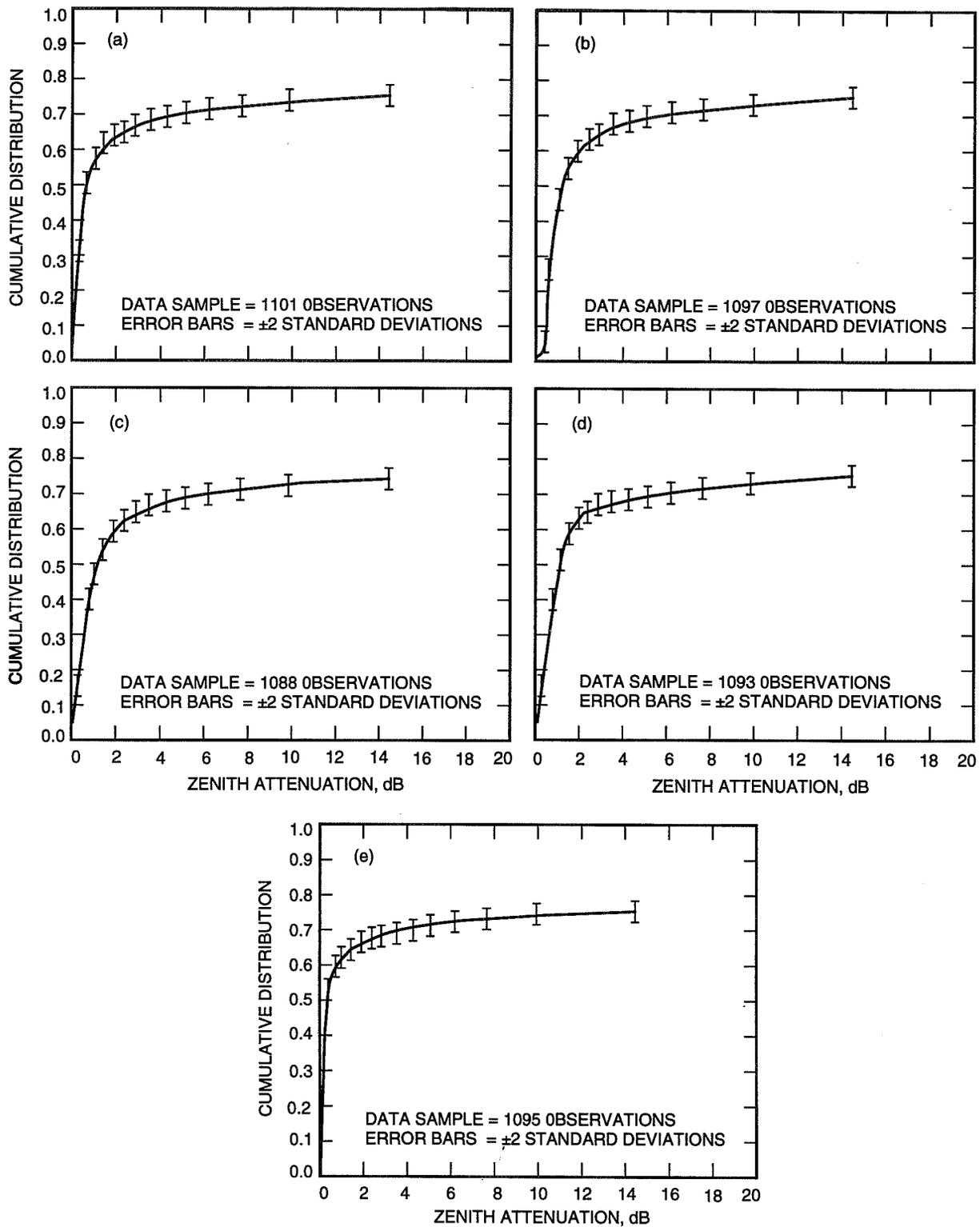


Fig. 2. Mount Lemmon weather model cumulative attenuation probability from May 18 to June 30, 1994: (a) 860-nm filter; (b) 532-nm filter; (c) V (532-nm) filter; (d) R (680-nm) filter; and (e) I (860-nm) filter.

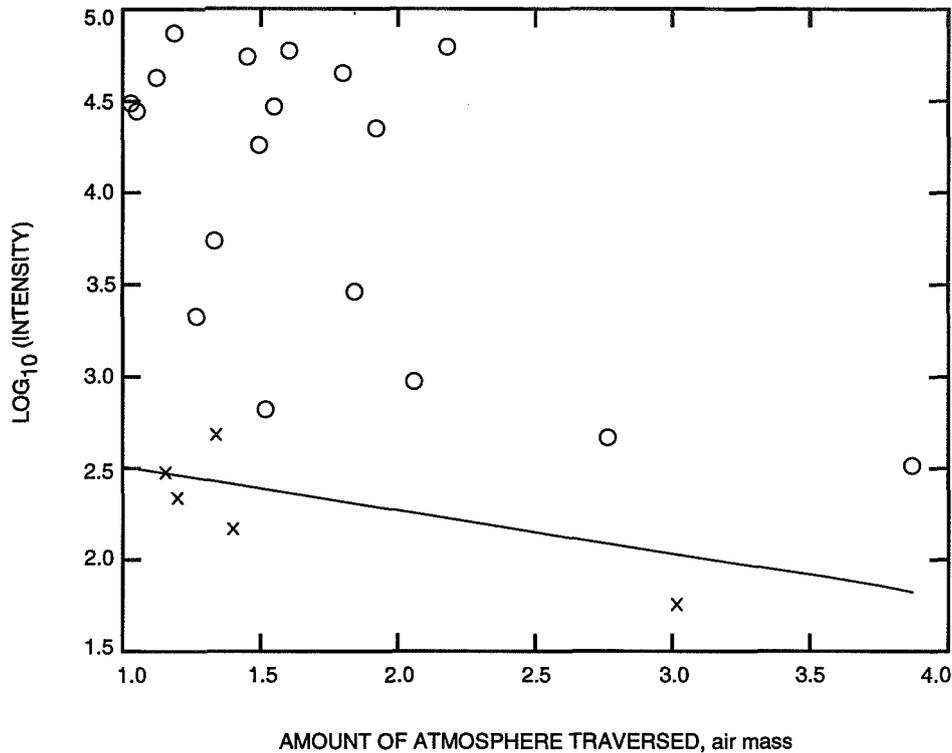


Fig. 3. Least-squares fit with extremely low data points.

The first procedure finds those data with intensities representing observations that do not belong to the calibration data set. In Fig. 3, the line drawn on the graph indicates the least-squares fit with the data points marked "X," which are subsequently eliminated. These excluded data values are extremely low, and the resultant least-squares fits would tend to bypass the majority of the data set. Therefore, only the outlier data would be close to these fit, and, hence, should be eliminated before testing the data set for less gross outliers.

The second procedure is based on the "jackknife" statistical technique [5,6] in which one observation is deleted from the data set, and the weighted least-squares solution is determined. This is repeated for every observation, with the result that data that do not belong to the data set will unduly influence the weighed least-squares results and can be identified. If data are deleted, the procedure is repeated on the reduced data set until no further outliers are detected. One iteration of this process is shown in Fig. 4; the data point marked "X" has been identified as an outlier. The two lines on the graph indicate the least-squares fit with and without the data point at 6 air masses.

After all outliers have been removed, the atmospheric loss is calculated for the night. Figure 5 shows the final fit for the data taken the night of June 7-8 at Mount Lemmon. Extrapolating the line to zero air mass, we determine the best estimate of the absolute intensity for that night at the 532-nm filter.

Transmission is computed by the quotient of the measured intensity of a set of observations divided by the best estimate of the absolute intensity. This estimate is formed from all estimates of absolute intensity that are self-consistent. Consistency is tested by comparing the current measurement of I_0 with an updated value of I_0 using historical plus current data. If the data set under examination produces a consistent value of I_0 , then it is incorporated into the best estimate. This test is performed for every data set taken. The results from one night's observations at Mount Lemmon are shown in Table 1.

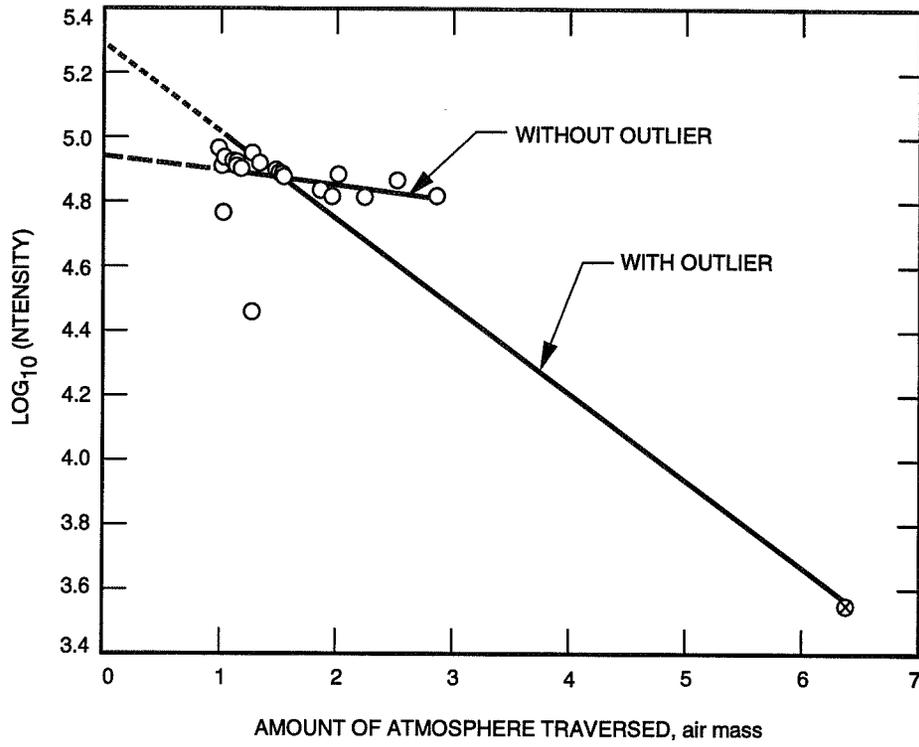


Fig. 4. Outlier removal using the jackknife statistical technique.

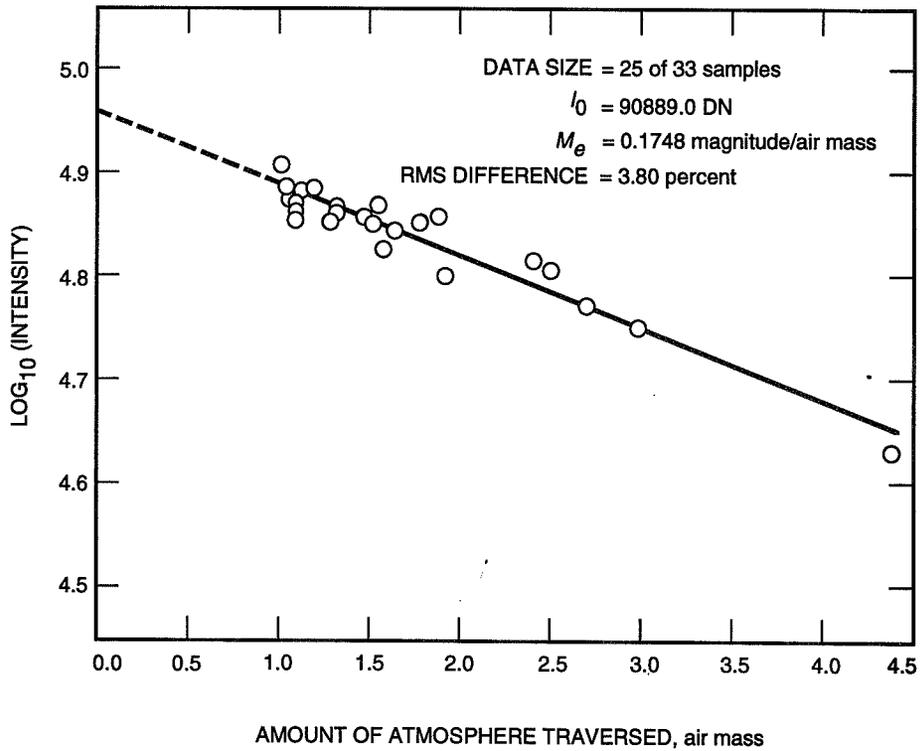


Fig. 5. Best fit of data points to determine I_0 .

For each filter, the calculated values of I_0 and m_e are given along with the relative rms error expressed in percent and the average transmission of the sky. The remainder of the columns show the results of eliminating data from the final calibration set. Starting with the total number of files, the data eliminated from the two procedures mentioned above are given in the #Xs and #tol columns, respectively. The #misc and #Os columns denote those values eliminated due to pointing at zenith angles greater than 85 deg and those data in which an intensity above background could not be detected, respectively.

Table 2 shows the old and updated values of I_0 along with the f-statistic and the degrees of freedom. The f-statistic algorithm is a test for consistency for the new and old I_0 . If the data are consistent with previous measurements, the data values for the night are included in the calibration of I_0 for that filter. Confidence in I_0 will increase as the number of observations included in the calibration increases. The I_0 's are then used to calculate the atmospheric loss for the night.

Table 1. Daily report for Mount Lemmon on the night of June 7-8, 1994.

Filter no.	I_0	m_e	Relative rms error, percent	Average transmission of the sky	#files	#misc	#Os	#Xs	#tol	#left ^a
6	841597	0.1201	8.8639	0.8953	33	1	1	1	5	25
5	649404	0.1142	3.9662	0.9002	33	0	2	3	8	20
4	370191	0.0630	10.4121	0.9437	33	0	1	1	4	27
3	90889	0.1748	3.7965	0.8513	33	0	2	4	2	25
2	25956	0.0522	10.2851	0.9530	33	1	1	1	3	27

^a #left = #files - (#misc. + #Os + #Xs + #tol).

Table 2. IO calibration for Mount Lemmon.

Filter no.	Cumulative I_0 (old)	Cumulative I_0 (new)	f-statistic	Degrees of freedom
6	768396	782163	1.7012	175
5	638065	639062	0.1011	162
4	210833	228700	0.2064	209
3	88646	89070	0.5346	163
2	28148	27672	3.2679	170

V. Performance

The AVM observatory at TMF has been operating officially since April 1, 1994. However, data taken between April 27 and June 22 have been removed from the database because, due to system malfunction, they were not representative of the site. As a result, the data are those for 40 percent of the period between January 29 and April 1, 1994, and for 32 percent of the period between April 1 and June 30, 1994. The system has been repaired and is now functioning routinely.

The AVM site at Mount Lemmon was operational 79 percent of the period between May 18 and June 30, 1994. Downtime outages at both sites occurred because of temporary system malfunction or maintenance periods.

VI. Conclusions

Preliminary data from the AVM sites at TMF and Mount Lemmon have provided data adequate to prepare preliminary cumulative distribution curves for the wavelengths measured. These curves can be used to predict link availability for the sites. A set of curves such as these will be produced each quarter to show variation in site characteristics by season. The data can be analyzed for all or part of a year as required.

Further studies include calculation of joint statistics for the two sites, as well as incorporation of daytime transmission data. The data can also be compared with satellite weather data to define the correlation between the two methods of clear-sky detection. Several scenarios can then be studied for a spacecraft-to-ground optical communications link.

References

- [1] K. Cowles, "Hardware Design for the Autonomous Visibility Monitoring Observatory," *The Telecommunications and Data Acquisition Progress Report 42-114*, vol. April-June 1993, Jet Propulsion Laboratory, Pasadena, California, pp. 295-301, August 15, 1993.
- [2] W. A. Hiltner, ed., *Astronomical Techniques*, vol. 2 of Star and Stellar Systems, Chicago: University of Chicago Press, p. 179, 1960.
- [3] C. W. Allen, *Astrophysical Quantities*, London: Athens Press, Section 55, pp. 124-125, 1976.
- [4] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, 3rd ed., New York: McGraw-Hill Book Company, Chapter 11, Section 2.1, 1974.
- [5] B. Efron, *The Jackknife, the Bootstrap and other Resampling Plans*, Conference Board of Mathematical Science-NSF Regional Conference Series in Applied Mathematics, 1982.
- [6] C. Z. Mooney and R. D. Duval, *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Newbury Park: SAGE Publications, p. 22, 1993.

Field Verification of the Wind Tunnel Coefficients

W. K. Gawronski

Ground Antennas and Facilities Engineering Section

J. A. Mellstrom

Avionic Systems Engineering Section

Accurate information about wind action on antennas is required for reliable prediction of antenna pointing errors in windy weather and for the design of an antenna controller with wind disturbance rejection properties. The wind tunnel data obtained 30 years ago using a scaled antenna model serves as an antenna industry standard, frequently used for the first purpose. The accuracy of the wind tunnel data has often been challenged, since they have not yet been tested in a field environment (full-sized antenna, real wind, actual terrain, etc.). The purpose of this investigation was to obtain selected field measurements and compare them with the available wind tunnel data. For this purpose, wind steady-state torques of the DSS-13 antenna were measured, and dimensionless wind torque coefficients were obtained for a variety of yaw and elevation angles. The results showed that the differences between the wind tunnel torque coefficients and the field torque coefficients were less than 10 percent of their values. The wind-gusting action on the antenna was characterized by the power spectra of the antenna encoder and the antenna torques. The spectra showed that wind gusting primarily affects the antenna principal modes.

I. Introduction

The wind tunnel tests of antenna wind loading were conducted more than 30 years ago,^{1,2,3} and they serve as an antenna industry source for the determination of antenna pointing errors [1]. The wind tunnel results presented in Footnote 3, the CP-6 Memorandum, are considered the most reliable since the antenna model rather than the dish-only model (as in Footnotes 1 and 2) was used in the experiment, and the ground effects were accounted for. The wind tunnel model of the antenna in the CP-6 Memorandum had a reflector diameter of 0.46 m (18 in.), which gives a scale factor of 1 to 75 when compared with the DSS-13 reflector of 34 m. Holes were drilled in the reflector on the outer 25 percent of its radius to obtain a 25-percent porosity in that area. The CP-6 Memorandum presented two sets of data, "metric"

¹ N. L. Fox and B. Layman, Jr., "Preliminary Report on Paraboloidal Reflector Antenna Wind Tunnel Tests," JPL Interoffice Memorandum CP-3 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1962.

² N. L. Fox, "Load Distributions on the Surface of Paraboloidal Reflector Antennas," JPL Interoffice Memorandum CP-4 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1962.

³ R. B. Blaylock, "Aerodynamic Coefficients for Model of a Paraboloidal Reflector Directional Antenna Proposed for a JPL Advanced Antenna System," JPL Interoffice Memorandum CP-6 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1964.

(alidade attached at the elevation bearing and not attached at the azimuth bearing) and “nonmetric” (alidade not attached at the elevation bearing and attached at the azimuth bearing). In this comparison, metric data were used (the difference between the two sets of data was within 10 percent).

The tunnel tests need verification, since they were run on small models of reflectors, as in Footnotes 1 and 2, or small models of antennas, as in Footnote 3, and they were conducted in wind tunnels where conditions are not necessarily the same as in the actual antenna environment. However, the repetition of such wind tunnel tests in the field would be a costly (expensive instrumentation) and time consuming (lack of controlled environment) procedure. Therefore, instead of a complex measurement of wind pressure distribution on the antenna structure, a simplified approach was chosen. It consisted of measuring elevation and azimuth torques generated by wind on the DSS-13 antenna and of comparing them with the torques obtained in the wind tunnel on the antenna model. The torques represented the summary wind loads on the antenna and thus served as the accuracy indicator of the wind tunnel tests.

In addition, wind gusting on the antenna was measured, with the gusting action characterized by the power spectra of the pointing error and the torques. The spectra reflected the frequency contents of the antenna response to the wind.

II. Torque Coefficients

Antenna torques due to wind depend on the antenna size, represented by the dish diameter, D ; wind velocity, v ; porosity of the antenna dish; the distance, d , of the dish vertex from the elevation axis (see Fig. 1); the elevation angle of the antenna, β_e ; and the wind direction with respect to the antenna dish, yaw angle β_a . By rescaling the torques, one can make the torques independent of antenna size and wind speed. These dimensionless and scaled torques, called torque coefficients, c_t , are defined in Footnote 1 as

$$c_t = \frac{T_w}{p_d A D} \quad (1)$$

where T_w is the on-axis torque, N-m (lb-ft); p_d is the dynamic pressure, N/m² (lb/ft²); A and D are the antenna dish frontal area, m²(ft²), and diameter, m (ft), respectively; and $A = \pi D^2/4$.

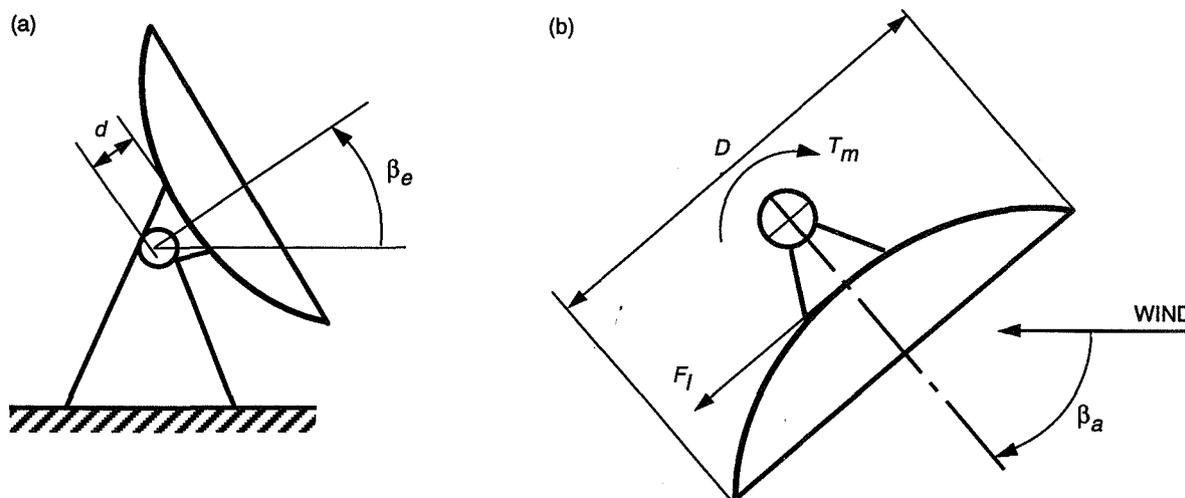


Fig. 1. Antenna configuration with respect to wind: (a) side view and (b) top view.

A. Obtaining Torque Coefficients From Field Data

The dynamic pressure depends on wind speed, v km/hr (mph), as follows:⁴

$$p_d \hat{=} \alpha_p v^2 \quad (2)$$

where $\alpha_p = 0.0478$ (in SI units) or 0.00256 (in English units), and is illustrated in Fig. 2. For the 34-m antenna, $AD = 30,874 \text{ m}^3$ ($1.0903 \times 10^6 \text{ ft}^3$), which when combined with Eq. (2) gives Eq. (1) as follows:

$$c_t = \alpha_t \frac{T_w}{v^2} \quad (3)$$

where $\alpha_t = 6.7761 \times 10^{-4}$ (in SI units) or 3.586×10^{-4} (in English units), and T_w is in N-m (lb-ft) and v in km/hr (mph).

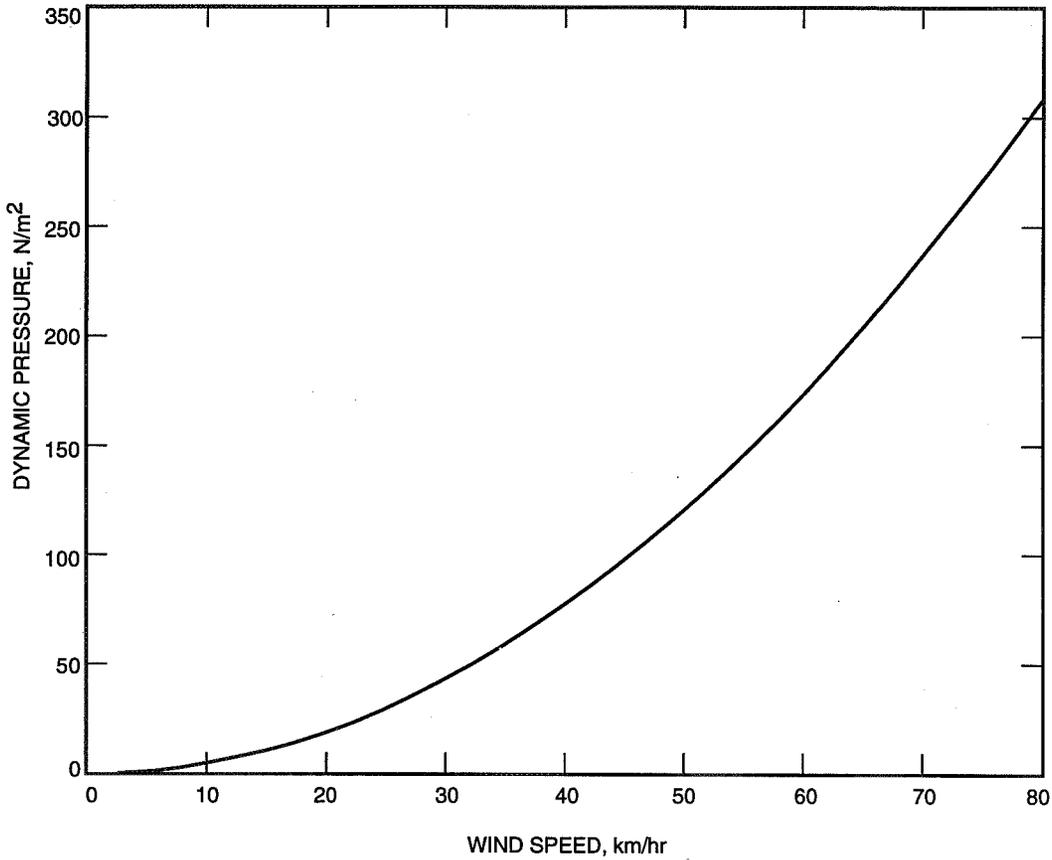


Fig. 2. Wind dynamic pressure on the antenna versus wind speed.

In the field, the torque was measured while slewing the antenna with the constant rate of 0.05 deg/sec . In this case, the measured torque, T_m , consisted of the torque generated by wind, T_w , and the friction torque, T_f ; thus,

$$T_w = T_m - T_f \quad (4)$$

⁴ N. L. Fox and B. Layman, Jr., op. cit.

The friction torque, T_f , was measured when slewing the antenna with the same rate and with the wind speed close to zero. Hence, the torque coefficient from Eq. (3) is now

$$c_t = \alpha_t \frac{T_m - T_f}{v^2} \quad (5)$$

The torque coefficients depend on the porosity of the antenna dish; the distance, d , of the dish vertex from the elevation axis; the elevation angle of the antenna; and the yaw angle. The distance, d , has the same effect on the torque coefficients for the DSS-13 antenna and the CP-6 Memorandum model. To prove it, note that the total azimuth axial torque is a sum of the dish torque, T_d , and the torque, T_l , generated by the lateral force, F_l (see Fig. 1). From this figure, one obtains

$$T_l = F_l d \cos \beta_e \quad (6a)$$

or

$$c_{tl} = c_{fl} \cos \beta_e \frac{d}{D} \quad (6b)$$

where β_e is the elevation angle, c_{tl} is the lateral torque coefficient, and c_{fl} is the lateral force coefficient,

$$c_{tl} = \frac{T_l}{p_d A D}, \quad c_{fl} = \frac{F_l}{p_d A} \quad (7)$$

The ratio $d/D = 0.14$ for the CP-6 Memorandum case and 0.11 for the DSS-13 antenna; thus, they are close enough so that from Eq. (6b) it follows that the lateral force coefficient is about the same in the CP-6 experiment and the DSS-13 antenna experiment.

The porosity of the DSS-13 antenna dish is about 25 percent of the outer 25-percent radius, similar to the CP-6 model.

The torque coefficient was tested with respect to the two remaining “free” parameters: elevation angle, β_e , and yaw angle, β_a . Thus, the torque coefficient was a function of β_e and β_a , i.e., $c_t = c_t(\beta_e, \beta_a)$.

B. Fitting Wind Tunnel Results to the Field Data

In the following, the field and wind tunnel data were compared. The torque coefficients were obtained for the range of yaw angles, from 0 to 180 deg, with the elevation angle fixed; thus, $c_t = c_t(\beta_a, i)$, where i is the sample number. They were also measured for a range of elevation angles, from 10 to 89 deg, with the yaw angle fixed, so that $c_t = c_t(\beta_e, i)$. Denote the torque coefficient from the CP-6 Memorandum $c_{t6}(\beta_a)$ if it depends on yaw angle or $c_{t6}(\beta_e)$ if it depends on elevation angle. The new coefficients, c_{t6new} , are obtained as a linear combination of c_{t6} ,

$$c_{t6new}(\beta) = s_1 c_{t6}(\beta) + s_2 \quad (8)$$

where $\beta = \beta_a$ or $\beta = \beta_e$, and the coefficients s_1 and s_2 are determined such that the error, ε , between the CP-6 Memorandum and the field data,

$$\varepsilon(s_1, s_2) = \sum_{i=1}^n (c_t(\beta, i) - c_{t6new}(\beta))^2 \quad (9)$$

where $\beta = \beta_a$ or $\beta = \beta_e$, is minimal. The parameter s_1 is the scaling coefficient, and s_2 is the shifting coefficient. The parameter s_1 scales the CP-6 Memorandum curve c_{t6} to best fit the field data. The parameter s_2 shifts the field data to compensate for undetermined friction forces.

The windy data were collected on two days: January 24, 1994, and March 22, 1994. The sampling frequency was 40 Hz, and the antenna azimuth rate was 0.05 deg/sec. The wind speed ranged from 25 to 65 km/hr (15 to 40 mph). Rotating the antenna 360 deg in azimuth with a fixed elevation angle (at 10 or 60 deg), the azimuth torques were measured.

On March 22, 1994, the torques were measured for the average wind speed of 60 km/hr (37 mph) and an elevation angle of 10 deg. The wind direction was 251 deg with respect to the azimuth encoder zero position. The yaw angle is the difference between the antenna azimuth position and the wind direction. The plot of the azimuth torques versus yaw angle is shown in Fig. 3(a). The plots were obtained by averaging torques every 5 sec, and by subtracting the friction torque, which was measured for a non-windy day at 17.9 N-m (15.8 lb-ft). This plot is expected to be antisymmetric, and a small departure from the antisymmetry is caused by the varying wind speed during this experiment. The mean wind speed (a linear fit to the wind speed data) is shown in Fig. 3(b). By rescaling the torques according to Eq. (5), the torque coefficient was obtained (Fig. 4, dots). The CP-6 Memorandum torque coefficient was best fitted to this data (Fig. 4, solid line). It required scaling $s_1 = 1.016$ and shifting $s_2 = 0.0066$. That means that the wind tunnel data are less than 2 percent apart from the field data. The friction torque estimation is about 10 percent error, since the shift $s_2 = 0.0066$ corresponds to 1.8 N-m (1.6 lb-ft) friction torque. Two other experiments for an elevation angle of 10 deg required scaling of 1.07 and 1.11; thus, the accuracy of the tunnel data for this case can be considered to be less than 10 percent.

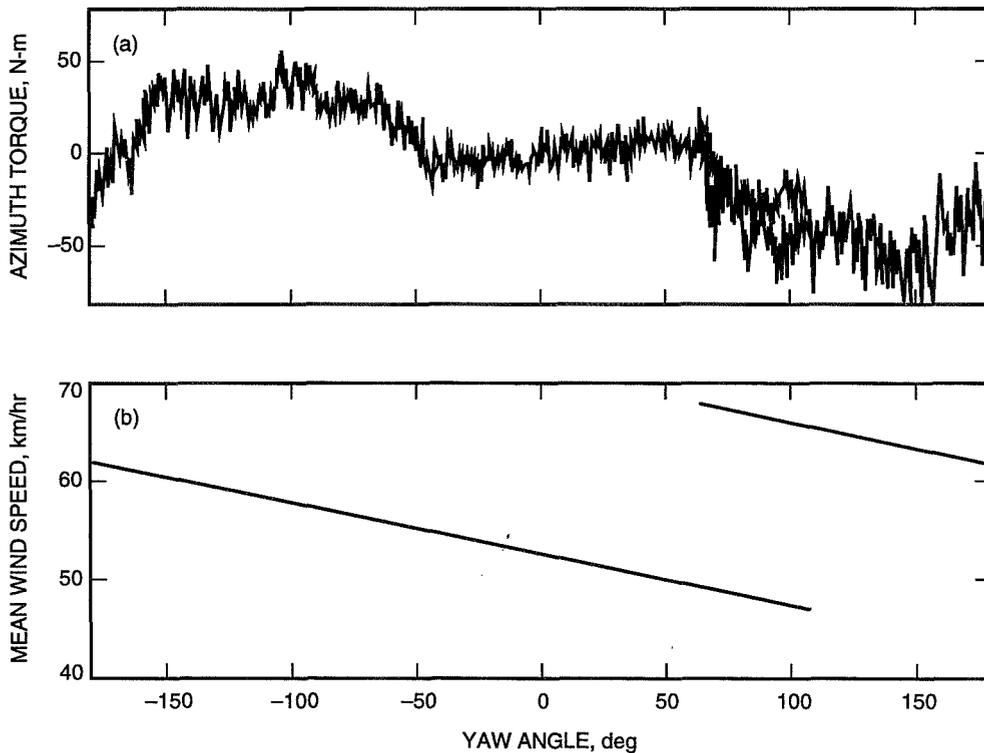


Fig. 3. Elevation angle 10 deg: (a) azimuth torque and (b) mean wind speed versus yaw angle.

A similar experiment was conducted on January 24, 1994, for an elevation angle of 60 deg. The averaged field-measured torques (dots) and CP-6 Memorandum data (solid line) are shown in Fig. 5. For this case, $s_1 = 1.08$ and $s_2 = 0.0008$; thus, the difference between the field and the CP-6 Memorandum data was within 8 percent, and the friction torque correction was about 1 percent.

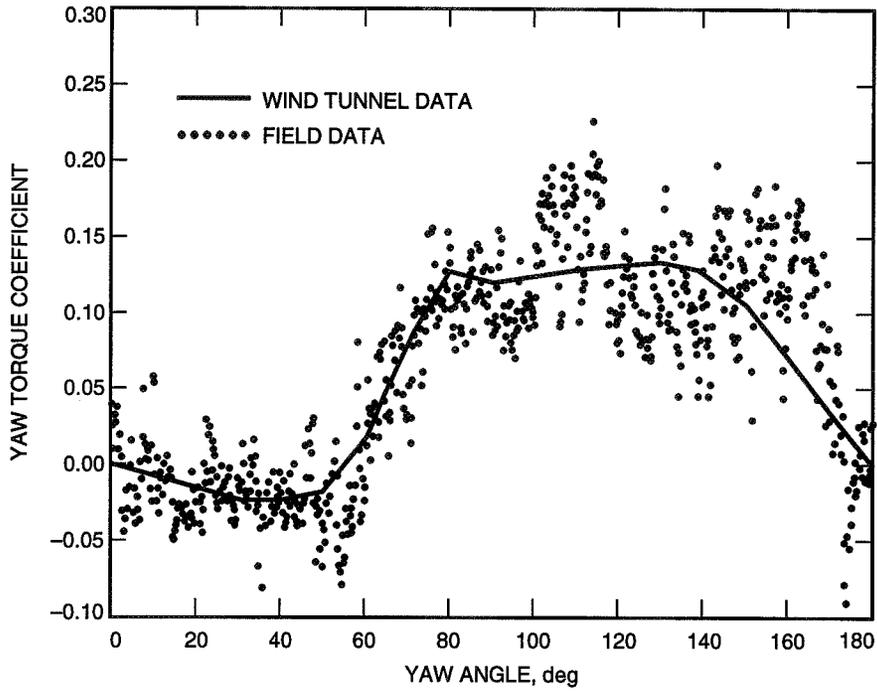


Fig. 4. Yaw torque coefficient from wind tunnel data and field data for elevation angle 11 deg.

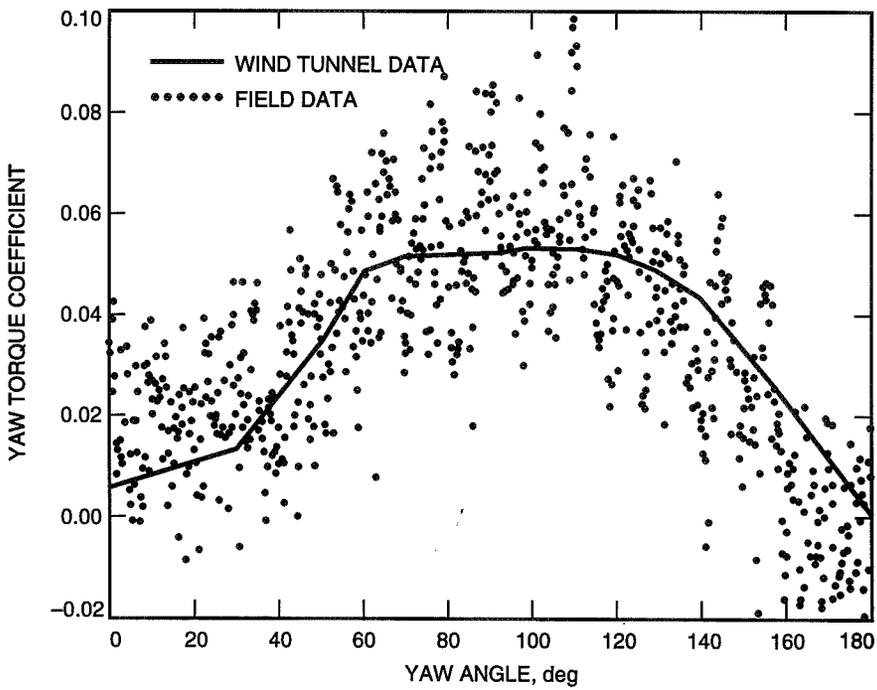


Fig. 5. Yaw torque coefficient from wind tunnel data and field data for elevation angle 60 deg.

The elevation (pitch) torque coefficients were **verified for winds** blowing from behind the antenna (a yaw angle of 180 deg). The torques were measured on January 24, 1994, for elevation angles ranging from 10 to 89 deg. The friction and unbalanced torques were obtained from Ahlstrom and Mellstrom.⁵ The friction torques are 5.2 N-m (4.6 lb-ft), and the unbalanced torques are 2.0 N-m (1.8 lb-ft), with opposite direction to the friction torques (the antenna dish was driven down). The torques are plotted in Fig. 6(a). These torques were scaled to obtain the torque coefficients, and the CP-6 Memorandum torque coefficients were fitted to the field data [see Fig. 6(b)]. The fitting coefficients were $s_1 = 0.87$, $s_2 = -0.013$, showing that the CP-6 and field measurements are within a 13-percent error margin. A slightly larger s_2 was due to poorer friction torque estimation (friction torques in Footnote 5 were obtained for different rates, and the unbalance torque was not known exactly).

III. Power Spectra Analysis

The torque coefficients characterize the wind steady-state pressure on the antenna. In order to have insight into the wind-induced dynamic of the antenna, the power spectra of the torques and the encoder output were obtained.

The resonance peaks in the spectra carry information about flexible modes excited by the wind. These modes can be identified from the elevation and azimuth closed-loop transfer functions of the antenna, shown in Fig. 7. The principal azimuth flexible-mode frequency is 1.66 Hz. Two other azimuth flexible modes are at 3.20 and 4.23 Hz. The principal elevation flexible-mode frequency is 1.95 Hz, and the other two are 3.16 and 5.18 Hz. The principal elevation mode is a bending mode of the antenna [2]. Note also that the second azimuth mode and the second elevation mode are probably the same one, because their frequencies are almost identical.

Figure 8 shows the power spectra of the azimuth encoder and azimuth torque for a 180-deg yaw angle (wind from behind the antenna), for elevation angle 10 deg, and for wind speeds 34 and 60 km/hr (21 and 37 mph). The plots show a dramatic increase of power spectra with wind speed and the excitation of the first azimuth flexible mode (1.66 Hz). The torque spectra show additional excitation of the higher-frequency (4.23 Hz) azimuth mode.

Figure 9 shows the dependency of the azimuth encoder and azimuth torque power spectrum on wind speed, when the yaw angle is 90 deg and the elevation angle is 10 deg. For low frequencies, this side wind has almost the same effect for 37 and 45 km/hr (23 and 28 mph) speeds.

Figure 10 presents the azimuth encoder and torque spectra for a 180-deg yaw angle, a 60-deg elevation angle, and a 42 km/hr (26 mph) wind. For this angle, the torque spectrum shows, besides the already observed azimuth modes (1.66 and 4.23 Hz), the first elevation mode of 1.95 Hz excited.

Figure 11 presents the spectra for a 180-deg yaw angle, a wind speed of 40 km/hr (25 mph), and elevation angles of 20 and 78 deg. For low frequencies, the spectra for a 20-deg elevation are larger than the spectra for a 78-deg elevation angle, while for high frequencies the situation is reversed.

The spectra show that most of the wind energy is accumulated in the low frequencies; hence, an improved controller can compensate for this larger portion of the wind disturbance. They also show that the principal azimuth and elevation flexible modes are significantly excited, followed by the azimuth 4.23-Hz mode.

⁵ H. Ahlstrom and J. Mellstrom, "Antenna Friction Unbalance Measurements," JPL Interoffice Memorandum 3324-91-012 (internal document), Jet Propulsion Laboratory, Pasadena, California, March 6, 1991.

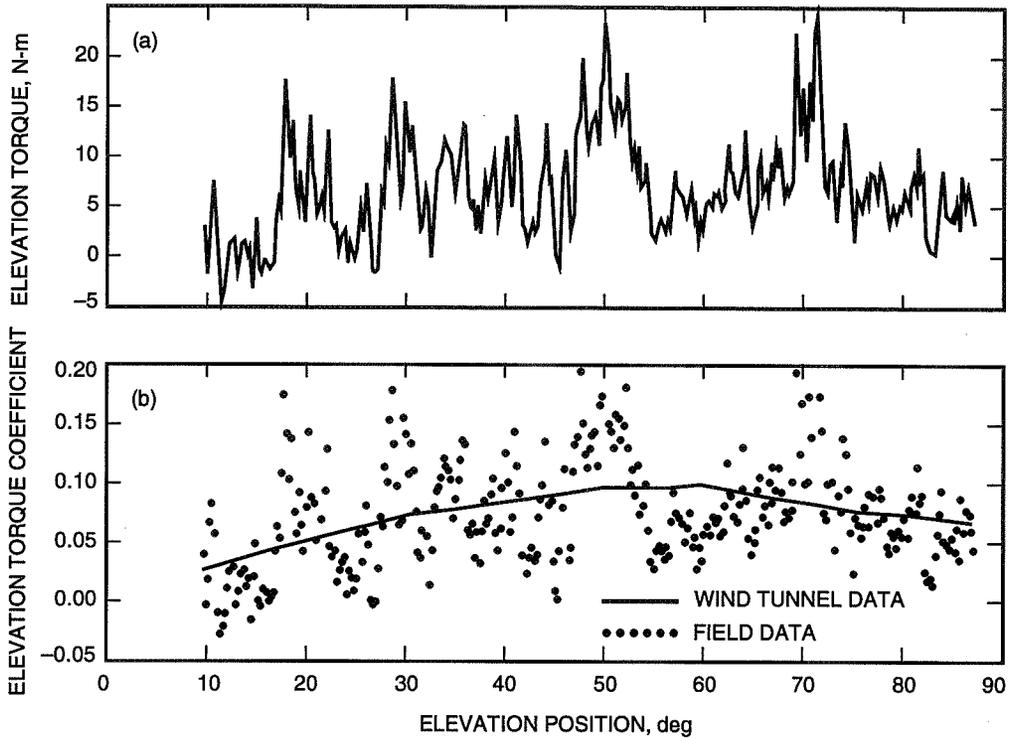


Fig. 6. Elevation angle 60 deg: (a) elevation torque and (b) elevation torque coefficient from wind tunnel data and field data.

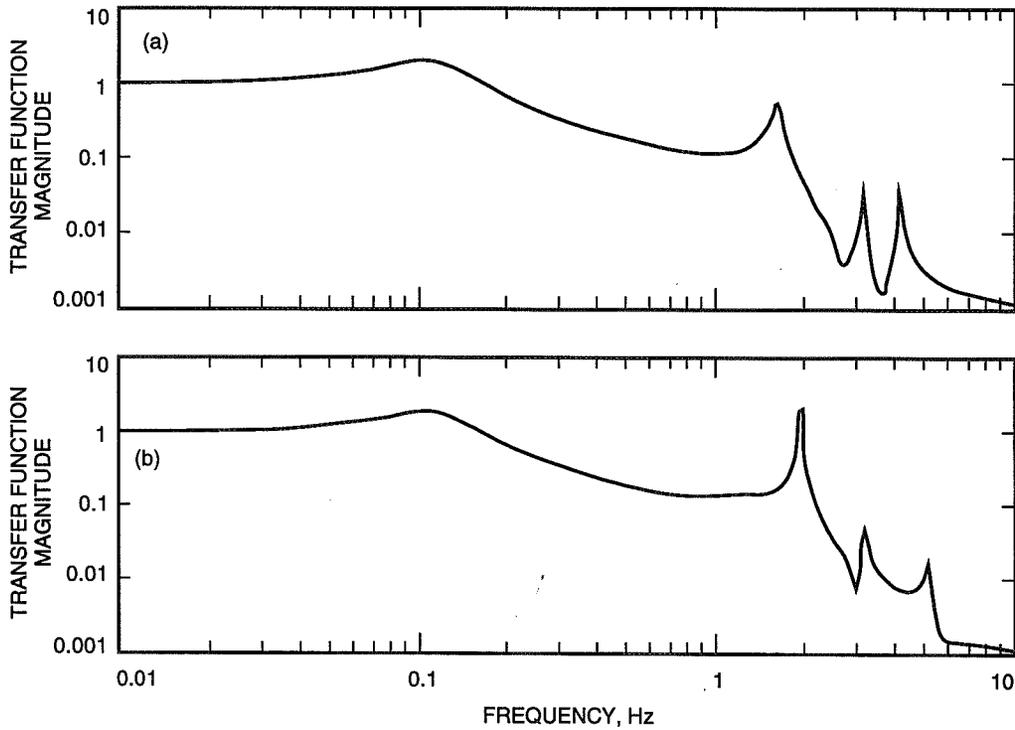


Fig. 7. Magnitudes of the transfer function of the DSS-13 servo: (a) in azimuth and (b) in elevation.

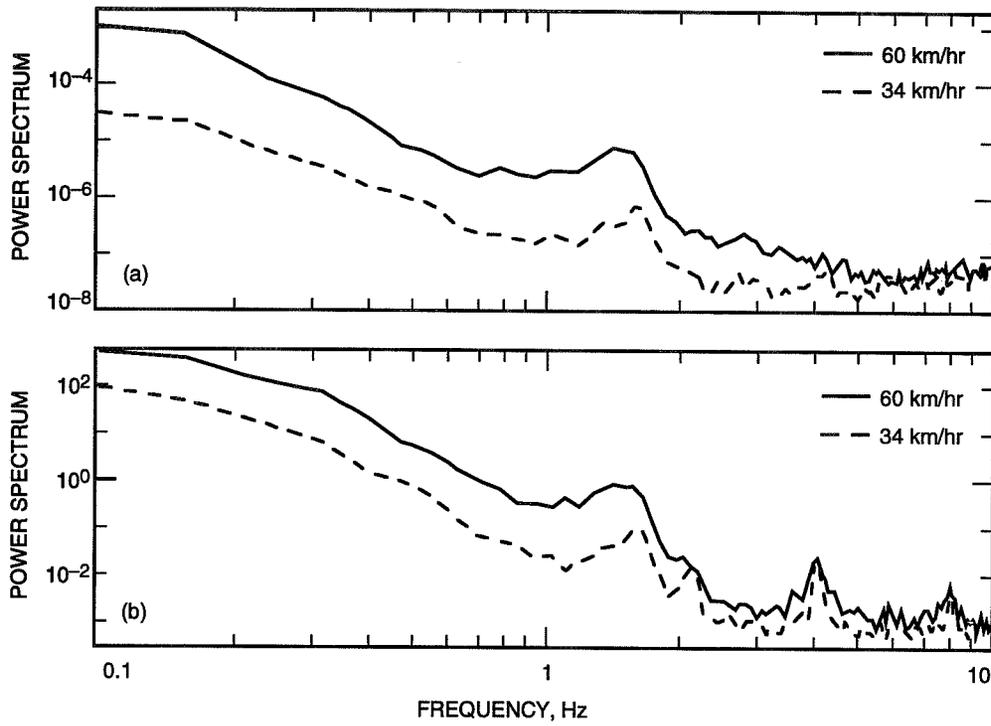


Fig. 8. Power spectra for yaw angle 180 deg and elevation angle 10 deg: (a) azimuth encoder and (b) azimuth torque.

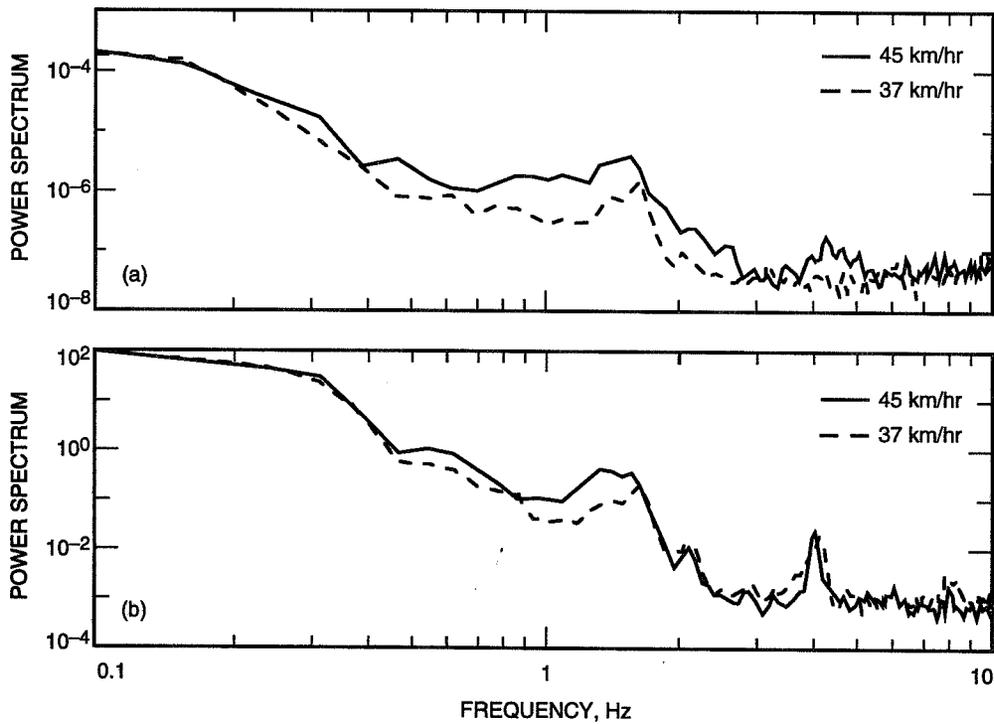


Fig. 9. Power spectra for yaw angle 90 deg and elevation angle 10 deg: (a) azimuth encoder and (b) azimuth torque.

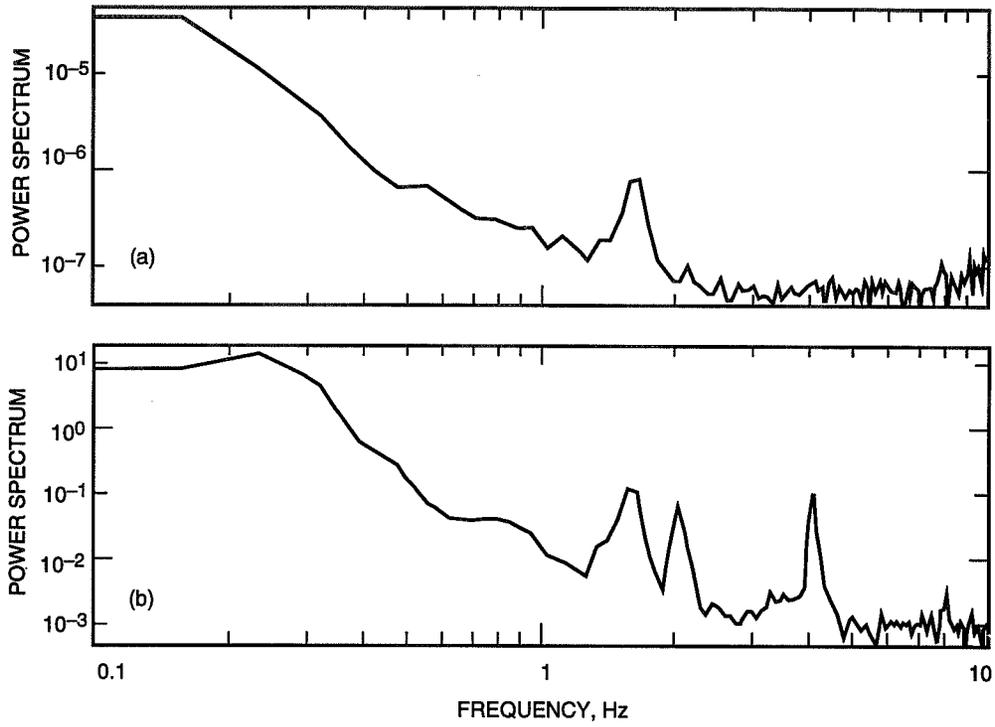


Fig. 10. Power spectra for yaw angle 180 deg and elevation angle 60 deg: (a) azimuth encoder and (b) azimuth torque.

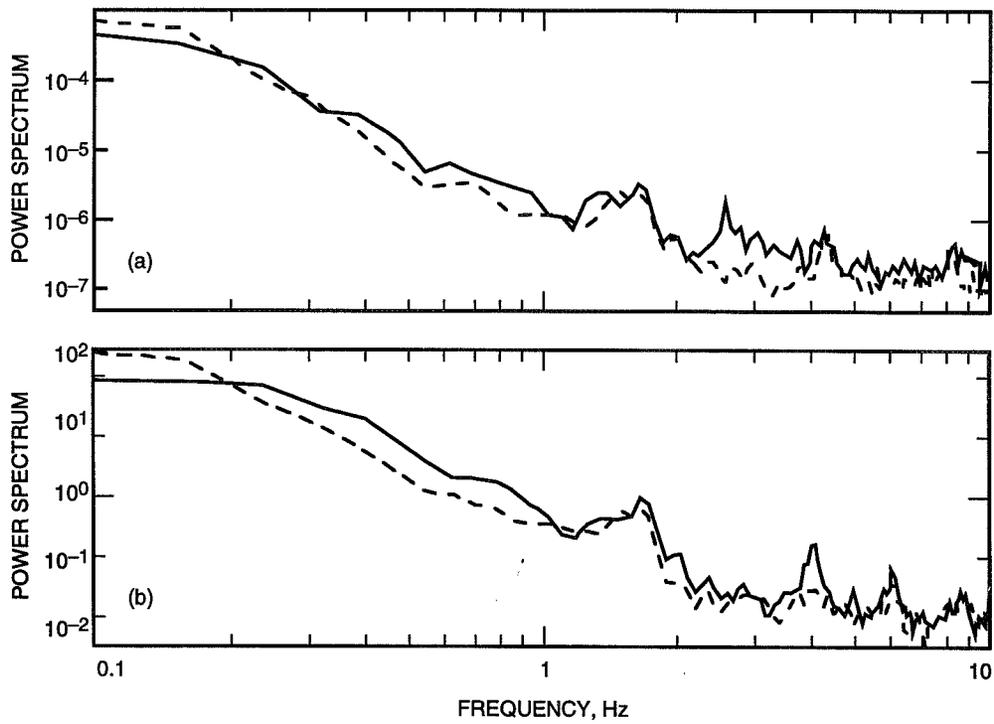


Fig. 11. Power spectra for yaw angle 180 deg and elevation angle 78 deg (solid line) and 20 deg (dashed line): (a) elevation encoder and (b) elevation torque.

IV. Conclusions

The comparison of selected field measurements with the wind tunnel data shows good agreement for the steady-state winds. The differences between the wind steady-state torque coefficients at the DSS-13 antenna and the wind tunnel data were less than 10 percent of the field-measured coefficients. The wind gusting action on the antenna, characterized by the power spectra of the antenna encoder and the antenna torques, showed the wind gusting primarily affects the antenna principal modes.

References

- [1] W. Gawronski and B. Bienkiewicz, "Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances," *The Telecommunications and Data Acquisition Progress Report 42-108, vol. October-December 1991*, Jet Propulsion Laboratory, Pasadena, California, February 15, 1992.
- [2] W. Gawronski and J. A. Mellstrom, "Beam-Waveguide Antenna Servo Design Issues for Tracking Low Earth-Orbiting Satellites," *The Telecommunications and Data Acquisition Progress Report 42-115, vol. July-September 1993*, Jet Propulsion Laboratory, Pasadena, California, pp.135-152, November 15, 1993.

1108 77

33279 p. 15

November 15, 1994

358482

Tracking the Galileo Spacecraft With the DSCC Galileo Telemetry Prototype

T. T. Pham and S. Shambayati
Telecommunications Systems Section

D. E. Hardi
Communications Systems Research Section

S. G. Finley
Tracking Systems and Applications Section

On day of the year (DOY) 062, 1994, a prototype of the Deep Space Communications Complex Galileo Telemetry subsystem successfully tracked and processed signals from the Galileo spacecraft, under fully suppressed-carrier modulation. The demonstration took place at Goldstone, employing the 70-m antenna and the 34-m high-efficiency antenna. This article presents the findings from that demonstration. Specific issues are the system performance in terms of signal-to-noise (SNR) degradation and the arraying gain. Validation of the test results is via symbol-error-rate measurement and the standard symbol SNR. The analysis is also extended to include characterization of the signal received from Galileo.

I. Introduction

On day of the year (DOY) 062, 1994, a prototype of the Deep Space Communications Complex (DSCC) Galileo Telemetry (DGT) subsystem successfully demonstrated signal acquisition and telemetry processing with the Galileo spacecraft. The DGT is being developed to support the Galileo mission, in response to the failure of the spacecraft X-band (8.415-GHz) high-gain antenna. The equipment will be supporting the orbital tour operation from May 1996 to December 1997, during which Galileo will repeatedly encounter the Jovian satellites Ganymede, Callisto, and Europa. Among the benefits the DGT offers, compared to other telemetry subsystems currently available in the Deep Space Network, are

- (1) A capability to record samples of the telemetry signal for later reprocessing over troublesome data gaps, thus enhancing the data return.
- (2) A capability to perform full-spectrum combining. Full-spectrum combining refers to combining the telemetry samples from multiple antennas prior to carrier, subcarrier, and symbol demodulation [1]. In doing so, the signal-to-noise ratio (SNR) in the receiver tracking loops is enhanced. This results in either a smaller telemetry loss or the ability to track the signal at a lower signal level.

- (3) An increase in coding gain with the use of concatenated $(14,1/4)$ convolutional and four-level redundancy Reed-Solomon codes [2].
- (4) A capability to track signals with suppressed carrier, allowing all, instead of partial, signal energy to be put into telemetry data.

The demonstration was planned with the following objectives in mind: (1) to verify the tracking and telemetry processing capability of the DGT with the actual spacecraft, (2) to verify that tracking at small bandwidths, in the order of a milli-Hertz, is feasible, and (3) to verify arraying capability.

This article presents the results of the DOY 062 demonstration, which took place at Goldstone, using the 70-m and 34-m high-efficiency (HEF) antennas. The analysis will focus on system performance in terms of signal-to-noise ratio (SNR) degradation, acquisition delay, and the supportable tracking bandwidths. The array gain will also be considered. In addition, the analysis will extend to studying the characteristics of the Galileo signal, such as its frequency stability and the spacecraft transmitted power. The effect of antenna gain variation due to spacecraft motion on the observed symbol error rate will also be examined.

Section II provides a brief description of the test configuration, both for the ground and aboard the spacecraft. Section III presents the data analysis. A summary is given in Section IV.

II. Configuration

Description of the test configuration is divided into two segments: on the ground and aboard the spacecraft. Deficiencies in the test configuration that prevent a full verification of the capabilities needed for actual support in the 1996 era are also pointed out.

A. Spacecraft Configuration

The Galileo spacecraft was configured in a memory-readout (MRO) mode so that known data could be used as a reference to validate the performance of the ground equipment. Data segments of 640 bits each were extracted from the Attitude and Articulation Control System (AACS) memory and packed into an 800-bit frame, along with a header and real-time engineering data. For a given frame, there were 88 unknown bits associated with real-time engineering data. The remaining 712-bit data segment was known and its equivalent representation in the symbol domain was used in the correlation of the received data stream.¹

The test was conducted with two data rates available at the time of demonstration: 10 bits/sec with the NASA standard $(7,1/2)$ convolutional code and 40 bits/sec uncoded. This configuration allowed the DGT to be tested at different data rates. In future operation during the orbital tour, it is expected that a rate change will occur several times during the track to take advantage of the improved link margin occurring at high elevation. There are two differences in the data flow in the current configuration, compared to that in the 1996 era:²

¹ S. Shambayati, "Day 62 DGT Demo Symbol Error Rate Results," JPL Interoffice Memorandum 3393-94-5501 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 26, 1994.

² These two characteristics will be removed in future operations with the upload of new flight software in March through May 1996. Under the new configuration, data flow will remain in one path via the buffered memory.

- (1) The symbol stream data path was broken whenever the data rate was toggled between 10 and 40 bits/sec.³ As indicated in Fig. 1, the memory-readout data could come either from the buffered memory during 10 bits/sec or from the low-level memory during 40 bits/sec. A switch in front of the subcarrier modulator helped to select the appropriate input. At the time of change in data rate, the switching likely caused a loss of lock in the receiver on the ground.
- (2) The transmitted data sequence to be used as reference for the measurement of symbol error rates was different between the two data rates. This point is illustrated in Fig. 2. The memory data were clocked out at a rate of 8 bits/sec. By the time header information and real-time data were included, the readout was equivalent to 10 bits/sec. At 10 bits/sec, each AACCS frame was encoded by the (7,1/2) encoder and became twice as large. At 40 bits/sec transmission, each AACCS frame was repeated four times. This exception caused an increase in the complexity of the analysis program that computed symbol error rate (SER).

The spacecraft was configured such that the received signal would emulate as much as possible the levels expected in the orbital tour. The low-power transmitter, which is supposedly 4.8 dB lower than the standard transmitter, was used to reduce the signal level. The spacecraft transmitted data with both residual and fully suppressed carrier modulation. Suppressed-carrier modulation (i.e., 90-deg modulation index) was used at 40 bits/sec. At the lower data rate of 10 bits/sec, the carrier modulation was switched to residual. The modulation index was set at 46 deg (closest to the ideal 45 deg) to maintain a relatively constant symbol SNR level across the data-rate change. Under this residual carrier modulation, the DGT could still emulate the suppressed-carrier tracking mode by ignoring the carrier component. This data set can also be used to isolate any failure associated with suppressed-carrier tracking.

B. Ground Configuration

Before getting into the description of ground configuration, a brief digression to the general architecture of the DGT is needed. A complete DGT configuration, such as the one to be deployed in Canberra in 1996, is depicted in Fig. 3. The DGT consists of two channels: one based on a Block V Receiver (BVR) and the other based on a Full Spectrum Recorder (FSR). Telemetry processing, up to the decoder output, is done independently by each channel. The parallel architecture of the system reduces the risk of equipment failure, thus increasing the data return.⁴ This article focuses specifically on the FSR channel.

In the FSR channel, the FSR directly processes the analog 295-MHz signal from the very long baseline interferometry (VLBI)/Radio Science Downconverter (VRD). First, the signal is downconverted to 64 MHz, 8-bit sampled at 256 MHz, then further digitally downconverted to 16 MHz. The samples are time delayed for the purpose of arraying. Individual harmonic components of the square subcarrier, up to the 7th harmonic, are extracted and written to data files on disk for the next assembly to use. The minimal sampling rate of the final product is set to be at least 10 samples per symbol to minimize telemetry processing loss. The Full Spectrum Combiner (FSC) corrects for the frequency and phase difference among different data streams and combines them. The Buffered Telemetry Demodulator (BTD) then performs the carrier and subcarrier demodulation, and symbol extraction on the samples from the FSC or the FSR, in the case of nonarraying. The resulting symbols are decoded in the Feedback Concatenated Decoder (FCD). If both FSR and BVR channels are used, the two FCD outputs are merged and the best product is delivered to the Project.

³ "Module GLL-3-280, Functional Requirement Galileo Orbiter Telemetry Measurements and Data Formats," *Galileo Orbiter Functional Requirements*, Galileo Project Document 625-205 (internal document), Jet Propulsion Laboratory, Pasadena, California, p. 13, January 13, 1986.

⁴ A more detailed description can be found in T. Pham, *DSCC Galileo Telemetry Subsystem, Functional Design and Software Requirements Document*, TDA/DSN Document 834-43, JPL D-11226 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 21, 1993.

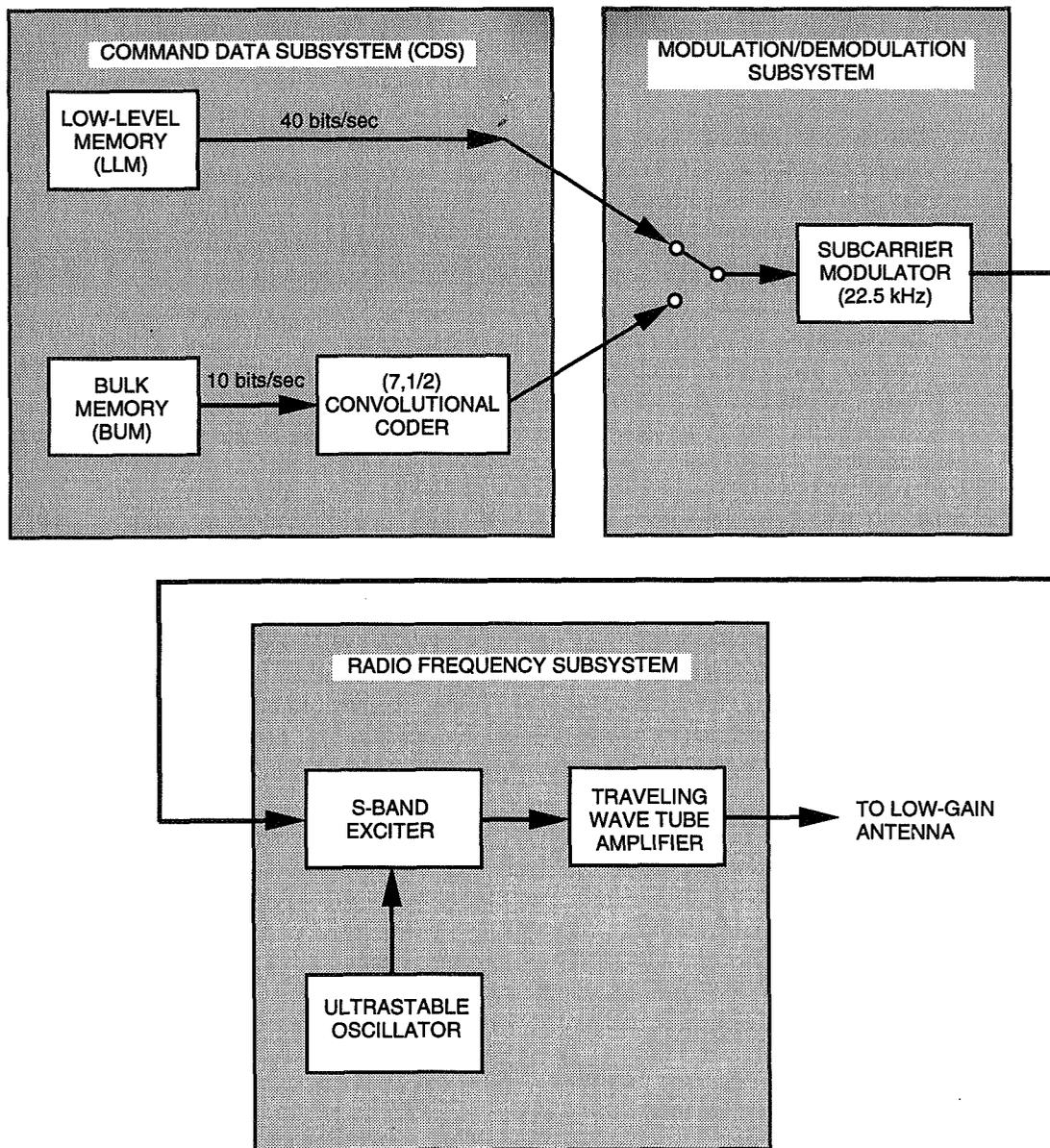


Fig. 1. Data flow onboard the Galileo spacecraft.

Figure 4 presents the DSN ground configuration on DOY 062. The Galileo signal was received at the 70-m (DSS-14) and 34-m HEF (DSS-15) antennas. Even with the use of the low-power transmitter, the Galileo signal level was in the range of 0- to 1-dB symbol SNR, which was still significantly higher than the expected future operating level of -5 dB.⁵ To simulate a lower SNR level, additional noise from the noise diode in the Microwave Precision Power Monitor (PPM) assembly was continuously injected into the low-noise amplifier (LNA) input at DSS 14. Different noise levels were added at different times to emulate different SNR conditions. The amplified S-band (2.3-GHz) signals were converted to 295 MHz IF by the VRD. The IF Distribution Assembly then distributed the signals to the FSRs. The FSR samples

⁵ This -5 -dB level corresponds to the SNR threshold of the new concatenated (14,1/4) convolutional and four-level redundancy Reed-Solomon codes for which the decoded bit error rate is less than 10^{-7} . The operational strategy planned for the orbital tour dictates that the symbol SNR level be kept constant at the decoder threshold. The data rate is adjusted accordingly, based on the power availability in the link budget, to maximize the data return.

were stored on tapes and brought back to JPL for further processing in the FSC and BTM. Once properly demodulated, the BTM symbols were compared against the reference data stream and the symbol error rates were generated. Detailed description on the symbol-error-rate analysis can be found in Footnote 1. Note that the demodulated symbols from the BTM were not processed through the FCD. The reasons were that the low signal level would have resulted in a high bit error rate, in the vicinity of 50 percent for the (7,1/2) convolutional code, and that the (7,1/2) code is not going to be used in normal operations during the orbital tour.

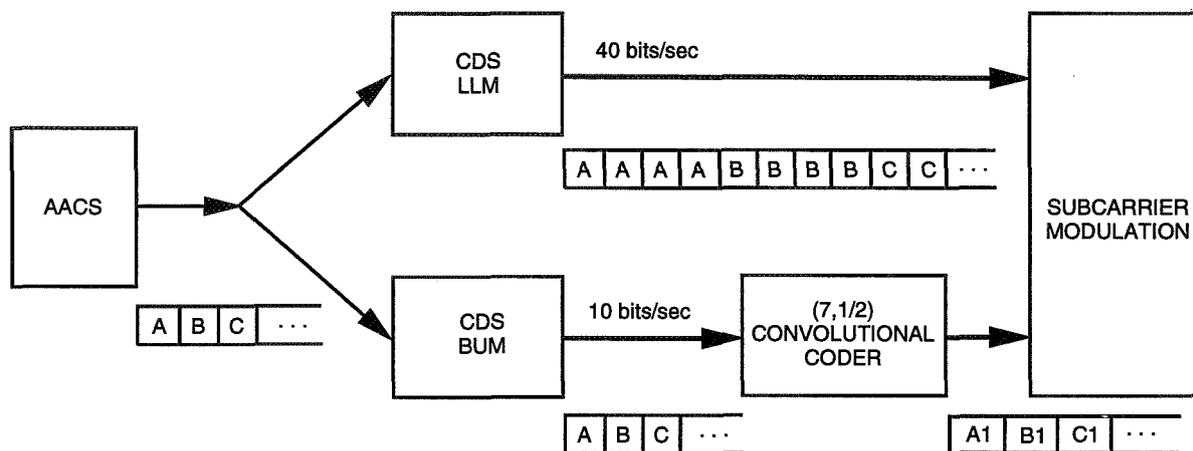


Fig. 2. Data sequence of 10 and 40 bits/sec.

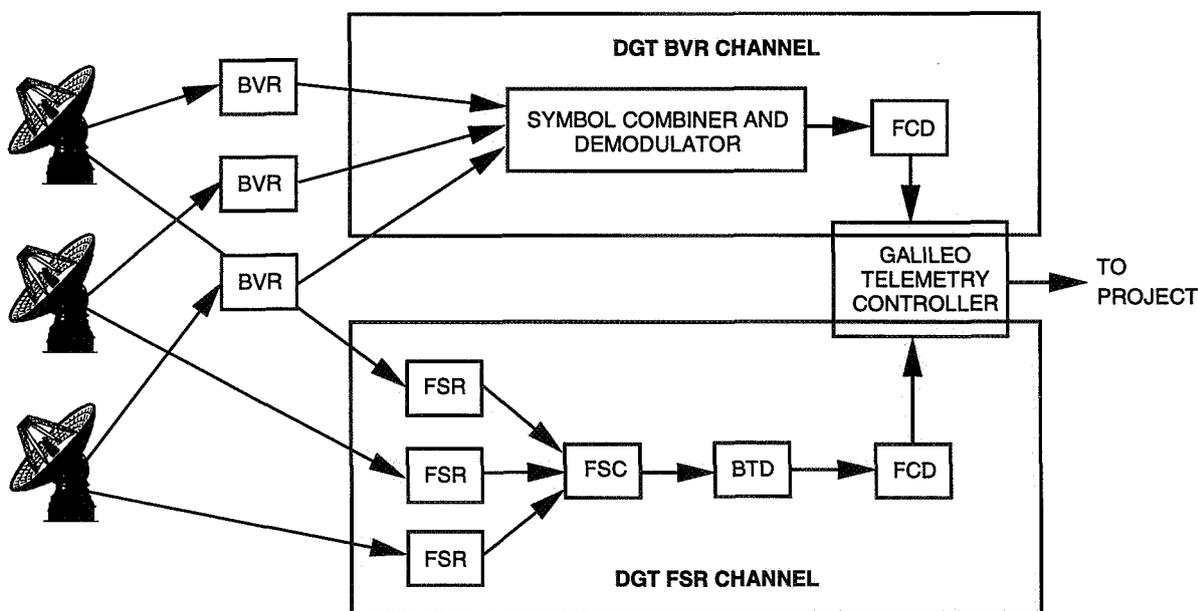


Fig. 3. DGT architecture.

Also indicated in Fig. 4 are monitor data provided at different stages of processing. The FSRs and FSC reported the data power-to-noise spectral density ratio (P_d/N_o) using the fast Fourier transform (FFT). These measurements were translated to symbol SNR (E_s/N_o) using the predicted data rate for an easy and consistent comparison with the BTM estimation. The BTM provided symbol SNR (E_s/N_o). The analysis program produced measured symbol error rates, which were available only over the memory readout period.

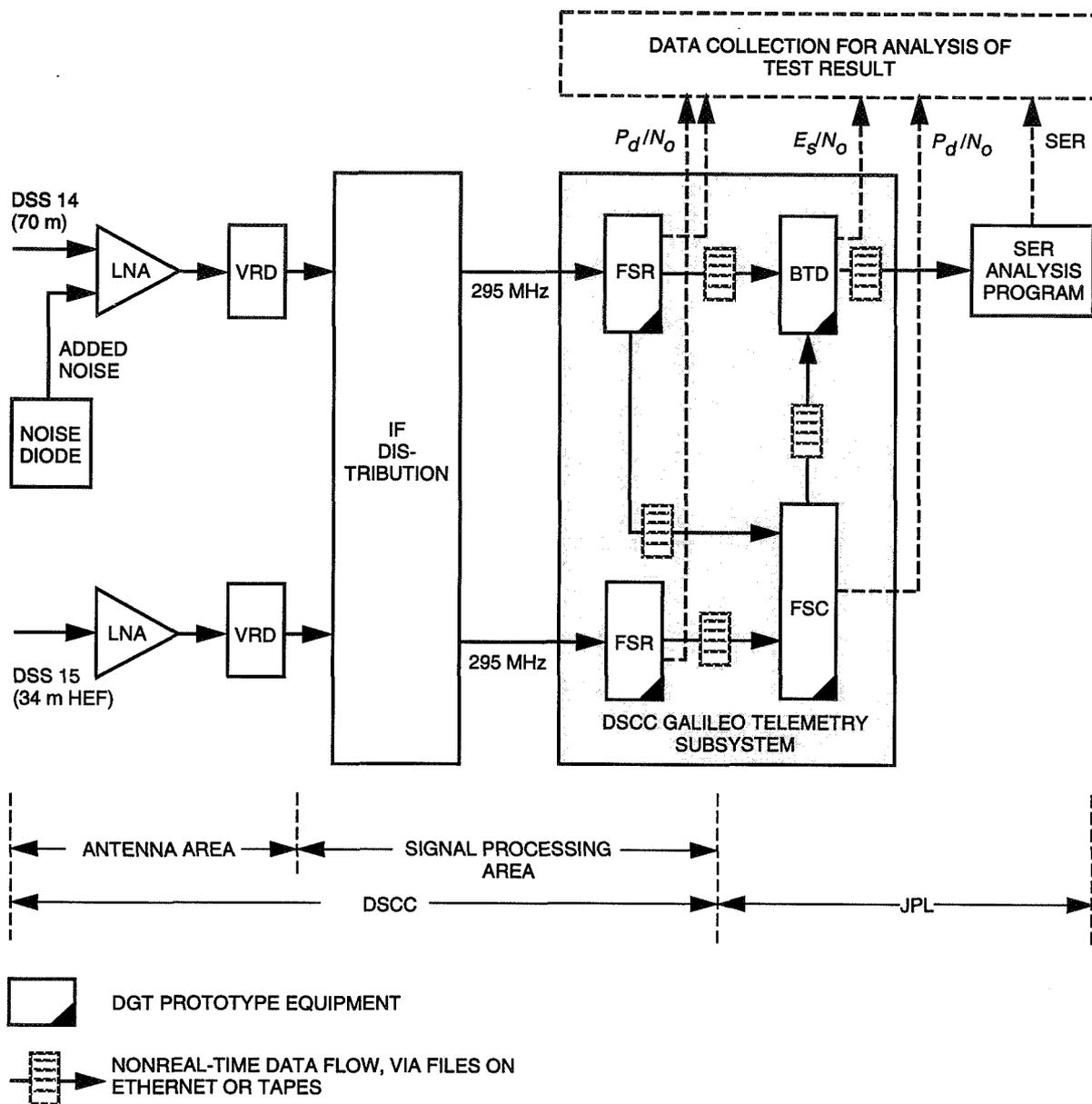


Fig. 4. Test configuration in the DSN.

C. Operational Timeline

Figure 5 presents the operational timeline events on the DOY 062 track. On the spacecraft side, there are four different configurations. The two outer configurations (involving the beginning and end of the track) were of standard operations, for the purpose of performance verification prior to and after the test. The two inner configurations involved memory readout and were the focus of the test. On the ground, noise injection was done in four stages: none at the beginning and at the end of the track, 8-K noise added from 10:49 to 13:16 Greenwich mean time (GMT), and 46-K noise added from 13:16 to 14:35 GMT. This resulted in three nominal signal conditions with symbol SNRs of about 0.5 dB, -0.5 dB, and -4.2 dB over the memory readout region.⁶ (Near the end of the track, a 7.3-dB level was also available

⁶ These levels are referenced to the FSR input or, equivalently, the antenna input. They include power from all subcarrier harmonics.

when the spacecraft was reconfigured for normal operation with the high-power transmitter). To properly track the signal, the receiver bandwidths in the BTD were set such that at least 15-dB-Hz loop SNRs were ensured. Most of the track was done with 0.25-Hz carrier, 0.09-Hz subcarrier, and 0.01-Hz symbol bandwidths. During the period where symbol SNR dropped to -4.7 dB-Hz (from 13:16 to 14:35 GMT), a set of narrower loop bandwidths (0.05 Hz for carrier, 0.02 Hz for subcarrier, and 0.003 Hz for symbol) was used. The FSR configuration also varied during the track. The prototype FSR was limited to capturing two spectral components of the signal, either the two subcarrier harmonics (1st and 3rd) or only the 1st harmonic and the carrier. Residual carrier data, from 11:11 to 12:14 GMT, were used for the validation purpose. Over this period, the SNR in telemetry data dropped by 0.46 dB. As the FSR switched in (at 11:03 GMT) and out (at 12:14 GMT) of the carrier-capturing mode, an approximate 9-min data outage occurred as a result of reconfiguration.

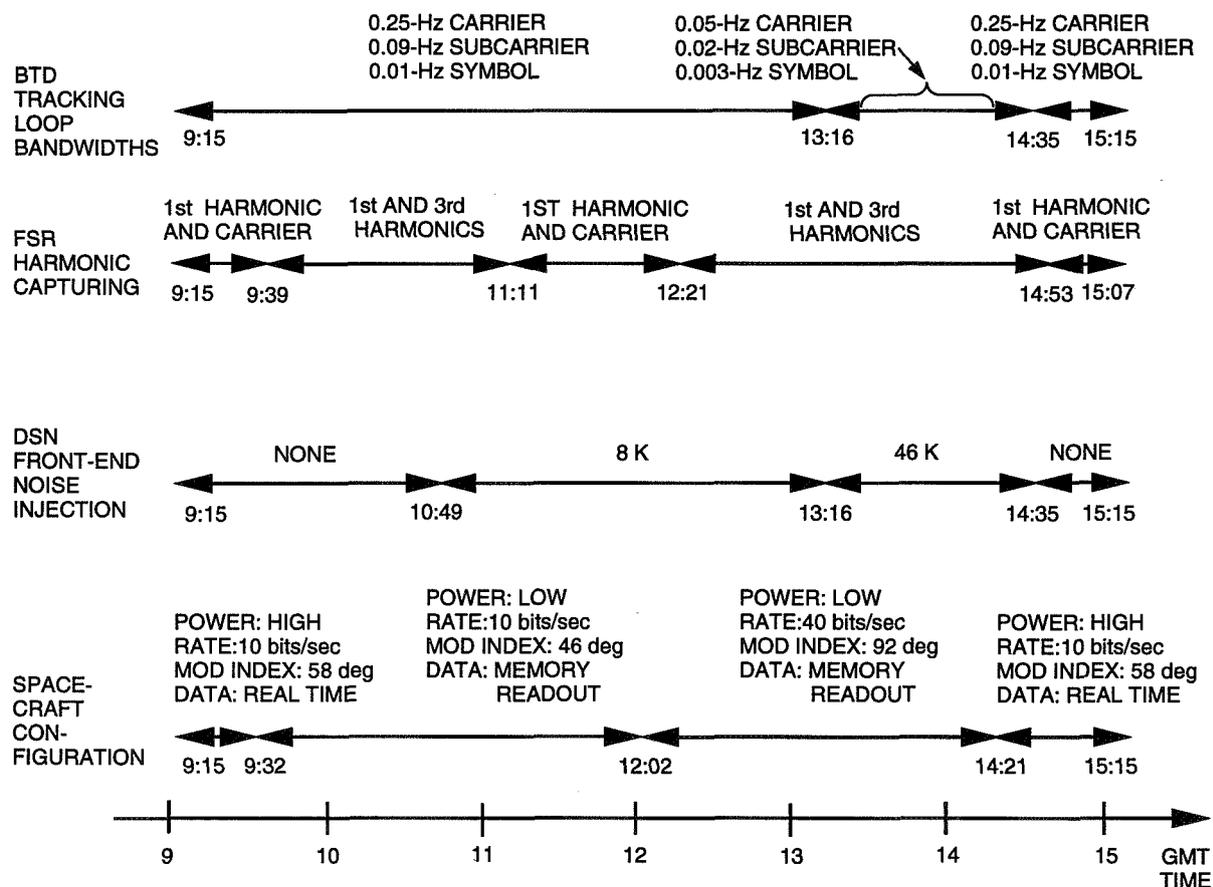


Fig. 5. Timeline events on spacecraft and DSN configuration.

III. Data Analysis

In this section, we focus on performance of the system. Particular attention is given to the SNR degradation and array improvement, acquisition time, the existence of data gaps caused by abrupt changes in signal conditions, and the feasibility of tracking at a small bandwidth. Also examined is the stability of the carrier and subcarrier frequencies. Any differences, if they existed, between the expected and observed signal characteristics will also be pointed out.

Predicted symbol SNR was obtained from the Galileo S-Band Analysis Program (GSAP).⁷ The antenna-referenced predicts were then translated to such observables as FSR, FSC, and BTM outputs using the degradation model presented in Fig. 6. In this model, the expected SNR degradation in the FSR was 0.99 dB for the 1st-harmonic capture and 0.53 dB for the 1st- and 3rd-harmonic captures. This degradation included the outer harmonic loss and the 0.08-dB degradation due to filter mismatch. For the FSC, the gain was expected to be equal to the theoretical gain minus a 0.1-dB loss due to thermal jitters. In the BTM, the expected degradation was 0.1 dB above the theoretical loss in the carrier, subcarrier, and symbol tracking loops. The theoretical loss in the BTM was a function of the symbol SNR, symbol rate, and tracking loop bandwidths. The extra 0.1 dB accounted for any filter mismatch.

The observed and predicted performances seen at different points in the system are shown in Figs. 7, 8, and 9. Figure 7 focuses on the symbol SNR observed at the FSR and BTM output under the single antenna condition (DSS 14 only). Figure 8 illustrates the array gain referenced to DSS 14. The symbol SNRs observed at DSS 14 and the combined symbol SNRs at DSS 14 and DSS 15 are shown. A 1-min integration window for symbol SNR estimation was used in Figs. 7 and 8. Figure 9 presents the symbol error rate for the single-antenna configuration. Each SER report corresponds to one AACSS frame. That is, the integration windows for the correlation between the reference and received data were 1600 symbols at a 10 bits/sec data rate, and 3200 symbols at 40 bits/sec. Several conclusions can be drawn from these charts.

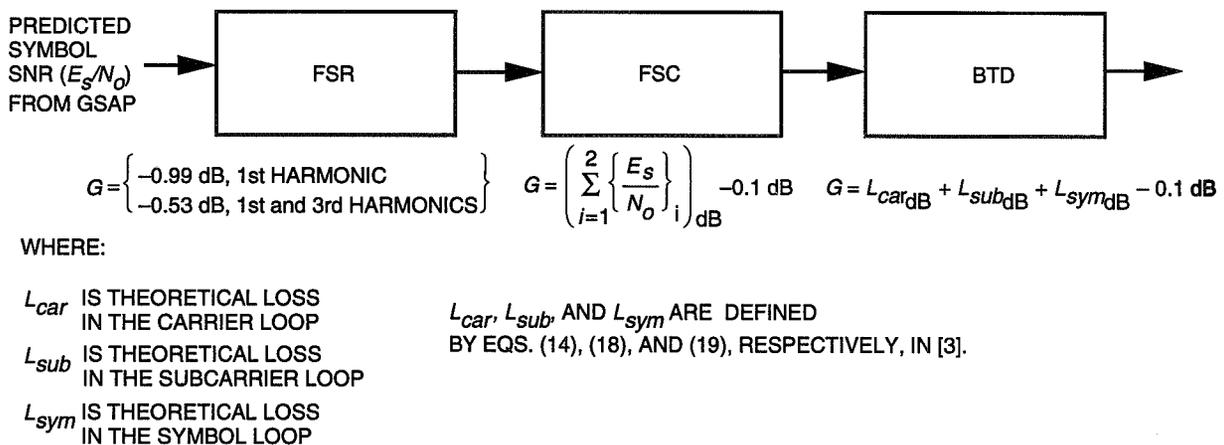


Fig. 6. Model of gain/loss profile in the DGT.

A. Predicts Versus Measurement

In Fig. 7, the given predicts were referenced to the FSR and BTM outputs. The predicted loss in the BTM, i.e., the difference between the predicted FSR and BTM symbol SNRs, was influenced by three factors: (1) setting of tracking loop bandwidths, (2) input symbol SNR, which was a function of the number of captured harmonics, noise injection on the ground, and selection of the high- or low-power transmitter onboard the spacecraft, and (3) the data rate.

Both the FSR and FSC estimated the SNR based on FFT spectral analysis of the cross product of the upper and lower sideband of the first subcarrier harmonic.⁸ The BTM, on the other hand, estimated

⁷D. Bell, "Summary of Final Updates to GSAP 3.5 to Create GSAP 3.6," JPL Interoffice Memorandum 3392-94-087 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 10, 1994.

⁸D. Rogstad, "Processing in the Full Spectrum Recorder and the Full Spectrum Combiner," personal communication through unpublished notes, Tracking Systems and Applications Section, Jet Propulsion Laboratory, Pasadena, California, 1993.

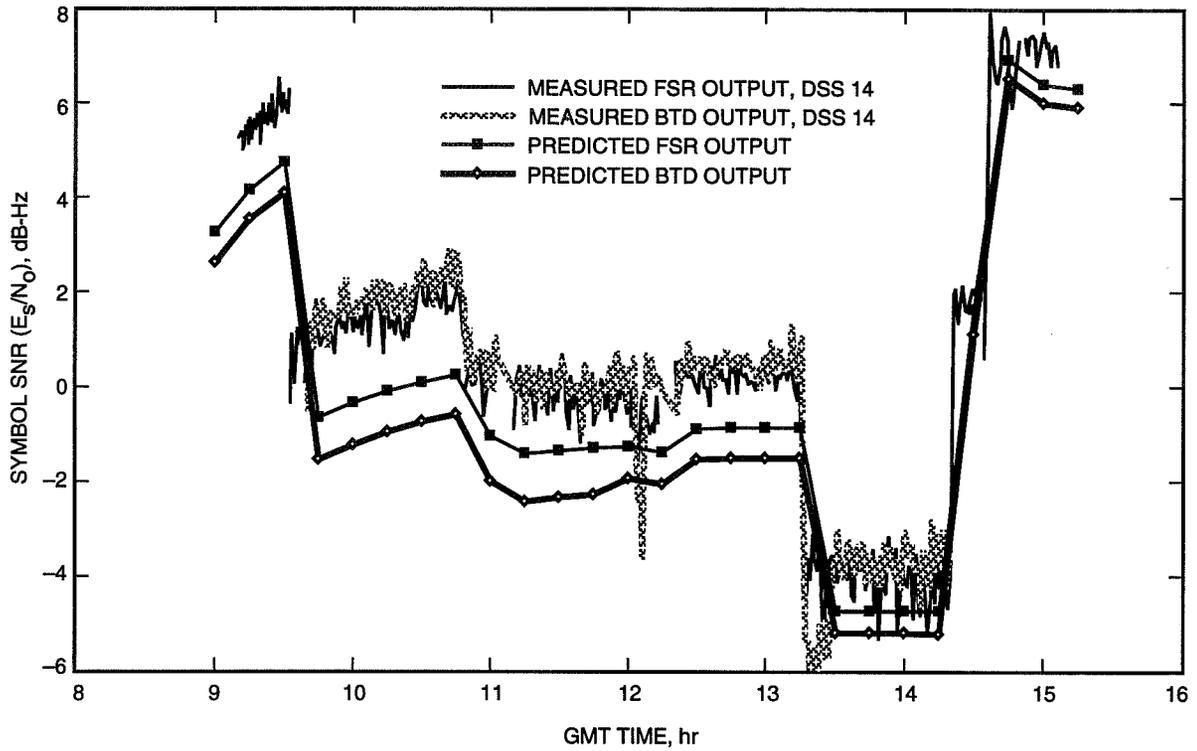


Fig. 7. Result from FSR and BTM, single antenna (DSS 14), DOY 062, 1994.

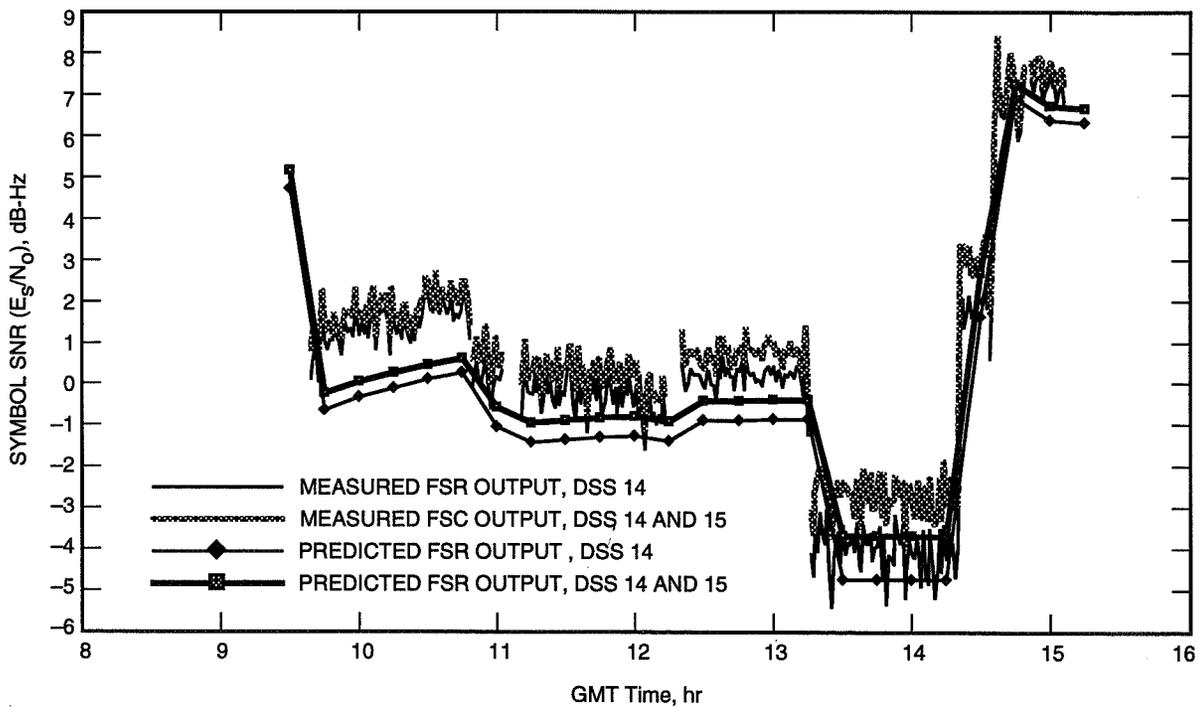


Fig. 8. Illustration of arraying gain, DOY 062, 1994.

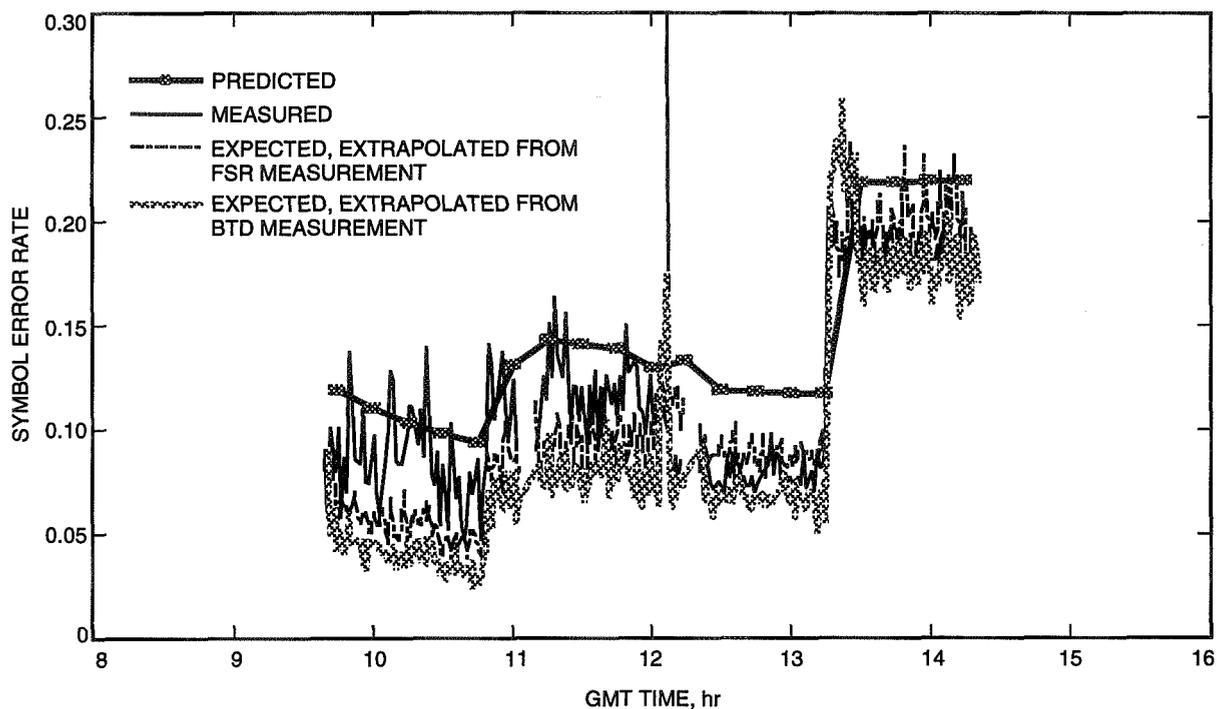


Fig. 9. Symbol error rate at DSS 14, DOY 062, 1994.

symbol SNR (E_s/N_o) using the split-symbol-moment-estimation (SSME) algorithm [4]. The two symbol SNR profiles from the FSR and BTM closely follow each other. This agreement established a general consistency between the two estimations. However, the measured SNRs from the BTM were about 0.3 dB higher than those from the FSR, instead of 0.6 to 0.9 dB lower. This difference is being sorted out and correction will be made prior to the DGT formal delivery to DSN Operations.

The SERs determined by the SER Analysis Program, shown in Fig. 9, were in the approximated region of the reported signal level. This fact lends support to the belief that the DGT was operating properly. The SER profile was essentially an inversion of the symbol SNR profile. Also indicated in Fig. 9 are the expected SERs based on the symbol SNRs reported by the BTM and FSR (with proper accounting of BTM loss). During the 40-bit/sec period, the measured SERs lie between the BTM and FSR measurements. During the 10-bit/sec period, the measured SERs appeared to be high and had a large variation. This problem is being further investigated.

When the predicts and measurements were compared, the observed SNRs were found to be much higher. The positive residual was attributed to an error in the GSAP prediction. Specifically, GSAP used prelaunch calibration data in computing the antenna gain as a function of cone angle. This calculation was about 1 dB lower than that experienced inside the 10-deg cone angle. Subsequently, the GSAP's database was updated to remove this discrepancy.⁹

The residuals also varied over time, as seen in Table 1. Three sets of residuals are presented, from the FSR, BTM, and SER measurements. At the beginning of the track, the residual SNRs were as large as 2.0 dB. Toward the end of the track, they progressively reduced to 1.0 dB. Since there was no noise injection at the start of the track, one might suspect that the variation was related to the added noise. That is, contribution of the added noise, which was measured during the precalibration period, was

⁹D. Bell, op. cit.

perhaps overestimated. However, this was ruled out due to the fact that the residual remained small at the end of the track when all added noise was removed, i.e., post 14:35 GMT. It is not known at this point what caused this phenomenon.

Table 1. Residual symbol SNR, observed at the FSR and BTD.

Configuration	Residual symbol SNR, dB (measured-predicted)		
	FSR	BTB	SER
From 9:39 to 10:49 GMT (low-power transmitter, 10 bits/sec, no added noise)	1.9	2.9	0.7
From 10:49 to 12:02 GMT (low-power transmitter, 10 bits/sec, 8-K noise added)	1.5	2.3	0.6
From 12:02 to 13:16 GMT (low-power transmitter, 40 bits/sec, 8-K noise added)	1.2	2.0	1.6
From 13:16 to 14:21 GMT (low-power transmitter, 40 bits/sec, 46-K noise added)	1.0	1.8	1.2
From 14:37 to 15:06 GMT (high-power transmitter, 10 bits/sec, no noise added)	1.1	- ^a	- ^a

^a Not available.

B. Acquisition Time in the BTB

The acquisition in the BTB, at the time of the demonstration, was limited to carrier-phase and subcarrier-phase acquisitions. A determination of frequency offsets was done off-line, either via self-determination or via FSR detection. Phase acquisition in the BTB was found to be less than 1 min. The in-lock time was defined as that time when the estimated symbol SNRs first achieved stability.

During the track, the BTB was subjected to three disruptions. The first occurred at 11:13 GMT with the introduction of the 8-K noise diode, which resulted in a 1.2-dB drop in symbol SNR. The second was caused by a data-rate transition from 10 to 40 bits/sec at 12:02 GMT. The last took place at 13:15 GMT when the 46-K noise diode was injected into the system. Due to a significant change in the symbol SNR of 3.7 dB, the BTB overreacted in the last event and took longer to recover. The recovery times associated with the three disruptions were 4, 5, and 14 min, respectively. Is this a point of concern? Before answering this question, let us take a look at what is expected in future operations. As previously mentioned in Section II, the data path will not be disrupted by rate changes. So, the disruption associated with the rate change is not a point of concern. The impact of the SNR change in the first and third disruptions, however, may be problematic. The expected distribution of data-rate changes results in a 0.97- to 1.8-dB change in the SNR for 74 percent of the time of the entire orbital tour, a 3.0-dB change for 25 percent of the time, and a 4.2- to 4.7-dB change for 1 percent of the time.¹⁰ These levels of change are not much different from the test conditions. Therefore, more consideration will need to be given to this area to ensure proper operations during actual mission support.

¹⁰ These values are based on A. DiCicco, "Distribution of Galileo Data Rate Changes," JPL Interoffice Memorandum GLL-MOT-94-092 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 10, 1994.

One might also notice that there were two empty data segments. They occurred between 11:02 and 11:11 GMT, and 12:13 and 12:21 GMT. These gaps were the results of FSR configuration changes in the number of subcarrier harmonics to be captured.

C. Tracking With Small Loop Bandwidths

For the receiver to properly track the phase of the incoming signal, the loop SNR must be at least 15 dB-Hz. This requirement translates to a reduction in the loop bandwidth as the signal received from the spacecraft becomes weaker and weaker. In addition, the narrower the loop, the less phase jitter and, thus, less degradation. For the Galileo mission, this means more link margin. There is another way of looking at this: With smaller loop bandwidth, the loop SNR exceeding the 15-dB threshold can be realized more quickly. This advantage in turn enables successful tracking at lower elevation, thus lengthening the tracking time. However, the loop can be narrowed only to the point where the frequency instability of the signal becomes a problem.

The minimum bandwidths currently available in the DSN telemetry systems are 1 Hz for the carrier, 30 mHz for the subcarrier, and 30 mHz for the symbol. The result from this demonstration indicates that a tracking bandwidth of 0.25 Hz to 50 mHz for the carrier, 20 mHz to 10 mHz for the subcarrier, and 10 mHz to 3 mHz for the symbol are supportable. For example, by narrowing these bandwidths to 0.1 Hz for the carrier, 10 mHz for the subcarrier, and 5 mHz for the symbol, the signal can be tracked at an SNR at least 8 dB lower and with a degradation of 0.8 dB less, assuming the configuration of a fully suppressed carrier, a 40-symbol/sec data rate, and a -4.7 -dB symbol SNR.

D. Array Performance

Figure 8 shows the measured and predicted array gain in the FSC. Without any noise in the front end, the predicted SNR level at DSS 15 was about 9.5 dB below that at DSS 14. The predicted gain for this configuration was 0.38 dB. As additional noise was injected into the system, the SNR level at DSS 14 got smaller, and the relative contribution from the DSS 15 antenna became more pronounced. The predicted gain increased to 0.47 dB during a period of 8-K added noise, and 1.06 dB during the 46-K noise-injection period.

Table 2 summarizes both the measured and predicted average gain for each of the main configurations. The result indicates that the arraying gain was fully realized. Notice that during the period of the 46-K noise injection, the measured gain of 1.33 dB was much higher than the predicted gain of 1.06 dB. This discrepancy is believed to be an artifact, due to the fact that the low signal level at the single antenna (DSS 14) resulted in erroneous measurements. These inaccurate measurements corresponded to large variation that often fell below the -5 -dB level seen in Fig. 8.

Table 2. Array performance.

Configuration	Array gain, dB	
	Predicted	Measured
From 9:39 to 10:49 GMT (no added noise)	0.38	0.34
From 10:49 to 13:16 GMT (8-K noise added)	0.47	0.54
From 13:16 to 14:35 GMT (46-K noise added)	1.06	1.33
From 14:35 to 15:06 GMT (no added noise)	0.34	0.36

E. Observed Spacecraft Signal Characteristics

The carrier and subcarrier frequency detected by the FSR appeared to be stable. The carrier residual frequency varied from 0.82 to 0.94 Hz. The subcarrier frequency was found to be 0.3 Hz below the expected 22.5 kHz. Its variation ranged from 0.34 to 0.32 Hz. This stability performance was consistent with the fact that the BTD was able to maintain tracking at a small bandwidth, as discussed earlier.

The effect of antenna gain variation on the observed symbol error was also examined. This variation was due to the 3 rotations/min of the spacecraft. Because there were unknown real-time data imbedded in the received telemetry frames, special care was given to the spectral analysis of the SER. Instead of applying an FFT directly on the whole measured SER data set, FFT processing on individual frames was needed. Over the 712-bit known data segment, 16 SER calculations were performed with an integration of roughly 44 symbols. The 16 samples were FFT transformed. Then an average FFT was computed over several frames. The final result was plotted in Fig. 10. The result demonstrated a clear tone at 0.05 Hz (a 20-sec period or 3 rotations/min) and 0.1 Hz (10-sec period). The detection of a faster oscillation with a 10-sec period was consistent with an earlier finding.¹¹ These oscillations in the SER confirmed the need to have a link margin sufficient to meet the overall bit error-rate requirement.

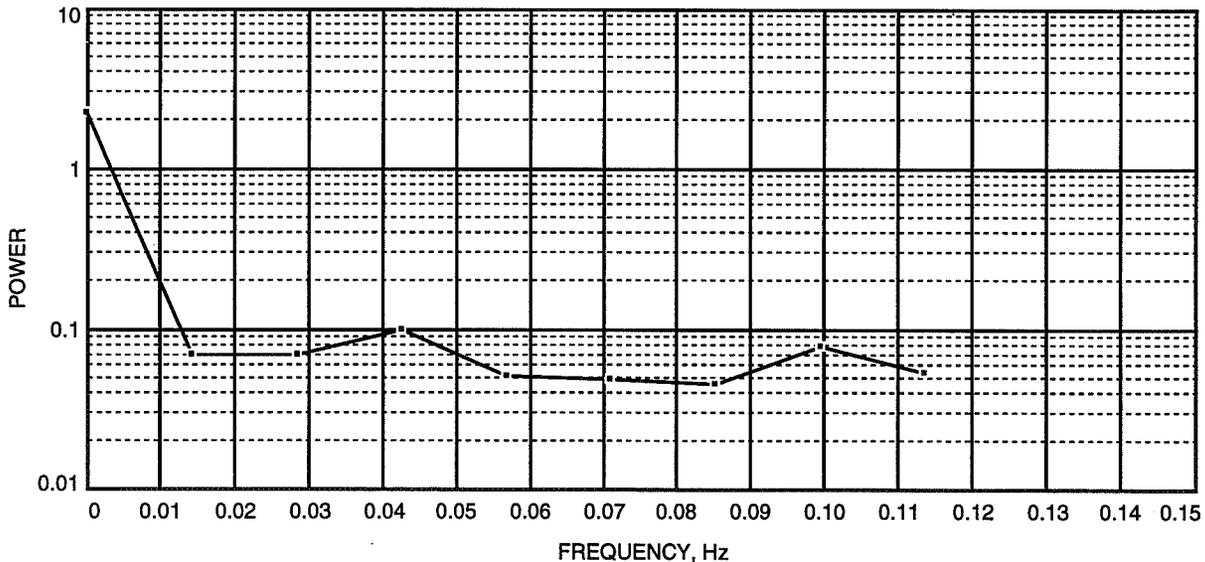


Fig. 10. Power spectral density of symbol error rate, DOY 062, 1994.

An abnormality related to the spacecraft transition from the high- to the low-power transmitter was detected. The two transitions occurred at 9:32 GMT, from high to low, and at 14:21 GMT, from low to high. Table 3 presents the expected and observed SNR changes for each transition. It was found that the measured change for both transitions was 0.6 dB less than expected. Notice that in conjunction with the power change, there was also a change in the modulation index and a data-rate change; however, these factors, as discussed next, were not responsible for the discrepancy. First, notice that a change in data rate was not common at both transitions; rather it occurred only at the second transition. This and the fact that good symbol synchronization was achieved at the specified data rate indicate that the data rates were well known and that the data-rate change was not the cause of the problem. Second, the uncertainty associated with the actual setting of the modulation indexes was found to be quite small, within the accuracy of the measurement. The modulation indexes extracted from the FSR measurement

¹¹ D. Watola, "DOY 173/124 Galileo LGA Downlink Data Analysis," JPL Interoffice Memorandum 331-93.5-030 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 20, 1993.

of relative power in the carrier and data components were within 2 deg of the expected values. For instance, the measurement indicated a modulation index of 59.7 deg versus 58 deg expected, 47.3 deg versus 46 deg expected, and 59.9 deg versus 58 deg expected, over the three periods of 9:15 to 9:30 GMT, 11:11 to 12:02 GMT, and 14:53 to 15:07 GMT, respectively. This agreement between observed and expected modulation indexes eliminated the uncertainty in the modulation index setting from the list of possible suspects. As a result, the discrepancy between the expected and observed change in the power levels across the transition were likely a result of switching between the low-power and high-power transmitters. The gain difference between the two transmitters was only 4.2 dB instead of the expected 4.8 dB.

Table 3. Discrepancy in SNR as the spacecraft transitions between normal and test configurations.

Configuration	Performance, dB	
	Predicted	Measured
First transition (mod index: 58 → 46 deg; power: P → P - 4.8 dB)	-6.23	-5.6
Second transition (mod index: 92 → 58 deg; power: P → P + 4.8 dB data rate: 40 bits/sec uncoded → 10 bits/sec coded)	6.38	5.8

IV. Conclusion

In conclusion, the demonstration has shown that the functions that are required for future Galileo support were realized in the DGT prototype. Proper data sampling of selected harmonics in the FSR, suppressed-carrier tracking in the BTB, and the ability to track at milli-Hertz bandwidths were demonstrated. The observed SNR degradation in the subsystem was found to be reasonable, with the inconsistency in the reported measurements from the FSR and BTB noted. Full spectrum combining was carried out with a measured gain within 0.1 dB of the expected.

In addition, it was discovered that the transition from the high-power to the low-power transmitter did not result in an expected power change. The difference between the measured and expected change was 0.6 dB. The impact of the gain variation on the received signal due to spacecraft rotation was observed as an oscillation in the measured symbol error rate. Although the measurement was made with essentially uncoded symbols, there should be some impact on the output of the future (14,1/4) convolutional and Reed-Solomon codes. The confirmation of this oscillation lends support to the conservative approach that the Galileo S-Band Analysis Program is taking, in terms of reserving some link margin to compensate for this effect.

The demonstration also pointed to a large discrepancy between predicts and measurements. This knowledge helped to update the prediction model accordingly, resulting in better mission operation planning.

Acknowledgments

This analysis would not have been possible without the dedication and support of many people. We would like to express our sincere thanks to Roger Lee, Dave

Fort, Kay Cloud, and Andre Jongling for staying up all night supporting the DOY 062 track. Obviously, data collection would not have been realized without extensive development effort from the FSR team led by Dave Rogstad and including Elliott Sigman, D'Arcy Tyrell, Maria Carlton, Ron Baalke, Robert Navarro, Robert Proctor, Leslie White, and Ismael Cruz. Our gratitude is also directed to Robert Lee, who spent many hours helping with the data processing, and to other BTM members, namely, Sami Hinedi, Kimberly Simpson, Haiping Tsou, Biren Shah, Tim Dry, and Simpson Million, who are supporting the BTM development. A contribution from Robert Kahn and Jeff Piero in prototyping the FSC and processing the combined data was valuable in determining the array performance. Excellent DSN operations support from Larry Bracamonte with noise injection, Glyn Martinez with the hardware installation, John Rothrock with predicts generation, and Michelle Andrews with general network operation support was greatly appreciated. Thanks to Richard Benninger for facilitating the equipment installation at the DSCC and to Joseph Statman, the Task Manager, for planning the demonstration and providing all necessary support to make it happen. We also would like to acknowledge the support from the Galileo Project in accommodating our requested test configuration, notably Eileen Theilig, Vickere Blackwell, Leif Harke, and Dave Allestad. And last but not least, our thanks to Steve Townes for generously agreeing to referee this article despite his schedule constraint and for many good writing suggestions.

References

- [1] D. Rogstad, "Suppressed Carrier Full-Spectrum Combining," *The Telecommunications and Data Acquisition Progress Report 42-107*, vol. July-September 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 12-20, November 15, 1991.
- [2] S. Dolinar and M. Belongie, "Enhanced Decoding for the Galileo S-Band Mission," *The Telecommunications and Data Acquisition Progress Report 42-114*, vol. April-June 1993, Jet Propulsion Laboratory, Pasadena, California, pp. 96-111, August 15, 1994.
- [3] A. Mileant and S. Hinedi, "Overview of Arraying Techniques in the Deep Space Network," *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October-December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 109-139, February 15, 1991.
- [4] B. Shah and S. Hinedi, "Performance of the Split-Symbol Moments SNR Estimator in the Presence of Inter-Symbol Interference," *The Telecommunications and Data Acquisition Progress Report 42-98*, vol. April-June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 157-173, August 15, 1989.

514-32

33280

p.19

TDA Progress Report 42-119

110878

November 15, 1994

358481

Computation of Reflected and Transmitted Horn Radiation Patterns for a Dichroic Plate

J. C. Chen

Ground Antennas and Facilities Engineering Section

A previous dichroic plate analysis has assumed that an ideal uniform plane wave illuminates the dichroic plate at a single angle of incidence. In fact, a horn radiates energy at the dichroic plate and illuminates it at many different angles. To model the horn and dichroic plate system, the horn pattern is represented as a group of plane waves traveling in different directions. The details of this analysis are presented in this article. The calculated and measured reflected radiation patterns show good agreement. The noise temperature predicted from the horn pattern model is shown to be more accurate than that from a simple plane-wave model.

I. Introduction

The dichroic plate under consideration is a metal plate perforated with identical apertures spaced on a periodic grid. The periodic characteristics of the dichroic plate simplify the analysis to the consideration of a single unit cell [1]. Computer codes have been developed to analyze the response of the dichroic plate assuming uniform plane-wave incidence (Fig. 1). In practice, energy radiated from a horn strikes the dichroic plate in many directions (Fig. 2). The transmission loss for energy incident at a nondesign angle is usually greater than that at the design angle. The horn pattern incidence analysis presented below will improve the understanding of dichroic plate performance for horn pattern incidence, allowing the computation of the transmitted and reflected patterns as well as the noise temperature contribution of the dichroic plate.

II. Analysis

The first step in modeling the horn and dichroic plate system is to consider the horn pattern as a group of plane waves traveling at different angles. Each plane wave has a certain amplitude and phase and strikes the plate at a different angle of incidence. The scattering matrix of the dichroic plate is calculated at the angle of incidence for each plane-wave component. Finally, the transmitted and reflected plane waves are summed up to form the transmitted and reflected radiation patterns.

The far-field horn pattern can be decomposed as a group of plane waves of different amplitudes and phases traveling at different directions. For example, the 26-dB Ka-band horn pattern at 32 GHz is considered to be a plane wave traveling at $\theta = 0$ deg and $\phi = 0$ deg, with an amplitude of 0 dB, plus a second plane wave traveling at $\theta = 1$ deg and $\phi = 0$ deg, with an amplitude of -0.137 dB, etc. (Fig. 3). Basically, the horn pattern is sampled as shown in Fig. 3. Each plane wave radiates from the horn in the

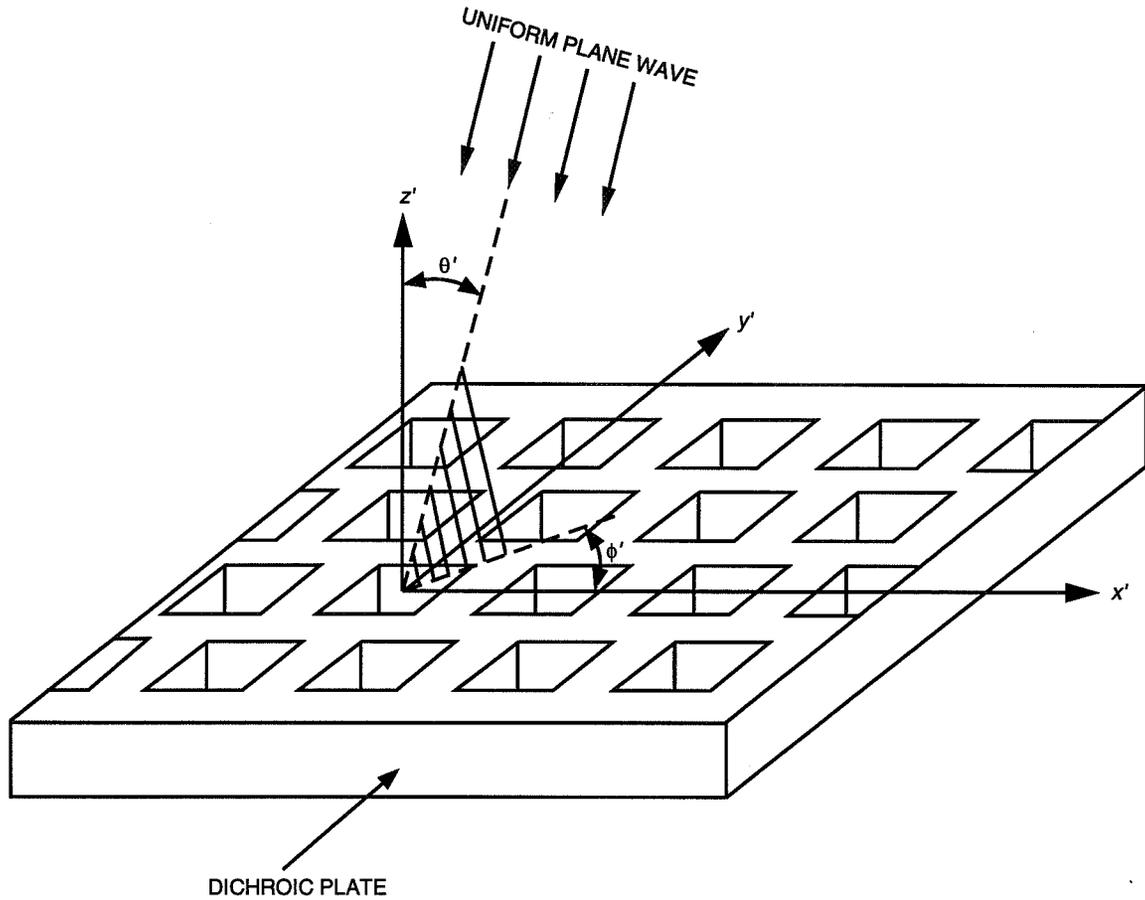


Fig. 1. The dichroic plate with uniform plane-wave incidence.

(θ, ϕ) direction with respect to the horn axis and strikes the dichroic plate at (θ', ϕ') with respect to the normal direction of the dichroic plate, as described in Appendix A.

The electric field $\vec{E}(\theta, \phi)$ radiated from a horn at angles (θ, ϕ) is given by

$$\vec{E}(\theta, \phi) = E_{\theta}(\theta) \sin \phi \hat{a}_{\theta} + E_{\phi}(\theta) \cos \phi \hat{a}_{\phi} \quad (1)$$

where $E_{\theta}(\theta)$ and $E_{\phi}(\theta)$ are the E - and H -plane patterns (amplitude and phase). The $E_{\theta}(\theta)$ and $E_{\phi}(\theta)$ can be computed theoretically or experimentally measured. The E -field in the dichroic plate analysis is represented by the TE and TM polarizations:

$$\vec{E}(\theta', \phi') = A_{TE} \hat{a}_{TE} + A_{TM} \hat{a}_{TM} \quad (2)$$

where θ' and ϕ' are the angles of incidence. To integrate the horn pattern with the dichroic plate analysis, coordinate transformations between the horn coordinates $(\hat{a}_r \hat{a}_{\theta} \hat{a}_{\phi})$ for spherical coordinates or $(\hat{a}_x \hat{a}_y \hat{a}_z)$ for Cartesian coordinates and the dichroic plate coordinates $(\hat{a}_{r'} \hat{a}_{\theta'} \hat{a}_{\phi'})$ or $(\hat{a}_{x'} \hat{a}_{y'} \hat{a}_{z'})$ are required and are presented in Appendix B.

The incident E -field [Eq. (1)] can be rewritten in a matrix form as

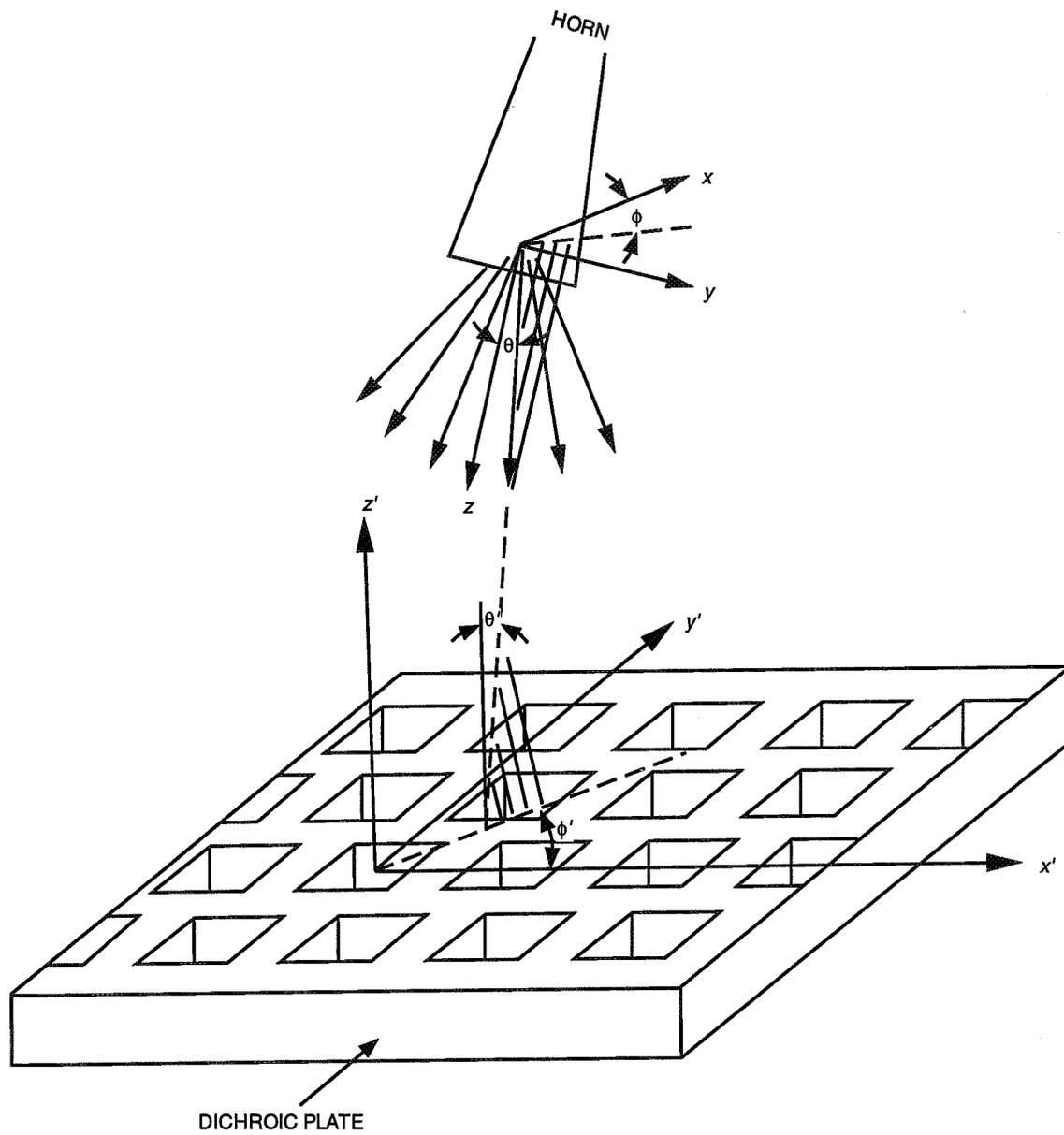


Fig. 2. The dichroic plate with horn pattern incidence.

$$\vec{E}^{inc} = [E^{inc}] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} \quad (3)$$

Applying the coordinate transformations from the horn spherical coordinate to the dichroic plate Cartesian coordinate, the E -field becomes

$$\vec{E}^{inc} = [E^{inc}] [P] [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (4)$$

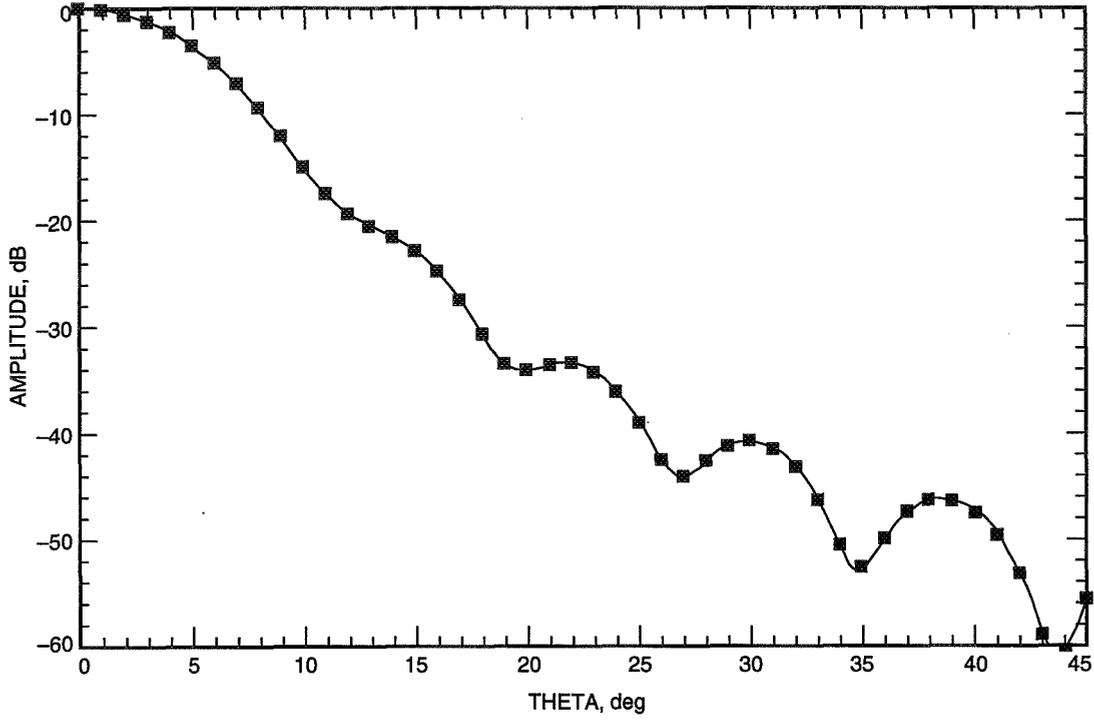


Fig. 3. The 26-dB horn pattern at 32.0 GHz sampled at 1-deg increments.

where $[P]$ is a 3×3 transformation matrix from spherical coordinates to Cartesian coordinates, and $[R_{\alpha\beta\gamma}]$ is a 3×3 Eulerian transformation matrix from the horn Cartesian coordinates to the dichroic plate Cartesian coordinates with Eulerian angles α , β , and γ (Appendix B) [2].

The E -field can also be represented by the TE and TM linear polarizations in the dichroic plate Cartesian coordinates:

$$\vec{E}^{inc} = [A_{TE} \quad A_{TM}] \begin{bmatrix} \hat{a}_{TE} \\ \hat{a}_{TM} \end{bmatrix} \quad (5)$$

where \hat{a}_{TE} and \hat{a}_{TM} are unit vectors of the TE and TM linear polarizations with angle of incidence θ' , ϕ' :

$$\hat{a}_{TE} = [-\sin \phi' \quad \cos \phi' \quad 0] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (6)$$

$$\hat{a}_{TM} = \frac{1}{\sqrt{1 + \tan^2 \theta'}} [-\cos \phi' \quad -\sin \phi' \quad \tan \theta'] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (7)$$

The A_{TE} and A_{TM} are the incident amplitude and phase on the TE and TM components of the field:

$$A_{TE} = \vec{E}^{inc} \cdot \hat{a}_{TE} \quad (8)$$

$$A_{TM} = \vec{E}^{inc} \cdot \hat{a}_{TM} \quad (9)$$

The incident E -field is changed from an E - and H -plane representation in the horn system to a TE and TM polarization representation in the dichroic plate system. Then the E -field is multiplied by the scattering matrices from the dichroic plate computer program:

$$\begin{bmatrix} B_{TE} \\ B_{TM} \end{bmatrix} = \begin{bmatrix} S_{TE,TE}^{21} & S_{TE,TM}^{21} \\ S_{TM,TE}^{21} & S_{TM,TM}^{21} \end{bmatrix} \begin{bmatrix} A_{TE} \\ A_{TM} \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} C_{TE} \\ C_{TM} \end{bmatrix} = \begin{bmatrix} S_{TE,TE}^{11} & S_{TE,TM}^{11} \\ S_{TM,TE}^{11} & S_{TM,TM}^{11} \end{bmatrix} \begin{bmatrix} A_{TE} \\ A_{TM} \end{bmatrix} \quad (11)$$

where B_{TE} and B_{TM} are the transmitted amplitude and phase, and C_{TE} and C_{TM} are the reflected amplitude and phase for the TE and TM polarizations. The $[S^{21}]$ and $[S^{11}]$ are 2×2 scattering matrices containing transmission coefficients and reflection coefficients of the dichroic plate, respectively. Therefore, the transmitted and the reflected E -fields are

$$\vec{E}^{tran} = [B_{TE} \quad B_{TM}] \begin{bmatrix} \hat{a}_{TE} \\ \hat{a}_{TM} \end{bmatrix} \quad (12)$$

$$\vec{E}^{refl} = [C_{TE} \quad C_{TM}] \begin{bmatrix} \hat{a}_{TE} \\ \hat{a}_{TM} \end{bmatrix} \quad (13)$$

Reverse coordinate transformations are applied to the transmitted and reflected E -field in the TE and TM mode representations to obtain the E -plane and H -plane representations. By summing the E -fields of the plane waves at each point of the horn pattern in the transmitted and reflected direction, the reflected and transmitted patterns through the dichroic plate are computed. The same analysis technique can also be applied for circularly polarized incidence. The procedure is discussed in detail in Appendix B.

III. Reflected Horn Patterns and Noise Temperature Prediction

The horn pattern analysis was tested using an X-/Ka-/KABLE (Ka-Band Link Experiment)-band dichroic plate that reflects X-band (8.4 to 8.6 GHz) and passes Ka-band [3]. The design priorities for this plate were (1) Ka-band downlink of 31.8 to 32.4 GHz, (2) Ka-band uplink of 34.2 to 34.7 GHz, and (3) KABLE frequencies of 33.6 to 33.7 GHz. The design assumed plane-wave incidence at an angle of $\theta' = 30$ deg and $\phi' = 0$ deg. The Ka-band horn is a 26-dB corrugated horn. Figure 4 shows the Ka-band horn pattern at 34.5 GHz. The x- and y-axes on the plot represent ϕ' and θ' , the angles of incidence on the plate. The curves are centered about $\phi' = 0$ deg and $\theta' = 30$ deg, where the central ray from the horn strikes the plate. The curves correspond to the $\theta = 0$ -, 5-, 10-, 15-, 20-, and 25-deg contours of the horn pattern, which have intensities of 1.0, 0.419, 0.0345, 0.0045, 0.0005, and 0.00005, respectively, at 34.5 GHz.

Each plane wave strikes the plate at a different angle, and the transmission coefficients decrease as the angle of incidence gets farther from the design angle. The transmission coefficient of the dichroic plate is multiplied by the corresponding intensity of the horn pattern (Fig. 5). Also included in the figure is a grating lobe curve for 34.5 GHz. Grating lobes are generated when the angle of incidence is equal to or greater than the grating lobe angles. The transmitted energy drops substantially when the angle of incidence is above the grating lobe curve.

The X-/Ka-/KABLE-band dichroic plate/Ka-band 26-dB horn combination was measured in the chamber at antenna range. The horn patterns were measured at $\phi = 0$ and 90 deg for the two orthogonal linear polarizations at 32.0 and 33.7 GHz. The calculated and measured reflected radiation

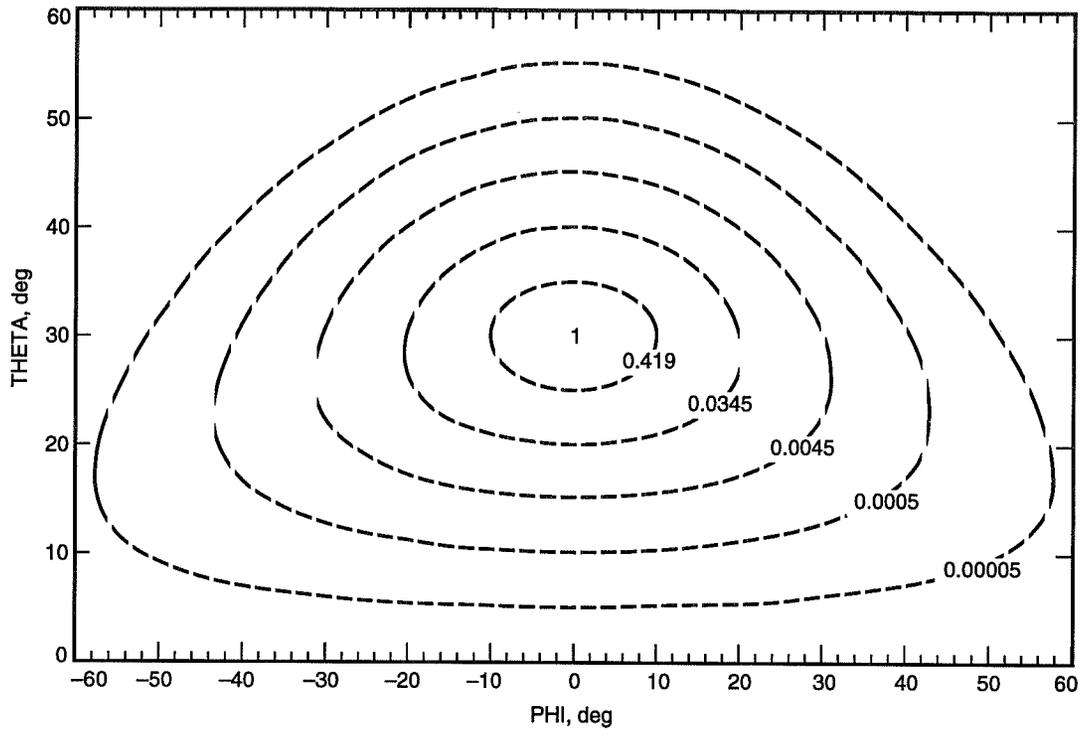


Fig. 4. Incident horn pattern at 34.5 GHz in the dichroic plate coordinates. The values shown are intensities.

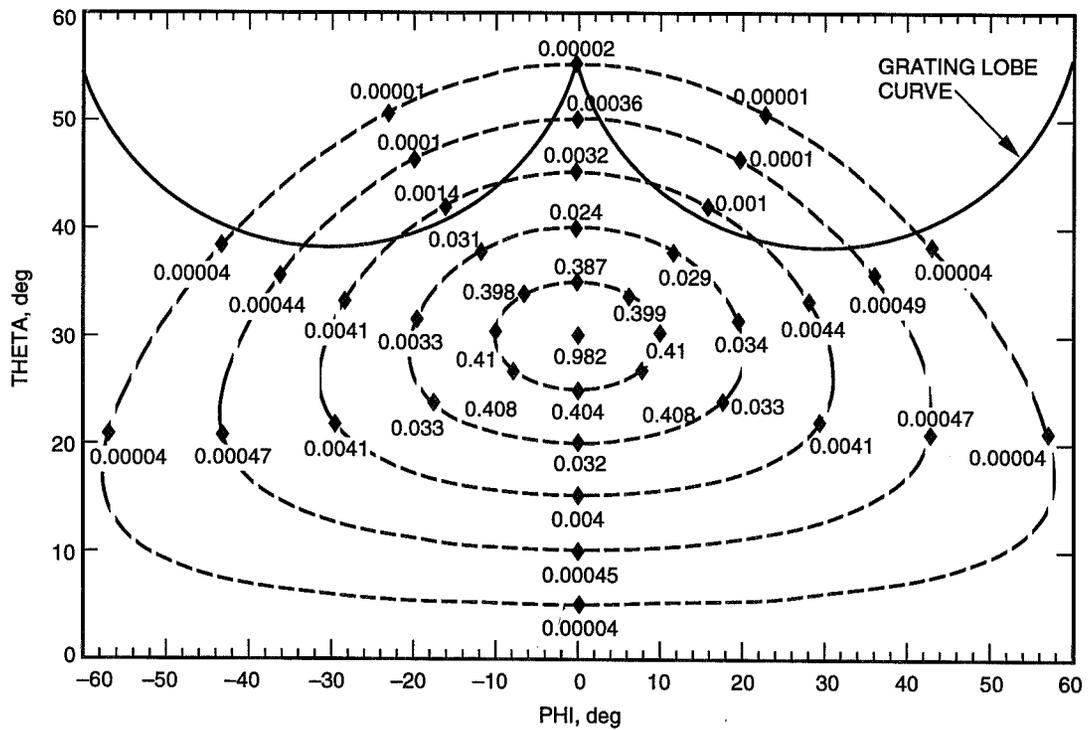


Fig. 5. Transmitted power through the dichroic plate with horn pattern incidence at 34.5 GHz. The values shown are intensities.

patterns are in good agreement (Figs. 6 through 13). The transmitted patterns were also calculated (Figs. 14 through 17).

The noise temperature was calculated for a pattern sampled at $\Delta\theta = 1$ deg up to 30 deg and for 32 ϕ -plane cuts. The noise temperature (NT) is the percentage of power lost multiplied by 300 K, which is the background temperature in the pedestal room at DSS 13:

$$NT = (1 - P^{tran}) 300 \text{ K} \quad (14)$$

where P^{tran} is the total transmitted power. The horn pattern model predicts an extra 0.74 K at 32.0 GHz over the plane-wave model (originally 1.34 K) and an extra 1.72 K at 33.7 GHz over the plane wave model (originally 7.57 K) (Table 1). The measured noise temperature is 10.5 K at the KABLE frequency band. The noise temperature at 33.7 GHz is higher than that at 32.0 GHz because the X-/Ka-/Kable-band dichroic plate was optimized for the Ka-band downlink. The loss at 34.5 GHz (uplink) increases from 0.087 to 0.121 dB. The loss increment calculated in dB at 34.5 GHz is even higher than that at 33.7 GHz, due to grating lobes at this frequency. The noise temperature includes the reflection loss, conductivity loss, and grating lobe loss, if any. The loss due to dichroic plate surface roughness is not included.

IV. Conclusion

The analysis of a dichroic plate including horn pattern incidence was shown to be a better model for a real system than the plane-wave model. The discrepancy between the calculated and measured noise temperature is reduced. The theoretical and experimental reflected horn patterns show good agreement. The analysis for the beam waveguide antenna system can now include the effect of the dichroic plate on the transmitted radiation pattern. The success of this technique leads one to believe that it is feasible to modify the local characteristic of the dichroic plate on a point-by-point basis to improve performance.

Table 1. Noise temperature^a contribution of the X-/Ka-/KABLE-band dichroic plate at the DSS-13 beam waveguide antenna.

Method	Frequency, GHz		
	32	33.7	34.5
Plane wave analysis, K	1.34	7.57	— (0.087 dB)
Horn pattern analysis, K	2.08	9.29	— (0.121 dB)
Measurement, K	— ^b	10.5 ^c	— ^d

^aThe noise temperature calculated includes reflection loss and conductivity loss. Surface roughness is not included.

^bTo be determined.

^cThe noise temperature was measured in the frequency range of 33.65 to 33.75 GHz.

^dNot applicable.

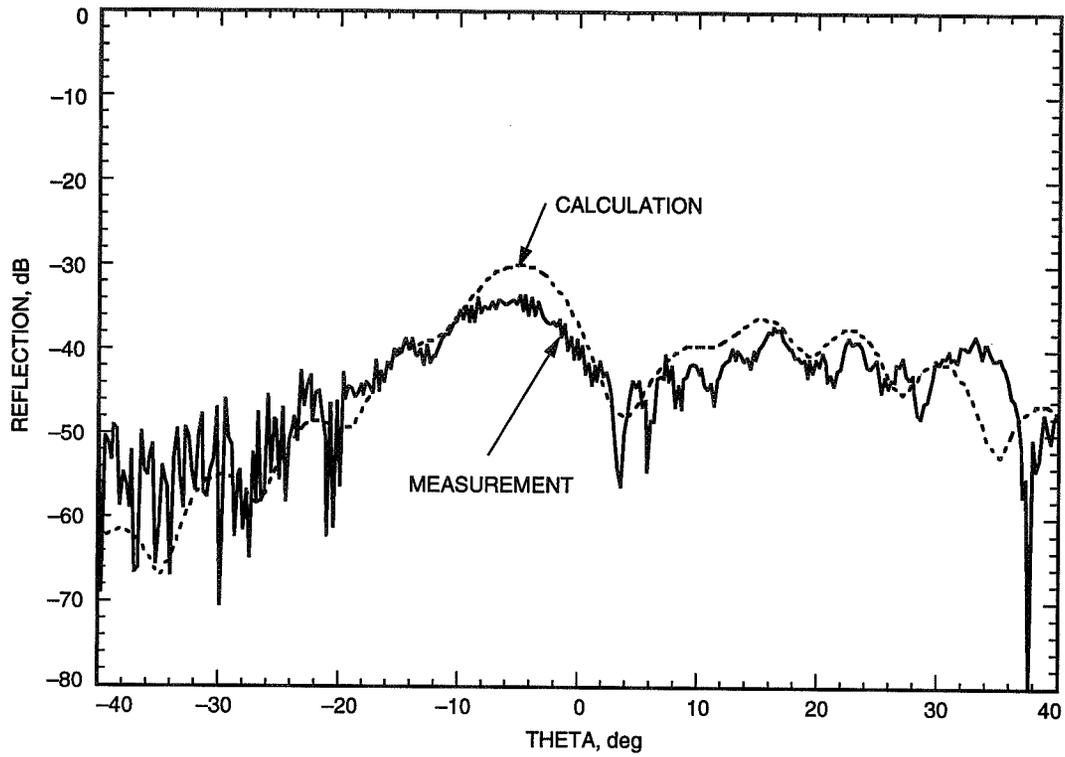


Fig. 6. Measured and calculated E_ϕ at $\phi = 0$ -deg cut at 32.0 GHz for linear polarization.

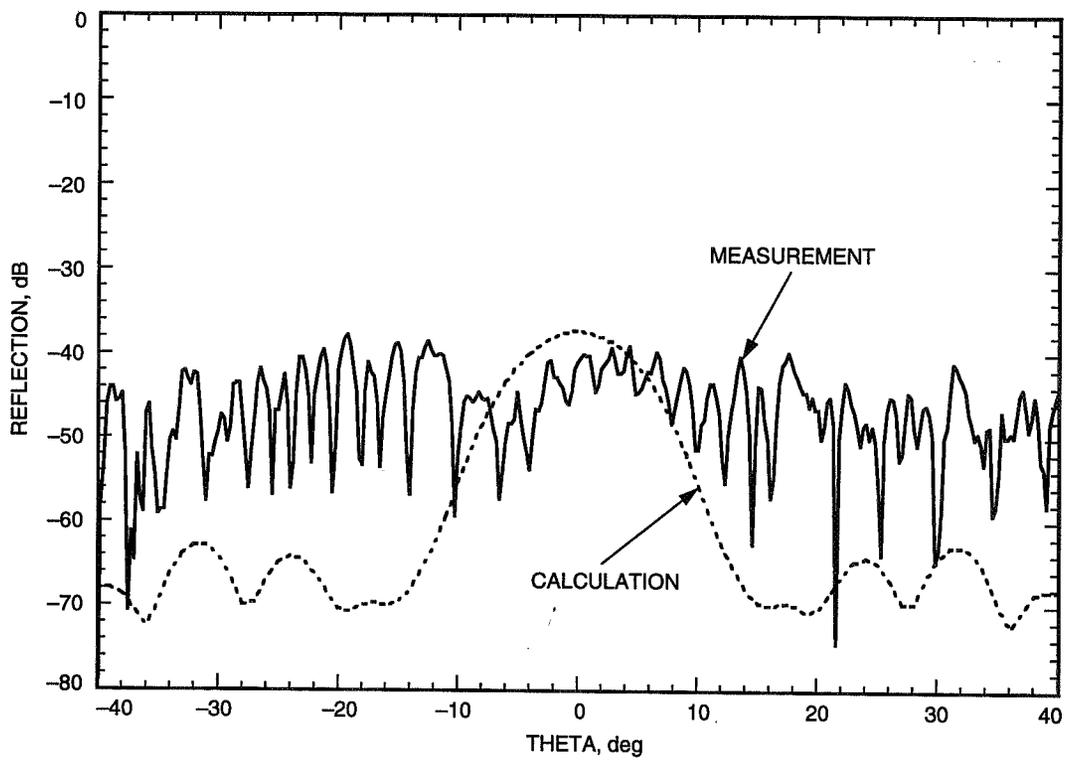


Fig. 7. Measured and calculated E_θ at $\phi = 90$ -deg cut at 32.0 GHz for linear polarization.

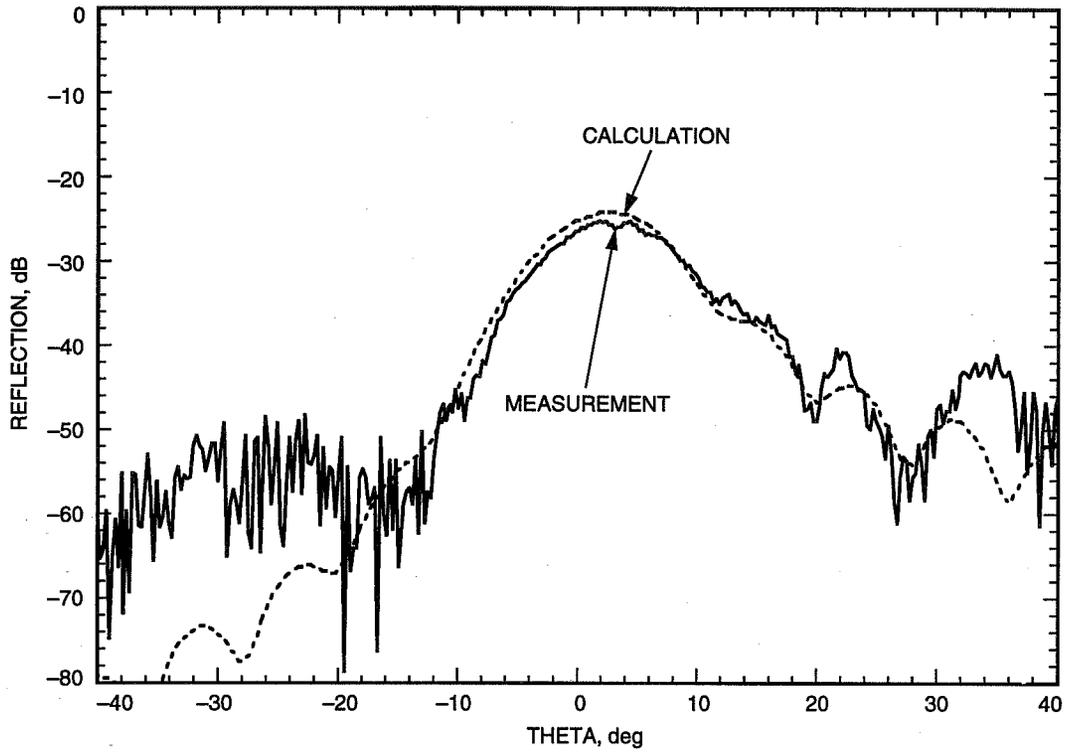


Fig. 8. Measured and calculated E_{θ} at $\phi = 0$ -deg at 32.0 GHz for orthogonal linear polarization.

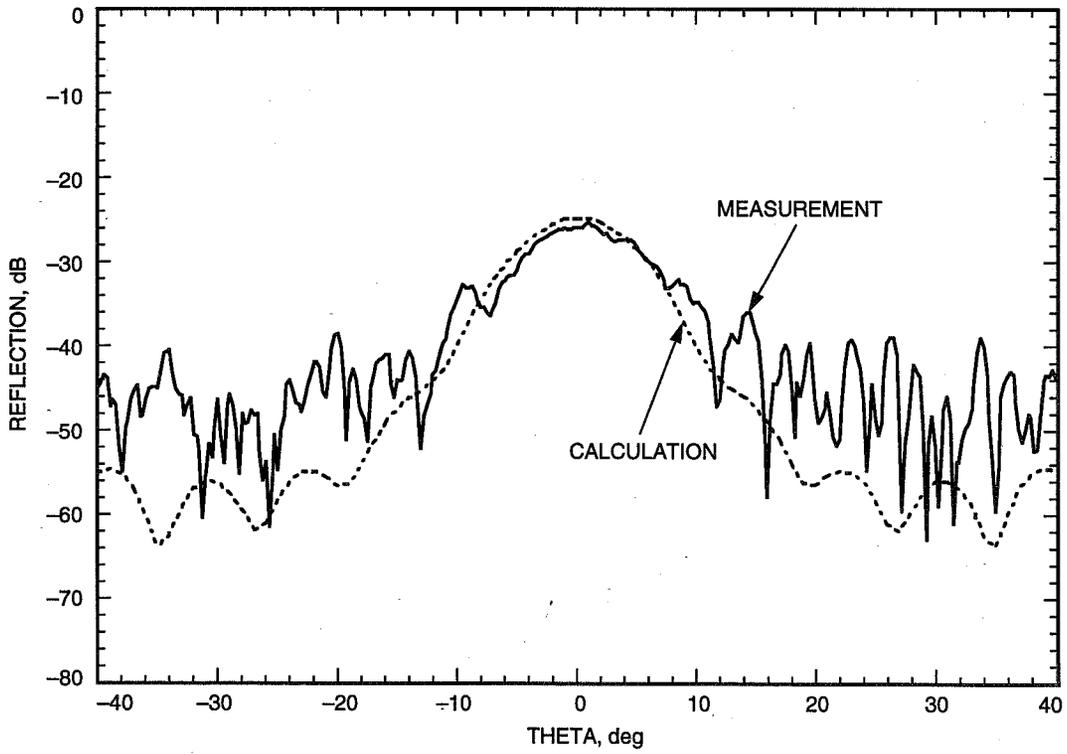


Fig. 9. Measured and calculated E_{θ} at $\phi = 90$ -deg at 32.0 GHz for orthogonal linear polarization.

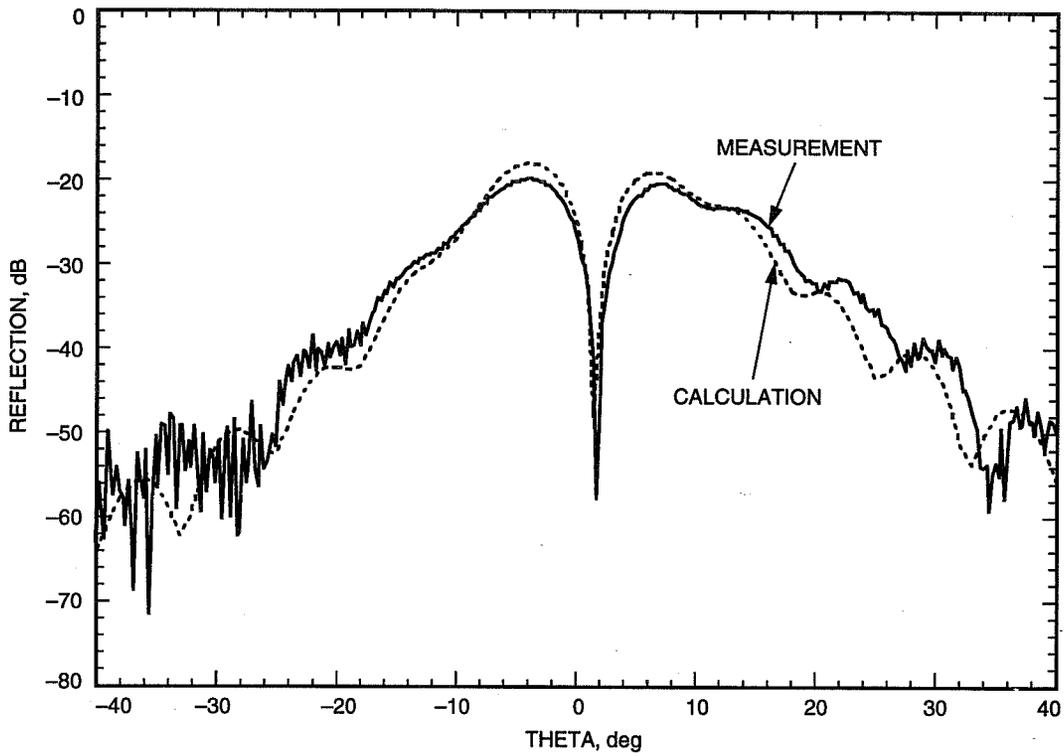


Fig. 10. Measured and calculated E_ϕ at $\phi = 0$ deg at 33.7 GHz for linear polarization.

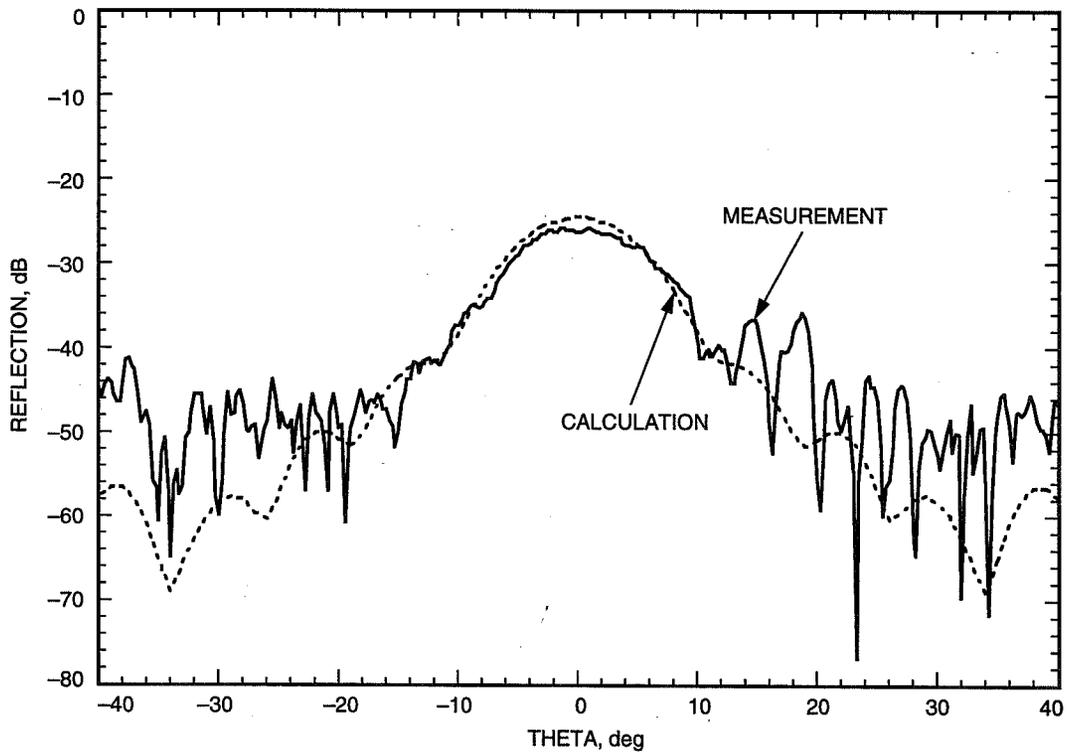


Fig. 11. Measured and calculated E_θ at $\phi = 90$ deg at 33.7 GHz for linear polarization.

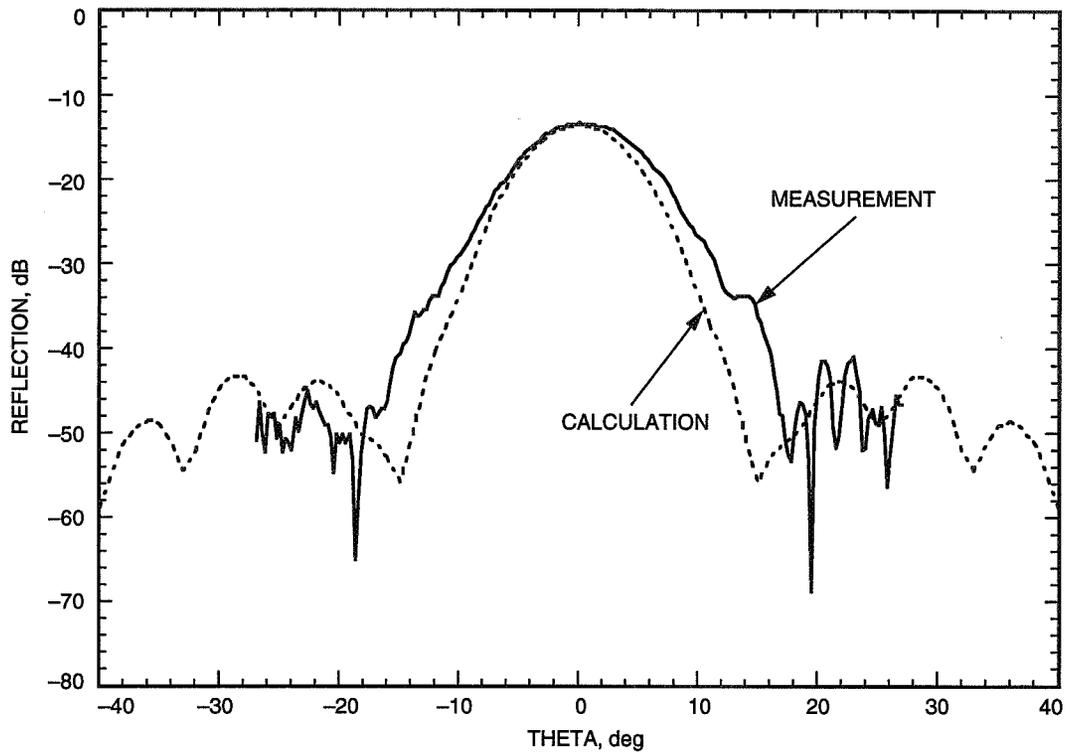


Fig. 12. Measured and calculated E_{θ} at $\phi = 0$ deg at 33.7 GHz for orthogonal linear polarization.

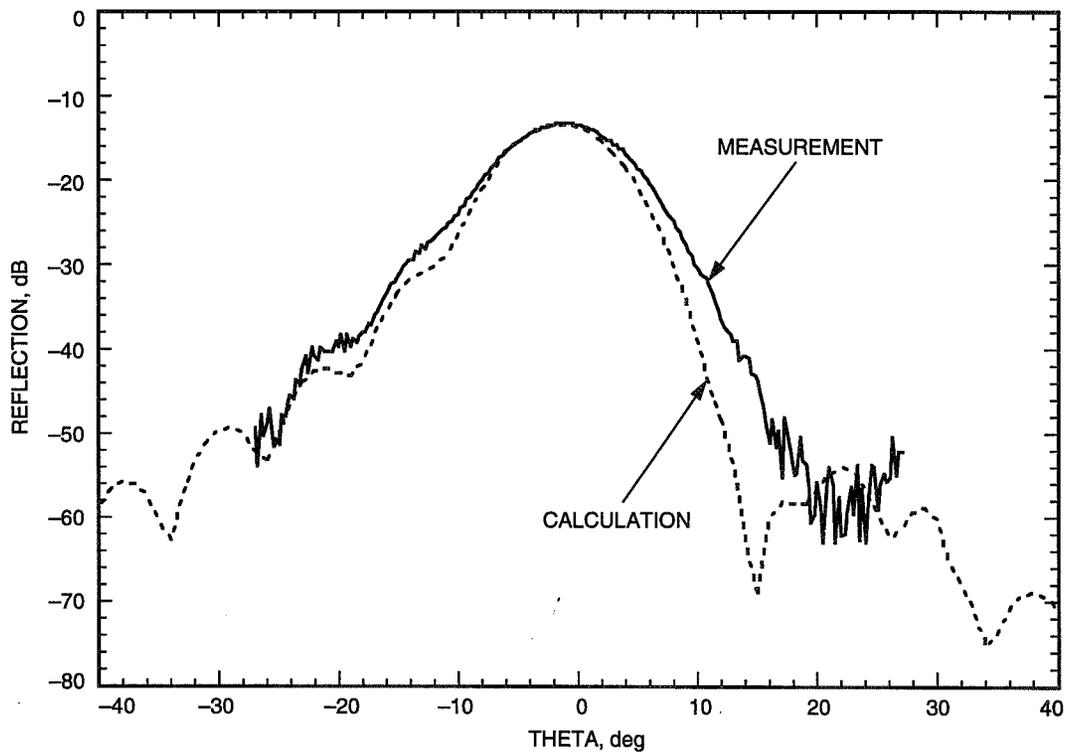


Fig. 13. Measured and calculated E_{ϕ} at $\phi = 90$ deg at 33.7 GHz for orthogonal linear polarization.

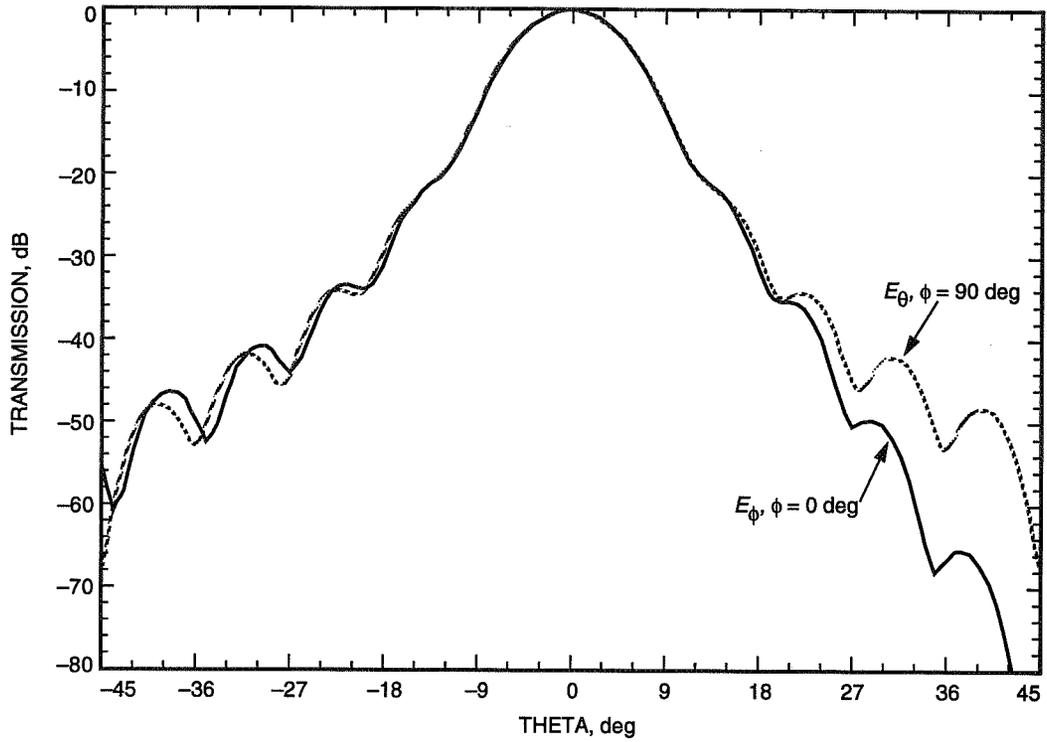


Fig. 14. Calculated transmitted pattern for the X-/Ka-/KABLE-band dichroic plate at 32.0 GHz for linear polarization.

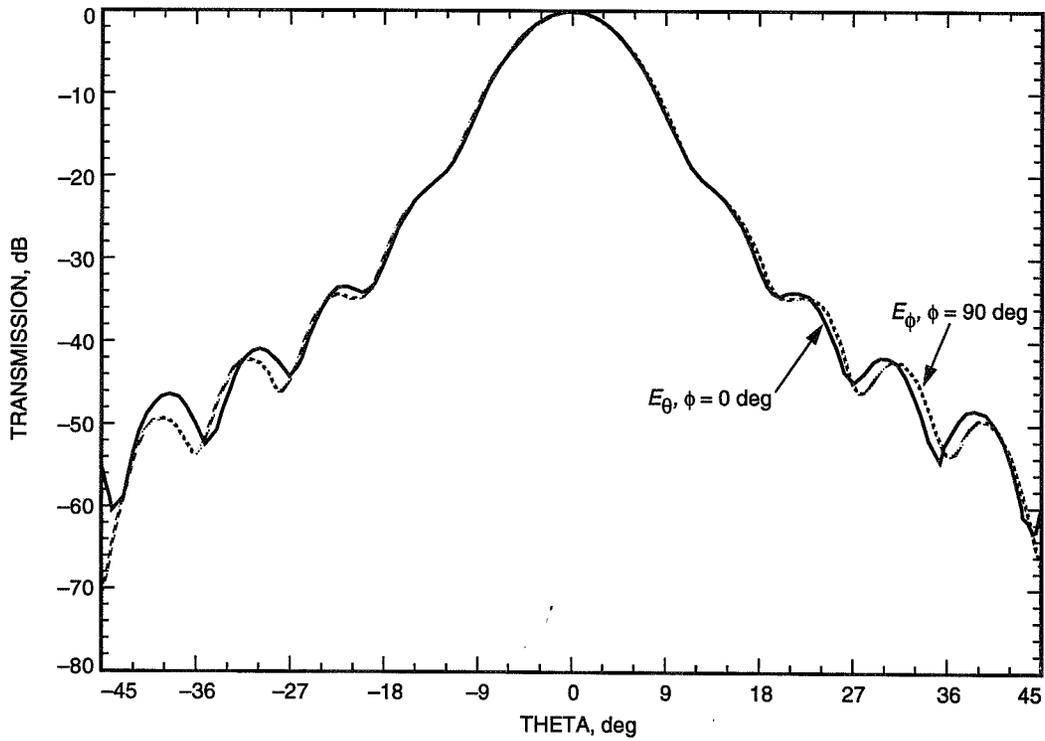


Fig. 15. Calculated transmitted pattern for the X-/Ka-/KABLE-band dichroic plate at 32.0 GHz for orthogonal linear polarization.

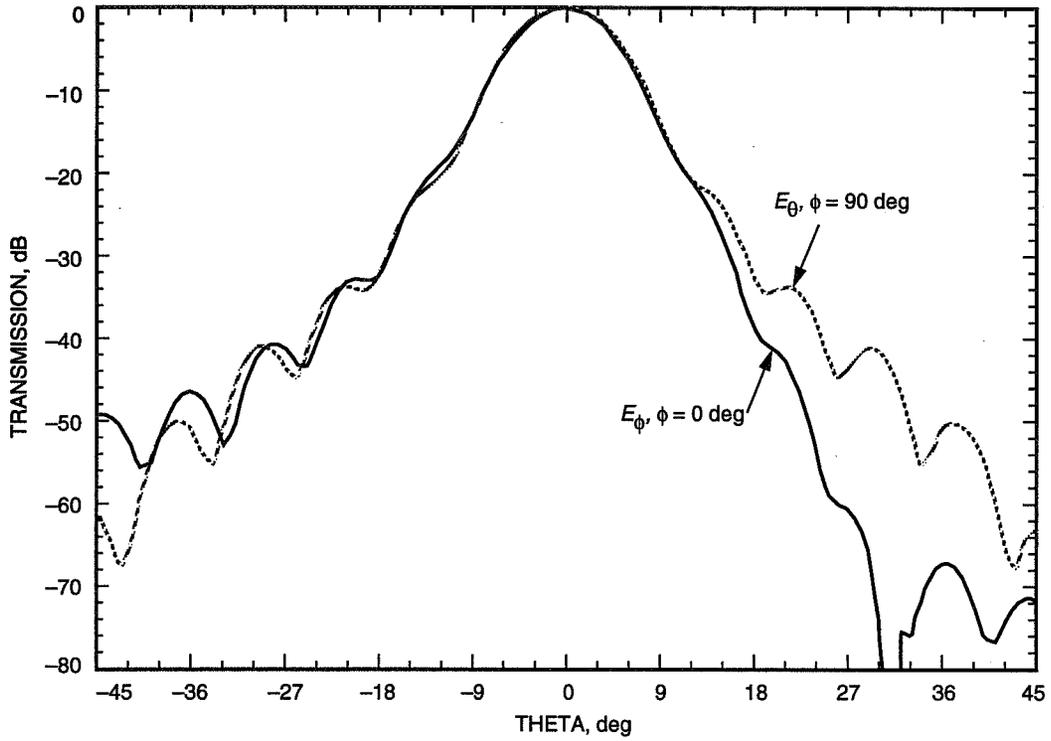


Fig. 16. Calculated transmitted pattern for the X-/Ka-/KABLE-band dichroic plate at 33.7 GHz for linear polarization.

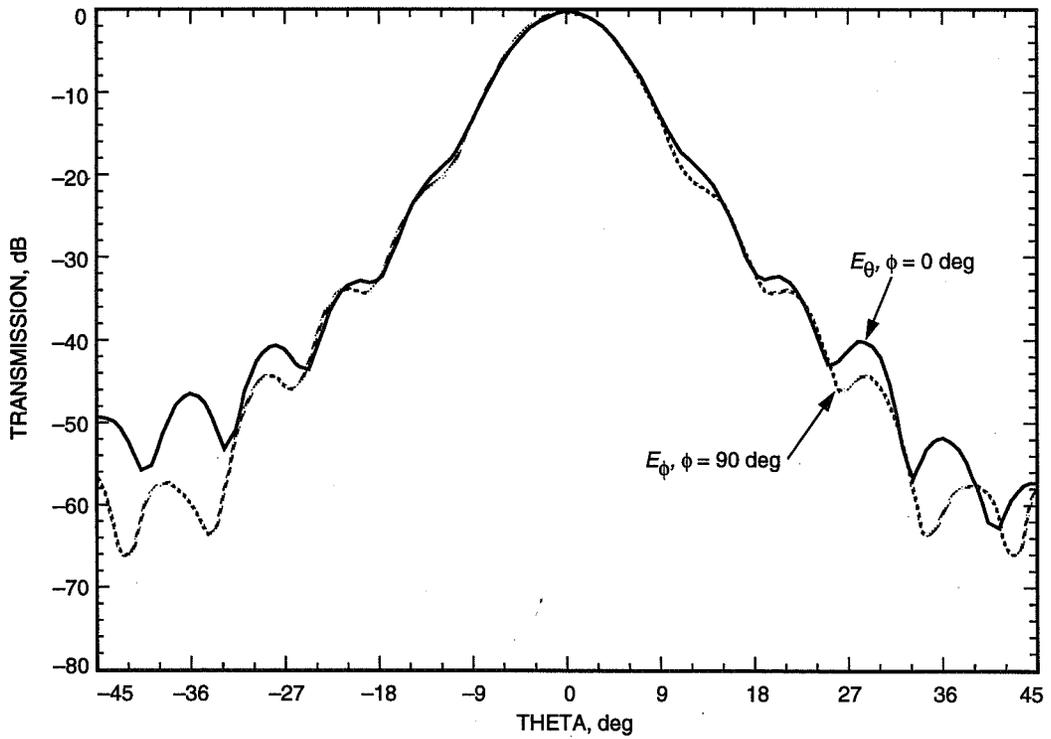


Fig. 17. Calculated transmitted pattern for the X-/Ka-/KABLE-band dichroic plate at 33.7 GHz for orthogonal linear polarization.

Acknowledgments

The author would like to acknowledge D. J. Hoppe for the technical discussions and J. Withington for the experimental patterns. The Cray supercomputer used in this investigation was provided by funding from the NASA Offices of Mission to Planet Earth, Aeronautics, and Space Science.

References

- [1] J. C. Chen, "Analysis of a Thick Dichroic Plate with Rectangular Holes at Arbitrary Angles of Incidence," *The Telecommunications and Data Acquisition Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 9–16, February 15, 1991.
- [2] J. B. Marion, *Classical Dynamics of Particles and Systems*, 2nd ed., West Philadelphia, Pennsylvania: Saunders College Publications, pp. 384–385.
- [3] J. C. Chen, "X-/Ka-Band Dichroic Plate Design and Grating Lobe Study," *The Telecommunications and Data Acquisition Progress Report 42-105*, vol. January–March 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 21–30, May 15, 1991.

Appendix A

Computation of Angles of Incidence on a Dichroic Plate for the Horn Radiation Pattern

A plane wave traveling in the direction (θ, ϕ) in the horn coordinate system will strike the dichroic plate at an incident angle (θ', ϕ') in the dichroic plate coordinate system.

Suppose the energy radiated from the horn is in the direction \vec{r} :

$$\vec{r} = [1 \quad 0 \quad 0] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} \quad (\text{A-1})$$

We may transform the spherical coordinate system to the Cartesian coordinate system using

$$\begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} = [P] \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} \quad (\text{A-2})$$

where $[P]$ is the following 3×3 matrix:

$$[P] = \begin{bmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ \cos \theta \cos \phi & \cos \theta \sin \phi & -\sin \phi \\ -\sin \phi & \cos \phi & 0 \end{bmatrix} \quad (\text{A-3})$$

Transforming the horn coordinate system $(x \ y \ z)$ to the dichroic plate coordinate system $(x' \ y' \ z')$ using Eulerian angles $\alpha, \beta,$ and $\gamma,$ we have

$$\begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} = [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{A-4})$$

where $[R_{\alpha\beta\gamma}]$ is the following 3×3 matrix:

$$[R_{\alpha\beta\gamma}] = \begin{bmatrix} \cos \gamma \cos \alpha - \cos \beta \sin \alpha \sin \gamma & \cos \gamma \sin \alpha + \cos \beta \cos \alpha \sin \gamma & \sin \gamma \sin \beta \\ -\sin \gamma \cos \alpha - \cos \beta \sin \alpha \cos \gamma & -\sin \gamma \sin \alpha + \cos \beta \cos \alpha \cos \gamma & \cos \gamma \sin \beta \\ \sin \beta \sin \alpha & -\sin \beta \cos \alpha & \cos \beta \end{bmatrix} \quad (\text{A-5})$$

Therefore,

$$\vec{r} = [1 \ 0 \ 0] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} \quad (\text{A-6})$$

$$= [1 \ 0 \ 0] [P] \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} \quad (\text{A-7})$$

$$= [1 \ 0 \ 0] [P] [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{A-8})$$

$$= [x' \ y' \ z'] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{A-9})$$

where

$$x'^2 + y'^2 + z'^2 = 1 \quad (\text{A-10})$$

The \vec{r} vector can also be written in the dichroic plate system:

$$\vec{r} = [1 \ 0 \ 0] \begin{bmatrix} \hat{a}_{r'} \\ \hat{a}_{\theta'} \\ \hat{a}_{\phi'} \end{bmatrix} \quad (\text{A-11})$$

$$= [1 \ 0 \ 0] [P_{\theta', \phi'}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{A-12})$$

$$= [\sin \theta' \sin \phi' \quad \sin \theta' \cos \phi' \quad \cos \theta'] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{A-13})$$

Comparing Eqs. (A-8), (A-9), and (A-13), we have

$$[\sin \theta' \sin \phi' \quad \sin \theta' \cos \phi' \quad \cos \theta'] = [1 \quad 0 \quad 0] [P] [R_{\alpha\beta\gamma}] = [x' \quad y' \quad z'] \quad (\text{A-14})$$

The wave is traveling in the $-z'$ direction; therefore, the incident angle is the angle between the normal of the dichroic plate, $\vec{z}' = (0 \ 0 \ 1)$ and $-\vec{r} = (-x \ -y \ -z)$. The incident angles on the dichroic plate (θ', ϕ') are then

$$\theta' = \arctan \left(-\frac{\sqrt{x'^2 + y'^2}}{z'} \right) \quad (\text{A-15})$$

$$\phi' = \arctan \left(\frac{y'}{x'} \right) \quad (\text{A-16})$$

Appendix B

Computation of Reflected and Transmitted Electric Fields for a Dichroic Plate

I. The E -Field Represented by the E - and H -Plane Polarizations

The electric field $\vec{E}(\theta, \phi)$ radiated by the horn is traveling in the r direction with angles (θ, ϕ) :

$$\vec{E}(\theta, \phi) = E_\theta(\theta) \sin \phi \hat{a}_\theta + E_\phi(\theta) \cos \phi \hat{a}_\phi \quad (\text{B-1})$$

where $E_\theta(\theta)$ and $E_\phi(\theta)$ are the E - and H -plane patterns (amplitude and phase). The incident electrical field can be rewritten in matrix form:

$$\vec{E}^{inc}(\theta, \phi) = [0 \quad E_\theta^{inc}(\theta) \sin \phi \quad E_\phi^{inc}(\theta) \cos \phi] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} = [E^{inc}] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} \quad (\text{B-2})$$

where $[E^{inc}]$ is a 1×3 matrix.

The spherical coordinates $(r \ \theta \ \phi)$ are transformed to Cartesian coordinates $(x \ y \ z)$ in the horn coordinate system using

$$\begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} = [P] \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} \quad (\text{B-3})$$

where $[P]$ is a 3×3 matrix given in Eq. (A-3). The E -field can be expressed as

$$\vec{E}^{inc} = [E^{inc}] [P] \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} \quad (\text{B-4})$$

The transformation from the dichroic plate coordinate system $(x' y' z')$ to the horn coordinate system $(x y z)$ using Eulerian angles α, β , and γ is shown below:

$$\begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} = [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{B-5})$$

where $[R_{\alpha\beta\gamma}]$ is a 3×3 matrix given in Eq. (A-5). Therefore, the E -field is

$$\vec{E}^{inc} = [E^{inc}] \begin{bmatrix} \hat{a}_r \\ \hat{a}_\theta \\ \hat{a}_\phi \end{bmatrix} \quad (\text{B-6})$$

$$= [E^{inc}] [P] \begin{bmatrix} \hat{a}_x \\ \hat{a}_y \\ \hat{a}_z \end{bmatrix} \quad (\text{B-7})$$

$$= [E^{inc}] [P] [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{B-8})$$

II. The E -Field Represented by the TE and TM Modes

The E -field can also be represented by TE and TM components in the dichroic plate coordinates $(x' y' z')$:

$$\vec{E}^{inc} = [A_{TE} \quad A_{TM}] \begin{bmatrix} \hat{a}_{TE} \\ \hat{a}_{TM} \end{bmatrix} \quad (\text{B-9})$$

where \hat{a}_{TE} and \hat{a}_{TM} are unit vectors of the TE and TM linear polarizations at angles of incidence θ', ϕ' :

$$\hat{a}_{TE} = [-\sin \phi' \quad \cos \phi' \quad 0] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} = [\Psi_{TE}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} \quad (\text{B-10})$$

$$\hat{\mathbf{a}}_{TM} = \frac{1}{\sqrt{1 + \tan^2 \theta'}} \begin{bmatrix} -\cos \phi' & -\sin \phi' & \tan \theta' \end{bmatrix} \begin{bmatrix} \hat{\mathbf{a}}_{x'} \\ \hat{\mathbf{a}}_{y'} \\ \hat{\mathbf{a}}_{z'} \end{bmatrix} = [\Psi_{TM}] \begin{bmatrix} \hat{\mathbf{a}}_{x'} \\ \hat{\mathbf{a}}_{y'} \\ \hat{\mathbf{a}}_{z'} \end{bmatrix} \quad (\text{B-11})$$

Note that

$$|\hat{\mathbf{a}}_{TE}| = 1 \quad (\text{B-12})$$

$$|\hat{\mathbf{a}}_{TM}| = 1 \quad (\text{B-13})$$

$$\hat{\mathbf{a}}_{TE} \cdot \hat{\mathbf{a}}_{TM} = 0 \quad (\text{B-14})$$

$$\hat{\mathbf{a}}_{TE} \cdot \hat{\mathbf{a}}_{r'} = 0 \quad (\text{B-15})$$

$$\hat{\mathbf{a}}_{TM} \cdot \hat{\mathbf{a}}_{r'} = 0 \quad (\text{B-16})$$

The A_{TE} and A_{TM} are the incident amplitude and phase in the directions of the TE and TM polarizations, respectively.

$$A_{TE} = \vec{E}^{inc} \cdot \hat{\mathbf{a}}_{TE} \quad (\text{B-17})$$

$$A_{TM} = \vec{E}^{inc} \cdot \hat{\mathbf{a}}_{TM} \quad (\text{B-18})$$

By Eqs. (B-8), (B-10), and (B-11), we have

$$A_{TE} = [E^{inc}] [P] [R_{\alpha\beta\gamma}] [\Psi_{TE}]^T \quad (\text{B-19})$$

$$A_{TM} = [E^{inc}] [P] [R_{\alpha\beta\gamma}] [\Psi_{TM}]^T \quad (\text{B-20})$$

where $[\Psi_{TE}]^T$ and $[\Psi_{TM}]^T$ are the transpose matrices of $[\Psi_{TE}]$ and $[\Psi_{TM}]$, respectively.

III. The Dichroic Plate Scattering Matrix

After the wave strikes the dichroic plate, the transmitted and reflected E -fields are given by

$$\vec{E}^{tran} = [B_{TE} \quad B_{TM}] \begin{bmatrix} \hat{\mathbf{a}}_{TE} \\ \hat{\mathbf{a}}_{TM} \end{bmatrix}$$

$$\vec{E}^{refl} = [C_{TE} \quad C_{TM}] \begin{bmatrix} \hat{\mathbf{a}}_{TE} \\ \hat{\mathbf{a}}_{TM} \end{bmatrix}$$

where B_{TE} and B_{TM} are the TE and TM components for the transmitted E -field, and C_{TE} and C_{TM} are the TE and TM components for the reflected E -field and are given by

$$\begin{bmatrix} B_{TE} \\ B_{TM} \end{bmatrix} = \begin{bmatrix} S_{TE,TE}^{21} & S_{TE,TM}^{21} \\ S_{TM,TE}^{21} & S_{TM,TM}^{21} \end{bmatrix} \begin{bmatrix} A_{TE} \\ A_{TM} \end{bmatrix} \quad (\text{B-21})$$

$$\begin{bmatrix} C_{TE} \\ C_{TM} \end{bmatrix} = \begin{bmatrix} S_{TE,TE}^{11} & S_{TE,TM}^{11} \\ S_{TM,TE}^{11} & S_{TM,TM}^{11} \end{bmatrix} \begin{bmatrix} A_{TE} \\ A_{TM} \end{bmatrix} \quad (\text{B-22})$$

where $[S^{21}]$ and $[S^{11}]$ are 2×2 dichroic plate scattering matrices containing transmission coefficients and reflection coefficients, respectively. The scattering matrix of the dichroic plate is calculated by a computer program using a plane-wave incidence model.

IV. Transforming Back to the E - and H -Plane Patterns

Applying the same coordinate transformation to $[E^{tran}]$, we have

$$\vec{E}^{tran} = [E^{tran}] [P] [R_{\alpha\beta\gamma}] \begin{bmatrix} \hat{a}_{x'} \\ \hat{a}_{y'} \\ \hat{a}_{z'} \end{bmatrix} = [B_{TE} \quad B_{TM}] \begin{bmatrix} \hat{a}_{TE} \\ \hat{a}_{TM} \end{bmatrix} \quad (\text{B-23})$$

or

$$[E^{tran}] [P] [R_{\alpha\beta\gamma}] = \begin{bmatrix} -B_{TE} \sin \phi' - B_{TM} \frac{\cos \phi'}{\sqrt{1+\tan^2 \theta'}} & B_{TE} \cos \phi' - B_{TM} \frac{\sin \phi'}{\sqrt{1+\tan^2 \theta'}} & B_{TM} \frac{\tan \theta'}{\sqrt{1+\tan^2 \theta'}} \end{bmatrix} \quad (\text{B-24})$$

$$[E^{tran}] = \begin{bmatrix} -B_{TE} \sin \phi' - B_{TM} \frac{\cos \phi'}{\sqrt{1+\tan^2 \theta'}} & B_{TE} \cos \phi' - B_{TM} \frac{\sin \phi'}{\sqrt{1+\tan^2 \theta'}} & B_{TM} \frac{\tan \theta'}{\sqrt{1+\tan^2 \theta'}} \end{bmatrix} [R_{\alpha\beta\gamma}]^T [P]^T \quad (\text{B-25})$$

Likewise, the reflected pattern is given by

$$[E^{refl}] = \begin{bmatrix} -C_{TE} \sin \phi' - C_{TM} \frac{\cos \phi'}{\sqrt{1+\tan^2 \theta'}} & C_{TE} \cos \phi' - C_{TM} \frac{\sin \phi'}{\sqrt{1+\tan^2 \theta'}} & C_{TM} \frac{\tan \theta'}{\sqrt{1+\tan^2 \theta'}} \end{bmatrix} [R_{\alpha\beta\gamma}]^T [P]^T \quad (\text{B-26})$$

V. Orthogonal Linear Polarization and Circular Polarization

The same method can apply to the orthogonal linear polarization and the circular polarization as well.

For the other linear polarization, the E -field is

$$\vec{E}(\theta, \phi) = E_{\theta}(\theta) \cos \phi \hat{a}_{\theta} + E_{\phi}(\theta) \sin \phi \hat{a}_{\phi} \quad (\text{B-27})$$

and for circular polarization,

$$\vec{E}(\theta, \phi) = \frac{1}{\sqrt{2}} E_{\theta}(\theta) (\cos \phi + j \sin \phi) \hat{a}_{\theta} + \frac{1}{\sqrt{2}} E_{\phi}(\theta) (\sin \phi + j \cos \phi) \hat{a}_{\phi} \quad (\text{B-28})$$

110879

33281 p. 9

358483

X-/Ka-Band Dichroic Plate Noise Temperature Reduction

W. Veruttipong and P. Lee

Ground Antennas and Facilities Engineering Section

The X-/Ka-band (8.4-GHz/32.0-GHz) dichroic plate installed at DSS 13 contributes an estimated 3 K to the system noise temperature at 32.0 GHz. Approximately 1 percent of the Ka-band incident field is reflected by the plate into the 300-K environment of the DSS-13 pedestal room. A low-cost, easily implemented method of reducing the noise temperature is presented. Using a curved reflector, the reflected field can be refocused into an 80-K cold load, reducing the noise temperature contribution of the dichroic plate by about 2 K.

I. Introduction

The X-/Ka-band (8.4-GHz/32.0-GHz) dichroic plate installed at the DSS-13 beam waveguide antenna contributes an estimated 3 K to the system noise temperature at 32.0 GHz. This amount is significant in a low-noise system where the estimated total noise temperature is only 27 K at an elevation of 90 deg and 35 K at an elevation of 30 deg. Approximately 1 percent of the Ka-band incident field is reflected by the plate into the 300-K environment of the DSS-13 pedestal room (Fig. 1). The reflected field is large since the current dichroic plate design software at JPL is based on an incident plane wave, while the actual incident field is either a spherical wave or a quadratic phase front, depending on whether M6 in Fig. 1 is a flat plate (spherical wave) or a curved mirror (quadratic phase front).

Several options have been suggested to reduce the dichroic plate noise temperature. One option is to redesign the plate to have a very small reflected field by using plane wave spectrum analysis to represent an incident spherical wave or quadratic phase front by the sum of plane waves incident at all angles, including imaginary angles. The plate would have a different hole size and shape for each incident angle. To design and analyze a dichroic plate using this technique is a very difficult problem, with uncertain results. A dichroic plate designed by this method may not have any better characteristics than the current design, because of approximations in the computation, fabrication tolerances, and alignment. Another method that may reduce the reflected field is to reposition the plate at the beam waist of the Ka-band field after it is reflected off a curved mirror at M6. At the beam waist, the phase distribution is fairly flat and the field resembles a plane wave. Since the dichroic plate is designed for a plane wave, the reflected field is expected to be smaller. This option would involve redesigning the layout of the X- and Ka-band feed systems, and could not be extended to similar dual-feed systems where M6 is not a curved mirror, such as the X-/Ka-band system at DSS 24 and the S-/X-band (2.3 GHz for S-band) systems at DSS 13 and DSS 24.

The option presented in this article is to reduce the effective temperature that the reflected field sees. The reflected field can be refocused with a curved mirror and directed to a small cold load (Fig. 2).

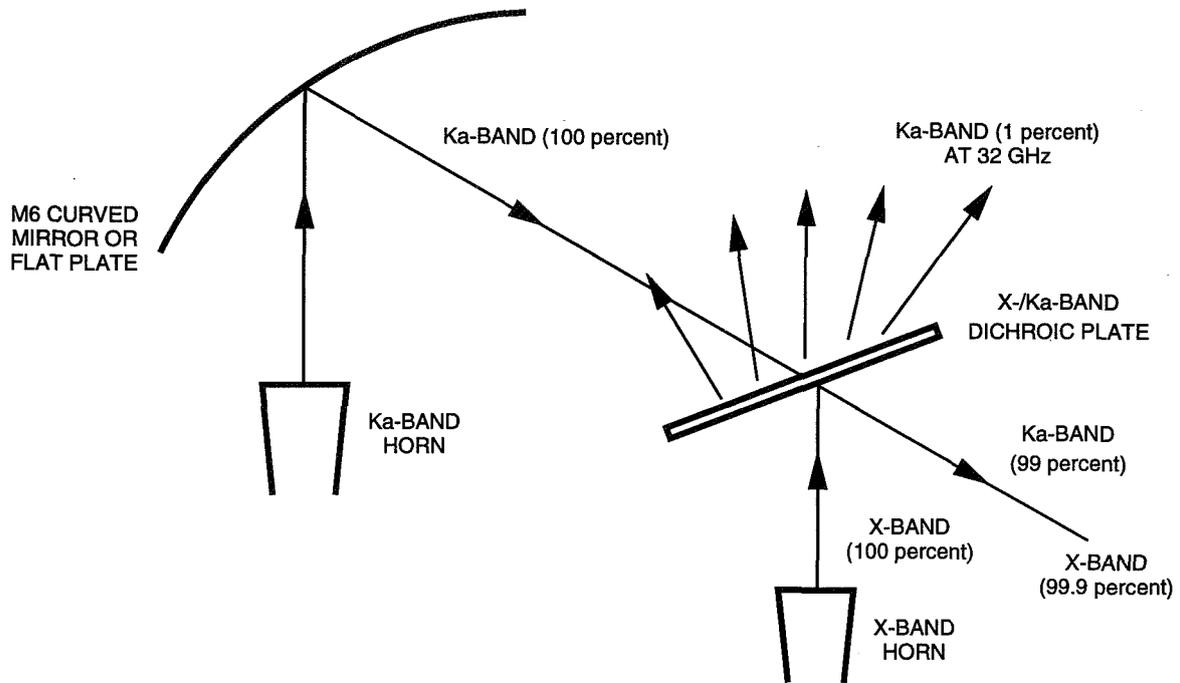


Fig. 1. X-/Ka-band feed layout with the dichroic plate.

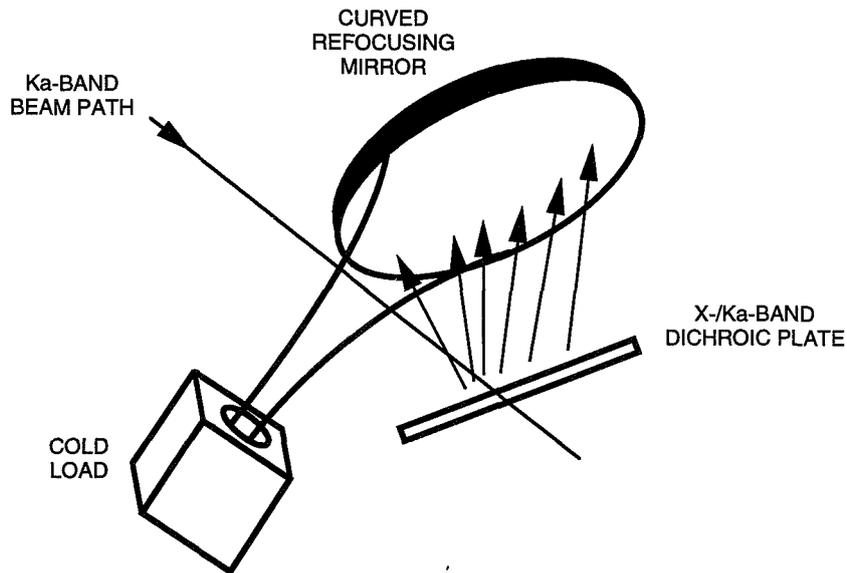


Fig. 2. Refocusing of the reflected field into a cold load.

This alternative is much less risky and less expensive than attempting to redesign the dichroic plate for incident spherical waves or quadratic phase fronts. It does not involve redesigning the feed system layout, as would the option of repositioning the dichroic plate at the beam waist. The concept can be extended to any dichroic plate feed system, whether or not M6 is a curved mirror. Even if the reflected field from the dichroic plate were reduced, this method could be used to further improve the noise temperature.

II. Measurements of the Scattered Field

To be able to focus most of the scattered field with a curved reflector, the amplitude distribution of the field must be confined to a reasonably sized envelope, and the phase distribution must be well behaved. Measurements using a 26-dBi Ka-band horn at 33.7 GHz with the X-/Ka-band dichroic plate were taken using the configuration shown in Fig. 3. The amplitude was found to be approximately confined within a 22.5-dBi envelope, and the phase was well behaved, as shown in Fig. 4. For this study, 33.7 GHz was used instead of 32.0 GHz because DSS 13 was configured to operate at 33.7 GHz and not at 32.0 GHz. The scattered patterns at 32.0 GHz and 33.7 GHz are similar, and the results are expected to be close for both cases. The measurement setup differs from the actual configuration at DSS 13. There is no curved mirror in the measurement setup. The distance from the horn phase center to the dichroic plate in the measurement setup is shorter than the distance from the beam waist of the field reflected off M6 to the dichroic plate in the DSS 13 setup. These differences are assumed to have little impact in the overall amplitude and phase distributions of the reflected field.

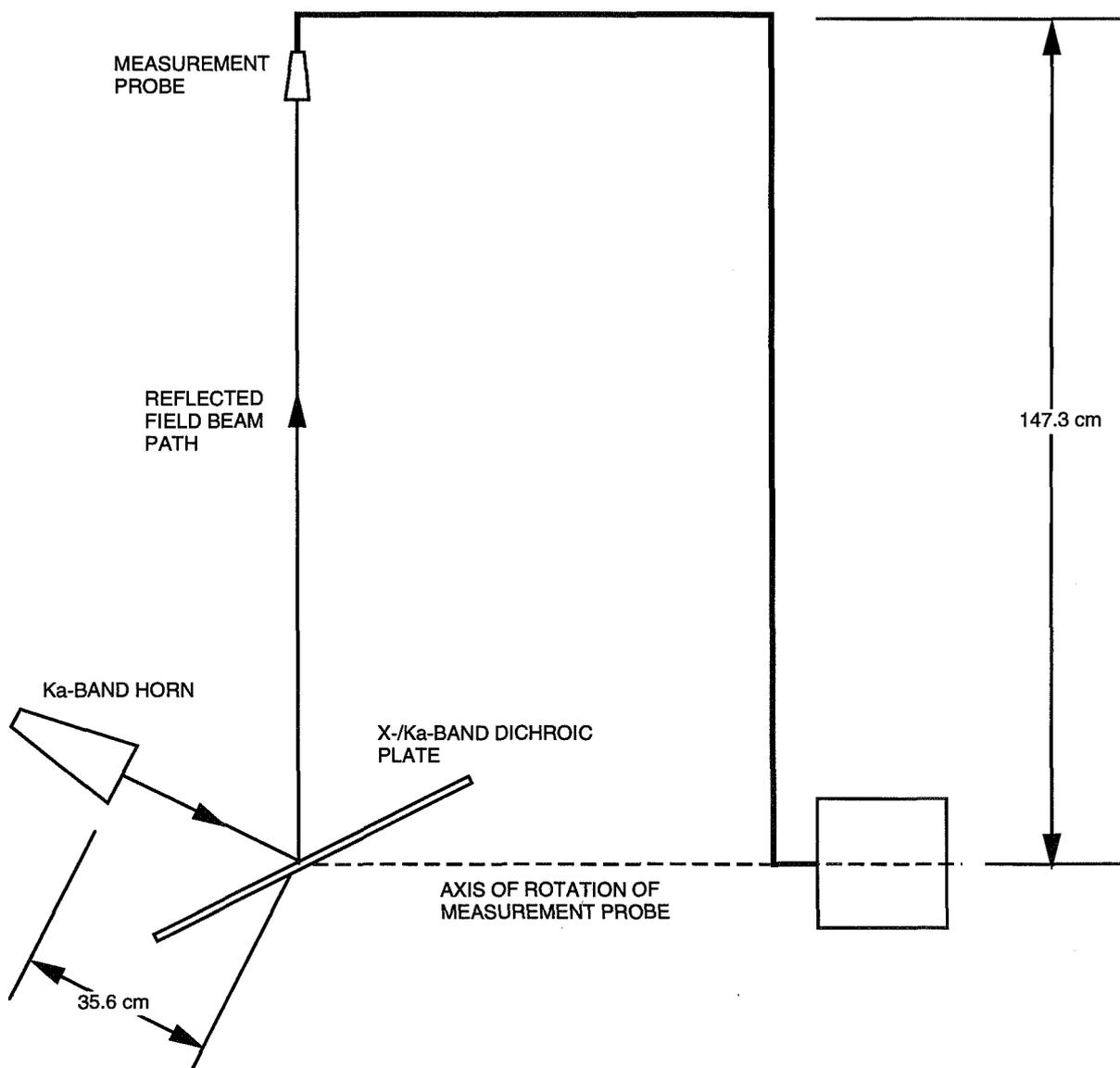


Fig. 3. Measurement configuration.

To use JPL physical optics (PO) programs¹ for performance analysis of the ellipsoidal reflector, a spherical wave expansion (SWE) of the measured data was utilized [1,2]. An SWE is a modal representation of the data, using vector spherical wave functions. With an SWE, the radiation patterns in the far field and near field can be calculated accurately, even though the original pattern was measured at only one distance. An SWE with 16 azimuthal modes and 160 polar modes was found to be sufficient to represent the measured data. Using the SWE representation of the field with a phase center calculation program, the phase center location of the radiation pattern can be found.

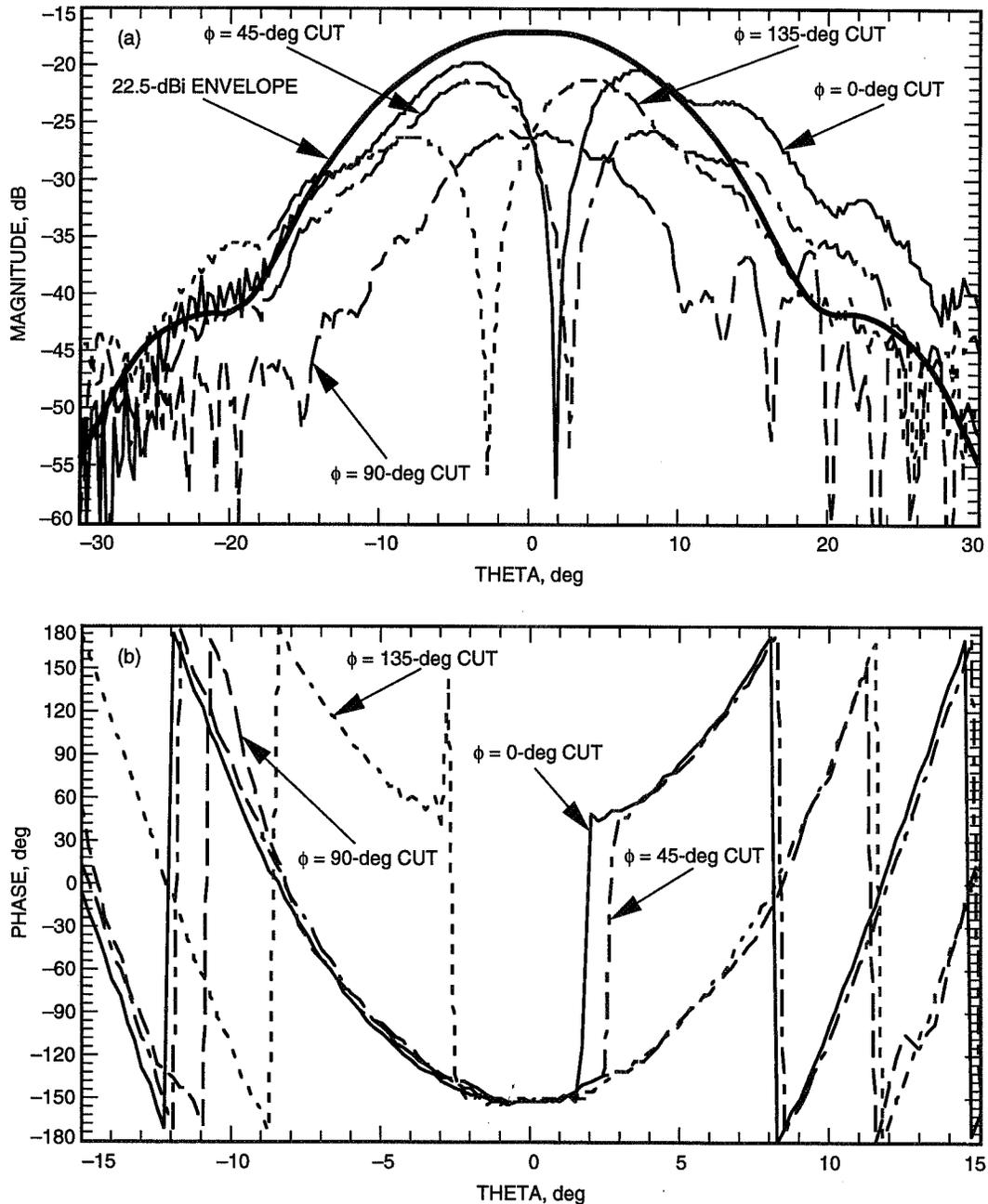


Fig. 4. Reflected field measured patterns of (a) amplitude and (b) phase.

¹ R. E. Hodges and W. A. Imbriale, "Computer Program POMESH for Diffraction Analysis of Reflector Antennas" (internal document), Jet Propulsion Laboratory, Pasadena, California, February 1992.

III. Ellipsoidal Reflector Design

The size of the ellipsoidal reflector depends on two factors. First, the reflector must be large enough to intercept most of the field scattered off the dichroic plate. Any part of the field that spills past the reflector will see the pedestal room temperature of 300 K. Second, the reflector must fit in the space behind the dichroic plate. Figure 5 shows the space available behind the dichroic plate for mounting an ellipsoidal reflector.

The curvature of the ellipsoidal reflector is determined by the desired positions of the two focal points of the ellipsoid and the location of the center of the reflector itself. The location of the center of the reflector was decided from the available space behind the dichroic plate. One focal point of the ellipsoid is placed at the location of the image of the phase center of the incident field. The other focal point of the ellipsoidal reflector depends upon the position and aperture size of the cold load. The cold load must be located where there is enough room for it, and where no other parts of the feed system will interfere in the radiation beam between the reflector and the load. To keep the cold load stable and the cost low, the aperture diameter of the load should be kept small. Placing the cold load at the beam waist of the focused field minimizes the aperture diameter. The ellipsoid geometry is given in Fig. 6. F1 is the approximate location of the phase center of the scattered field, and F2 is the approximate position of the cold load aperture.

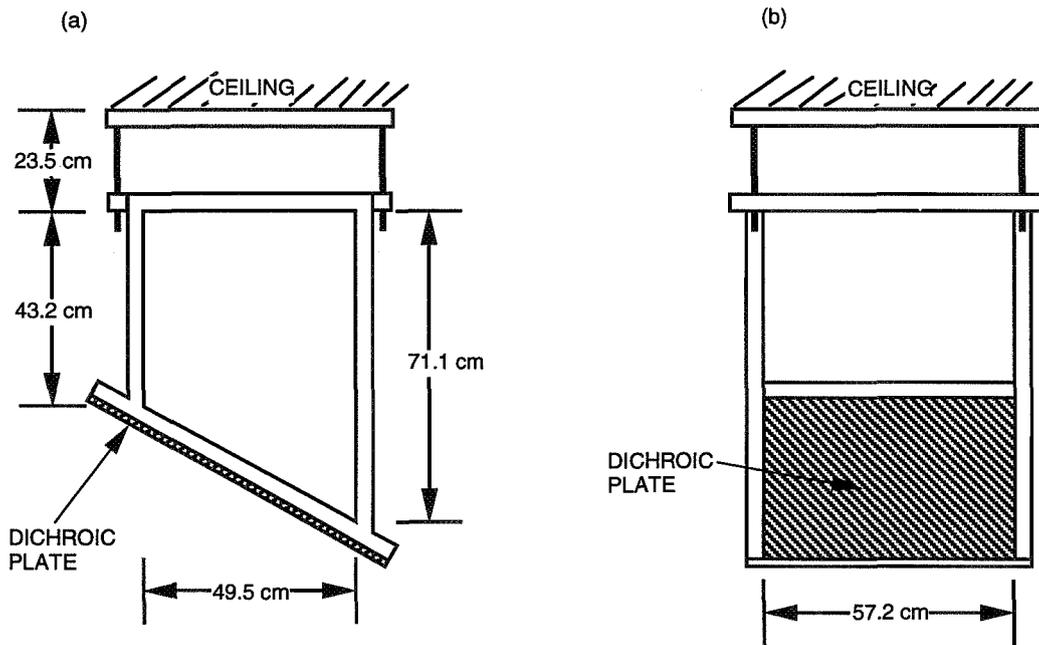


Fig. 5. Dichroic plate mounting frame: (a) side view and (b) back view.

IV. Cold Load Design

Using PO analysis computer programs, it was calculated that a cold load with a 12.7-cm aperture diameter located 106.7 cm away from the center of the reflector would intercept about 98 percent of the total refocused power. A cone mounted to the front of the cold load could direct more of the power to the load.

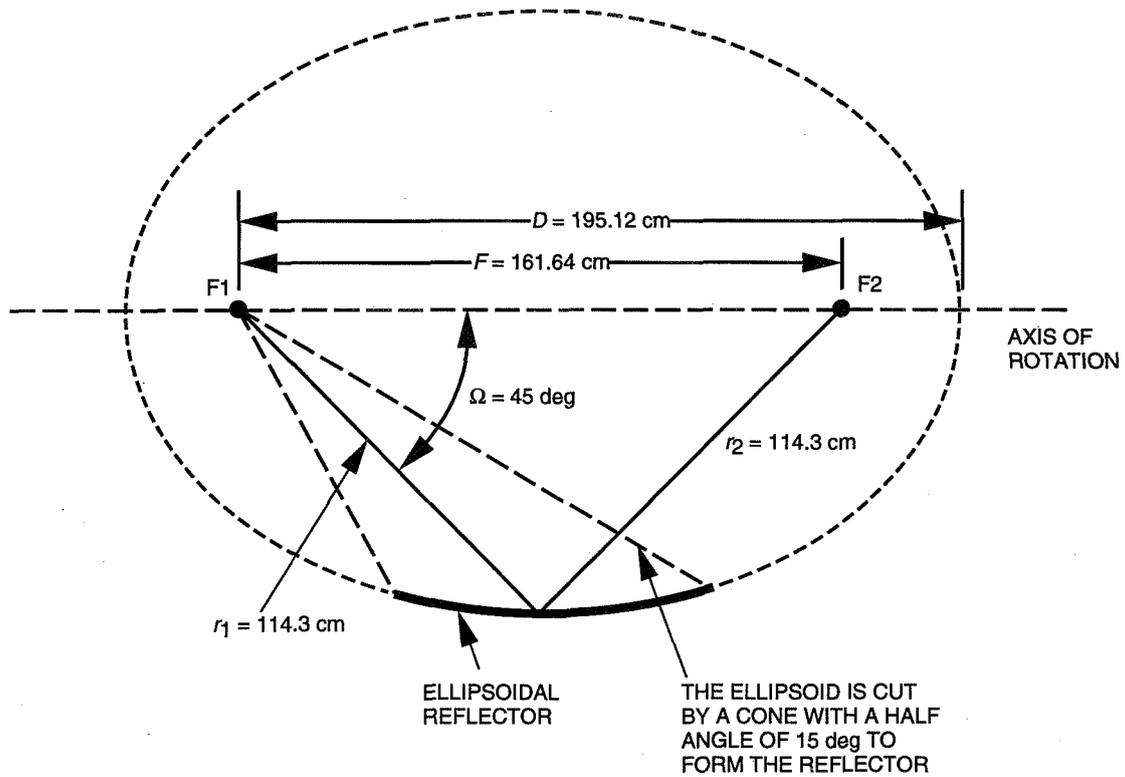


Fig. 6. Ellipsoidal reflector geometry.

Table 1. Calculation of the noise temperature contribution.

Calculation of power			
Source of power	Distribution of power		
Incident on dichroic plate	1% reflected		
Reflected by dichroic plate	95% refocused by reflector	5% not refocused by reflector	
Refocused by reflector	98% directed to cold load	2% not directed to cold load	
Total power, %	$(0.01)(0.95)(0.98) = 0.931$	$(0.01)(0.95)(0.02) = 0.019$	$(0.01)(0.05) = 0.05$
Calculation of total noise temperature			
Effective temperature, K	80 (cold load)	300 (pedestal room)	300 (pedestal room)
Noise temperature contribution	$(0.931\%)(80 \text{ K}) = 0.74 \text{ K}$	$(0.019\%)(300 \text{ K}) = 0.06 \text{ K}$	$(0.05\%)(300 \text{ K}) = 0.15 \text{ K}$
Total noise temperature, K	$0.74 + 0.06 + 0.15 = 0.95$		

V. Reduction in Noise Temperature

The current noise contribution from the dichroic plate has been estimated at 3 K at 32.0 GHz. Since the pedestal room temperature is 300 K, about 1 percent of the field incident on the dichroic is scattered. Calculations using PO analysis programs predict that 95 percent of the reflected field would be refocused by the reflector, and 98 percent of the refocused field would be directed into the cold load. As shown in Table 1, this results in a total noise temperature contribution of 1 K from the dichroic plate reflected field, for a reduction of 2 K.

VI. Conclusion

The noise temperature contribution of the X-/Ka-band dichroic plate at DSS 13 could be reduced from an estimated 3 K to about 1 K by refocusing the scattered field from the dichroic plate into a cold load. The concept is a low-cost, low-risk alternative to redesigning the dichroic plate to have a smaller reflected field, and is more flexible than repositioning the dichroic plate to the beam waist.

Acknowledgments

The authors thank J. Withington for the measurement data of the field reflected from the dichroic plate, and M. Esquivel for reviewing this article. The Cray supercomputer used in this investigation was provided by funding from the NASA Offices of Mission to Planet Earth, Aeronautics, and Space Science.

References

- [1] A. C. Ludwig, "Near-Field Far-Field Transformations Using Spherical-Wave Expansions," *IEEE Transactions on Antennas and Propagation*, vol. AP-19, no. 2, pp. 214-220, March 1971.
- [2] A. C. Ludwig, *Calculation of Scattered Patterns From Asymmetrical Reflectors*, JPL Technical Report 32-1430, Jet Propulsion Laboratory, Pasadena, California, February 15, 1970.

5/6-32

33282

110880

P-6
TDA Progress Report 42-119

November 15, 1994

358484

A Model for Analysis of TDA Budget Allocation

D. S. Remer

DSN Advanced Planning Office

and

Harvey Mudd College of Engineering and Science

H. R. Buchanan

Radio Frequency and Microwave Subsystems Section

There is an ever-increasing need to achieve greater efficiency in the operation of the Deep Space Network (DSN), i.e., increased productivity at reduced cost. One of the tools used in the course of a planning workshop (March 1994) on this subject was a methodology for budget allocation applicable to long-range planning. This article presents a model for analysis of the TDA budget allocation. For the 1994 through 1999 period, the percentage of the total TDA budget allocated to capacity and capability is being cut in half, whereas the percentage spent on efficiency of delivery will be increasing.

I. Introduction

This is the beginning of an effort to collect data for the DSN reengineering effort planned for the near future. This article describes the methodology used in partitioning Work Authorization Document (WAD) data into categories that are potentially useful for making future planning decisions. These data, covering fiscal year (FY) 1993 through FY 1999, were presented to management at the 1994 TDA & Code OT Planning Workshop held on March 15 to 17, 1994, in Arcadia, California.

The goal of this article is to see what the trends are and where (i.e., in what categories) we are spending our money. The 1993 data are derived from the actual *WAD Obligations Performance Report* for FY 1993, in which the allocated budget was about \$200M.

In the case of 1994 through 1999, data were taken from a WAD 93-3 (TDA, Level 5) computer printout dated February 22, 1994.

II. Definition of Categories

The first step was to define the categories for the WAD partitioning that we would use for the study. The categories follow; expanded definitions are presented in the Appendix.

- (1) Capacity and Capability: modifications and additions to the network to meet the customer's requirements, such as new antennas, new uplink/downlink frequencies, and increased gain/temperature (G/T).

- (2) **Efficiency of Delivery of Capacity and Capability:** the process of efficiently delivering the product through the operation and maintenance of the network, including sustaining functions.
- (3) **Technology and System Studies:** investment for the future in research and development, including feasibility demonstrations.
- (4) **Management:** strategic planning, customer interface, the allocation of resources, and establishing methods and standards to optimize the network.
- (5) **Reserves:** funds held in reserve in the WAD and not allocated to specific tasks.

III. Budget Allocation for the Years 1993 through 1999

WAD data for FY 1993 through FY 1999 were analyzed and categorized. The details were summarized by fiscal year for each budget category. Specific details are available from the authors.

In the case of FY 1993, the data source used was the 1993 *WAD (TDA) Obligations Performance Report* (Level 5) dated September 25, 1993.

The data source for FY 1994 through FY 1999 was the printed version of WAD 93-3 (TDA, Level 5) dated February 22, 1994. The programs included are listed in Table 1.

Table 1. Fiscal year 1994 through 1999 programs.

Title	NASA code no.
DSN Advanced Systems	310xx - XXXXXXX
Cost Reimbursables—NASA	31408 - XXXXXXX
DSN Systems	31430 - XXXXXXX
DSN Operations	31440 - XXXXXXX
Communications Systems	31510 - XXXXXXX
Communications Operations	31520 - XXXXXXX
Data Processing Systems	31550 - XXXXXXX
DSN Planning	31590 - XXXXXXX
Systems Development	31591 - XXXXXXX
System Engineering and Support	31630 - XXXXXXX

IV. Methodology

The WAD data for each task were analyzed and assigned to the proper budget category at as high a level as seemed appropriate. The source included data down to Level 5, which gives the titles of each separate account. This allows fine-grain identification of the category for each individual task.

Data for each fiscal year were processed on a separate spreadsheet for each budget category as defined above. The percentage of the grand total for each category was also calculated. The percentage gives a more meaningful indication of the annual trend since the total planned budget is generally increasing in

the period under study. Use of the spreadsheet offers a structured approach to tabulating and processing the data.

Let us discuss two examples of how we treated the data in doing the WAD partitioning for this article. The first example is in the FY 1994 budget. NASA code 31430-4101 covers *Program Management—Office, Financial and Administrative* at \$4.4M. When we looked at level-5 detail data (31430-410104, not shown in this article), we noted that it included “FY95 Forward Funding” at \$3.5M. We allocated this amount to Reserves for FY 1994. The remaining \$0.9M in 31430-4101 was allocated to Management for FY 1994. As a second example in the FY 1994 budget, NASA code 31430-60 covers *New Developments—Major Tasks—Multi Mission 1st User* at \$2.3M. Looking at the level-5 detail, we concluded that all the 31430-60 funds (\$2.3M) belong in the Capacity and Capability category for FY 1994.

V. Results

Table 2 summarizes the subtotals and totals of the WAD data as a percentage of the fiscal year grand total. A plot of the categories (in percent) versus fiscal year is given in Fig. 1. Figure 2 shows the FY 1994 and FY 1999 data in a pie-chart format to give another perspective of the trend over the decade of the 1990s. The annual trend of Capacity and Capability is generally downward, while the Efficiency of Delivery of Capacity and Capability is generally upward. Technology and System Studies, Management, and Reserves are fairly constant over the years. The most notable change is the reduction of Capacity and Capability from 27 percent of the TDA budget in 1994 to only 13 percent in 1999. This is a 50 percent reduction!

Table 2. Budget allocation.

Category	Percentage of budget ^a						
	1993	1994	1995	1996	1997	1998	1999
Capacity and Capability ^b	24	27	23	19	15	16	13
Efficiency of Delivery of Capacity and Capability ^c	61	54	57	62	65	64	67
Technology and System Studies ^d	12	10	10	9	10	10	11
Management ^e	3	4	3	3	3	3	3
Reserves	0	6	7	7	7	7	7
Total, percent	100	100	100	100	100	100	100

^a Budget percentages were taken from WAD 93-3.

^b Modification and additions to the network to meet the customer's requirements (new antennas, new uplink/downlink frequencies, increased G/T).

^c Operate and maintain the network, including sustaining functions.

^d Investment for the future in research and development, including feasibility demonstrations.

^e Strategic planning, customer interface, allocation of resources, methods, and standards to optimize the network.

Note that the FY 1998 data in Table 2 differ somewhat from the prevailing trend on Fig. 1 in the Capacity and Capability and Efficiency of Delivery of Capacity and Capability categories. This is due primarily to a 1998 major funding increase of \$12.5M for two major implementation tasks in the Capacity and Capability category.

VI. Summary

The overall Capacity and Capability expenditures as a percentage of the total WAD budget are decreasing while the percentage of Efficiency of Delivery of Capacity and Capability is increasing. The other categories are relatively small and constant. The purpose of documenting this methodology is so that others could replicate it or modify it in the future as major changes to the WAD occur. This work also produces some potential metrics that can be used in the future as we embark on a reengineering effort.

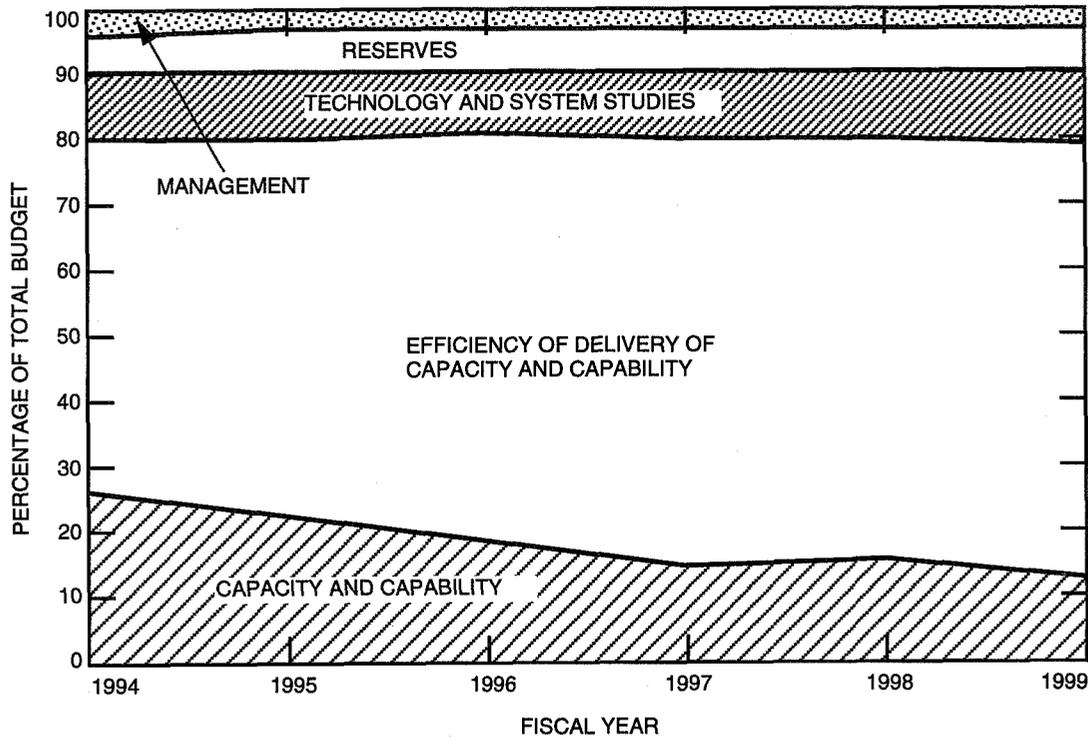


Fig. 1. Percentage of total budget by categories, 1994 through 1999.

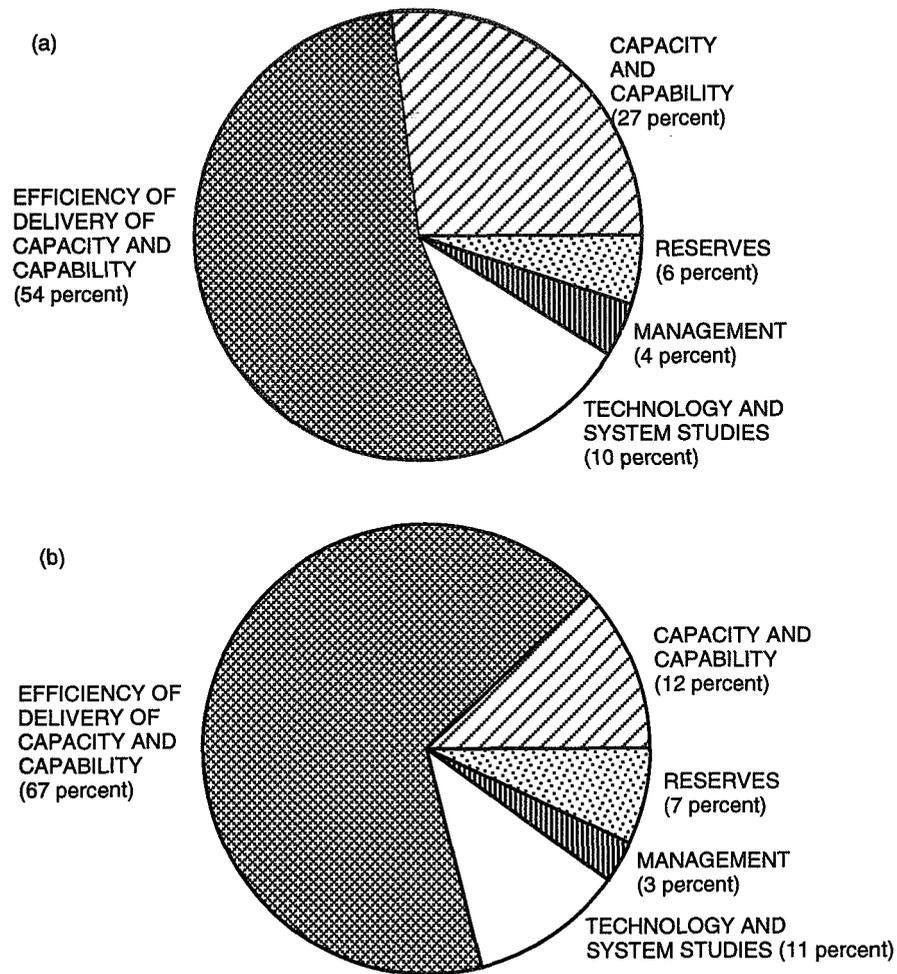


Fig. 2. WAD category allocations as a percentage of the grand totals: (a) FY 1994 and (b) FY 1999.

Appendix

Definition of Categories for Allocation of WAD Resources

I. Capacity and Capability

This category includes those modifications and additions to the Deep Space Communications Complexes that result in meeting the customer's requirements. These modifications and additions include the following:

- (1) Enhancements of link performance, such as increased gain-temperature ratio (G/T) of antennas, lower noise receiving systems, wider bandwidth/telemetry channels, and higher power transmitters.

- (2) New capability, such as **new uplink and/or** downlink frequencies.
- (3) Increased network capacity by the addition of new antennas and associated equipment.

II. Efficiency of Delivery of Capacity and Capability

This category includes the costs of the following:

- (1) Operating and maintaining the network equipment.
- (2) Operating the Maintenance Repair Facility and training the maintenance and repair staff.
- (3) Developing and implementing engineering change kits for maintaining “as-built” performance, and upgrading or adding new equipment that increases the original “as-built” performance.
- (4) Training the operators and managing operations facilities, such as calibration and standard laboratories, and transporting test instruments for periodic calibration.
- (5) Operations engineering, including system engineering, software database functions, project and station implementation, and scheduling.

It also includes those tasks that increase the tracking hours available for the customer. These tasks include the following:

- (1) Automation of network control, improving the operational processes to produce shorter calibration times, shorter turnaround times between missions, faster maintenance response to equipment failures, and tighter monitoring of equipment status.
- (2) Implementation of additional equipment to reduce downtime. Examples are the addition of on-line spares **and** modification of equipment and/or components to *increase* the mean time between failures and *reduce* the mean time to restore service.

III. Technology and System Studies

This category includes those costs for research and development (through the supporting research and technology program), other applicable technology development tasks, and the resources for feasibility demonstrations that precede implementation of performance upgrades.

IV. Management

This category includes costs for management concerns such as strategic planning, customer interface, allocation of resources, and the implementation of methods and standards to optimize the effectiveness of the DSN.

V. Reserves

This category includes major reserve funds that are not clearly allocated to specific tasks and cannot be assigned to any of the above categories at the time the WAD is generated.

277-32
33283
P-14

110881

358485

RF Optics Study for DSS-43 Ultracone Implementation

P. Lee and W. Veruttipong
Ground Antennas and Facilities Engineering Section

The Ultracone feed system will be implemented on DSS 43 to support the S-band (2.3 GHz) Galileo contingency mission. The feed system will be installed in the host country's cone, which is normally used for radio astronomy, VLBI, and holography. The design must retain existing radio-astronomy capabilities, which could be impaired by shadowing from the large S-band feed horn. Computer calculations were completed to estimate system performance and shadowing effects for various configurations of the host country's cone feed systems. Also, the DSS-43 system performance using higher gain S-band horns was analyzed. A new S-band horn design with improved return loss and cross-polarization characteristics is presented.

I. Introduction

The Ultracone is an ultralow-noise S-band (2.3-GHz) feed system that will be implemented on DSS 43, the 70-m DSN antenna in Australia, for the Galileo S-band contingency mission. The Ultracone will have better performance than the S-band polarization diverse (SPD) design that is currently installed on DSS 43, allowing for higher data transmission rates from the Galileo low-gain S-band antenna. The performance improvements are achieved by making the Ultracone a receive-only, fixed-polarization system with copper components. The SPD design is a transmit-and-receive, polarization-diverse system. The Ultracone is predicted to lower the system noise temperature by 2.5 to 3.0 K. The requirements for the design are a maximum system noise temperature of 12.5 K at zenith and a minimum antenna and microwave gain of 63.3 dBi at a 45-deg elevation, referenced to the maser input flange.¹

The Ultracone will be installed in the DSS-43 host country's cone. This cone is used primarily by radio astronomy, VLBI, and holography. The Ultracone feed system cannot easily be added without interfering with the already existing feed systems. One of the Ultracone task's functional requirements is to retain existing host country radio-astronomy capabilities. This article presents the RF optics study that examined options for implementing the Ultracone feed system without impacting radio-astronomy capabilities, while giving the best S-band performance possible.

¹ *Galileo S-Band Level C Review* (internal document), Jet Propulsion Laboratory, Pasadena, California, December 1992.

II. 70-m Antenna Geometry

The DSS-43 70-m antenna geometry is shown in Fig. 1.² The antenna has a symmetric main reflector and an asymmetric subreflector, which can be set only at discrete azimuthal stops. The reflector system focuses to a ring slightly above the top of three feed cones, which house the feed systems. Located on the focal ring are points that correspond to the subreflector stops. The phase centers of the feed systems are positioned at these focus points. Figure 2 shows a top view of the host country's cone that will be used to house the Ultracone feed system. It contains three focal points. F1 is the proposed location of the Ultracone, the radio-astronomy 22-GHz dual-feed system is currently located at F2, and F3 is used for radio astronomy, VLBI, and holography.

III. Improving S-Band Performance

A. Feed Horn Gain

The 70-m antenna optics were designed for the maximum gain-to-noise-temperature ratio (G/T) at 8.45 GHz using a 22.5-dBi feed horn.³ For the original implementation of 2.295 GHz on the 70-m antennas, a 22.5-dBi horn was also used. At 2.295 GHz, the rear spillover past the main reflector is higher than at 8.45 GHz since the longer wavelength has a wider diffraction pattern off the subreflector. In a low-temperature system, such as the Ultracone, this spillover is a major contributor to the overall noise temperature. One method of reducing the rear spillover is to use a higher gain S-band feed horn.

The effects of increasing the horn gain were analyzed using JPL physical optics (P.O.) computer programs.⁴ Increasing the horn gain to 23.1 dBi was found to reduce the rear spillover by about one-half, while only slightly decreasing the overall gain of the antenna. The horn gain was increased by adding an extension 26 cm long, which increases the aperture diameter from 66.2 to 72.3 cm. The G/T improvement is about 0.11 dB/K at 30-deg elevation, 0.14 dB/K at 45-deg elevation, and 0.23 dB/K at 90-deg elevation. Figure 3 shows the antenna performance versus horn aperture diameter at 2.295 GHz, and Tables 1 and 2 show the overall gain and noise temperature budgets with the 22.5-dBi and 23.1-dBi horns, respectively.⁵ The G/T performance at 22 GHz was also analyzed using higher gain horns. Since the rear spillover past the main reflector is already very low at 22 GHz, the improvement in noise temperature is very slight. The resulting decrease in system gain, due to underillumination of the main reflector, offsets the improvement in noise temperature, causing the overall G/T to decrease.

B. New Horn Design

To further improve system performance at S-band, the feed horn was redesigned for better return loss and cross-polarization characteristics. Computer analysis of the new design predicts better than a -39-dB return loss over the transmit band (2.025 to 2.210 GHz) and receive band (2.2 to 2.3 GHz), compared to a -31- to -41-dB return loss over the same bands for the JPL standard S-band horn design.⁶ The maximum predicted peak cross-polarization level is better than -40 dB for the new design, compared to -35 dB for the standard design. Figure 4 shows the computer-predicted return loss and cross-polarization levels for the new horn and the standard horn.

² A. G. Cha, *Physical Optics Analysis of NASA/JPL Deep Space 70-m Antennas*, JPL D-1853 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 1984.

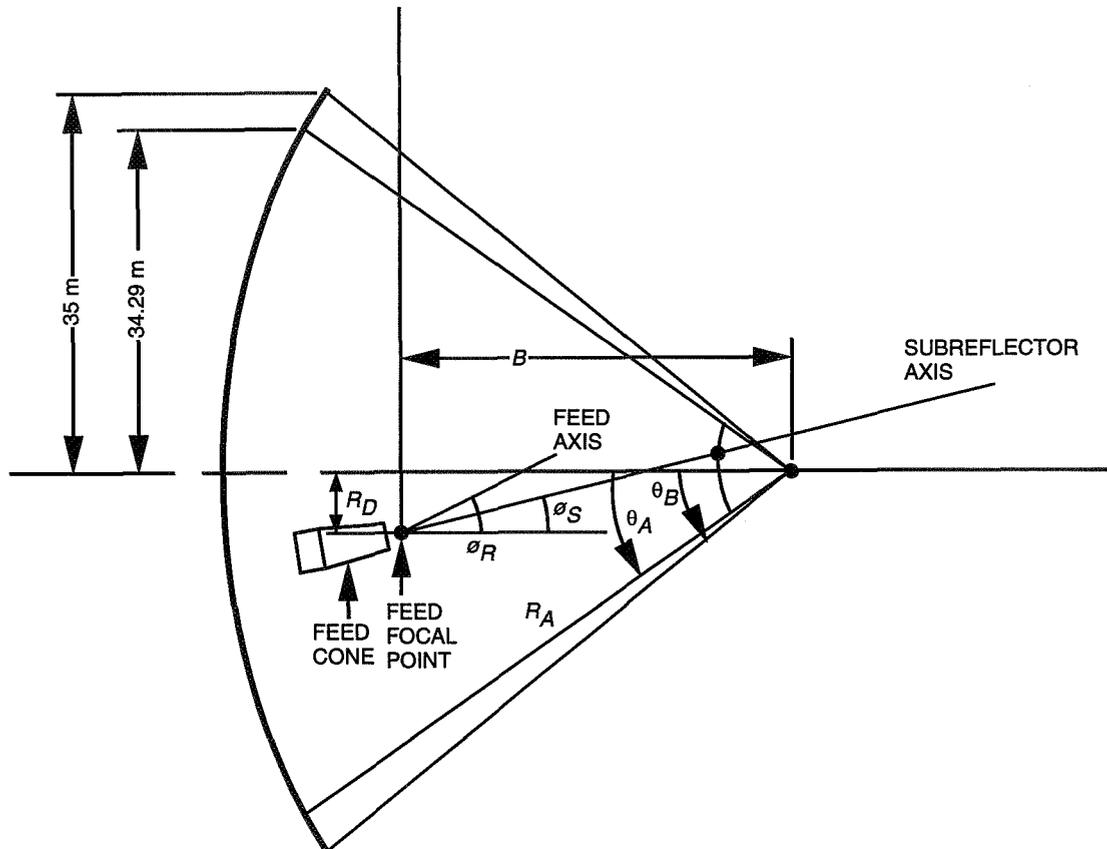
³ W. F. Williams, *Considerations for Determining the Shaped Main Reflector for the DSN 70-Meter Upgrade Program*, JPL D-1875 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 1, 1984.

⁴ R. E. Hodges and W. A. Imbriale, *Computer Program POMESH for Diffraction Analysis of Reflector Antennas* (internal document), Jet Propulsion Laboratory, Pasadena, California, February 1992.

⁵ D. Trowbridge and S. Petty, personal communication, Radio Frequency and Microwave Subsystems Section, Jet Propulsion Laboratory, Pasadena, California, August 5, 1994.

⁶ "Horn Assy, Wide Band, Dual Mode," JPL Drawing 9449420, Rev. E (internal document), Jet Propulsion Laboratory, Pasadena, California, March 29, 1979.

The gain of the new horn is 23.1 dB at 2.295 GHz. Removing the final section of the horn results in a gain of 22.5 dB at 2.295 GHz, making it compatible with the standard S-band horn gain. Theoretical and measured far-field copolarization patterns are shown in Fig. 5 at 2.295 GHz in the receive band and at 2.025 GHz in the transmit band. The measured patterns were taken on a far-field range on the mesa at JPL. The theoretical patterns were calculated using a circular waveguide mode-matching program, CWG.F [1,2], and a circular-aperture far-field radiation program, CRAD.F [3].



- $B = 15.451 \text{ m}$
- $R_D = 108.01 \text{ cm}$
- $\theta_R = 5.73722 \text{ deg}$: ACTUAL FEED TILT FROM MAIN REFLECTOR AXIS, PRESENT TRICONE FEED
- $\theta_S = 5.26000 \text{ deg}$: FEED TILT AT WHICH SUBREFLECTOR EDGE WOULD SUBTEND EQUAL ANGLES AT FEED PHASE CENTER, DEFINED AS SUBREFLECTOR OR AXIS FOR FABRICATION
- $\theta_A = 62.2626 \text{ deg}$: (TO 34.29-m OPTICAL BOUNDARY)
- $\theta_B = 63.31 \text{ deg}$: (TO 35 m)
- $R_A = 38.74 \text{ m}$

Fig. 1. DSS-43 70-m antenna geometry.

Two major design changes were made between the standard S-band horn and the new Ultracone horn. First, the corrugations were oriented perpendicular to the axis of the horn instead of perpendicular to the sidewall (Fig. 6). The JPL computer program used to analyze the horn (CWG.F) is a circular-waveguide mode-matching program. It can model the type of corrugations in the new horn design exactly, but can

only approximate the corrugations in the older design. It has been shown [4] that horns with either type of corrugations perform well. The orientation has little effect on fabrication costs.

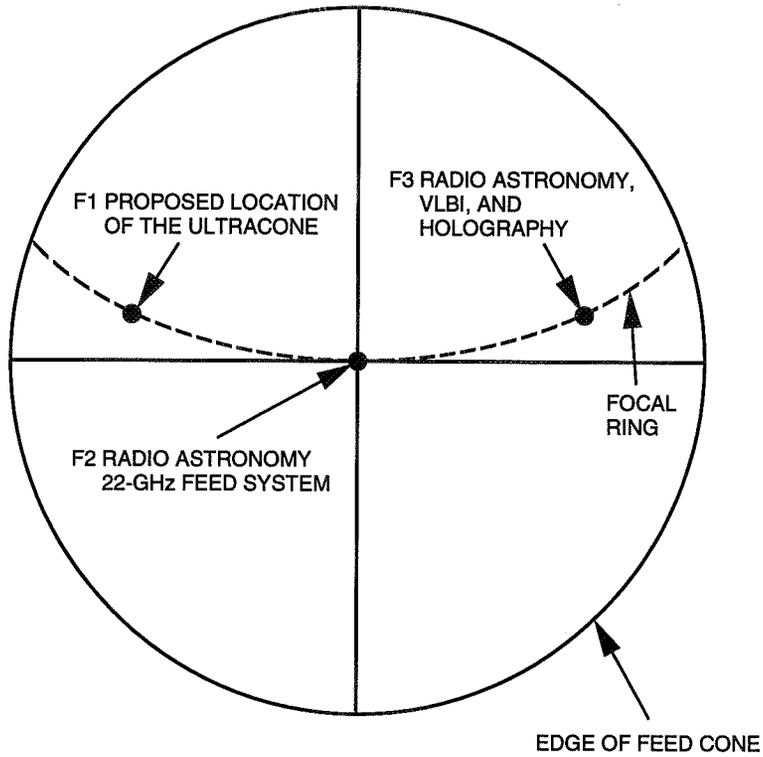


Fig. 2. Top view of DSS-43 host country's feed cone.

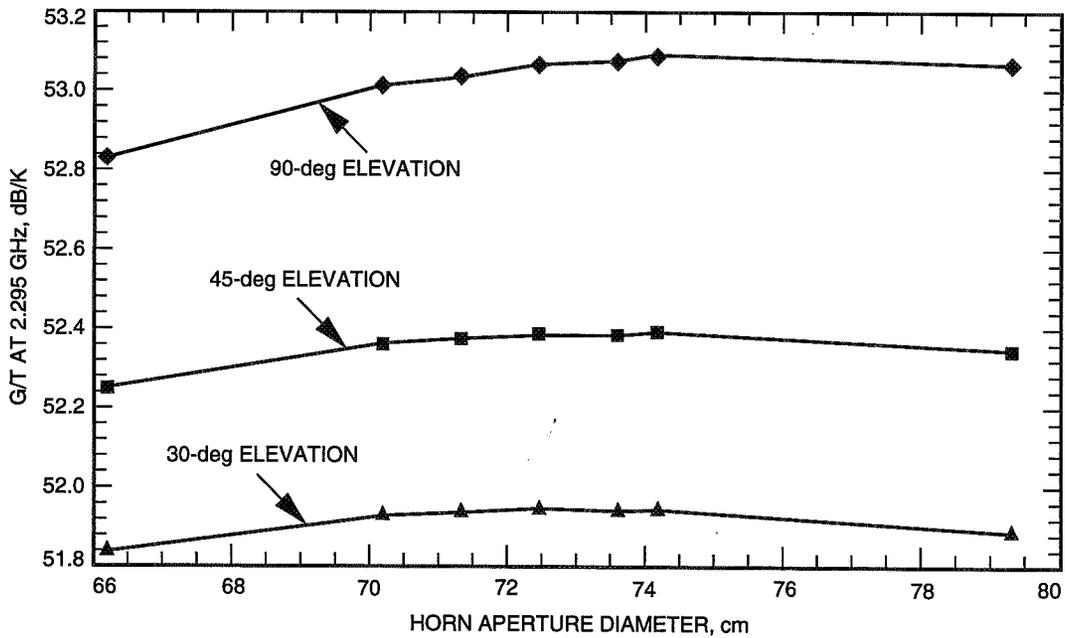


Fig. 3. DSS-43 G/T versus feed-horn aperture diameter at 2.295 GHz.

Table 1. Gain and noise temperature budgets at 2.295 GHz for a 22.5-dBi horn.

Noise temperature budget				
Element	Noise temperature, K			Comments
	90 deg	45 deg	30 deg	
Cosmic background	2.7	2.7	2.7	Effective black body
Atmosphere	1.9	2.6	3.6	
Subreflector forward spillover	0	0	0	
Main reflector rear spillover	1.42	1.07	0.96	240 K, 0.59 percent
Main reflector resistive	0.08	0.08	0.08	
Main reflector gap leakage	0.1	0.1	0.1	
Subreflector resistive	0.07	0.07	0.07	
Quadripod scatter	2.3	3.8	4.2	
Subtotal	8.57	10.42	11.71	Noise at feed aperture
Modified subtotal	8.53	10.38	11.66	Noise at preamplifier input -0.01855 dB at 20 deg C
Waveguide resistive	1.25	1.25	1.25	
Coupler injection	0.093	0.093	0.093	
Low-noise amplifier	2.036	2.036	2.036	
Follow up	0.037	0.037	0.037	
Total noise temperature	11.94	13.79	15.07	

Gain budget				
Element	Efficiency			Comments
	90 deg	45 deg	30 deg	
P.O. computed subtotal	0.90212	0.91003	0.90485	Includes F1 field: Forward spill: 0.9636 Rear spill: 0.9941 Central blockage: 0.988
Main reflector resistive	0.99977	0.99977	0.99977	
Main reflector panel leakage	1	1	1	Calculated
Main reflector gap leakage	1	1	1	Calculated
Main reflector surface rms	0.997	0.997	0.997	
Subreflector resistive	0.99977	0.99977	0.99977	
Subreflector surface rms	0.99986	0.99986	0.99986	
Waveguide resistive	0.99574	0.99574	0.99574	-0.01855 dB at 20 deg C
Waveguide voltage standing-wave ratio	0.991	0.991	0.991	
Quadripod blockage	0.922	0.922	0.922	
Total efficiency	0.8178	0.8250	0.8203	
Overall gain, dB	63.65	63.69	63.66	64.524 dB is 100 percent at 2.295 GHz

Computed G/T				
Element	G/T			Comments
	90 deg	45 deg	30 deg	
Computed G/T, dB/K	52.88	52.29	51.88	

Table 2. Gain and noise temperature budgets at 2.295 GHz for a 23.1-dBi horn.

Noise temperature budget				
Element	Noise temperature, K			Comments
	90 deg	45 deg	30 deg	
Cosmic background	2.7	2.7	2.7	Effective black body
Atmosphere	1.9	2.6	3.6	
Subreflector forward spillover	0	0	0	
Main reflector rear spillover	0.70	0.53	0.47	240 K, 0.29 percent
Main reflector resistive	0.08	0.08	0.08	
Main reflector gap leakage	0.1	0.1	0.1	
Subreflector resistive	0.07	0.07	0.07	
Quadripod scatter	2.3	3.8	4.2	
Subtotal	7.85	9.88	11.22	Noise at feed aperture
Modified subtotal	7.81	9.83	11.17	Noise at preamplifier input -0.01855 dB at 20 deg C
Waveguide resistive	1.25	1.25	1.25	
Coupler injection	0.093	0.093	0.093	
Low-noise amplifier	2.036	2.036	2.036	
Follow up	0.037	0.037	0.037	
Total noise temperature	11.23	13.25	14.59	

Gain budget				
Element	Efficiency			Comments
	90 deg	45 deg	30 deg	
P.O. computed subtotal	0.89467	0.90252	0.89738	Includes F1 field: Forward spill: 0.9716 Rear spill: 0.9971 Central blockage: 0.988
Main reflector resistive	0.99977	0.99977	0.99977	
Main reflector panel leakage	1	1	1	Calculated
Main reflector gap leakage	1	1	1	Calculated
Main reflector surface rms	0.997	0.997	0.997	
Subreflector resistive	0.99977	0.99977	0.99977	
Subreflector surface rms	0.99986	0.99986	0.99986	
Waveguide resistive	0.99574	0.99574	0.99574	-0.01855 dB at 20 deg C
Waveguide voltage standing-wave ratio	0.991	0.991	0.991	
Quadripod blockage	0.922	0.922	0.922	
Total efficiency	0.8110	0.8182	0.8135	
Overall gain, dB	63.61	63.65	63.63	64.524 dB is 100 percent at 2.295 GHz

Computed G/T				
Element	G/T			Comments
	90 deg	45 deg	30 deg	
Computed G/T, dB/K	53.11	52.43	51.99	

The second major design change was to replace the large matching groove in the throat section of the horn with a series of 10 slots tapered from approximately one-half wavelength to one-quarter wavelength in depth (Fig. 6). The tapered section provides a better match over a wider band than the single matching groove. The groove depths used in the tapered section are listed in Table 3. To accommodate the deeper slots of the taper at the input, the outside dimensions of the first horn section were enlarged, and the cold-water distribution hardware, which is not used, was removed.

Three minor design changes were made. First, the one-quarter wavelength slots are no longer constant throughout the entire horn. They are adjusted according to the horn radius to provide an effective electrical length of one-quarter wavelength [5]. Second, except for the input matching section, the slots

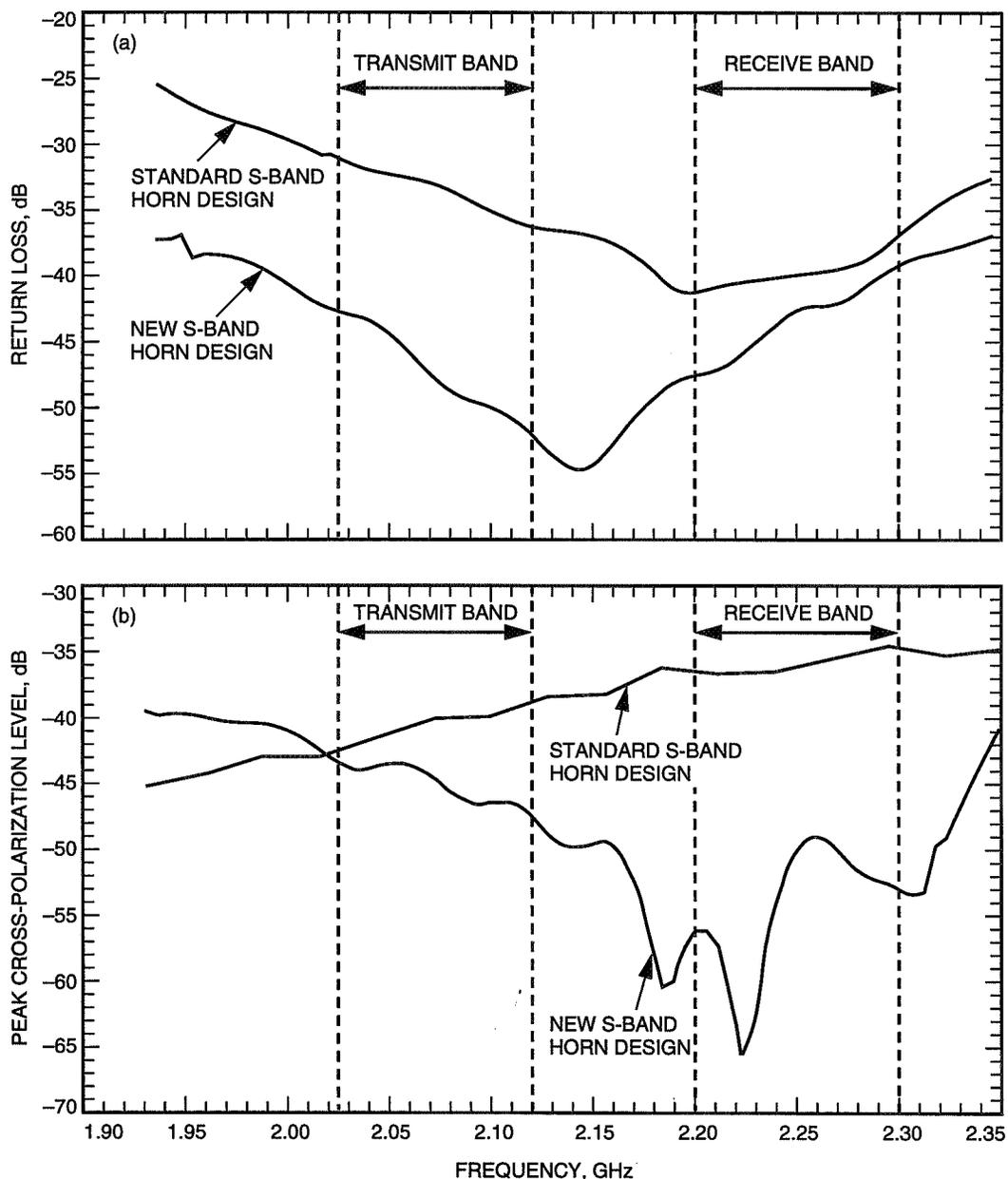


Fig. 4. Predicted performance of a standard S-band horn and the new S-band horn designs: (a) return loss and (b) peak cross-polarization levels.

are not as deep as those in the standard horn, making the outside dimensions of the horn slightly smaller. Reducing the depth of these slots improved the cross-polarization performance. Third, the new 22.5-dB horn is 1.7 cm longer than the standard 22.5-dB horn.

C. Noise Shield

Adding a noise shield around the edge of the main reflector dish was briefly considered. Physical optics programs were used to estimate the spillover past the main reflector dish with and without a noise shield. A 1-m shield would decrease the S-band noise temperature by 0.8 K at zenith, and a 2-m shield would provide a 1.1-K decrease at zenith. The cost of the noise shield was roughly estimated to be several hundred thousand dollars, which was deemed too expensive. No further investigation was made.

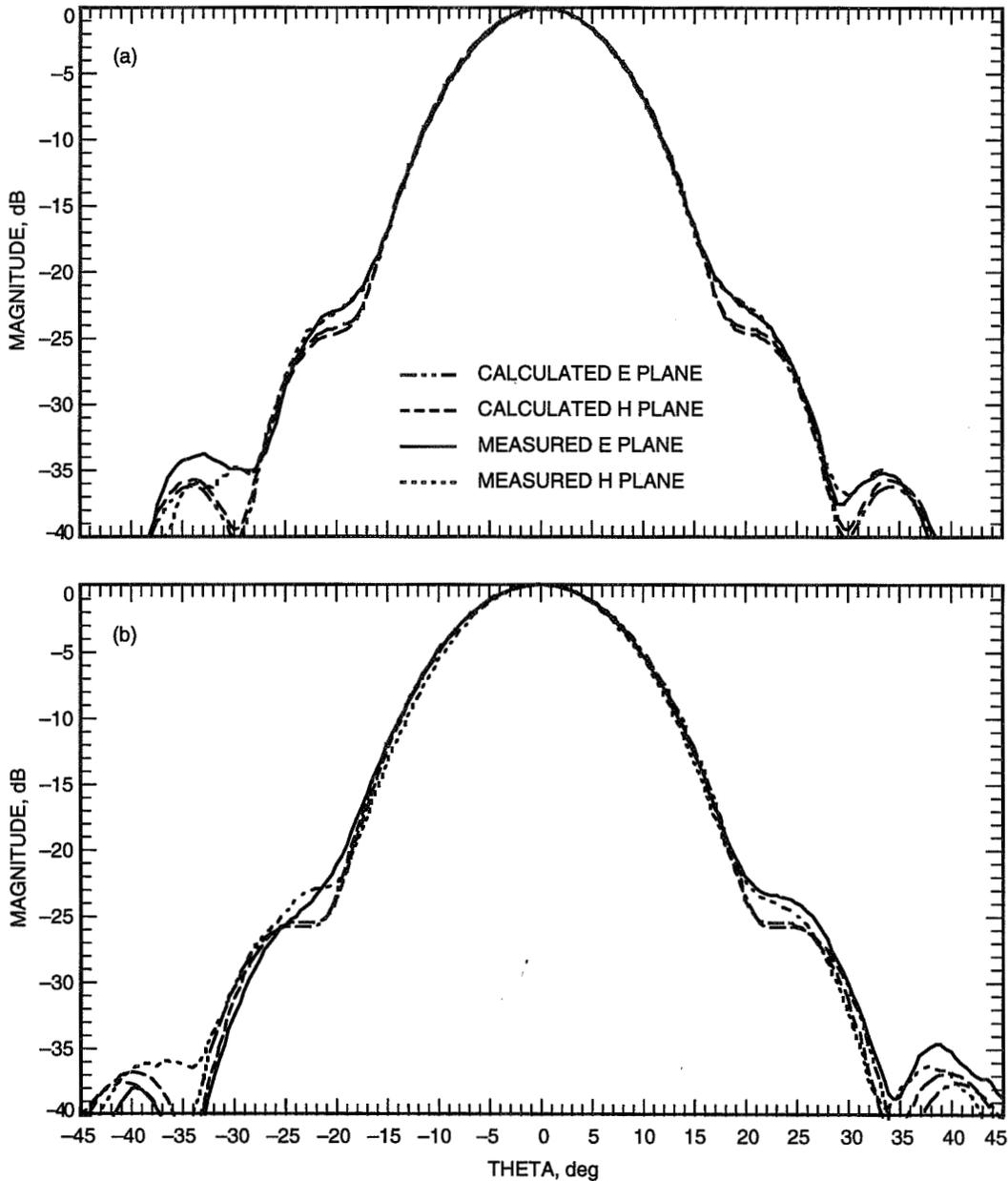


Fig. 5. Copolarization patterns of the new S-band horn at (a) 2.295 GHz and (b) 2.025 GHz.

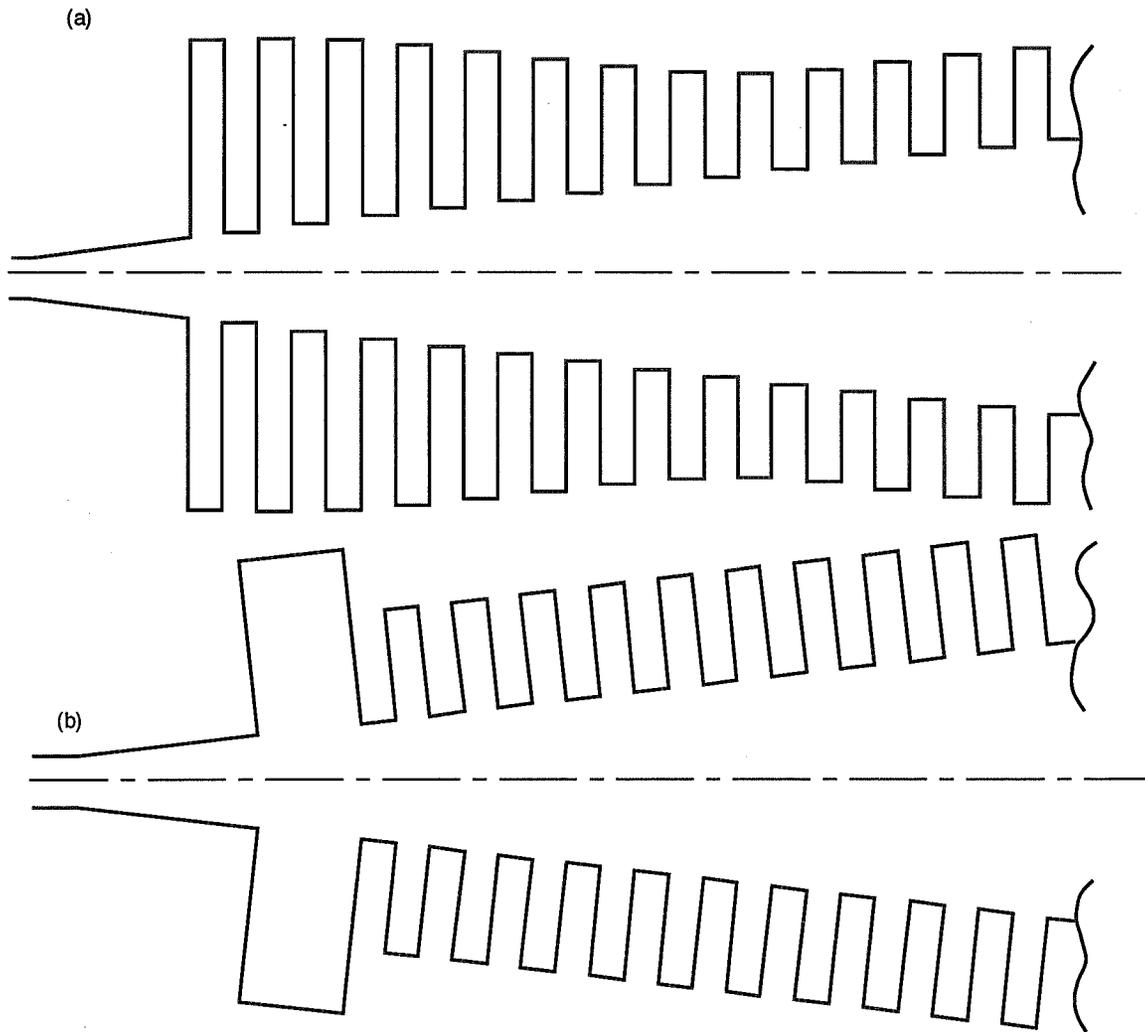


Fig. 6. Throat section of (a) the new S-band horn design and (b) a standard S-band horn.

IV. Preserving 22-GHz Performance

Installing the Ultracone horn in the F1 position and the 22-GHz radio-astronomy horn in the F2 position would create shadowing and interference problems between the two horns. By using a 23.1-dB S-band horn instead of the 22.5-dB horn, the shadowing and interference makes the 22-GHz system virtually unusable (Fig. 7). Options considered to avoid these problems were moving the S-band horn down along the feed axis, moving the 22-GHz horn up along the feed axis, moving the 22-GHz horn along the focus ring off F2, and adding a “stovepipe” extension to the top of the S-band horn to lower the feed and increase the horn gain. Moving the S-band horn along the focus ring, away from the 22-GHz horn, was not possible because of space constraints in the feed cone. In all cases, it was assumed that the S-band horn aperture flange would be trimmed to the minimum possible diameter.

A. Horn Relocation

The blockage of the 22-GHz horn by the S-band horn was estimated for several different positions of the two horns. The overall system performance was analyzed as the horns were moved off the feed focal

points. The analysis was completed using a combination of P.O. programs and the Geometric Theory of Diffraction (GTD) programs.

One possible way to reduce blockage of the 22-GHz horn was to move the S-band horn down along the feed axis. A computer analysis showed that the S-band performance is very insensitive to the horn location along the feed axis, but because of space constraints in the cone, this could not be done. The possibility of moving the 22-GHz horn up along the feed axis was examined. This defocuses the system by moving the phase center of the horn off the focal point of the main reflector and subreflector system. Part of this defocusing can be compensated for with subreflector displacement, if enough subreflector movement is available. Figure 8 shows the calculated gain loss from moving the horn, both with and without subreflector compensation, without considering losses due to blockage. With the 23.1-dB S-band horn, there is always significant blockage of the 22-GHz radiation pattern. With the 22.5-dB S-band horn, a 7.6-cm movement would prevent any appreciable blockage and would correspond to approximately a 0.2-dB loss in performance if subreflector compensation of 0.4 cm were used. If subreflector compensation could not be used, the loss from defocusing would be 0.6 dB.

Another approach investigated was to move the 22-GHz horn away from the S-band horn along the feed focal ring. Since the subreflector is designed to stop only at discrete locations along the focal ring, the horn would be defocused as it was moved from F2 towards F3 (Fig. 9). Figure 10 shows the computed gain loss from the horn movement, without considering losses due to blockage. Approximately 12.7 cm of movement, which would probably be the minimum movement possible to reduce the amount of blockage to acceptable levels, would cause about a 1.3-dB loss at 22 GHz due to defocusing, which is unacceptable.

Four specific cases were examined in more detail. The first case left the 22-GHz horn and the 22.5-dB S-band horn in their standard positions. The edge taper of the 22-GHz horn radiation pattern at the closest point of interference with the S-band horn was -11.5 dB. The radiation blocked by the S-band horn was integrated and found to be 2.2 percent of the total radiation, which would lead to approximately 0.1 dB of loss at 22 GHz. The second case was the same as the first, but with the 23.1-dB S-band horn. The pattern edge taper at the point of interference increased to -0.5 dB, which is totally unacceptable. The third case put the 22.5-dB S-band horn in its standard position with the 22-GHz horn moved 15.24 cm along the focal ring. The blockage in this case was insignificant. The fourth case was the same as the third, but with the 23.1-dB S-band horn. The pattern edge taper at the closest point of interference was -16.5 dB. The total power blocked was found by integration to be 0.8 percent, which would correspond to approximately a 0.035-dB loss.

Table 3. Taper section groove depths.

Slot number	Slot depth, cm
1	7.107
2	6.928
3	6.585
4	6.135
5	5.609
6	5.048
7	4.499
8	4.020
9	3.676
10	3.544

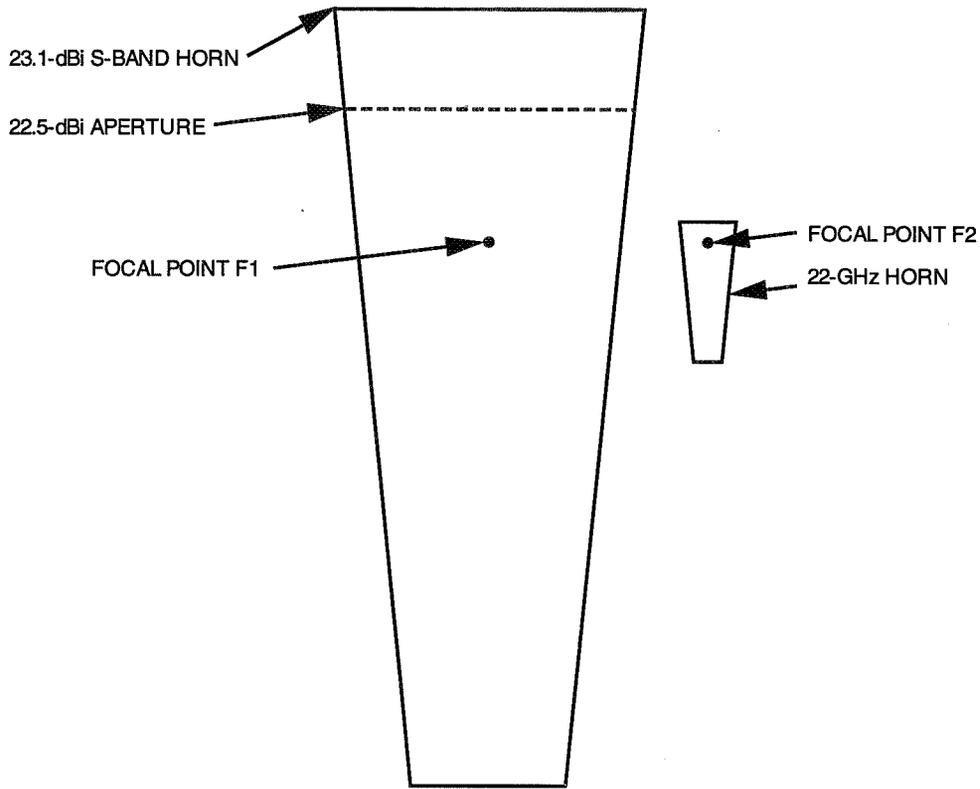


Fig. 7. Interference between the 23.1-dBi gain S-band horn, located at F1, and the 22-GHz horn, located at F2.

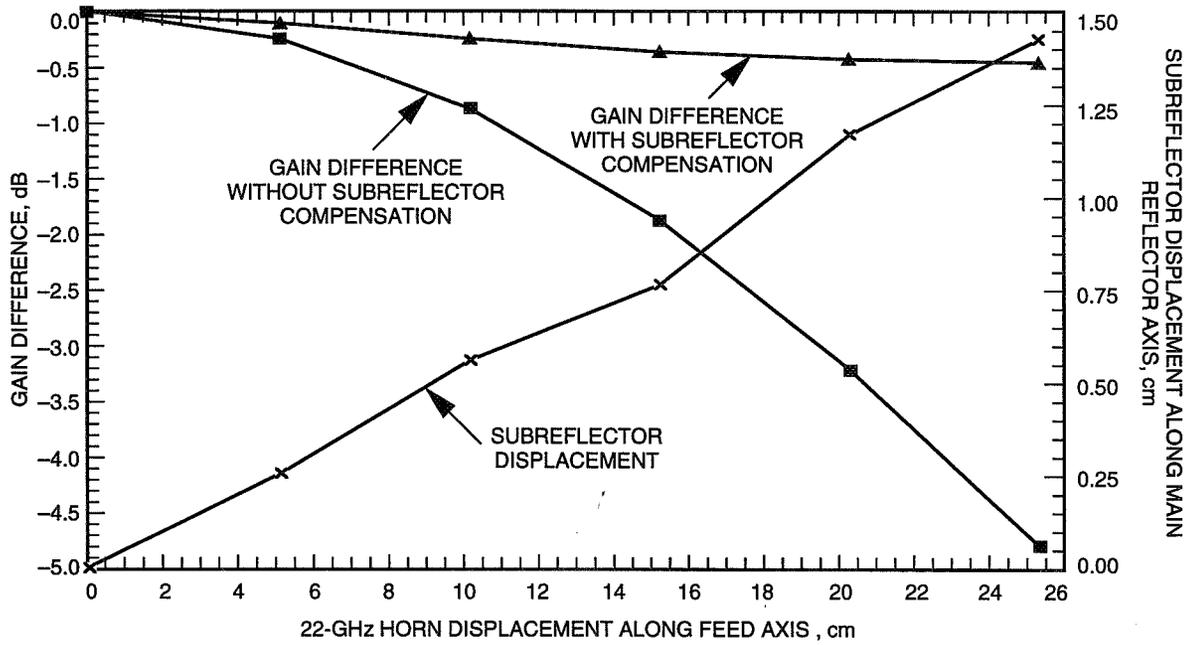


Fig. 8. Gain loss versus 22-GHz movement along the feed axis towards the subreflector, with and without subreflector compensation.

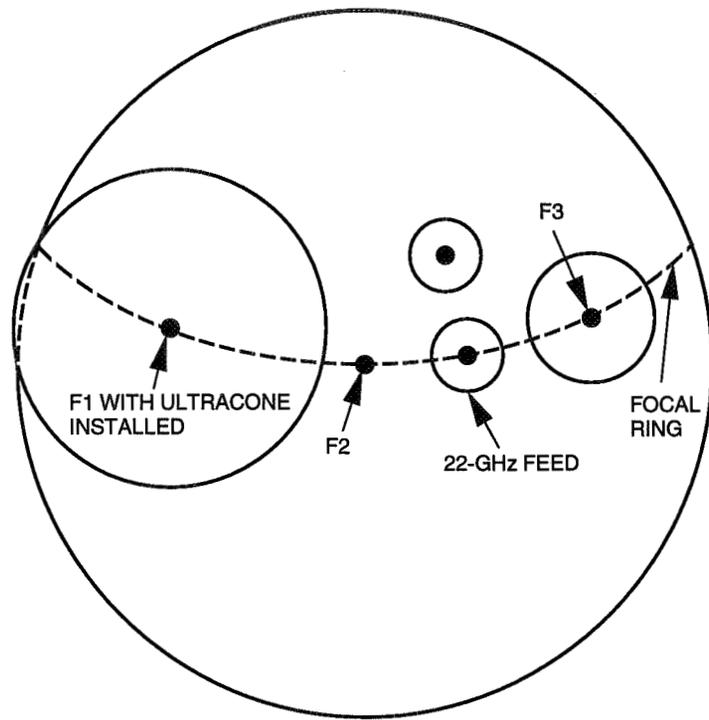


Fig. 9. Movement of the 22-GHz horn away from F2 towards F3.

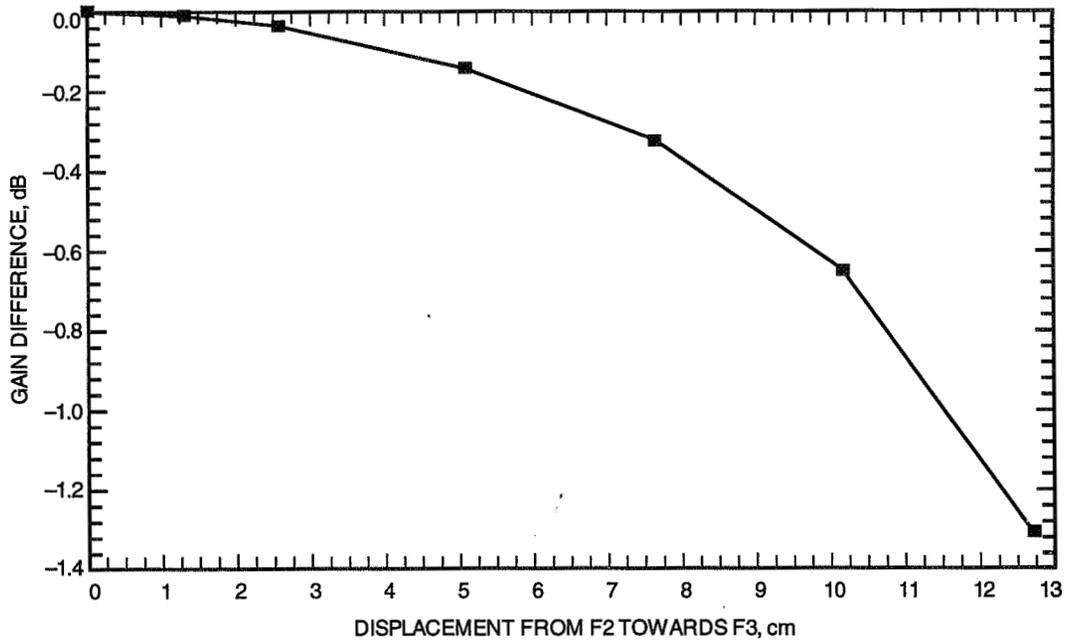


Fig. 10. Computed gain loss of the 22-GHz horn versus movement towards F3.

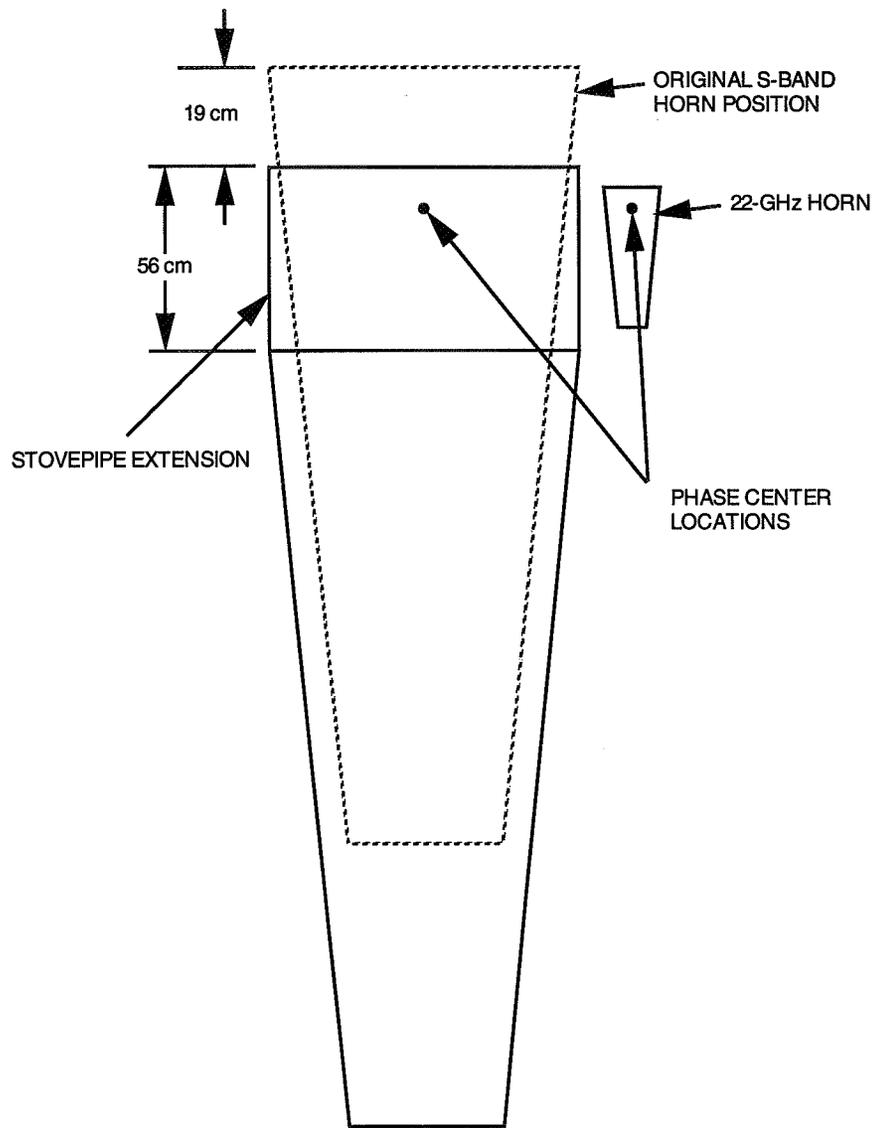


Fig. 11. Stovepipe configuration.

B. Stovepipe Design

One design alternative that was considered was a “stovepipe” extension [6] for the S-band horn. This design consists of adding a corrugated extension with no flare angle to the top of the horn. The extension moves the phase center of the feed horn to near the aperture of the horn and increases the gain of the horn. This would allow for a higher gain S-band horn while decreasing the shadowing of the 22-GHz horn by retaining the same aperture diameter and lowering the horn along the feed axis (Fig. 11). Using the data provided in [6], it was estimated that to move the aperture of the feed horn down 19 cm along the feed axis, an extension of 55.9 cm would be required to keep the phase center of the horn at the focal point. This would mean that the feed system would have to be moved down in the cone by 74.9 cm. This option was no longer considered when it was discovered that space in the feed cone did not allow the S-band horn to be lowered.

V. Conclusion

The RF study concluded that the S-band performance of the 70-m antenna could be improved by increasing the gain of the feed horn from 22.5 to 23.1 dB. A new, 23.1-dB horn was designed with better return loss and cross-polarization patterns than the standard JPL S-band horn. The Ultracone implementation will use the new feed horn in the F1 position of the host country's cone. The 22-GHz radio-astronomy horn that was originally in the F2 position must be relocated to reduce the blockage by the S-band horn to acceptable levels. Computer analysis showed that moving the 22-GHz horn 15.24 cm along the focal ring from F2 toward F3, as shown in Fig. 9, will reduce the blockage to 0.8 percent, which corresponds to a 0.035-dB loss.

The final design configuration chosen for implementation involved redesigning the subreflector positioner so that the subreflector can be moved to any angle. This allows the system to focus on any point on the focal ring instead of only three discrete points in each feed cone. The 22-GHz feed will be moved along the focal ring by 15.24 cm, and the subreflector will be repositioned so that no losses due to defocusing will be incurred.

Acknowledgments

The authors thank D. Trowbridge and S. Petty for providing data for gain and noise budget estimates, H. Reilly and R. Cirillo for the feed horn antenna measurements, O. Casanova for providing valuable information and support, and P. Stanton for reviewing this article. The Cray supercomputer used in this investigation was provided by funding from the NASA Offices of Mission to Planet Earth, Aeronautics, and Space Science.

References

- [1] D. Hoppe, "Scattering Matrix Program for Circular Waveguide Junctions," *Cosmic Software Catalog*, 1987 edition, NASA-CR-179669, NTO-17245, NASA's Computer Software Management and Information Center, Athens, Georgia, 1987.
- [2] G. L. James, "Analysis and Design of TE₁₁-to-HE₁₁ Corrugated Cylindrical Waveguide Mode Converters," *IEEE Transactions on Microwave Theory and Techniques*, vol. MTT-29, no. 10, pp. 1059-1066, October 1981.
- [3] S. Silver, *Microwave Antenna Theory and Design*, New York: McGraw-Hill, pp. 336-337, 1949.
- [4] P. Clarricoats and A. Olver, *Corrugated Horns for Microwave Antennas*, London: Peter Peregrinus Ltd., pp. 120-122, 1984.
- [5] B. Thomas, G. James, and K. Greene, "Design of Wide-Band Corrugated Conical Horns for Cassegrain Antennas," *IEEE Transactions on Antennas and Propagation*, vol. AP-34, no. 6, pp. 750-757, June 1986.
- [6] F. Manshadi and R. Hartop, "Compound-Taper Feedhorn for NASA 70-m Antennas," *IEEE Transactions on Antennas and Propagation*, vol. 36, no. 9, pp. 1213-1216, September 1988.

518-32
33284
110882
358486

JPL-ANTOPT Antenna Structure Optimization Program

D. M. Strain

Ground Antennas and Facilities Engineering Section

New antenna path-length error and pointing-error structure optimization codes were recently added to the MSC/NASTRAN structural analysis computer program. Path-length and pointing errors are important measures of structure-related antenna performance. The path-length and pointing errors are treated as scalar displacements for statics loading cases. These scalar displacements can be subject to constraint during the optimization process. Path-length and pointing-error calculations supplement the other optimization and sensitivity capabilities of NASTRAN. The analysis and design functions were implemented as "DMAP ALTERs" to the Design Optimization (SOL 200) Solution Sequence of MSC-NASTRAN, Version 67.5.

I. Introduction

For many years, the principal antenna structural design tool of the Ground Antennas and Facilities Section has been the JPL-IDEAS Design Optimization Program [1]. It has proved very useful both as a design tool and as a test-bed for numerous analytical and optimization techniques. It has also taken advantage of numerous unique features of the FORTRAN-V compilers of the UNISYS machines. While this was expedient, it has proved increasingly restrictive. Costs have increased, maintainability has suffered, and the code is not "portable."

JPL-IDEAS depends on a limited repertoire of finite elements, both for computational efficiency and to effect the optimization algorithms. The NASTRAN structural analysis computer program has traditionally been used to check IDEAS results and for analyses that required a more extensive element library. Having to maintain parallel finite-element models was at best cumbersome and time consuming. With the advent of design optimization capabilities in MSC/NASTRAN, and the rumored imminent departure of the UNISYS platforms, it was decided to implement the essential features of IDEAS directly into MSC/NASTRAN.

II. Analysis Capabilities

A. Surface Accuracy Calculations

Antenna primary surface accuracy can be represented as the weighted root-mean-square (rms) error of a paraboloid best fitting the deformed reflector surface in a half-path-length sense [2,3]. A subset of the finite-element model grid points is chosen to represent the primary-reflector panel support points. Each of these grid points is assigned a weighting function proportional to the reflector panel area it supports.

The following is a summary of the path-length error calculations for one loading case. Matrix and vector quantities are noted with their dimensions in square brackets, where N is the size of the weighted grid subset. Quantities designated as "raw" refer to errors with respect to the undeflected, nominal paraboloid; those designated as "fit" are errors with respect to the best-fitting paraboloid. The other symbols are defined as

f = nominal focal length of main reflector

$\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ = spatial coordinates of the main reflector support points [$N \times 1$]

$\mathbf{U}, \mathbf{V}, \mathbf{W}$ = translation displacement components [$N \times 1$]

\mathbf{wgt} = diagonal matrix of surface panel weighting function [$N \times N$]

$\sum \mathbf{wgt}$ = sum of the surface panel weights

ρ = raw path-length errors, with respect to nominal paraboloid [$N \times 1$]

$\mathbf{gX}, \mathbf{gY}, \mathbf{gZ}$ = direction cosines of surface normal vectors [$N \times 1$]

\mathbf{A} = basis functions of normal equations, $\mathbf{wgtAS} = \mathbf{wgt}\rho$ [$N \times 6$]

\mathbf{S} = solution of normal equations [6×1]

ϕ = fit path-length errors, with respect to best-fit paraboloid [$N \times 1$]

$$\rho_i = \frac{-2f\mathbf{X}_i\mathbf{U}_i - 2f\mathbf{Y}_i\mathbf{V}_i + 4f^2\mathbf{W}_i}{\mathbf{X}_i^2 + \mathbf{Y}_i^2 + 4f^2}$$

$$rms_{raw} = \left(\frac{\rho_i^T \mathbf{wgt}_i \rho_i}{\sum \mathbf{wgt}} \right)^{1/2}$$

$$\mathbf{gX}_i = \frac{-\mathbf{X}_i}{\mathbf{X}_i^2 + \mathbf{Y}_i^2 + 4f^2}$$

$$\mathbf{gY}_i = \frac{-\mathbf{Y}_i}{\mathbf{X}_i^2 + \mathbf{Y}_i^2 + 4f^2}$$

$$\mathbf{gZ}_i = \frac{2f}{\mathbf{X}_i^2 + \mathbf{Y}_i^2 + 4f^2}$$

$$\mathbf{A}_i = [-\mathbf{gZ}_i\mathbf{gX}_i \quad -\mathbf{gZ}_i\mathbf{gY}_i \quad -\mathbf{gZ}_i\mathbf{gZ}_i \quad -\mathbf{Z}_i\mathbf{gZ}_i\mathbf{gZ}_i \quad (\mathbf{Z}_i\mathbf{gZ}_i\mathbf{gY}_i - \mathbf{Y}_i\mathbf{gZ}_i\mathbf{gZ}_i) \\ (-\mathbf{Z}_i\mathbf{gZ}_i\mathbf{gX}_i + \mathbf{X}_i\mathbf{gZ}_i\mathbf{gZ}_i)]$$

$$\mathbf{S} = [\mathbf{A}^T \mathbf{wgt} \mathbf{A}]^{-1} \mathbf{wgt} \rho$$

$$\phi = \rho - \mathbf{AS}$$

$$rms_{fit} = \left(\frac{\phi^T \mathbf{wgt} \phi}{\sum \mathbf{wgt}} \right)^{1/2}$$

The six rows of the solution vector \mathbf{S} are the $X, Y,$ and Z translational offsets of the vertex of the best-fitting paraboloid; the change in focal length of the best-fitting paraboloid; and the rotational offsets of the best-fitting paraboloid about the $x-$ and $y-$ axes.

The coefficient matrix ($[A^T \text{wgt} A]$), the right-hand side ($\text{wgt} \rho$), and solution (\mathbf{S}) are printed and other results of the path-length analysis are tabulated for each loading case.

B. Pointing Error Calculations

The pointing error is calculated as the sum of the error contributions from each of the optical components of the antenna: the primary reflector, the secondary reflector (S/R), and the microwave-feed phase center [4]. The errors due to the displacements of the primary reflector are calculated from the fitting parameters of the best-fit paraboloid (BFP). The motions of the secondary reflector are taken to be those of a grid point whose *motions* approximate the motions of the secondary. Similarly, the motions of the feed phase center are taken from the *motions* of a designated "feed" grid point. The spatial coordinates of these grids are not considered during pointing-error calculations. The symbols are defined as

- K = beam deviation angle of primary reflector
- K_2 = beam deviation angle of secondary reflector
- M = magnification factor
- A, C = Cassegrain geometry parameters, where hyperboloid eccentricity = C/A
- F = focal length of nominal paraboloid
- dX, dY, dZ = translational offsets of BFP
- dF = change in focal length of main reflector paraboloid ($f_{BFP} - f_{nominal}$)
- rX, rY = rotational offsets of BFP or secondary reflector

Angular errors due to primary (BFP) motion in a symmetric Cassegrain system are calculated as follows:

Motion component	X-axis error	Y-axis error
Lateral translation	$-K/F \times dY$	$K/F \times dX$
Rotation	$(1 + K) \times rX$	$(1 + K) \times rY$

Errors due to secondary reflector motion in a symmetric Cassegrain system are calculated as follows:

Motion component	X-axis error	Y-axis error
Lateral translation	$(K/F - K_2/F/M) \times dY$	$-(K/F - K_2/F/M) \times dX$
Rotation	$-(2 \times (C - A) \times K/F) \times rX$	$-(2 \times (C - A) \times K/F) \times rY$

Errors due to feed phase center motion in a symmetric Cassegrain system are calculated as follows:

Motion component	X-axis error	Y-axis error
Lateral translation	$(K/M - 1/F) \times dY$	$-(K/M - 1/F) \times dX$

No pointing error accrues from axial motions or focal-length change of the components in a symmetric system.

The pointing error about the elevation axis can be predicted relative to the elevation-axis position transducer. The rotation of the grid point, whose rotation represents that of the transducer, is subtracted from the total pointing error about the x-axis and is reported as the "corrected" pointing error.

The default antenna-pointing analysis calculations assume a symmetric Cassegrain configuration: a paraboloidal main reflector and a hyperboloidal secondary reflector with coincident foci, and a point source microwave feed located at the other focus of the hyperboloid. Simple descriptions of this geometry are sufficient input for this traditional DSN case. For non-Cassegrain or asymmetric (e.g., clear-aperture) antenna configurations, the pointing-error calculations shown above can be overridden for each of the motions of each of the optical components. Components whose override factors have values of 0.0 (the default) will be calculated as outlined above. If the factor is other than 0.0, the pointing error will be the factor times the specified displacement.

Pointing errors about each axis are tabulated for each loading case with the results of the path-length analysis.

III. New Input Data

Although many NASTRAN input data (over those required for a simple statics run) are required to characterize the antenna for the new program, the amount of additional work required of the user has been kept to a minimum by automating the data preparation. Four simple data generation programs are available, each of which is written in standard FORTRAN.

The identification of the main reflector surface panel-support grids and of grids that represent the motions of the secondary reflector, feed phase center, and elevation transducer is accomplished through the use of NASTRAN DMIG, USET, and PARAM bulk data cards [5].

Two other programs can be used to generate the "DESVAR" and "DVPREL1" NASTRAN cards, required for sensitivity and optimization, from a list of element property cards. One of the programs recognizes PROD, PBAR, PBEAM, and PSHELL cards and generates cards whose design variable-to-property relation is cross-sectional area for the bar elements and thickness for the shells. The second program recognizes PBAR and PBEAM cards and generates design variables and relations suitable for rectangular tubes. The outside dimension and wall thickness are the design variables. These programs can be adapted easily to process other properties or output other design variables.

These additional data cards need only be generated once for a particular antenna configuration. The small amount of additional input (6 through 16 PARAMeters describing antenna geometry and pointing coefficients and four data cards per design constraint) can be generated easily by hand.

Surface accuracy and pointing error responses are each calculated as scalar values for each design loading case. A NASTRAN scalar point is designated for each design constraint and used as a placeholder for these responses. Displacement design constraints are defined on the displacements of these scalar points.

IV. Program Flow

The new antenna analysis and sensitivity code have been implemented into MSC/NASTRAN, Version 67.5. The new code has been installed and tested on the JPL Cray Y-MP supercomputer and a Silicon Graphics (SGI) Crimson model workstation.

The general flow of the NASTRAN design optimization solution sequence (for statics) is

- (1) Read and collate input data.
- (2) Update geometric and structural matrices and tables.
- (3) Solve for static responses.
- (4) Determine sensitivity of objective and constraint functions to design variables.
- (5) Perform optimization.

The objective is usually to minimize structure weight, with constraints on antenna performance and element stresses. Steps (2) through (5) are iterated until the design problem converges.

The new antenna analysis and design code are inserted into the NASTRAN design optimization solution sequence in two places: at the end of statics analysis, step (3) above, and after the sensitivity analysis calculations, step (4) above.

At the end of step (3), the new code performs some input data checking, initializes and saves invariant antenna analysis data, and echoes various antenna description and processing parameters to the print file. At this point, all statics response quantities are available and static antenna analyses are performed. The resulting scalar quantities are inserted into the appropriate displacement vectors for each loading case.

Design sensitivity analysis [6], step (4) above, computes the rate of change in structural response to changes in design variables. Small perturbations to the design variables are approximated in the structural matrices, and response derivatives are explicitly solved for. The result, for the statics case, is a set of displacement derivative vectors for each design variable for each loading case. Path-length and pointing-error analyses are performed on each of these displacement vectors, as appropriate. The original statics errors are subtracted to give the change in antenna performance for variations in the design variable. These resulting scalar quantities are inserted into the displacement vectors for each loading case and passed to the optimizer.

The added antenna analysis and design codes do not modify or interfere with the operation of the original optimization code and are compatible with all the other sensitivity and design options in the design optimization solution sequence. The advent of geometry optimization in the next release of MSC/NASTRAN is especially intriguing. The method employed is extensible to constraints on other performance and operational measures, e.g., secondary reflector offsets or geometric theory of diffraction (GTD) performance analyses.

V. Effectiveness

JPL-ANTOPT is intended eventually to supplant the JPL-IDEAS program. Test design optimization runs have been made using a small verification model (Fig. 1) and the IDEAS design model for the new DSS-24 34-m antenna (Fig. 2). These models are summarized in Table 1. (The numbers in square brackets are the number of design variable linking groups in each element type. Path-length error constraints (P/L) are in mm rms. Pointing errors (PNT) are in arc-sec.)

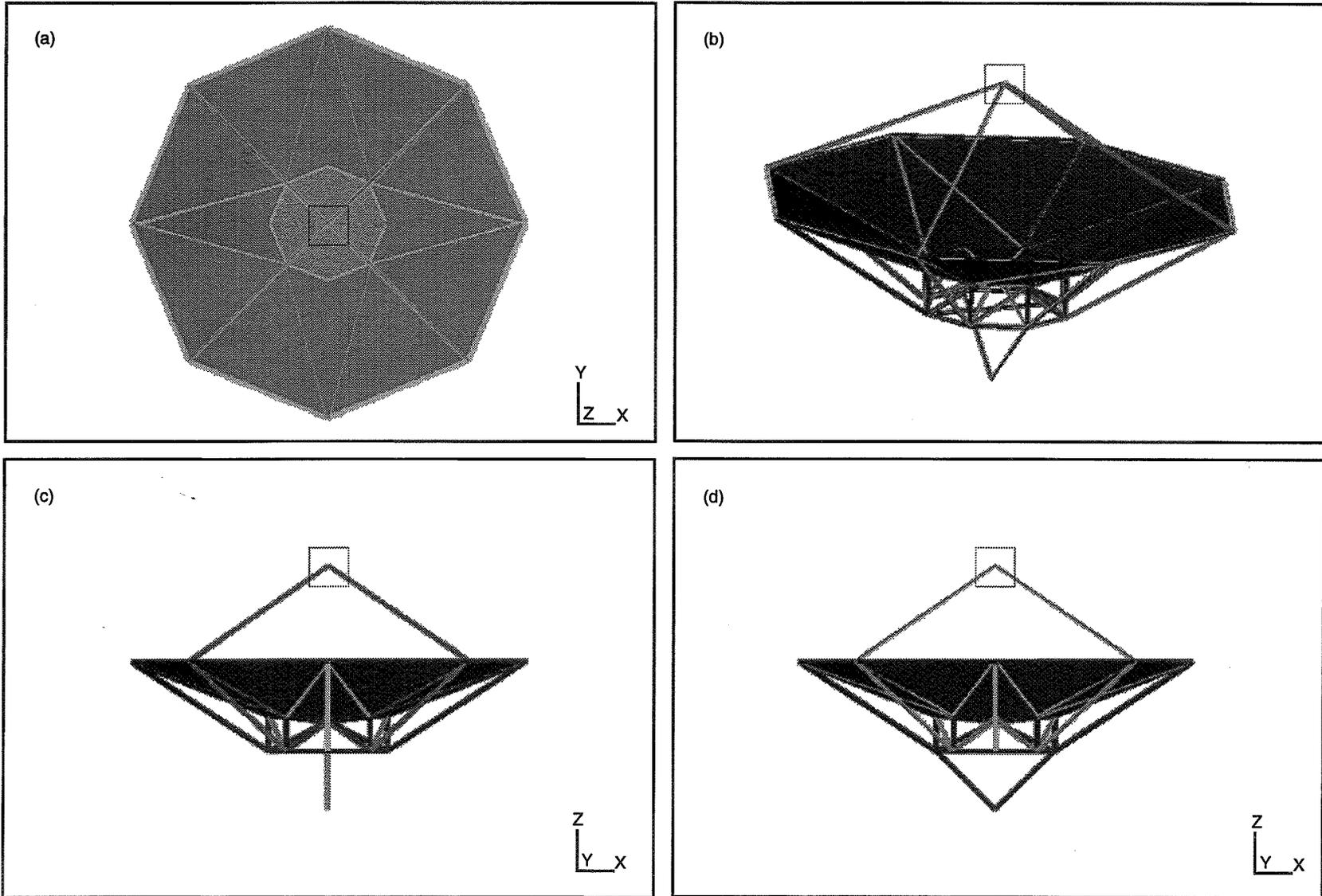


Fig. 1. Test model: (a) plan view, (b) partial rotation, (c) back elevation, and (d) side elevation.

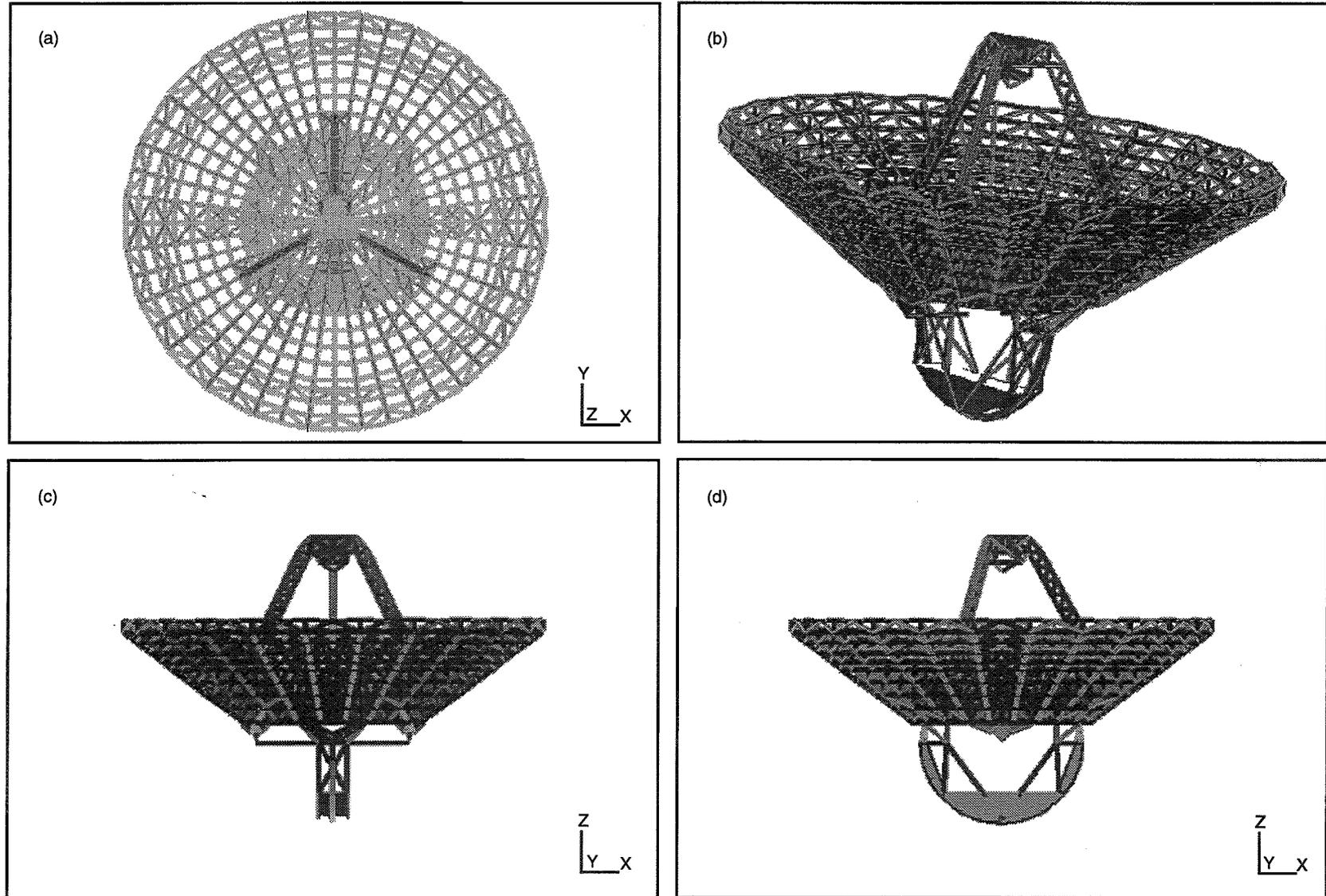


Fig. 2. The 34-m model: (a) plan view, (b) partial rotation, (c) back elevation, and (d) side elevation.

Table 1. Comparison model summary.

Component	Test model	34-m model
Grids	28	1473
Elements		
Rod	88 [12]	3876 [243]
Tri		106 [10]
Quad		562 [29]
Mass	18	1127
Design constraints		
Z-gravity P/L	0.254	0.381
Y-gravity P/L	0.254	0.381
Y-gravity PNT	50.0	900.0
Rim +Z force PNT	50.0	
S/R +Y force PNT	100.0	

Design constraint values were chosen to exercise the programs and have no other significance. For the baseline configurations, each constraint is violated by a factor of from 1.5 to 4. Stress and other side constraints were disabled. The default optimizer input parameters were used for both programs except that termination of design iterations due to meeting convergence criteria was prevented.

The design variables for each model are the cross-sectional areas of the rods and thicknesses of the plates. Minimum and maximum design variable limits were imposed: for the test model rod areas, 64.5 mm² and 6451.6 mm², and for the 34-m model, 322.6 mm² and 32258.0 mm² for the rod areas and 2.54 mm and 254.0 mm for the plate thicknesses. Some of the structure model was excluded from the optimization problem, including lumped masses that were used to account for the weight of the main reflector panels, secondary reflector, and counterweight; the remainder is deemed the variable structure.

Three comparative design optimization runs were performed using each model. The baseline configuration for the test model was uniform rod areas of 645.2 mm²; that of the 34-m model was 3225.8-mm² rod areas and 25.4-mm plate thicknesses. In addition, starting points of minimum and maximum design variable sizes were used. The geometry and element connectivity were not changed.

Figures 3 and 4 are plots of the variable and feasible structure weights as functions of design iteration history for the six optimization runs for each program. Feasible structure weight is the variable structure weight scaled by the largest ratio of performance measure to design constraint. For example, if a design constraint were violated by 10 percent at a particular iteration, the feasible weight would be 110 percent of the variable structure weight. The ratio of feasible to variable weights, then, is proportional to the worst constraint violation. The individual curves in the plots are designated I-WGT and I-Feas for variable structure weight and feasible weight as reported by the IDEAS program, and A-WGT and A-Feas for the weights reported by the ANTOPT program. (The differences between the initial feasible weights of the IDEAS and ANTOPT 34-m model runs is attributable to the differences in the weight and stiffness characteristics of the plate elements in the two programs.) It can be seen that the new program consistently achieves a feasible design in a few iterations.

JPL-ANTOPT comprises over 1800 lines of MSC/NASTRAN DMAP source code and has been accepted by NASA's Computer Software Management and Information Center (COSMIC) as program NPO-19349.

VI. Summary and Conclusion

The JPL-ANTOPT program performs microwave antenna performance analysis and structure optimization in an efficient manner that is easy for the structure designer to use. The new code is portable

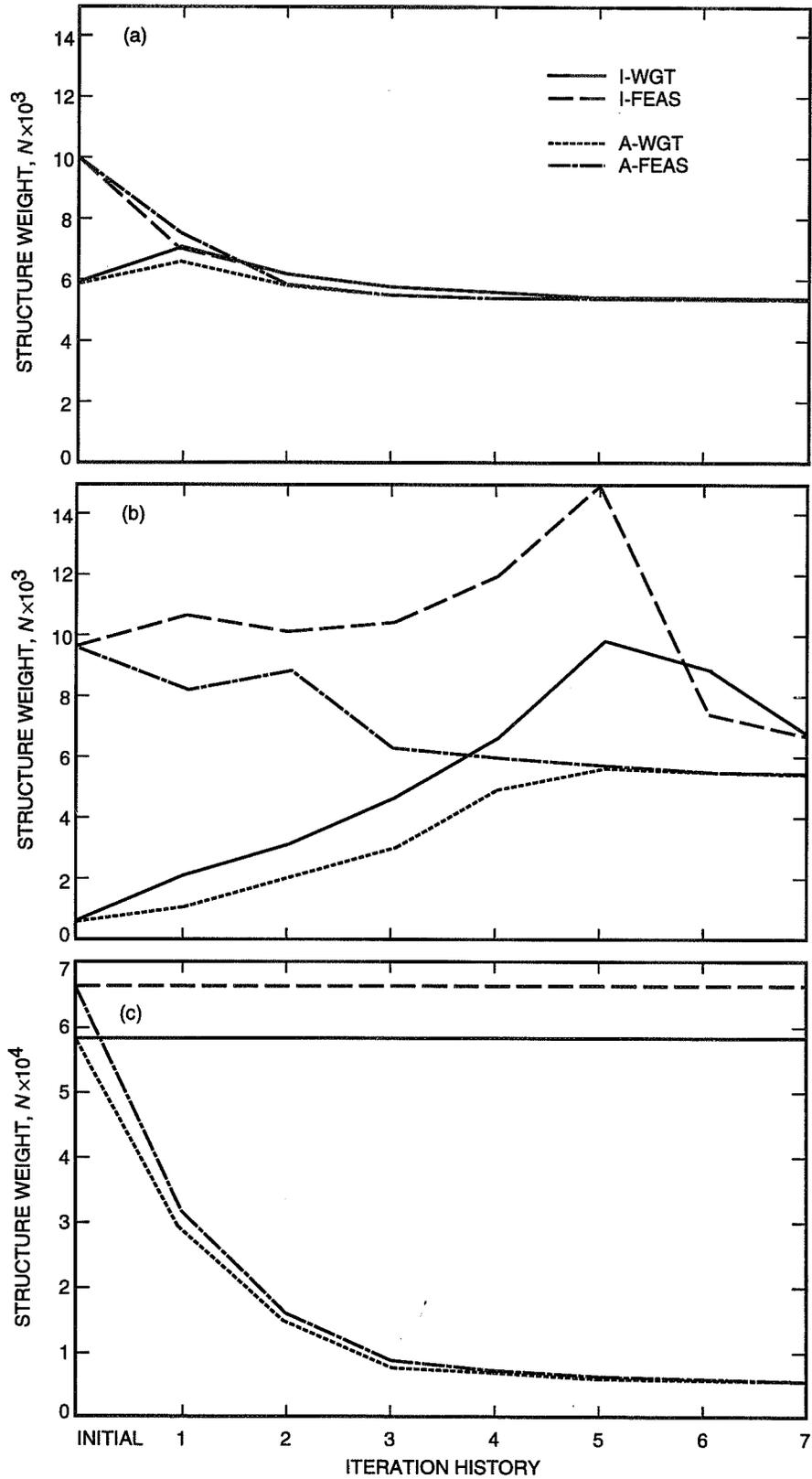


Fig. 3. Test model weight history: (a) initial = baseline; (b) initial = minimum; and (c) initial = maximum.

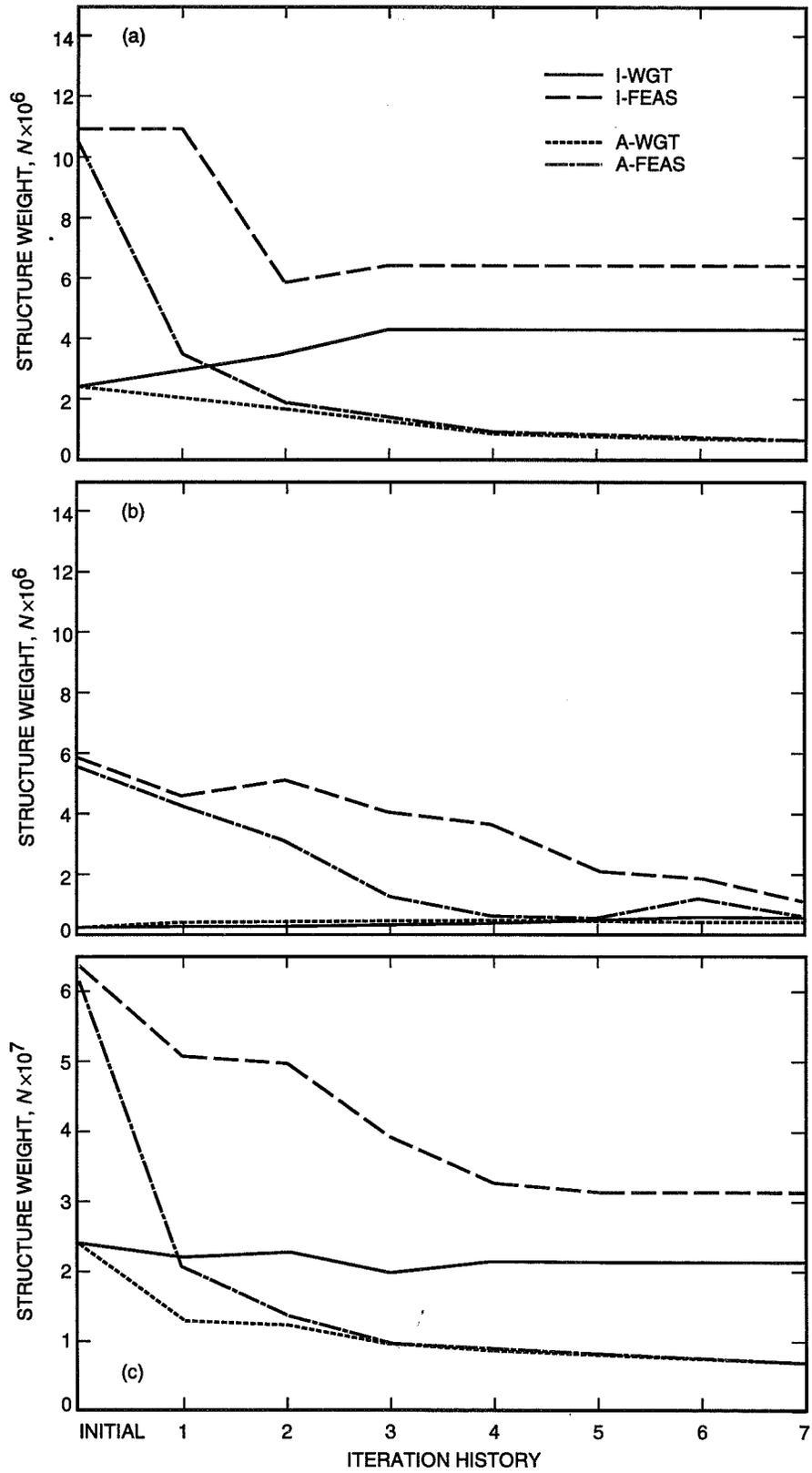


Fig. 4. The 34-m model weight history: (a) initial = baseline; (b) initial = minimum; and (c) initial = maximum.

across computer platforms and is compatible with other NASTRAN optimization capabilities. The method employed to introduce antenna performance measures into the NASTRAN optimization sequence is easily adaptable for use with other performance measures.

References

- [1] R. Levy and D. M. Strain, *User's Manual for the JPL-IDEAS Finite Analysis and Design Optimization Program*, COSMIC NPO-17783, NASA Software Technology Transfer Center, University of Georgia, Athens, Georgia, October 1988.
- [2] S. M. Barondess and S. Utku, *Computation of Weighted Root Mean Square of Path Length Changes Caused by the Deformations and Imperfections of Rotational Paraboloidal Antennas*, JPL Technical Memorandum 33-118, Jet Propulsion Laboratory, Pasadena, California, March 1, 1963.
- [3] M. S. Katow and L. Schmele, "Utku/Schmele Paraboloid RMS Best-Fit Program," *Computer Programs for Antenna Feed System Design and Analysis, Vol. 1: Programs and Sample Cases*, A. Ludwig, ed., JPL Technical Report 32-979, Jet Propulsion Laboratory, Pasadena, California, pp. 75-82, April 15, 1967.
- [4] J. Ruze, *Small Displacements in Parabolic Reflectors*, Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, Massachusetts, February 1, 1969.
- [5] *MSC/NASTRAN Quick Reference Guide, Version 67*, MacNeal-Schwendler Corporation, Los Angeles, California, 1992.
- [6] G. J. Moore, *MSC/NASTRAN Design Sensitivity and Optimization User's Guide, Version 67*, MacNeal-Schwendler Corporation, Los Angeles, 1992.

Errata

Due to a *TDA Progress Report* software problem, the plots in Figs. 1, 3, 5, 6, 7, and 11 of S. R. Tyler's "A Trajectory Preprocessor for Antenna Pointing" that appeared in *The Telecommunications and Data Acquisition Progress Report 42-118*, vol. April-June 1994, pp. 139-159, August 15, 1994, were distorted. Also, some numbers in the x-axis of Fig. 11 were transposed. The correct figures are provided below.

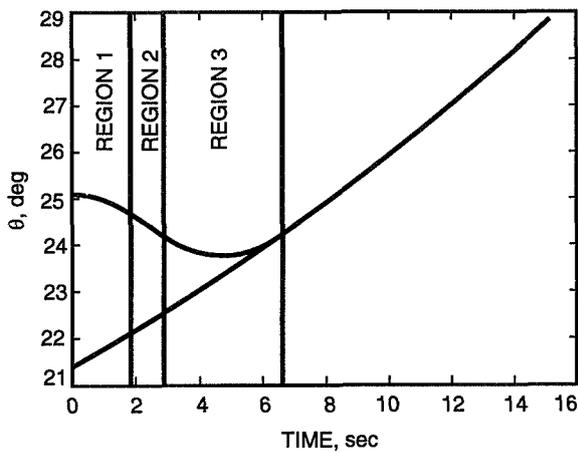


Fig. 1. Matching of antenna and target positions for constant accelerations (elevation only).

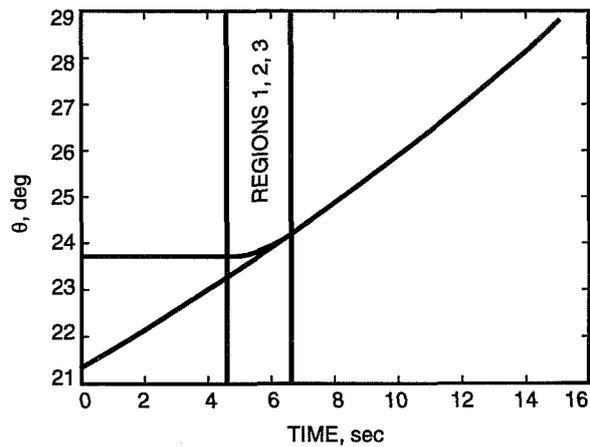


Fig. 3. Matching of antenna and target positions for elevation only (chasing target).

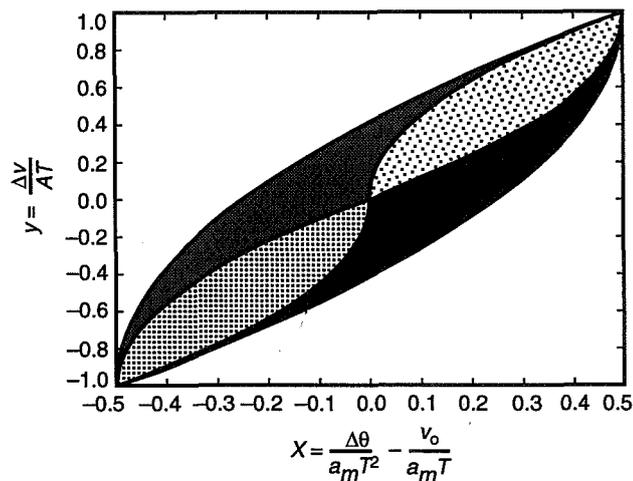


Fig. 5. Area of valid solutions ($a_m = a_{max}/2$ and $a_o = a_f = 0$).

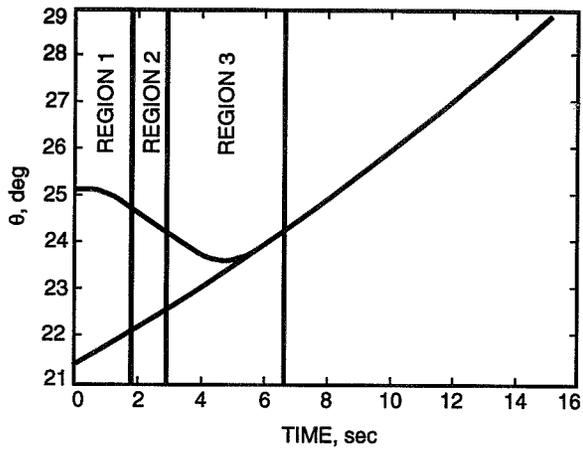


Fig. 6. Matching the target position for raised cosine acceleration (elevation only).

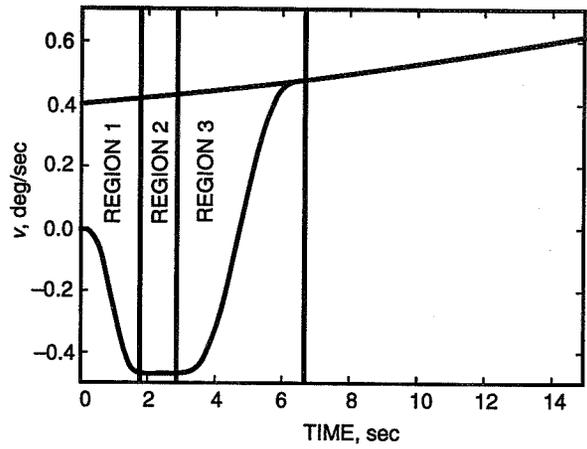


Fig. 7. Matching the target velocity for raised cosine acceleration (elevation only).

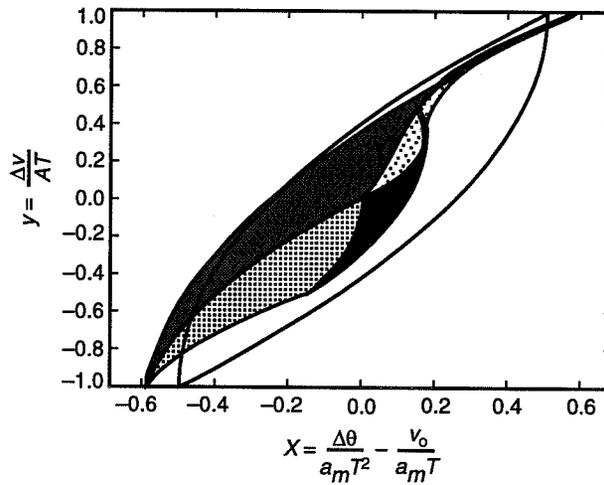


Fig. 11. Shrinkage of region of valid solutions for $a_f = -a_0 = a_m$