

Applications of Correlator Chips in Radio Science

Jon B. Hagen
Cornell University
NAIC Lab 124
124 Maple Ave
Ithaca NY 14850
607-255-5274

Abstract - Spectral line observations in radio astronomy require simultaneous power estimation in many (often hundreds to thousands) of frequency bins. Digital autocorrelation spectrometers, which appeared thirty years ago, are now being implemented in VLSI. The same architecture can be used to implement transversal digital filters. This was done at the Arecibo Observatory for pulse compression in radar observations of Venus.

1 Introduction

This paper reviews applications and implementations of the digital correlator, a signal processor long used for radio spectrometry and increasingly used for digital filtering. Most radio engineers intuitively define points on the power spectrum (power density v.s. frequency) of a signal to be the average voltages from a set of square-law detectors following a bank of band-pass filters. A formal definition is provided by the Wiener-Khinchin theorem: the power spectrum of a signal is the Fourier transform of the signal's autocorrelation function (ACF). In autocorrelation spectrometry the intensive on-line averaging is done in a correlator; the Fourier transform is a one-time operation done prior to further data analysis. (Occasionally experimental ACFs are compared to theoretical ACFs without ever transforming from the time domain to the frequency domain).

2 Basic Correlator Architecture

The architecture in question is extremely simple; the chip contains N identical arithmetic modules or "lags" which operate in parallel, independently forming products from two data streams and accumulating the sum of these products. One of the data streams, the "immediate" data is common to every lag. The other data stream, the "delayed" data, is supplied from successive taps of a shift register contained on the chip. Each clock pulse advances the shift register data and adds the new lagged products to the contents of each accumulator. Each module thus calculates a point on the autocorrelation function, i.e. the average product of the signal voltage at time T and the signal voltage at time $(t-T)$. An output multiplexer provides access to the individual accumulators. The basic block diagram is shown in Figure 1.

To measure a power spectrum via the autocorrelation function, the immediate and the delayed inputs are both tied to the same signal (the undetected baseband output voltage of the radio astronomy receiver). This type of correlator was introduced in radio astronomy by

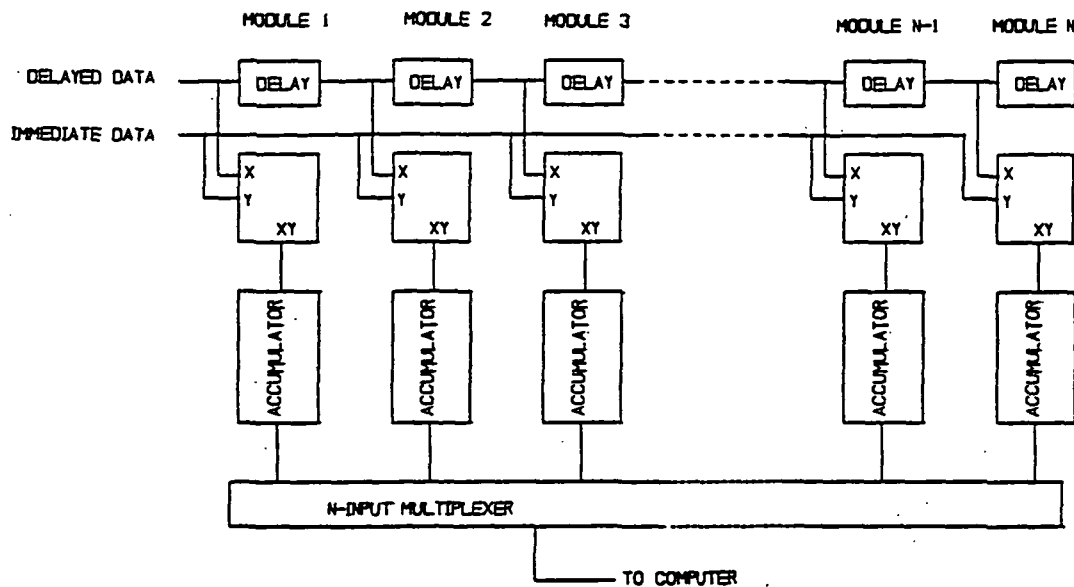


Figure 1: Basic Correlator Architecture

Sander Weinreb (1) in 1963 and has been the workhorse for spectral analysis. It is discussed further in Section 4 after a review of the basic principles of spectrometry.

3 Spectrometry in Radio Astronomy

Though a radio astronomer occasionally deals with objects that are sources of spectrally flat white noise, the more interesting and challenging sources have considerable spectral structure. Such sources include molecular clouds with emission or absorption lines, natural masers, and dispersed pulses from rapidly rotating neutron stars (pulsars). The spectrum can have enough structure to warrant hundreds to thousands of points of resolution. Each point requires its own radiometer, i.e. bandpass filter plus square-law detector or an equivalent such as one module of a correlator. This kind of multiplex spectrometer (meaning that all points are measured simultaneously) is always necessary in radio astronomy because the received signals are noiselike random processes and time averaging is needed at each filter/detector or correlator module. If the required accuracy is, say, 1% or one part in one hundred, it is necessary to average about 1002 or 10,000 samples of V^2 . Sampling theory shows that the time needed to gather the required number of independent samples will be inversely proportional to the bandwidth. These radiometry considerations show why an ordinary laboratory spectrum analyzer - a scanning receiver with a single bandpass filter - is not suitable for this work as it would increase the required observation time in proportion to the number of frequency channels. Such increases in observing time are simply not available; radio astronomers are usually trying to dig small spectral features out of background noise (the sum of cosmic noise, atmospheric noise, antenna noise, and amplifier noise) and, even with multiplex spectrometers, integration times run to dozens or even hundreds of hours since the signal-to-noise ratios (before averaging) are typically in the -30 to -60 dB range.

4 Digital Correlators

The basic correlator module can be implemented partly or wholly in analog circuitry but all-digital implementations have been favored from the start. Autocorrelation is not the only way to do digital spectrometry; the Fast Fourier Transform (FFT), familiar since the late sixties, is, in principle, a more appealing algorithm since it ultimately must be more efficient - arithmetic operations are done at a rate proportional to $N \log(N)$ rather than N^2 where N is the number of points on the spectrum. Surely, for N large enough, the FFT will be better. Yet for N smaller than, say, 10,000, and at sampling rates, i.e. bandwidths, high enough that the digital circuitry is running at full speed, it seems that the correlator architecture has two advantages.

The first advantage comes from the simplicity of the architecture. The modules run independently. Apart from the shift register that distributes the delayed data, there are no data interconnections between modules. The correlator is easily expandable. To get more points on the spectrum one cascades more of the same modules. Conversely, a machine with many modules can be subdivided into several smaller correlators which can independently analyze several different signals. The simplicity of the basic module thus provides flexibility at the system level.

The second advantage is that correlators can use very coarse quantization of the input signal. Instead of digitizing the signal with 8 to 16 bits, as is often required in digital signal processing, it has been common to digitize to only one bit; the amplitude is ignored and the signal's polarity determines whether the value of the bit is a 1 or a -1. As one would expect, such a distorted representation of the signal produces a distorted version of the desired correlation function, but, as long as the signal is a Gaussian random process, the distortion can be removed perfectly after the integration is finished. The correction procedure is very simple: divide all the points (lags) by the first point (zero lag), multiply them by $\pi/2$ radians, and then replace them by their trigonometric sines. This relationship between the correlation function of the input signal and the correlation of the 1-bit representation was first described by J.H. VanVleck in a wartime study of overmodulation (clipping) for radar jammers.

Obviously 1-bit arithmetic greatly simplifies the multipliers and accumulators that form the correlator modules. Each multiplier reduces to a single exclusive-or gate and each accumulator reduces to a counter. (The same simplification doesn't work for an FFT processor; while the data could be represented with one bit, the trigonometric coefficients still need to be multibit and, except for the first stage, the butterfly operations still use multibit \times multibit multiplication). Simple one-bit correlation, however, does have its price: more integration time, by a factor of $\pi^2/4$ or 2.47, is required if the averaging is to be as effective as if the signal had been multibit. Going from a 1-bit digitizer to a 2-bit digitizer reduces the 2.47 integration time penalty to only 1.29 but obviously complicates the multiplier/accumulator circuitry. A three-level digitizer ("1.6 bits") is often used; it permits multiplier/accumulator circuitry almost as simple as the 1-bit case and yields an integration time penalty of 1.51. In all of these quantization schemes the distortion of the ACF can be corrected [2].

Some of the integration time penalty can be recovered. When the signal is oversampled by a factor of 2, i.e. sampling at a rate of 4 times the bandwidth, the factors of 2.47, 1.29, and 1.51 reduce to 1.82, 1.14, and 1.26, respectively. Very little is gained by going to still

faster sampling rates.

5 Digital Filtering

The canonical transversal digital filter uses a tapped delay line (shift register). Successively delayed samples of the signal are taken from these taps and multiplied by constant weighting coefficients. The weighted signals are combined in a single multi-input adder. Each new input sample pushes data down the shift register to produce a new output value from the adder. The correlator architecture described above provides another way to implement the transversal filter. The input signal provides the immediate data to an N-lag correlator. The filter coefficients are generated serially (e.g. from a ROM addressed by a counter) and form the delayed "data." After each input sample one correlator module will contain the desired convolution of the last N data values and the N coefficients. Its accumulator contents are latched into an output register and its accumulator is zeroed. After the next input sample, the output register is loaded from the adjacent correlator module. The process continues indefinitely, with the correlator modules being read out and rezeroed cyclicly.

At the Arecibo Observatory an 8-bit \times 1-bit correlator was used this way as a transversal filter to produce radar maps of the surface of Venus. The c.w. transmitted signal is phase-coded with a long (4095, say) pseudo-noise sequence to get the bandwidth needed for range resolution. The received signal (radar echo) is decoded by the transversal filter. The filter coefficients are the plus ones and minus ones of the code so the code generator itself provides these coefficients in the required serial sequence. With the code properly phased, the first lag module of the correlator produces the signal reflected from the front cap of the planet. The next lag module produces the signal corresponding to the next range ring, etc [3]. The lags are read in order and the resulting data sequence is identical to what would have been obtained by transmitting a powerful short pulse (pulse length equal to the baud length) and using no decoder. Good range resolution, essentially equal to the baud length, is a consequence of the very sharp peak resulting when the pseudo-noise code is convolved with itself. The code sequence must be long enough to avoid aliasing other ranges in each decoder channel, i.e. the code length is deeper than the radius of the planet. In this application our correlator had fewer lag modules than the number of bauds in the code sequence - an economic constraint. The code was phased such that the ranges of interest were decoded by the implemented modules; other ranges landed in absent modules (which were, of course, just ignored). This trade-off of ranges for less hardware would not have been possible with the canonical delay line transversal filter.

6 VLSI Implementations

Correlators were implemented first with discrete transistor logic and then with all the standard families of integrated logic elements. A gate-array correlator chip designed by Albert Bos (4) has 16 lags running at 50 MHz. More recently, Brian Von Herzen (5) produced a VLSI chip with 320 lags that runs at 250 MHz. John Canaris (6), of the VLSI design group at the University of Idaho (now at the University of New Mexico), is designing a 1024-lag

correlator chip with 3-level quantization to run at 100 MHz. A commercial product, the Zoran ZR33891 Digital Filter Processor is an 8-lag correlator chip that does $9\text{-bit} \times 9\text{-bit}$ multiplications at a 20 MHz rate and has 26-bit accumulators. As explained above, radio astronomy applications favor designs use much coarser quantization in order to have many more lags and higher speed.

7 Multi-chip Expansions

The basic correlator chip shown in Figure 1 can be used in a variety of ways to build expanded systems. Perhaps the most obvious is series expansion (simple cascading), which only requires that the chip have an output pin from its last delay element to provide the delayed data for the following chip. Time multiplexing is a parallel expansion technique that provides both higher speed and more lags. For example, suppose our N -lag correlator chip is able to accept data at only half the required rate. We can split the data stream into two half-speed streams, the even samples and the odd samples, and then use four chips to calculate the two possible autocorrelation functions (ACFs) and two possible cross-correlation functions (CCFs). When the integrated results are dumped to the computer, the sum of the ACFs gives us N even lags and the sum of the CCFs gives us N odd lags. By going from one to four chips we've doubled the speed and doubled the number of lags. Generalizing, we can divide data into N streams and use N^2 correlators to form all the possible lagged products. Another interesting form of parallelism, distributed arithmetic (from the distributive law of multiplication), was described by Zohar (7) as a way by which signals with many bits can be correlated using simple 1-bit, 2-bit, etc, correlator chips. Since signed numbers complicate matters, consider first a digitizer that encodes positive integers into, say, standard 8-bit binary representation. Group the eight bits into four 2-bit digits. We could multiply two such 4-digit numbers by separately multiplying each digit of one by each of the digits of the other. This could be done in parallel by 4^2 or sixteen 2-bit multipliers. Sixteen 2-bit correlators could independently form sub-products and accumulate them during an integration cycle. The results would be read out, appropriately weighted, and added to accomplish $8\text{-bit} \times 8\text{-bit}$ correlation. This technique will work with signed numbers if we have either

1. a number system where the place weights carry signs, or
2. a number system where the digit values possible for each place are signed.

An example of the first case is the negabinary system (radix -2) where the digit values are just 0 and 1 but the place weights are 1, -2, 4, -8, etc. An example of the second case is the balanced ternary system where the values are -1, 0 and 1 and the place weights are 1, 3, 9, 27, etc. For the negabinary system we would need 1-bit correlators with AND gates as multipliers rather than the EX-OR gates of the standard 1-bit correlator. For the balanced ternary system the popular 3-level correlator is just the right correlator element. Four 3-level (1.6-bit) correlators can run in parallel to correlate 9-level (3.2-bit) numbers. Correlators using many bits provide wide instantaneous dynamic range, especially if the quantization levels are exponentially spaced rather than uniformly spaced.

8 Summary

Digital VLSI correlator chips using coarse quantization and having many lag units are simple building blocks for spectrometers, convolvers, and digital filters. Chips can be paralleled to increase the number of channels, speed, and dynamic range.

References

- [1] Weinreb, Sander, "A Digital Spectral Analysis Technique and its Application to Radio Astronomy," Technical Report 412, MIT Research Laboratory of Electronics, August, 1963.
- [2] Hagen, J.B. and D.T. Farley, "Digital-correlation techniques in radio science," *Radio Science*, Vol. 8, Numbers 9.9, pp 775- 784, August-September 1973.
- [3] Hagen, Jon B. "Communication Techniques in Radio Physics and Astronomy," *IEEE Communications Magazine*, October 1986 - Vol. 24, No. 10, pp 16-20.
- [4] Bos, Albert, "A High Speed 2-Bit Correlator Chip for Radio Astronomy," *IEEE Transactions on Instrumentation and Measurement*, Vol. 40, No. 3, June 1991.
- [5] Von Herzen, Brian, "VLSI Partitioning of a 2-Gs/s Digital Spectrometer," *IEEE Journal of Solid-State Circuits*, Vol. 26, No. 5, May 1991.
- [6] Canaris, John, "NASA SERC Digital Correlator Projects" Fourth NASA Symposium on VLSI Design, October 1992.
- [7] Zohar, Shalhav, "The Role of the 1 Number System in Multibit Hardware Correlators," *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 37, No. 10, October 1989.