Comparing Statistics of Maximum Likelihood, Gardner, and Band Edge Filter Timing Recovery

Fred J. Harris, Ph.D.
ECE Department
University of California
La Jolla, CA, USA
fjharris@eng.UCSD.edu

Chris Dick
Signal Processing Division
Xilinx Corp.
San Jose, CA, USA
chris.dick@xilinx.com

Uma S. Jha, Ph.D.
Space and Airborne Systems
Raytheon Co.
El Segundo, CA, USA
uma.s.jha@raytheon.com

***Abstract*- The signal delivered to a receiver is a distorted, time jittered, noisy version of a waveform created by the transmitter as a weighted sum of time offset known bandwidth limited symbol wave shapes. The receiver forms a locally generated estimate of the symbol clock frequency and phase aligned with the optimal sample time instances of the receiver's correlator wave shape. Many techniques have been developed for timing recovery, the most well-known being Maximum Likelihood, Gardner, and Band Edge. We examine and compare important statistics of the three processes and discuss their influence on the timing synchronization process.**

Keywords—timing recovery, band edge, Gardner, maximum likelihood

I. Introduction

In a QAM modem, as shown in Figure 1, the modulation process starts with periodic presentation of data d(n), b-bit binary words that are mapped to one of 2^b ordered pairs of the form $x_k + j\ y_k$. The x_k and y_k are the amplitudes assigned to the scaled h(n) of the I(n), In-phase, and the scaled h(n) of the Q(n), Quadrature phase, signal paths respectively. We accomplish this by scaling the sampled data impulses presented to the square-root Nyquist shaping filters h(n). The filters perform two tasks: i) shape the output time series to limit the signal's bandwidth, and ii) increase the output sample rate to M times the input symbol rate. Typically, M is usually set to 2 or 4 but can be significantly larger for a digital IF or for a Software Defined Radio (SDR) with a fixed high sample rate. The 1-to-M up-sampling recognizes that the two-sided bandwidth of the shaping filter is greater than its symbol rate and requires the higher sample rate to satisfy Nyquist. Further, raising the sample rate widens the Nyquist zone and hence the interval between spectral replicates at the filter output. This permits lower order, lower cost, analog filters following the DACs that perform the filtering of the spectral artifacts remaining after the DACs form their analog outputs.

In our model, the channel following the DAC performs the necessary output signal conditioning tasks including analog smoothing filters, quadrature up-converter mixers, intermediate frequency filter, final RF frequency up-conversion mixer, final RF filter, output driver and power amplifier and transmitter antennae that couples the signal to the physical channel. After propagation through the channel, the signal is collected by the receiver antennae, RF filtered, amplified, and down converted by an RF mixer, IF filtered, further down converted by an IF quadrature down converter, and low pass filtered and finally converted to the sampled data domain by the input ADCs.

This is where our synchronization problems start. Assuredly the clocks at the transmitter and the receiver are not the same frequency nor the same phase once they are frequency locked. As shown in Fig. 1, the timing recovery system must contain an interpolator to compute samples of the matched filter output at the desired time sample locations from the time samples formed by the free running ADC clock. The interpolator is controlled by a phase locked loop (PLL) that locks its time sample locations to desired attributes of the received signal. A problem is there is not a periodic component in the received signal to which the PLL can lock. Even though the presentation rate of the shaping filter waveform is controlled by a periodic clock the power spectrum of the composite waveform is continuous and does not contain spectral lines at the symbol rate. The modulated signal must be subjected to a nonlinearity to form a spectral line related to the symbol clock. A PLL can then extract the discrete line from the noise like signal contribution called modulation noise. In practice, various combinations of prefiltering prior to the nonlinear operation are used to reduce the modulation noise bandwidth presented to the PLL.

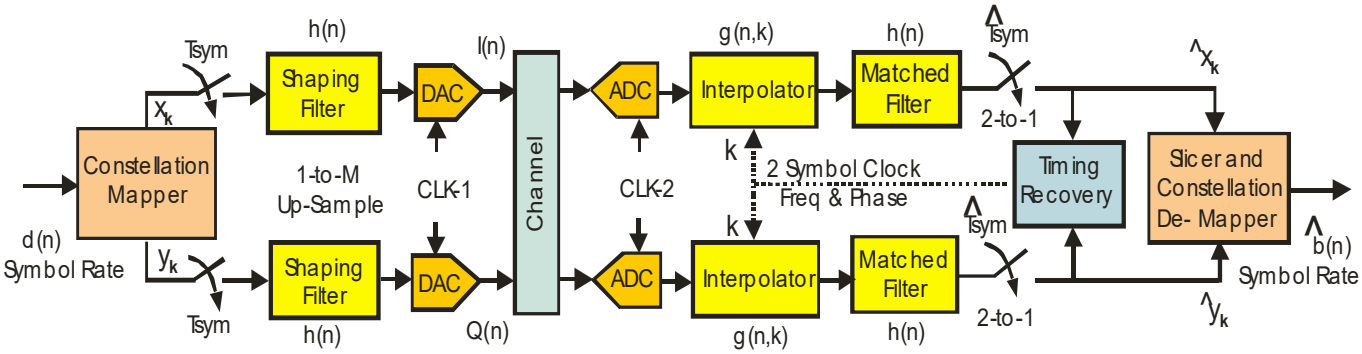

Figure 1. Block Diagram of Modulator and Demodulator Emphasizing Clocks at Input to Shaping Filter and Output of Matched Filter

II. Review of Timing Error Detectors

Most of the methods to generate the spectral line fall into one of three major categories of timing recovery. These are, maximum likelihood (ML) approximations, transition (or zero crossing) methods, and minimum variance schemes. The ML approximations seek the peak of the correlation function, or matched filter (MF) output and locates the peak by forming the derivative MF at the test points. The derivative MF is zero at the peak location and the zero slope identifies the peak position. The MF derivative to the left side of a positive valued peak is positive, telling us that we are climbing up a hill, that the peak is in front of us and we should move forward. Similarly, the MF derivative to the right side of the positive valued peak is negative, telling us that we are climbing down a hill, that the peak is behind us and that we should move backwards. The slopes to either side of a negative peak have the reverse polarity and we must alter the guidance rules. We do so by qualifying the derivative MF with the MF amplitude by forming the product $\dot{y}(\tau)\, y(\tau)$ and then adjust the time offset τ to drive the average value of the $\dot{y}$ y product to zero when $\dot{y}(\tau)=0$. When we perform the product of the MF and derivative MF impulse responses we obtain the s-curve. In the past, the derivative MF was approximated by the difference between an early and late MF. This is the standard ad-hoc approximation to the ML timing estimator shown in Eqs. (1) where $\nu(\tau)$ is MF output normalized by its peak value and further scaled by its signal to noise ratio. Note that at low SNR the $\tanh(\nu(\tau)$ is approximately $\nu(\tau)$ and at high SNR the $\tanh(\nu(\tau)$ is approximately $\mathrm{sgn}(\nu(\tau))$, which match except for the SNR scaling, the common product approximations, $\dot{y}(\tau)\, y(\tau)$ and $\dot{y}(\tau)\, \mathrm{sgn}[y(\tau)]$.

$$\nu(\tau) = \frac{2E_b}{N_0}\frac{y(\tau)}{y_{pk}} \tag{1a}$$

$$\Lambda_{\ln}(\tau) = \sum_k \ln\left[\cosh[\nu(\tau)]\right] \tag{1b}$$

$$\frac{d}{d\tau}\Lambda_{\ln}(\tau) = \sum_k \nu(\tau)\tanh[\nu(\tau)] = 0 \tag{1c}$$

The loop that drives the product $\dot{y}(\tau)\, y(\tau)$ to zero when $\dot{y}(\tau)$ is zero is the ML loop. By reversing the sign of the feedback signal, the same loop can also drive the $\dot{y}(\tau)\, y(\tau)$ to zero when $y(\tau)$ is zero. To distinguish this mode from the ML mode, Gardner called this option the minimum likelihood timing loop. In a QPSK signal the zero-amplitude level corresponds to the mid-point of transitions between modulation levels. Many transition detectors operate at two samples per symbol, with the second sample used to detect the transition level. We can understand the operation of a two sample per symbol transition detector by examining the negative going transitions in an eye diagram as shown in Fig. 2. Transitions are detected by a conditional test of three successive samples. The test verifies that the first sample is positive valued and that the third is negative valued. If these samples coincide with the proper sample location, the average value of the mid sample is zero. On the other hand, if the test samples were taken early with respect to the maximum eye opening, then the average value of the mid sample on the transition will be positive. Similarly, if the test samples were taken late with respect to the maximum eye opening, then the average value of the mid sample on the down sloping transition will be negative. If the test samples are all same polarity, there has not been a transition and the mid sample contains no information about the timing the sample positions relative to the symbol times. The Gardner timing error detector forms the product shown in Eq. (2), where det(y(n-1)) is the detection level ±1. The difference of the two detection levels is either ±1 or 0. The difference approximates the derivative term $\dot{y}(\tau)$ at the sample value $y(\tau)$ for the $\dot{y}(\tau)\, y(\tau)$ product.

$$p(n) = [\det(y(n-1)) - \det(y(n+1))]\cdot y(n) \tag{2}$$

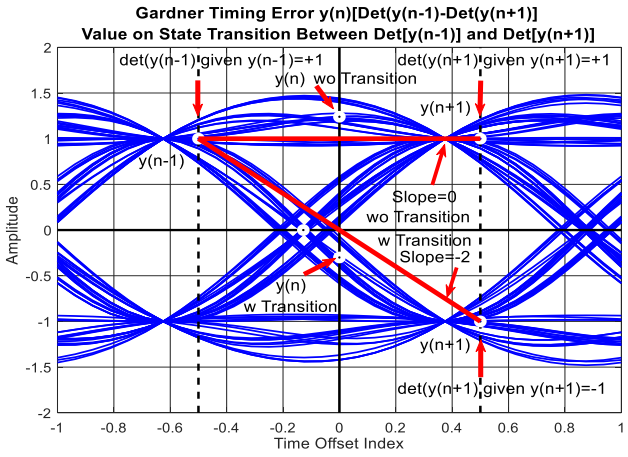

Figure 2. BPSK Eye Diagram and Sample Points for Detecting Transition Sample Values in Gardner Timing Error Detector

A property of a cyclo-stationary QAM signal with periodic random symbol modulation is that, while it's spectrum contains no spectral lines, the conjugate product of the signal with itself contains spectral lines at symbol rate. These spectral lines are formed by the interaction of the band edge conjugate spectra during the spectral correlation caused by the time domain product. This is demonstrated in Fig. 3 where the top subplot shows the spectrum of a QPSK signal and the center subplot shows the spectrum of the signal's conjugate product with itself. We clearly see the spectral line as well as the DC line since the magnitude square has a large DC component. We can suppress the DC by first differentiating the signal and then form the conjugate product of the signal and its derivative. The third subplot show the spectrum of this product with it spectral line but without the DC line. Now we seek to suppress the large spectral contribution between the spectral lines.

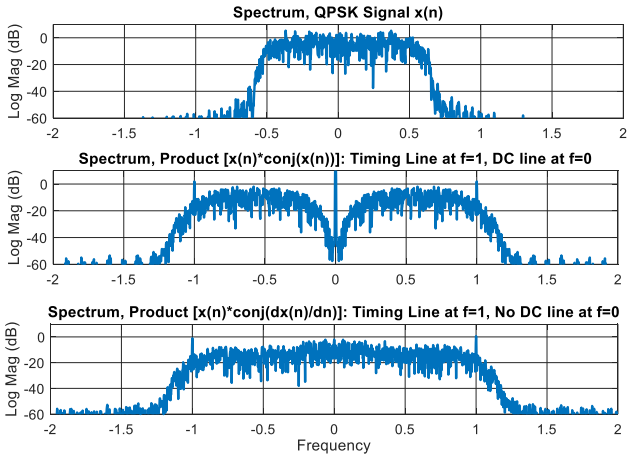

Figure 3. Spectra of QPSK Signal s(n), of Conjugate Product s(n) and s(n), and of Conjugate Product s(n) and its derivative ds(n)/dn.

The timing lines can be emphasized while suppressing the central modulation spectral mass with a pair of band edge filters. As shown in Fig. 4, we form a pair of band edge filters BE_{POS} and BE_{NEG} for the positive and negative frequency band edges and form BE_{SUM} and BE_{DIFF} as the sum and difference respectively of the two band edge output series. Finally, we form the conjugate product of the BE_{SUM} and BE_{NEG} series. The spectrum of the imaginary part of this product term contains the timing line without the modulation spectrum. We address the question: "How much of the modulation noise spectrum is present in the ML and the Gardner time error detectors?"

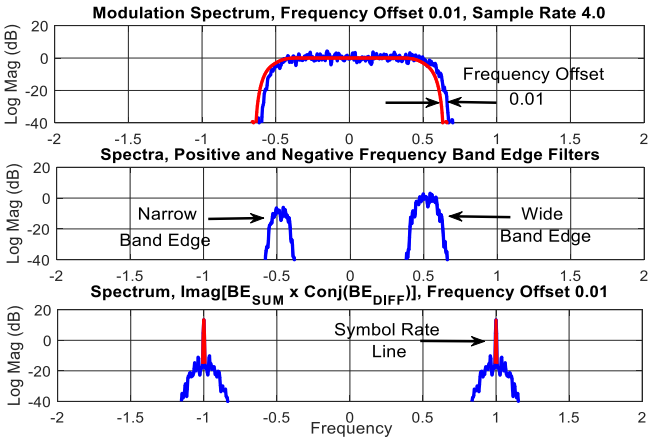

Figure 4. Spectra of QPSK Signal, Small Frequency Offset 0.01, of two BE Filter Signals and of Imaginary Part Conjugate Product [BE_{SUM} x BE_{DIFF}].

III. Spectral Lines and Timing Error Detector S-Curves

We first introduce the spectrum of the timing error detectors applied to a QPSK signal for three different values of excess bandwidth α. We will then examine the S-curves and the modulation noise variance of the three timing error detectors for the selected α values. Figure 5 presents the QPSK signal eye diagrams for the three α values. Take note of the reduced variation of transitions between states with increased excess bandwidth. Figures 6a, 6b, and 6c present these spectra. We can easily see the spectra of the modulation noise accompanying the desired timing lines formed by the ML and Gardner timing error detectors. Carful examination of the tone shows that its amplitude increases with increased excess bandwidth. This suggests that the timing recovery loops might prefer the larger excess bandwidth which offers the larger amplitude tones which enable the loop to maintain lock at lower SNR.

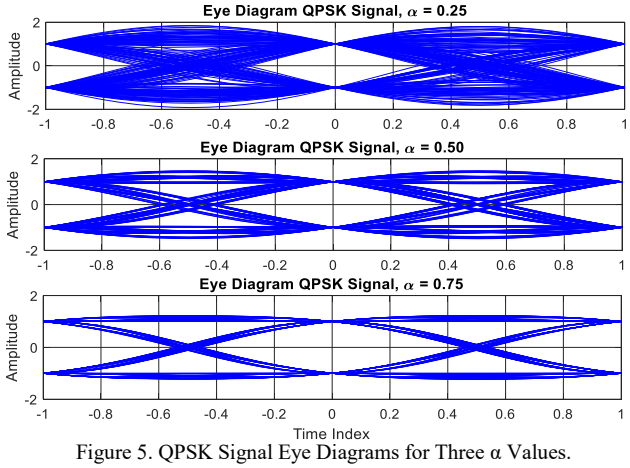

Figure 5. QPSK Signal Eye Diagrams for Three α Values.

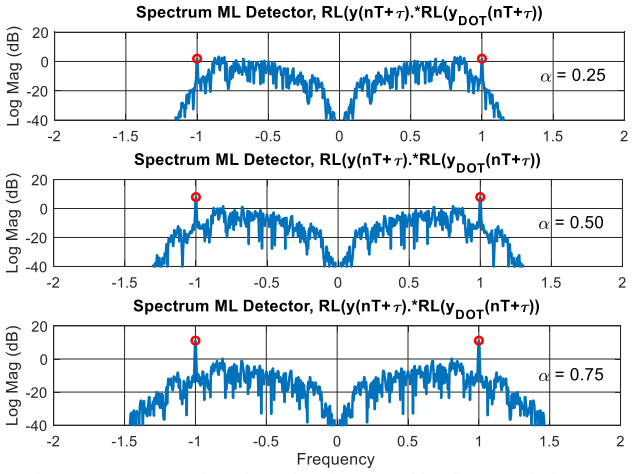

Figure 6a. Spectra of Product Signals Formed by the ML Timing Error Detector, y(n) and derivative y(n) for the Three α Values.

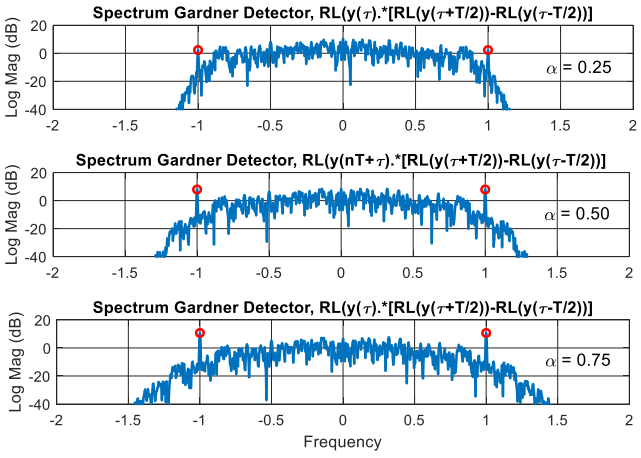

Figure 6b. Spectra of Product Signals Formed by the Gardner Timing Error Detector, y(n) [det(y(n-1)-det(y(n+1)] for the Three α Values.

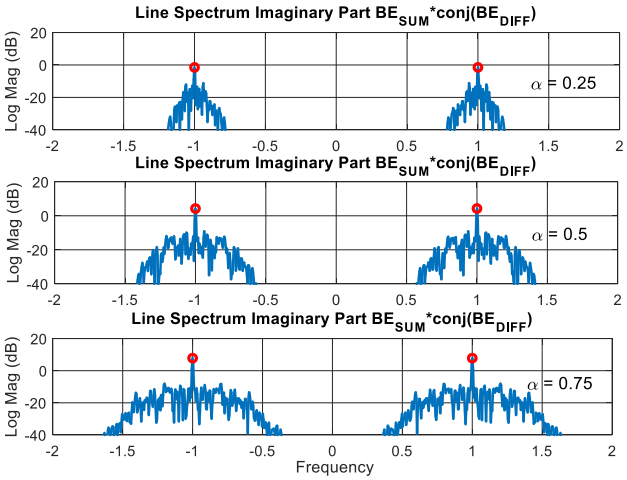

Figure 6c. Spectra of Product Signals Formed by the Band Edge Filter Timing Error Detector, BE_{SUM} Conj[BE_{DIF}] for the Three α Values.

Figures 7a, 7b, and 7c present the S-curves formed by the three timing error detectors for the 3 α values. Each of the S-curve figures is accompanied by a curve showing the modulation noise variance contours for the three α values. We note that

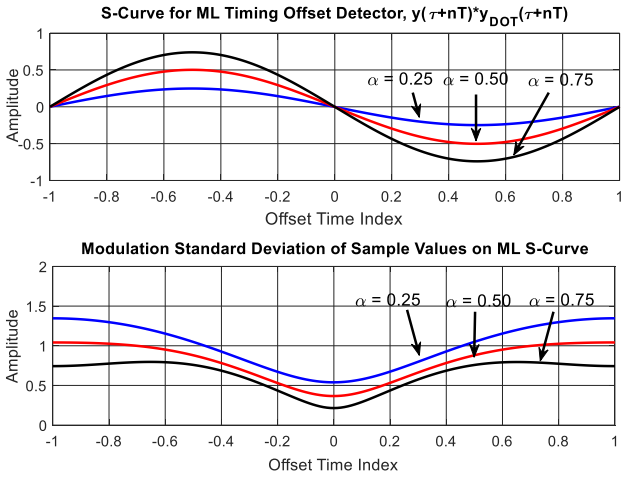

Figure 7a. S-Curve Formed by the ML Band Edge Filter Timing Error Detector, and Variance as function of Timing Offset for the Three α Values.

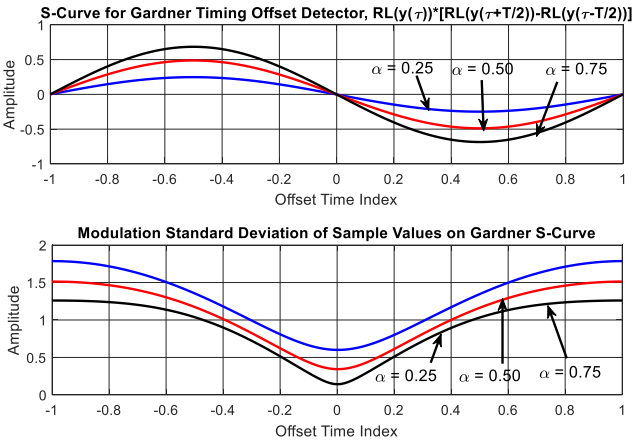

Figure 7b. S-Curve Formed by the Gardner Band Edge Filter Timing Error Detector, and Variance as function of Timing Offset for the Three α Values.

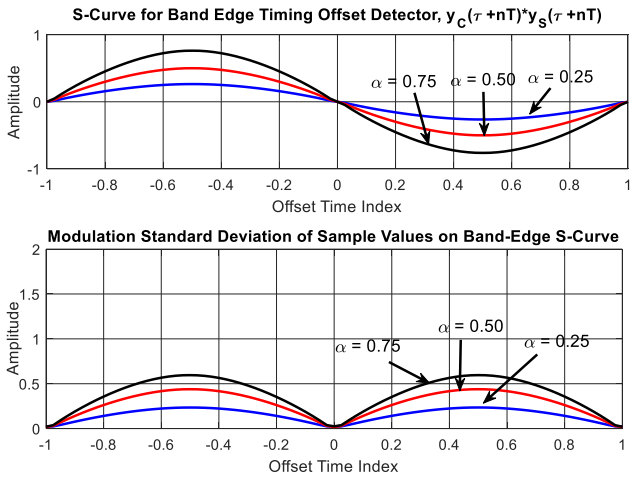

Figure 7c. S-Curve Formed by the Band Edge Filter Timing Error Detector, and Variance as function of Timing Offset for the Three α Values.

the noise level increase when the timing detector operates offset from the desired timing position. What we see here is the noise due to the state transitions as the signal moves between signal states. These transitions are the modulation noise in the timing error detectors. We then observe that, as we noted from the spectral plots, larger values of excess bandwidth increase the amplitude swing of the S-curves. Comparing the amplitude of the S curves in the three timing error detectors we see they are essentially the same size. They are all harvesting the cyclostationary energy component in the received signal and they all do so equally well. The primary difference in the detector, is not how much tone signal they extract, but how much noise they gather while extracting that tone.

For ease of comparison, we have collected the variance profiles of the timing error detectors and placed them in Fig. 8. What we learn here is that for α=0.25, the narrow excess BW, the ML detector has smaller modulation noise levels than the Gardner detector. and that at α=0.50 the two detectors have the same levels of modulation noise. What is impressive is the significantly reduced levels of modulation noise of the band edge filter timing error detector. The reduction is more than an order of magnitude for α=0.25. We also note that the band edge filter noise level at its minima does not change with different amounts of excess bandwidth as it does for the ML and Gardner detectors. Interesting that the noise levels at the minima of the ML and Gardner detectors decrease with increased bandwidth. This reduction is expected as we noted as in Fig. 5 we noted the reduced variation in transitions between states.

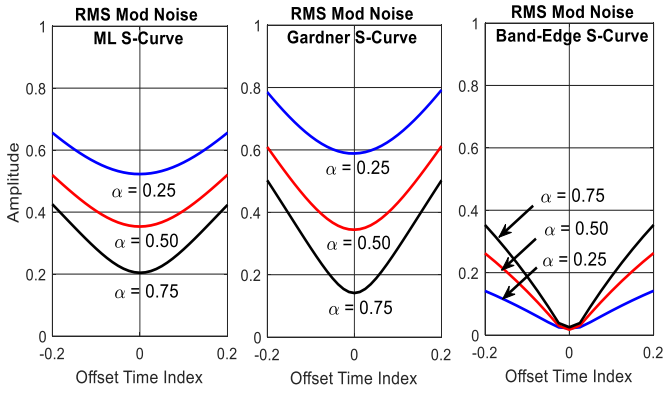

Figure 8. Variance Curves of Timing Error Detectors Showing Increased Noise Levels with Timing Offsets from Optimum Timing position

IV. Closing Comments

In this paper we reviewed and examined the modulation noise associated with the well-known ML and Gardner timing error detectors. These detectors form the product of the amplitude and derivative, y(t) ẏ(t), at the time sample test location to guide a PLL to the desired time sample location. The ML loop drives to the sample location for which ẏ(t) is zero, and the Gardner loop drives to the sample location for which y(t) is zero. These points correspond to the maximum eye opening and the zero crossing of eye transitions between states. Our intuition led us to believe that since minimum variance occurs at the maximum eye opening we expect lower variance for the ML timing detector. As it turns out, our intuition was faulty. We discovered that over a range of small excess BW, ML is better than Gardner, over a range of medium excess bandwidth the two are comparable, and over a range of large excess BW Gardner is better than ML. We also observed that the amplitude of the timing lines increases, and timing variance decreases for larger excess BW. This was an expected result due to increased energy in the excess BW and due to reduced variation in transitions between states respectively. We examined the spectra of y(t) ẏ(t), the product term and found modulation noise dominated the spectrum and would have to be rejected by subsequent processing to access the timing lines. Gardner spectra for a range of α are seen in Fig. 9. We included, for comparison, the corresponding band-edge filter spectra. This detector option suppresses the modulation noise prior to forming the timing line rather than after. The spectra for the band-edge version are seen in Fig. 10.

Some comments about the subplots in Fig. 11. Here we have the eye diagrams of the signal and of the time series obtained from the band edge filters. We see here that the zeros of the band edge filter eye diagrams are aligned with the peak and zero crossings of the signal eye diagram. The product of the band-edge signals, shown in the bottom subplot, is seen to be a variable amplitude sinewave phase locked to the modulation clock. The average of these random amplitude sinewaves is the band-edge S-curve. The lesson we learn here is that the timing information we harvest from the received signal resides in the band edges of the signal. We should avoid being cavalier with shaping filter's excess BW. The timing recovery Genie lives on the energy in that excess BW.

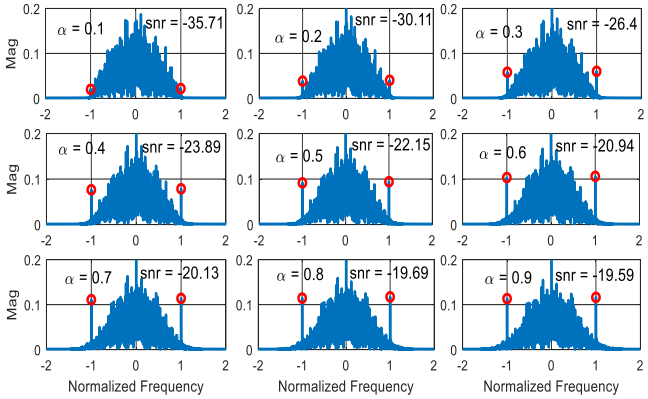

Figure 9. Spectra: Gardner Timing Line and Modulation Noise

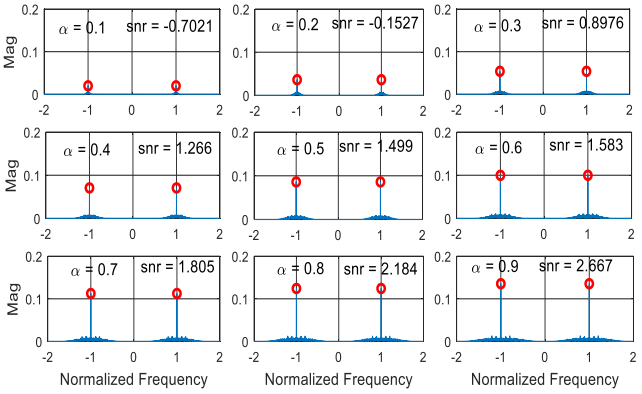

Figure 10. Spectra: Band-Edge Timing Line and Modulation Noise

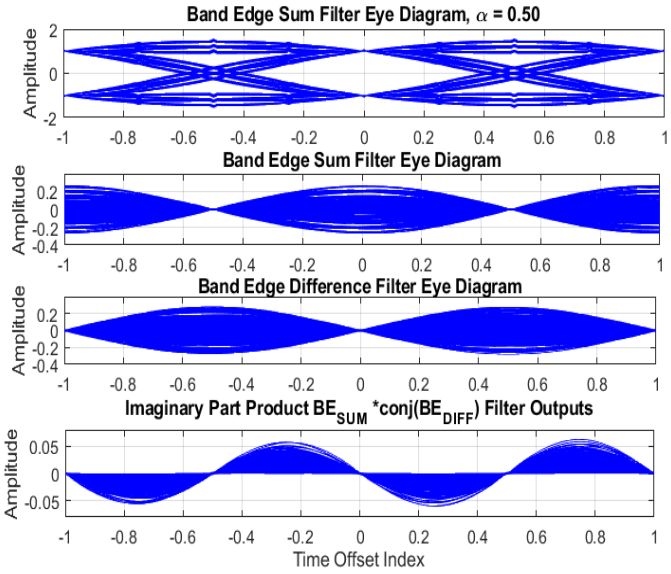

Figure 11. Eye Diagrams of Signal and of BE_{SUM} and BE_{DIFF} and S-Curve Formed by Imaginary Part of Conjugate Product [BE_{SUM} x Conj(BE_{DIFF})].

Our final comment, is that we did not address the performance of the time error detectors when the constellation is rotating due to frequency offsets. The timing line formed by the ML and Gardner detectors exhibits significant reduction in ampli-

tude. Frequency offsets and spinning constellations have negligible effect on the band-edge timing line amplitude. In fact, the real part of the conjugate cross product contains a DC line with amplitude proportional to the frequency offset and is normally used to operate a frequency locked loop.

V. References

[1] Heinrich Meyer, Marc Moeneclaey, and Stephan Fechtel, Chapter 8, "Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing", Wiley Inter-science, 1998.

[2] Richard Blahut, Chapter 8.5, "Modem Theory: An Introduction to Telecommunications", Cambridge University Press, 2010.

[3] fred harris, "Band Edge Filtering and Processing for Timing and Carrier Recovery", COMCON-7, Athens Greece, 28-June, 2-July 1999.

[4] fred harris, "Band Edge Filters: Characteristics and Performance in Carrier and Symbol Synchronization", WPMC-2010, Recife, Brazil, Oct 11-14, 2010.

[5] fred harris, Elettra Venosa, Xiaofei Chan, and Chris Dick, "Band Edge Filters Perform Non-Data Aided Carrier and Timing Synchronization of Software Defined Radio QAM Receivers", WPMC-2012, Taipei, Taiwan, Sept 24-27,2012.

[6] fred harris, Elettra Venosa, Xiaofei Chan, and Chris Dick, "Non-Data Aided Carrier and Timing Synchronization via 2-Samples per Symbol Band-Edge Filters", 46-th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 5-7 November 2012.