

Let's Assume the System is Synchronized

fred harris

San Diego State University & Aalborg University

fred.harris@sdsu.edu

Abstract

We arm our graduate students with the best tools and our best vision to enable them to invent, to improve, and to create the wireless world of tomorrow. To better compete and succeed in the coming decades we give them the best education in signal processing and communication systems. We do, don't we? Perhaps not! This paper calls attention to a gaping void and suggests that we fill that educational gap.

1. Introduction

It is amazing how many papers on radio systems, networks, error correcting codes, and related topics contain a version of the sentence "Let's assume the system is synchronized." Alright, let's assume the system is synchronized. But I have a few questions: Who did it? How did they do it? Who will do it in the next decades as many of us retire from the field? An important one is; where are they acquiring the skills required to negotiate and navigate the future physical layers?

This brings us to the question of "What do we mean by synchronize"? Its etymology starts with Chronos, (also Khronos and Chronus) the ancient Greek Immortal Being (Χρόνος) personified in the modern age as Father Time. We thus form synchronize from the Greek prefix *syn*, meaning "together with" and *chronos* which we interpret as "time".

With the industrial revolution and the ascendancy of the high speed railroad came the need to synchronize clocks in adjacent towns in order to maintain arrival and departure timetables. This helps protect trains from cornfield collisions. Today, the higher speed transport and commerce of communication signals places an even greater premium on the measurement

of time and the alignment of remote clocks and oscillators.

When discussing the importance of synchronization in my modem design class I present the needle point shown in figure 1 and remind the students that Momma's middle name is synchronizer! If the radio is not synchronized no other system can operate! Not the matched filters, not the equalizer, not the detectors, not the error correcting codes, not the decryption, not the source decoding!

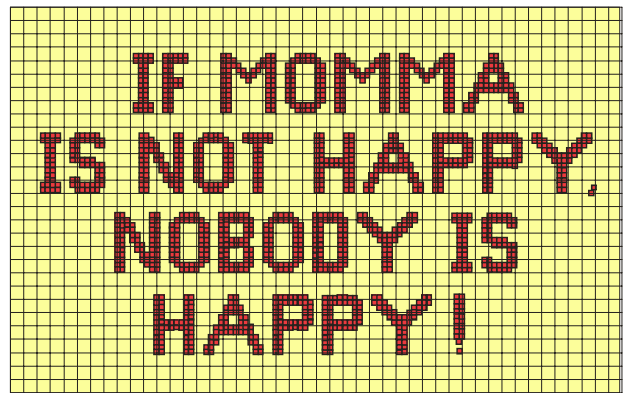


Figure 1. The Synchronizers' Needle Point

At the waveform level, synchronization entails the frequency and phase alignment of remote oscillators for carrier acquisition and tracking, for modulation symbol timing, for chip alignment and hopping boundaries of spread spectrum modulation overlays.

What have we missed by assuming the system is synchronized? We skipped a challenging and most interesting part of the process. We have skipped the task of estimating, in a short time interval, the unknown parameters of a known signal in the presence of noise. We have replaced the task of processing a noisy wave shape for the easier task of processing a binary

data stream with unknown errors. We have also skipped making Momma happy!

2. Source of the Problem

Figure 2 presents a very simple model of a communication system. Here input bits are subjected to a number of transformations at the modulator with matching inverse transformations at the demodulator. In total, the input bit stream to the modulator is converted to the radio frequency wave form at the modulator output. A communication system is seen through the eyes of the beholder and different eyes see different systems. Some see the system from the Shannon perspective, as a discrete modulator and demodulator connected to a discrete channel as shown in figure 3. In Shannon's model, the discrete modulator and demodulator perform a set of transformations on discrete sequences as shown in figure 4.

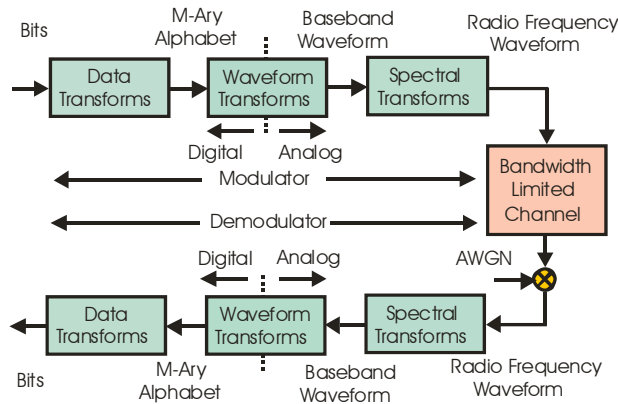


Figure 2. Simple Communication System

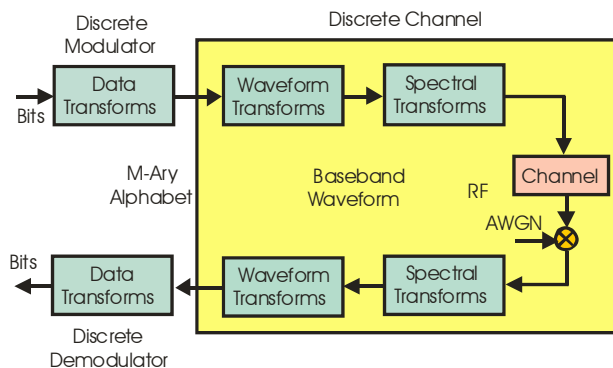


Figure 3. Discrete Communication System

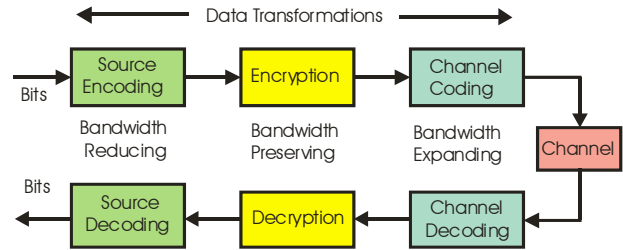


Figure 4. Shannon's Communication System

What Shannon did was brilliant: his model abstracted the system. His model separated the communication system from the physical system and the discipline blossomed under this separation. Truly remarkable advances in communication systems can be traced to developments in this model space. I fear that since there are no wave shapes in the Shannon model we have lost sight of their importance to the process. The model's success inadvertently relegates the wave shapes in the physical system to a secondary status in the academic community.

At the other extreme of the model space is the hardware model that emphasizes the signal conditioning and signal processing of communication waveforms exiting the modulator and entering the demodulator. Such a model is shown in figure 5. In this model the discrete transformations emphasized in the Shannon model are located in the processing blocks titled Bit-Map at the input to the transmitter and at the output of the receiver.

The transmitter of the physical layer model explicitly shows the shaping filter, the up-converter, and the output power amplifier which perform the base band spectral shaping, the linear RF spectral transformations, and non-linear spectral transformations respectively. We note the non-symmetry of the transmitter and the receiver. The receiver contains many more subsystems than does the transmitter. These subsystems are seen to be servo control loops that participate in the signal conditioning required to demodulate the input waveform. These loops estimate the unknown parameters of the known input signal and invoke corrective

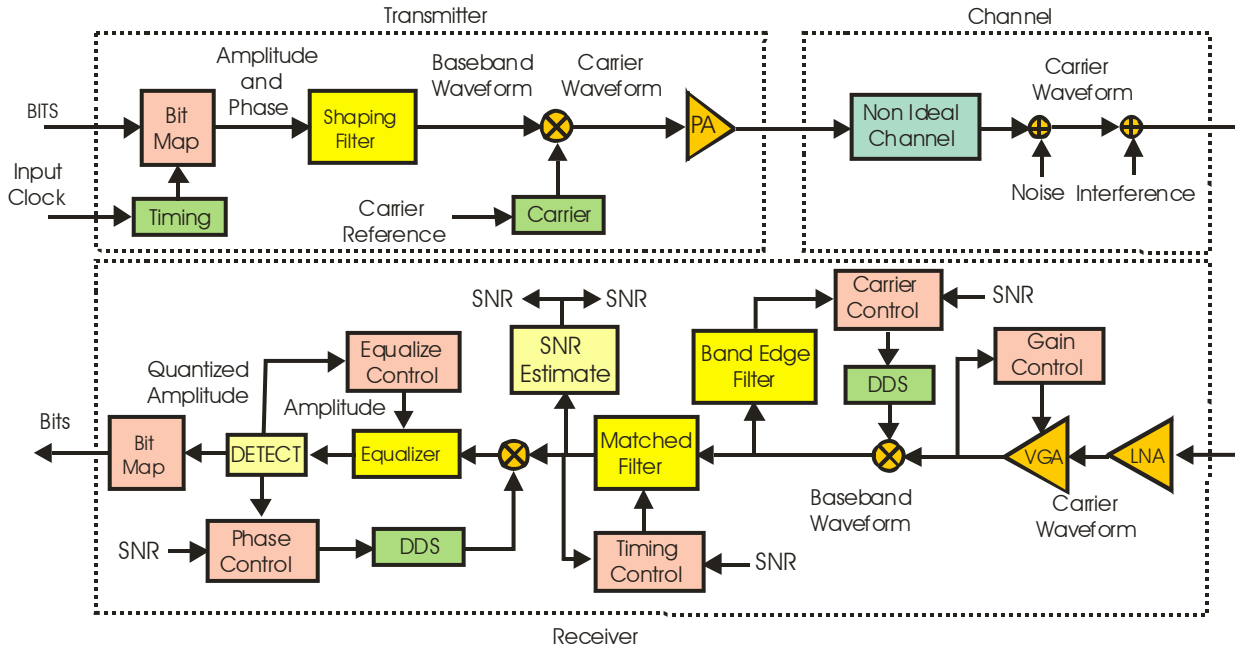


Figure 5. Physical layer model of Transmitter and Receiver.

signal processing and signal conditioning operations to ameliorate their degrading effects on the demodulation process.

A partial list of these processes includes the following subsystems: i.) an AGC loop to estimate and remove the unknown channel attenuation, ii.) a carrier recovery loop to estimate and remove the unknown frequency offset between the input signal's nominal and actual carrier frequency, iii.) a timing recovery loop to estimate and remove unknown time offsets between the receiver sampling clock and the optimal sample positions of the matched filter output series, iv.) an equalizer loop to estimate and remove unknown channel distortion responsible for inter-symbol interference, v.) a phase recovery loop to estimate and remove unknown carrier phase offset between input signal and local oscillator, and vi.) an SNR estimator to supply important side information to the just enumerated subsystems.

The closing question in this section in this section is “Where do your students learn the science and engineering embedded in the subsystems of figure 5? Where do they learn the new DSP based techniques that have supplanted legacy analog designs? We now exam-

ine an overview of some of the signal processing tasks performed during the synchronization process.

3. Phase Lock Loop, Timing Recovery:

The function of the various loops in the receiver is to estimate the various unknown parameters of the received noisy signal. The estimate $\hat{\theta}$ is based on observations of a sampled data sequence $y(n)$ derived from the output of a matched filter that acts to reduce the effects of the additive noise.

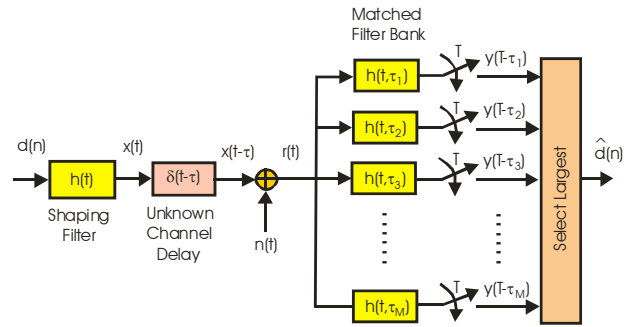


Figure 6. Bank of Matched Filters Parameterized over Time Delay Variable τ .

The estimate can be obtained from a bank of matched filters parameterized over the un-

known variable such a time delay τ . This option is shown in figure 6. The outputs of the filter bank at specific symbol time increments nT are subjected to a detector and are averaged to obtain stable statistics. The smoothed outputs are compared and the filter with the largest output magnitude is the one matched to the signal time delay τ_k .

In modern receivers, the filter bank is available to the receiver as the paths of an M-path polyphase filter. Rather than operate all the paths simultaneously, they are operated sequentially in response to side information which guides a state machine to the peak of the correlation function. This side information is the slope at the output of each hypothesized filter selection. The system selects any filter in the bank and tests the hypothesis this it is the correct filter. It does that by forming the derivative at the test point. In legacy receivers the derivative was estimated from early and late matched filters bracketing the test point in question. In modern receivers the derivative is formed by a derivative matched filter bank. The derivative

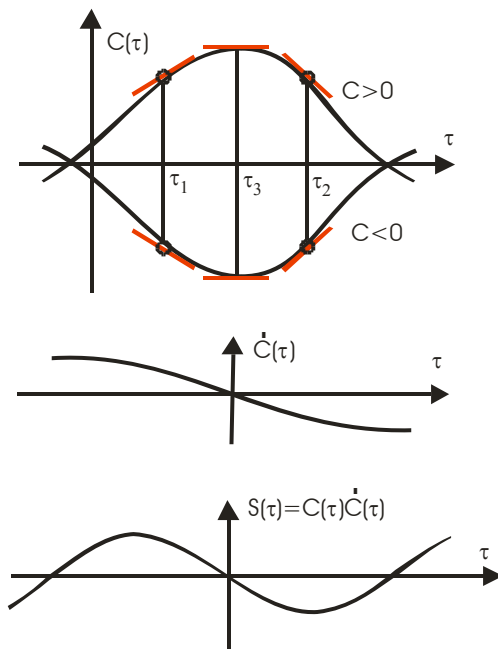


Figure 7. Correlation Function showing Slope at Various Test Points and Detector S-Curve formed as product $c(\tau)\dot{c}(\tau)$.

at selected points of a correlation function is shown in figure 7 for positive valued and for negative valued correlations. Note that for a positive valued correlation a positive slope indicates the peak is ahead of the test point and a negative slope indicates the peak is behind the test point. Since the slope has the reverse polarity when the correlation value has a reverse polarity the information residing in the slope must be conditioned on the polarity of the amplitude. The amplitude conditioning of the slope is embedded in a detector S-curve formed as the product of the amplitude and the slope.

The state machine is designed to move the hypothesis test point in the direction that sets $c(\tau)\dot{c}(\tau)$ to zero. Note this happens at two locations! If $\dot{c}(\tau)$ is zero, we are at the peak of the correlation function and the system is a maximum likelihood estimator and if $c(\tau)$ is zero, we are at the zero of the correlation function and the system is denoted the Gardner or minimum likelihood estimator. The state machine operates as a servo system called a phase locked loop whose block diagram is shown explicitly in figure 8. The figure shows the standard components of the PLL timing loop. These include the matched filter bank and the derivative matched filter bank as well as the loop filter which averages through the received noise and modulation noise and the phase accumulator that selects the candidate hypothesis filter from the filter bank. Since the loop contains 2-integrators it is a type-2 loop, able to track a ramp input, a frequency offset, with zero steady state error.

Two other components shown in the block diagram, derived from the ML equations, are particularly interesting. These are the SNR scale factor ($2E_b/N_0$) that serves to tell the loop the quality of the signal it is processing. If the signal sample has a low SNR the input to the loop filter should be scaled in proportion to its quality. This of course reflects the philosophy of all matched filter processes. The second component is the TANH that conditions the

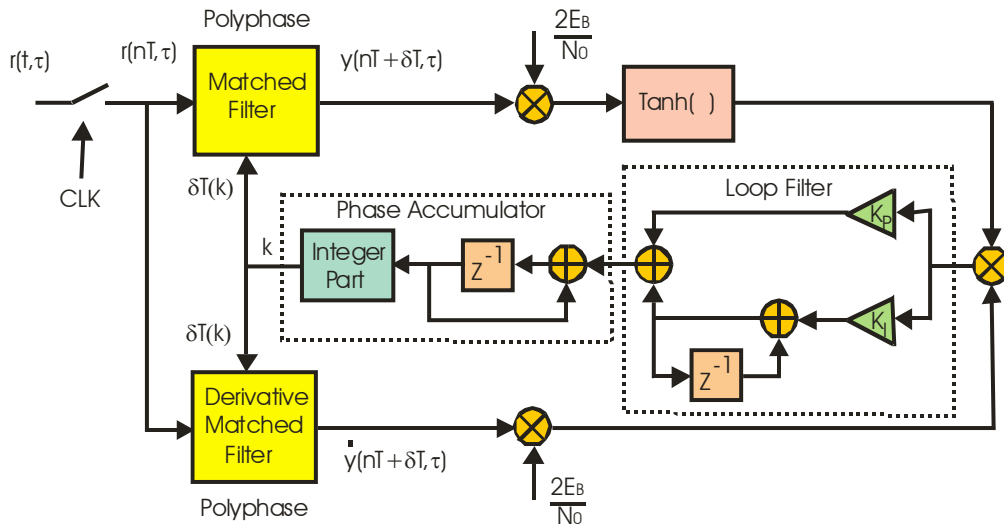


Figure 8. DSP Based Polyphase Filter Bank Maximum Likelihood Timing Recovery Loop

amplitude from the matched filter as it interacts with the derivative. At high SNR the TANH faults to the sign of the input signal as a conditional correction to the derivative information.

At low SNR the TANH defaults to a unity gain applied to the matched filter output to avoid possible errors in the sign decision of the amplitude as the conditional correction of the derivative. The SNR gain term in the loop filter throttles the loop bandwidth in response to the SNR. At low SNR the loop filter reduces its bandwidth so that it has to average over a longer interval to obtain stable control signals. Conversely, at high SNR the loop filter increases its bandwidth since it can obtain stable control signals by averaging over shorter time intervals.

The problem with these two components is that most receivers do not have real time background SNR estimators operating to feed the quality assessment of data samples to the loop filter. Thus most receivers replace the TANH with its small SNR gain and operate as sub-optimal systems. We can do better than that!

Figure 9 shows the transient response observed at the output of the phase accumulator as it moves from an initial filter path to the correct filter path in the polyphase filter bank. The red curve is the accumulator content while the

blue overlay curve is the integer part of the accumulator that defines the index pointer to the selected path of the M-path filter. Figure 10 shows the eye diagram of the signal at the input and output of the timing recovery loop.

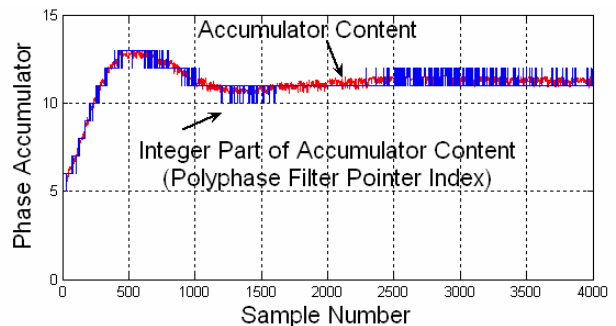


Figure 9. Transient Response of Phase Accumulator During Timing Acquisition

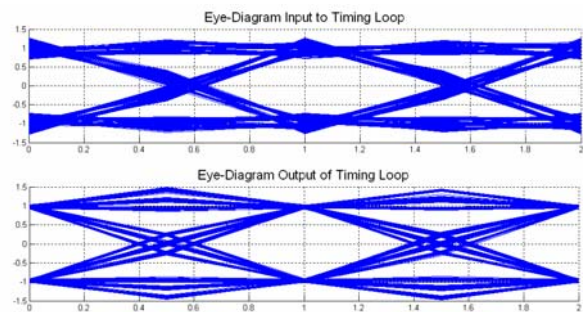


Figure 10. Eye Diagrams at Input and Output of Timing Recovery Loop

4. Phase Lock Loop, Phase Recovery:

In many modulators the signal formed by the shaping filter is amplitude and phase modulated in accord with the input bit mapping process. The amplitude and phase terms are represented in Cartesian coordinates and described as a complex base band signal. The quadrature components of the signal are up-converted or amplitude modulated on the quadrature components, the cosine and sine, of a radio frequency carrier.

At the demodulator the process is reversed and the radio frequency carrier is down converted by a pair of cosine and sine quadrature sinusoids. The frequency and phase of the up-converter and the down-converter oscillators do not match by virtue of manufacturing tolerances, age and temperature related drift, and Doppler offsets due to velocity vectors between platforms.

The signal obtained at the output of the quadrature down converter is monitored and applied to a phase detector to obtain a measure of the phase misalignment. Here we quickly describe the process by which the phase lock loop aligns the local oscillator phase with that of the phase of the received signals' underlying carrier. First we examine the phase detectors for a binary phase shift key BPSK signal. What do we do with the samples of the I – Q pair, which are time aligned with the correlation peaks of their matched filters, to obtain information about their phase θ ? What we do is examine a legacy solution formed by their product I-Q. As seen in figure 11, the product is proportional to the $\sin(\theta/2)$ which is the expression for an S-curve phase detector. The detector has two zero value references at 0° and at 180° which of course is responsible for the two-fold ambiguity in the synchronized phase lock.

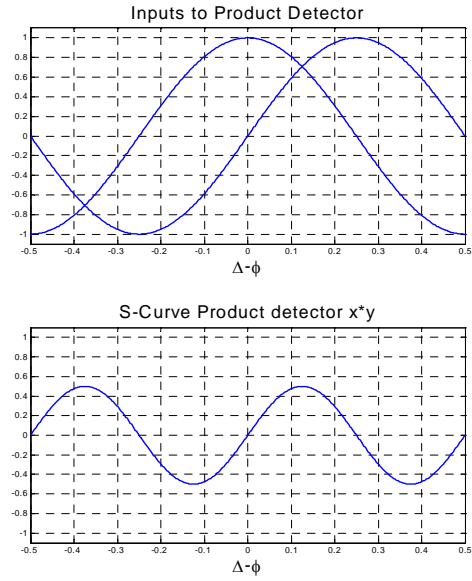


Figure 11. I-Q BPSK Phase Detector and S-Curve

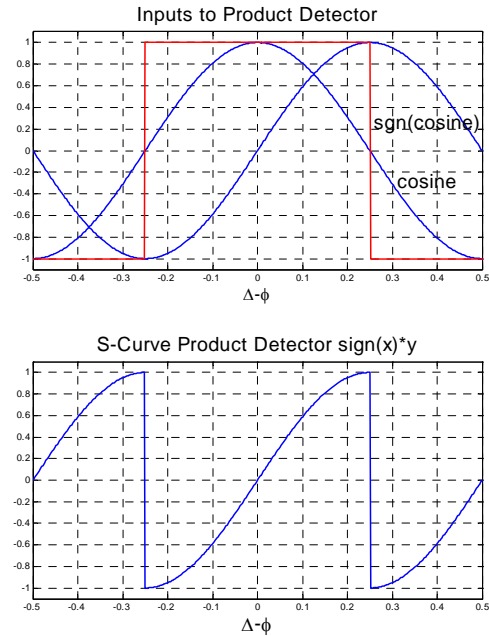


Figure 12. SGN(I)·Q BPSK Phase Detector and S-Curve

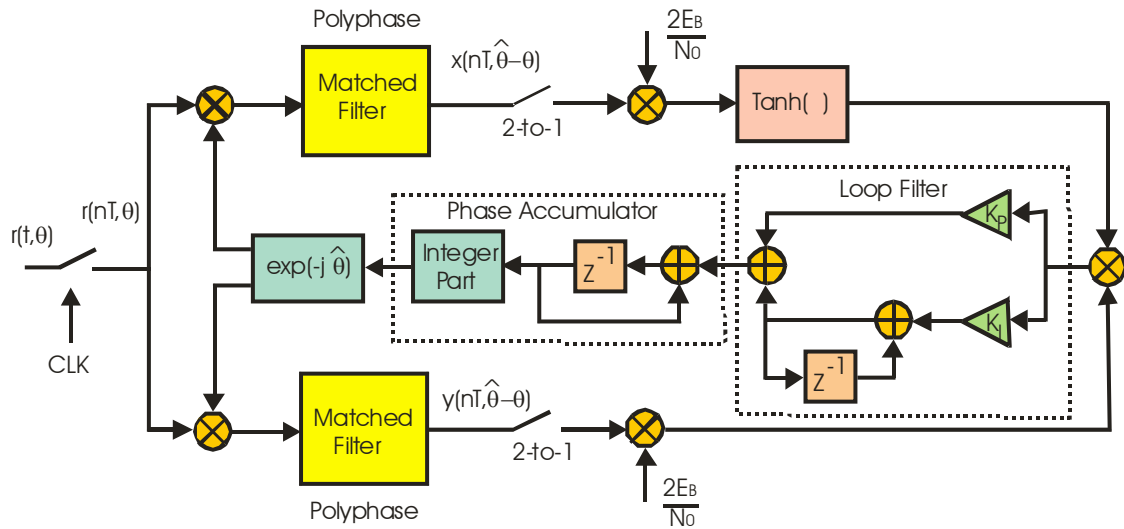


Figure 13. Maximum Likelihood Phase Recovery for BPSK Receiver

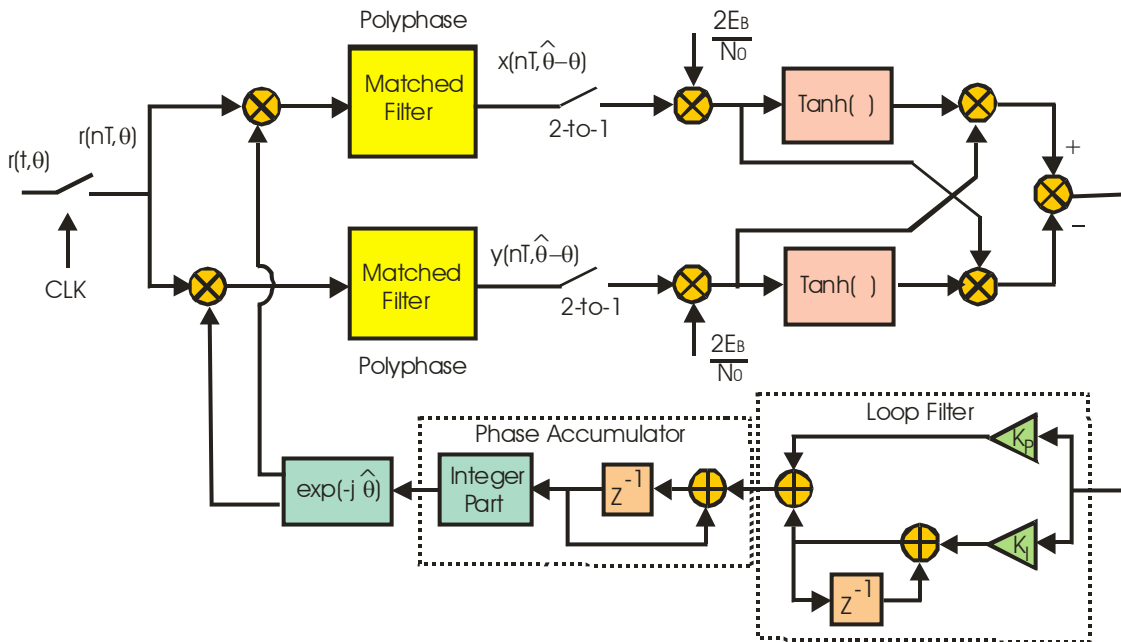


Figure 14. Maximum Likelihood Phase Recovery for QPSK Receiver

An alternate phase detector is the $\text{SGN}(I) \cdot Q$ which is shown in figure 12 where we see a more linear S-curve spanning a wider range of input phase offsets. These correspond to the small signal to noise ratio and the large signal to noise ratio approximations to the TANH described in the previous section. Thus it will come as no surprise to see the structure shown in figures 13 and 14 of a BPSK and a QPSK

receiver implementing Maximum Likelihood phase recovery.

Incidentally, we make an interesting observation with respect to figure 13 which is the ML phase estimator for the BPSK receiver. Suppose we leave off the SNR scaling factors from the two paths, and further replace the TANH of the upper path with a wire (or a unity gain path) and similarly replace the 2-to-1

down sampler on each path with a wire. What would we have? Give up? We would have a Costas loop phase recovery system. The Costas loop is embedded in many legacy receivers. Having discarded the items listed earlier we quickly conclude that the Costas loop is far from an optimum phase recovery process. In fact it is easy to verify that the sub-optimal loop performs quite poorly at low SNR. As we bravely move forward to the next decade, let us not carry legacy suboptimal solutions with us. Figure 15 presents the S curves for the ML phase detector of the QPSK receiver for a range of scaled SNR. Note the change in slopes at the zero crossings of the S-curve. These slope changes alter the loop gain of the PLL and are responsible for SNR dependent loop bandwidth of the acquisition system.

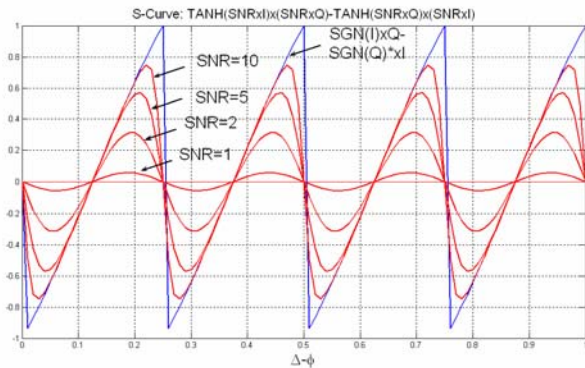


Figure 15. S-Curve for QPSK Phase Detector for Range of input SNR.

Figure 16 shows the trajectory of the QPSK constellation during Phase acquisition. The trajectory has the appearance of a comet with its trailing tail because the figure plotted the constellation points near the end of the trajectory in red to emphasize the final state of the trajectory.

As observed in the section on timing recovery, most receivers do not have real time background SNR estimators operating to feed the quality assessment of data samples to the loop filter. Thus most receivers replace the TANH with its large SNR gain and operate as

sub-optimal systems. We again note that we can do better than that!

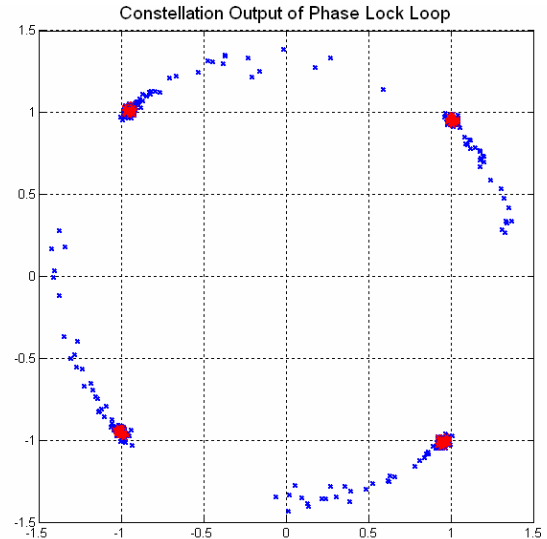


Figure 16. Constellation Trajectory while PLL Acquires Phase

5. Phase Lock Loop, Frequency Recovery:

If the frequency offset between the local oscillator used in the final down converter and the center frequency of the input signal is sufficiently small the phase detector and the phase lock loop of the previous section can acquire and de-spin the input signal. On the other hand, if the frequency offset is significantly larger than the bandwidth of the PLL loop filter the loop will not successfully acquire and de-spin the signal. In this event an acquisition aid must be invoked to assist the phase lock loop.

In legacy receivers a lock detector is interrogated after a preset time-out to see if the acquisition aid should be invoked. If invoked, the local oscillator is slowly swept through the likely range of frequency offsets till the loop acquires the signal. The acquisition occurs when the frequency offset is small enough for the frequency error signal to pass through the narrow bandwidth loop filter. The lock detector terminates the sweep upon detection of the acquisition.

In modern receivers a maximum likelihood frequency estimators perform the task of fre-

frequency acquisition. To reduce phase jitter due to the frequency lock loop noise injection, this loop is disabled when the system acquires final phase lock.

The frequency lock loop is based on a ML frequency estimator. When we take the derivative of the output of matched filter with respect to the unknown frequency offset we obtain the frequency derivative matched filter. The frequency derivative filter, often called the band edge filter, is a sensitive detector of frequency offsets of the input signal's spectral mass. How this is accomplished is visualized with the aid of figure 17.

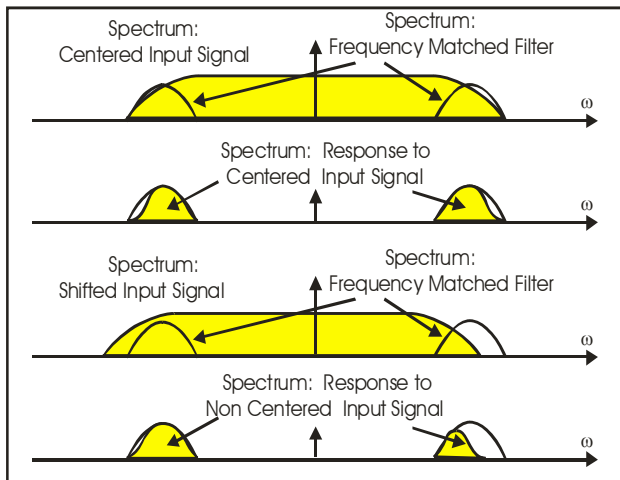


Figure 17. Frequency Centered and Offset Input Spectra Interacting with Band Edge Filters

Here we see two scenarios; One in which the input spectrum is centered between the two band edge filters and one in which the input spectrum has shifted towards one and away from the other band edge filter. When centered, the two band edge filters collect the same energy from the input spectrum and their average energy difference is zero. When offset, the two band edge filters collect different amounts of energy from the input spectrum and their energy difference contains a DC term proportional to the frequency offset. The energy difference of the band edge filters is formed as the difference of the conjugate products of the time series from each band edge filter. This difference, proportional to the frequency difference, is the input signal to the frequency lock loop filter. A block diagram of the frequency lock loop with its associated lock detector is shown in figure 18.

Figure 19 shows the phase profile of the frequency lock loop operating on an input signal with a frequency offset equal to 2.0% of the signal symbol rate. We see here the phase of the input carrier offset and the phase response of the loop. In steady state the two phase profiles have the same slope, hence same frequency but with a phase offset. We also see the band edge energy difference which returns to zero as the signal acquires frequency lock. The

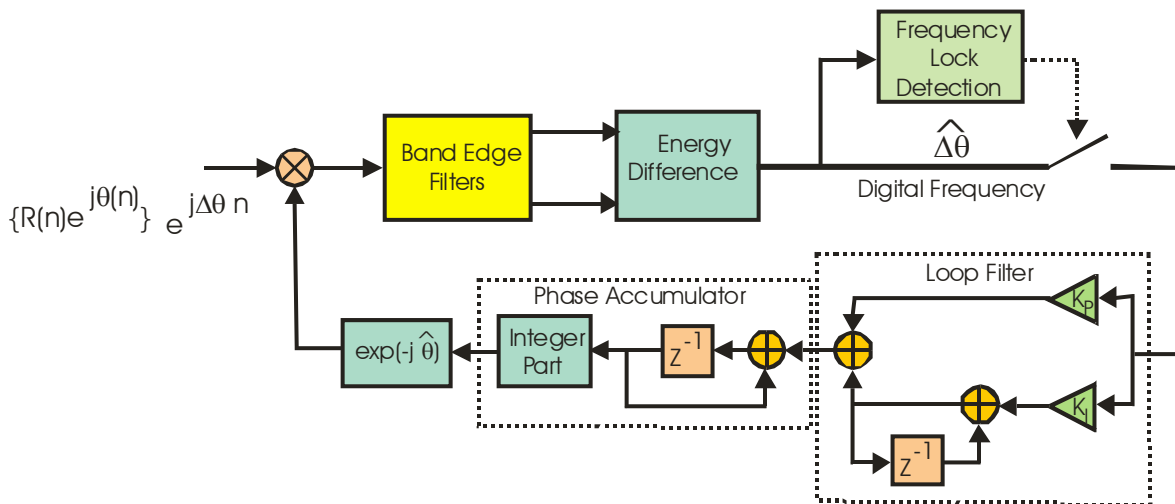


Figure 18. Maximum Likelihood Frequency Recovery for QAM Receiver

lower subplot of figure 19 shows the frequency offset of the loop DDS relative to the known input frequency offset.

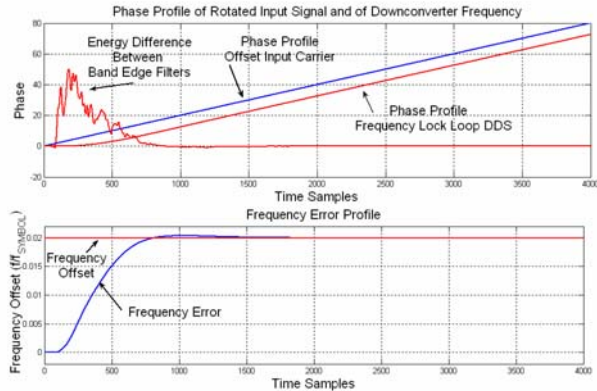


Figure 19. Phase Profiles of Input and Output of Frequency Lock Loop and Energy Difference Control Signal and Frequency Error Profile

6. Concluding Comments

We have commented on a gaping void in the communication sequences we teach to the next generation of communications and signal processing engineers. A red alarm flashes every time I see a paper starting with the comment "Let's assume the system is synchronized". I think we are doing the next generation of engineers a disservice by not presenting an important aspect of communications residing in the physical layer.

We have presented a very light overview of the synchronization process performed in many radio communication systems. The particular processes we examined were timing recovery, phase recovery, and frequency recovery. We avoided derivations and emphasized concepts easily illustrated with figures and discussions based on senior level undergraduate electrical engineering curriculum. What we have not discussed would fill a few small textbooks. We did not ask nor answer the question where does the energy reside in the modulation signal that the synchronizer accesses to perform its tasks. Do you know? It resides in the excess band-

width of the modulation spectrum. We might wonder what systems do not have excess bandwidth. We might and the answer is standard OFDM! How do we synchronize in OFDM? Answer: we acquire on preambles and track on pilots! What about other modulation formats such as Offset or Staggered Quadrature Amplitude Modulation (O-QPSK), or Vestigial Sideband (VSB), or Gaussian Minimum Phase Shift (GMSK), O-QPSK OFDM, or Shaped OFDM (S-OFDM), and on and on and on?

7. Bibliography

- [1]. Rice, Michael, "Digital Communications: A Discrete Time Approach", Prentice-Hall, 2009.
- [2]. Myer, Heinrich, Marc Monocle, and Stefan A. Bechtel, "Digital Communication Receivers: Synchronization, Channel Estimation, and Signal processing", John Wiley & Sons, 1998.
- [3]. harris, fred, "Multirate Signal processing for Communication Systems", Prentice-Hall, 2004.
- [4]. Proakis, John, "Digital Communications", Third Ed., McGraw-Hill, 1995.
- [5]. Ramanmurthy, Arjan and fred harris, "An All Digital Implementation of Constant Envelope, Bandwidth Efficient GMSK Modem using Advanced Digital Signal Processing Techniques", WPMC-06, Sept. 2006, San Diego, CA
- [6]. Ramanmurthy, Arjan and fred harris, "An all Digital Implementation of M-VSB Modem Using Advanced Digital Signal Processing Techniques", SPAWC-06, July 2006, Cannes, France.
- [7]. harris, fred, Chris Dick, and Michael Rice, "Digital Receivers and Transmitters Using Polyphase Filter Banks for Wireless Communications", Microwave Theory and Techniques, MTT, Vol. 51, No. 4, April 2003, pp 1395-1412.