Carrier Synchronization: Digital Costas' Loop Approach

Siddhartan Govindasamy

I. Overview of carrier synchronization

Consider quadrature communications system as shown in Figure 1. If the cosine signals at the transmitter and receiver are not synchronized (i.e. if there is a phase offset between the signals), the received signal cannot be decoded correctly unless the phase offset is corrected for. Here, we will see one approach for correcting timing offsets digitally, when the transmitted signal is a Binary-Phase-Shift-Keyeing (BPSK) signal. Note that a very similar approach is possible when the signals are 4-QAM, or QPSK. We start by developing an appropriate model for the phase offsets when the signals are sampled digitally, and then present an approach to correcting for the phase offset.

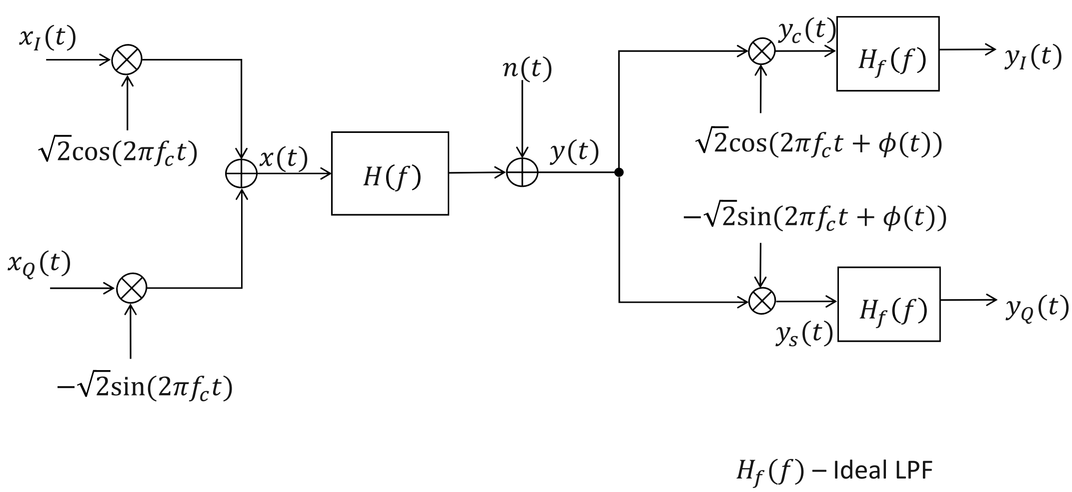

Fig. 1. Transmitter and Receiver of Quadrature System.

II. DT Flat fading with carrier offset

We have thus far seen that we can represent a quadrature modulated system (i.e. one with in-phase and quadrature phase components) using a baseband equivalent model as follows

$$y_b(t) = h_b * x_b(t)e^{-j\phi(t)} + n(t) \tag{1}$$

where $x_b(t) = x_I(t)+jx_Q(t)$, $y_b(t) = y_I(t)+jy_Q(t)$, $h_b(t)$ is a complex impulse response related to the true impulse response of the channel, and $\phi(t)$ is a phase offset between the sinusoids at the transmitter and the receiver. Under certain conditions, namely that the bandwidth of the transmitted signal is much lower than the *coherence bandwidth* of the channel, where coherence bandwidth refers to a range of frequencies over which a channel frequency response can be assumed to be constant, we can model the channel as a single multiplication This leads to the following simplification of the above equation

$$y_b(t) = hx_b(t)e^{-j\phi(t)} + n(t) \tag{2}$$

where h is now a complex scale factor. This model is called a flat fading model, since multiplying by a single scale factor in the time domain corresponds to a flat frequency response.

Suppose that $x_b(t)$ originates from a complex DT signal and $y_b(t)$ is sampled to produce a complex DT signal. We can then model our system as

$$y[k] = hx[k]e^{-j\phi[k]} + n[k] \tag{3}$$

where $x[k]$ and $y[k]$ are complex DT signals, h is a complex number, $\phi[k]$ is a DT signal related to $\phi(t)$ and $n[k]$ is a sampled version of the noise in the system.

This model, while simple, is very powerful and is used in many wireless communications systems. Additionally, it is worth noting that for a system whose implementation is all done in DT/digitally, the underlying properties of the system are not particularly important as the system designer only observes a system in (3). In other words, you can (for now at least) assume that your system is given by (3), without worrying too much about the underlying CT system.

III. Carrier Synchronization

Here we discuss the Costas Loop approach to carrier synchronization.

First lets assume that the channel follows the model in (3), and let us express the channel h, which is complex number, in exponential form, i.e.

$$h = |h|e^{j\angle h} \tag{4}$$

where $\angle h$ is the phase of the complex number h, and $|h|$ is its magnitude. Let us then use the following substitution

$$\psi[k] = \phi[k] - \angle h \tag{5}$$

$\psi[k]$ represents the combination of the phase offset due the oscillators at the transmitter and receiver not being matched, and the phase offset introduced by the channel (there is a minus sign on $\angle h$ to make notation a little simpler, it doesn't really matter as the angle is a random quantity which could be positive or negative). This leads to (3) being expressed as

$$y[k] = \left(|h|e^{j\angle h}\right) x[k]e^{-j\phi[k]} + n[k] \tag{6}$$

$$= |h|x[k]e^{-j(\phi[k]-\angle h)} + n[k] \tag{7}$$

$$= |h|x[k]e^{-j\psi[k]} + n[k] \tag{8}$$

Now, if you do the processing on a computer, you can collect a block of samples of $y[k]$ and estimate what $|h|$ is and divide it out from your signal. Define the normalized signal as $\bar{y}[k]$, which can be expressed as

$$\bar{y}[k] = x[k]e^{-j\psi[k]} + \frac{1}{|h|}n[k] \tag{9}$$

Now, we are in a position to correct for the phase offset. Suppose that we generate an estimate for $\psi[k]$ where the estimate at time k is $\hat{\psi}[k]$. Then we could multiply $\bar{y}[k]$ with this estimate to produce a phase compensated estimate of $x[k]$ which we will call $\hat{x}[k]$, as follows

$$\hat{x}[k] = \bar{y}[k]e^{j\hat{\psi}[k]} = x[k]e^{-j\psi[k]} \cdot e^{j\hat{\psi}[k]} + \frac{e^{j\hat{\psi}[k]}}{|h|}n[k] \tag{10}$$

Let's simplify our notation and write $\tilde{n}[k] = \frac{e^{j\hat{\psi}[k]}}{|h|}n[k]$ which is a scaled version of the original noise, and is hence still noise. The previous expression becomes

$$\hat{x}[k] = \bar{y}[k]e^{j\hat{\psi}[k]} = x[k]e^{-j\psi[k]} \cdot e^{j\hat{\psi}[k]} + \tilde{n}[k] = x[k]e^{-j(\psi[k]-\hat{\psi}[k])} + \tilde{n}[k] \tag{11}$$

If $\hat{\psi}[k]$ is an accurate estimate of $\psi[k]$, we will have

$$\hat{x}[k] = x[k] + \tilde{n}[k] \tag{12}$$

which effectively cancels out the effects of the phase offset due to the oscillators not being matched, and phase offsets introduced by the channel.

In the next section, we will look at one way to generate $\hat{\psi}[k]$, namely defining new signals $e[k]$ and $d[k]$ as follows

$$e[k] = -\left(\Re\left\{\hat{x}[k]\right\}\right)\left(\Im\left\{\hat{x}[k]\right\}\right) \tag{13}$$

$$d[k] = \beta e[k] + \alpha \sum_{\ell=0}^{k} e[\ell]\,, \tag{14}$$

and updating $\psi[k]$ as follows

$$\hat{\psi}[k+1] = \hat{\psi}[k] + d[k]. \tag{15}$$

The following section attempts to explain why this approach works. If you prefer not to get too much into the weeds, you can skip to the final section which describes the Costas Loop approach algorithmically.

IV. Digital Costas Loop

A. *Phase Estimate for BPSK*

Suppose that we are using Binary-Phase-Shift Keyeing (BPSK). In other words $x[k] = \pm 1$ (or some other value, but it helps to normalize to 1 here). We can use a digital Costas loop to correct for carrier frequency offsets. This method works best when the carrier frequency offsets are relatively small as it relies on a small angle approximation as you will see below. In a real system, if you wait long enough the random drift of the phase offset will eventually line up long enough for the algorithm to ”lock”, but this may take a long time if the oscillators are at very different frequencies from each other. Note that this is just one approach that is frequently used so it is by no means optimal in any sense. It is just an approach that someone very smart came up with 60 years ago.

To simplify analysis, let's start by neglecting the noise in (11), i.e. set $\tilde{n}[k] = 0$. Let's take the real and imaginary parts of the signal $\hat{x}[k]$, using the Euler formula for complex exponentials : $\Re\{e^{j\theta}\} = \cos\theta$ and $\Im\{e^{j\theta}\} = \sin\theta$.

$$\Re\left\{\hat{x}[k]\right\} = x[k]\cos(-\psi[k] + \hat{\psi}[k]) \tag{16}$$

$$\Im\left\{\hat{x}[k]\right\} = x[k]\sin(-\psi[k] + \hat{\psi}[k]) \tag{17}$$

Taking the product of the real and imaginary parts above, we get

$$\begin{aligned}
&\left(\Re\left\{\hat{x}[k]\right\}\right)\left(\Im\left\{\hat{x}[k]\right\}\right) \\
&= -(x[k])^2\cos(-\psi[k] + \hat{\psi}[k])\sin(-\psi[k] + \hat{\psi}[k]) \\
&= \frac{1}{2}\sin(-2\psi[k] + 2\hat{\psi}[k])
\end{aligned} \tag{18}$$

where we have used the identity $\cos(\alpha)\sin(\beta) = \frac{1}{2}\sin(\alpha-\beta)+\frac{1}{2}\sin(\alpha+\beta)$. If $2\psi[k]$ is small, we can use the approximation $\sin(\theta) \approx \theta$, in which case the former equation becomes an estimate of the negative of the phase error between the true phase offset and our estimate of the phase offset

$$e[k] = -\left(\Re\left\{\hat{x}[k]\right\}\right)\left(\Im\left\{\hat{x}[k]\right\}\right) \approx \left(\psi[k] - \hat{\psi}[k]\right) \tag{19}$$

This approach enables us to generate an estimate for the error without actually knowing what was transmitted (such approaches are called blind approaches).

Note that there are blind methods of generating the error signal $e[n]$ for 4-QAM. For higher order constellations, it is typical to send a known header at first before the data. So the timing synchronization consists of a non-blind portion, followed by a blind portion. Since the receiver knows the header, it can compute the error signal $e[n]$ during the header transmission. During data transmissions, the system can quantize the received, timing-corrected signal to the nearest valid constellation point, and used the quantized value to compute the phase error.

In the remainder of this section we will be analyzing an approach for driving this error to zero, which would mean that the phase offset between the transmitter and receiver has been corrected.

B. "Theory" of operation of Digital Costas Loop

To drive the phase error to zero, we will use the Costas Loop, and analyze what happens to the error in steady state under this approach.

Since the phase offset between transmitter and receiver drifts and jitters over time, and since the noise is actually non-zero, we will need to do some filtering to remove the effects of the noise and the jitter in $e[k]$. Let's suppose that we filter $e[k]$ using a filter with Discrete-Time frequency response $F(e^{j\Omega})$ and call the output $d[k]$. This filter is often called a *loop filter*. One possible choice for the loop filter (which is the one we will use here) is given by the following difference equation

$$d[k] = \beta e[k] + \alpha \sum_{\ell=0}^{k} e[\ell]. \tag{20}$$

Note that this approach corresponds to the loop filter having a proportional term ($\beta e[k]$) plus an "integral" term $\alpha \sum_{\ell=0}^{k} e[\ell]$.

With this difference equation, the frequency response of the loop filter can be shown to be

$$F(e^{j\Omega}) = \beta + \frac{\alpha}{1 - e^{-j\Omega}} \tag{21}$$

In the Digital Costas loop approach, $\hat{\psi}[k]$ is updated as follows

$$\hat{\psi}[k+1] = \hat{\psi}[k] + d[k], \quad \text{in other words} \tag{22}$$

$$\hat{\psi}[k] = \hat{\psi}[k-1] + d[k-1] \tag{23}$$

The frequency response of the system relating $d[k]$ and $\hat{\psi}[k]$ is found as follows

$$\hat{\Psi}(e^{j\Omega}) = e^{-j\Omega}\hat{\Psi}(e^{j\Omega}) + e^{-j\Omega}D(e^{j\Omega}) \tag{24}$$

Rearranging yields

$$\frac{\hat{\Psi}(e^{j\Omega})}{D(e^{j\Omega})} = \frac{e^{-j\Omega}}{1 + e^{-j\Omega}} \tag{25}$$

The relationship between $e[k]$ and $\hat{\psi}[k]$ is shown in the block diagram in Figure 2. From (19), $e[k] \approx \psi[k] - \hat{\psi}[k]$. Taking this approximation to be exact, we can redraw Figure 2 with $\psi[k]$ as the input and $\hat{\psi}[k]$ as the output as in Figure 3. We can further redraw Figure 3 with $\psi[k]$ as the input and the error $e[k]$ as the output as in Figure 4 The frequency response relating $e[k]$ and $\psi[k]$ can be found using some block-diagram algebra (if you are curious how, you can look up Black's formula). The frequency response is

$$\frac{E(e^{j\Omega})}{\Psi(e^{j\Omega})} = \frac{1}{1 + \frac{e^{-j\Omega}F(e^{j\Omega})}{1-e^{-j\Omega}}} = \frac{1 - e^{-j\Omega}}{1 - e^{-j\Omega} + e^{-j\Omega}F(e^{j\Omega})} \tag{26}$$

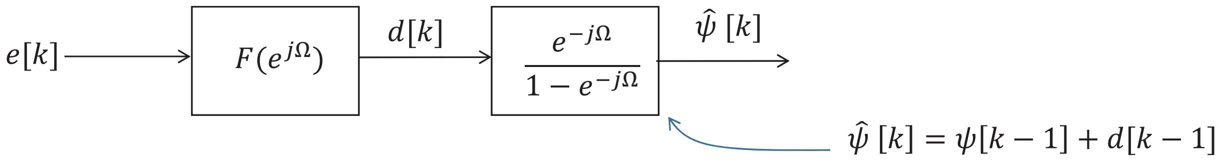

Fig. 2. Error to phase estimate block diagram

.

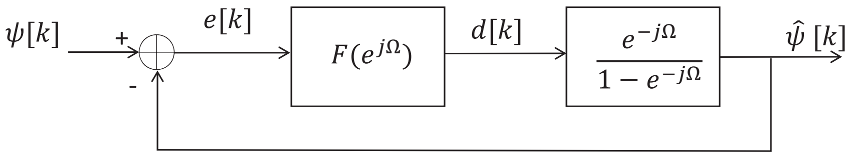

Fig. 3. Phase to phase estimate block diagram

.

Now, need to use a tool which is not covered in all Signals and Systems classes called the z-transform. The z-transform is related to the DTFT in the same way that the Laplace transform is related to the CTFT. The frequency response can be written in terms of the z transform, in which case it is called the transfer function. This is done by replacing $e^{j\Omega}$ with z as follows

$$\frac{E(z)}{\Psi(z)} = \frac{1}{1+\frac{z^{-1}F(z)}{1-z^{-1}}} = \frac{1-z^{-1}}{1-z^{-1}+z^{-1}F(z)} \tag{27}$$

We use the z transform here as it will enable us to use something called the Final Value Theorem to analyze the behavior of the error in steady state. You may have seen a version of

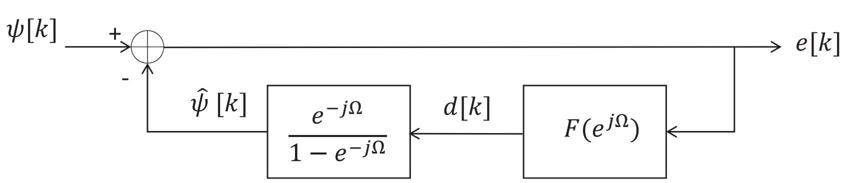

Fig. 4. Phase to error block diagram

.

the Final Value Theorem applied to Laplace Transforms. Ultimately, we want the error to be zero in steady state.

Moving terms around, we can write the z-transform of the error, $E(z)$ in terms of the z-transform of the phase offset $\Psi(z)$, and the transfer function of the loop filter $F(z)$ as follows

$$E(z) = \frac{1}{1 + \frac{z^{-1}F(z)}{1-z^{-1}}} = \frac{1 - z^{-1}}{1 - z^{-1} + z^{-1}F(z)}\Psi(z) \tag{28}$$

The final value theorem states that the steady state value of $e[k]$ can be found as below:

$$\lim_{k\to\infty} e[k] = \lim_{z\to 1}(z-1)E(z) \tag{29}$$

Therefore

$$\lim_{k\to\infty} e[k] = \lim_{z\to 1}(z-1)E(z) = \lim_{z\to 1} \frac{(z-1)(1-z^{-1})}{1 - z^{-1} + z^{-1}F(z)}\Psi(z) \tag{30}$$

$$= \lim_{z\to 1} \frac{(1-z^{-1})^2}{z^{-1} - z^{-2} + z^{-2}F(z)}\Psi(z) \tag{31}$$

$$= \lim_{z\to 1} \frac{(1-z^{-1})^2}{F(z)}\Psi(z) \tag{32}$$

We would like the quantity above to be zero, but it does depend on the loop filter used and the phase offset signal $\psi[k]$. Next we assume a simple model for the phase offset of $\psi[k] = f_\Delta k + \theta$. This model incorporates a frequency offset of f_Δ and a constant phase offset of θ. The z transform of $\psi[k]$ is found by looking up tables of z transforms, and it equals

$$\Psi(z) = \frac{f_\Delta z^{-1}}{(1-z^{-1})^2} + \frac{\theta}{1-z^{-1}}\,. \tag{33}$$

Additionally, we will assume that the loop filter takes the following form

$$F(z) = \frac{\alpha}{1-z^{-1}} + \beta = \frac{\alpha + \beta(1-z^{-1})}{1-z^{-1}} \tag{34}$$

with $\beta < 1$ and $\alpha < 1$. Note that the above corresponds to an accumulator (DT version of an integrator) with scale factor α, and a proportional term scaled by β. With this choice of $F(z)$, the steady state error is given by

$$\lim_{k\to\infty} e[k] = \lim_{z\to 1} \frac{(1-z^{-1})^3}{\alpha + \beta(1-z^{-1})}\left[\frac{f_\Delta z^{-1}}{(1-z^{-1})^2} + \frac{\theta}{1-z^{-1}}\right] \tag{35}$$

$$= \lim_{z\to 1} \frac{f_\Delta (1-z^{-1})z^{-1}}{\alpha + \beta(1-z^{-1})} + \frac{\theta(1-z^{-1})^2\theta}{\alpha + \beta(1-z^{-1})} = 0 \tag{36}$$

Therefore, with this choice of $F(z)$ and the model for the phase offset $\psi[k] = f_{\Delta}k + \theta$, the steady state phase error goes to zero.

With this approach, the steady state error goes to zero. In turns out that this combined system corresponds to a PI controller, in which case the steady-state error is removed. Intuitively, if there is any small residual error, it would accumulate over time in the accumulator (the summation term in the expression for $d[k]$ above), and cause $\hat{\psi}[k]$ to change. Typically we would want values of $\alpha \approx \beta/10$ (these are rules of thumb that come with experience).

Finally, since $\hat{\psi}$ is a phase angle, we need to wrap $\hat{\psi}[k]$ around between π and $-\pi$ to ensure it does not go unstable. This can be accomplished algorithmically

V. Digital Costas Loop Summary

The above approach for BPSK signals can be described pseudo-algorithmically as follows:

1) Estimate the magnitude of the channel by finding the root-mean-square value of the received samples – this works because the transmitted values are ± 1 and the noise is assumed to be small.
2) Divide the received signal by the estimate of the magnitude of the channel.
$$\bar{y}[k] = \frac{y[k]}{|h|} \tag{37}$$
3) Initialize $\hat{\psi}[0] = 0$, and $k = 0$.
4) Correct for the phase offset by multiplying the normalized received signal $\bar{y}[k]$ by $e^{j\hat{\psi}[k]}$, where $\hat{\psi}[k]$ is the estimated phase offset as follows
$$\hat{x}[k] = \bar{y}[k]e^{j\hat{\psi}[k]} \tag{38}$$
5) Compute an error signal as follows
$$e[k] = -\left(\Re\left\{\hat{x}[k]\right\}\right)\left(\Im\left\{\hat{x}[k]\right\}\right) \approx \left(\psi[k] - \hat{\psi}[k]\right) \tag{39}$$
6) Calculate the $d[k]$ signal which is the output of the loop filter when the input is $e[k]$.
$$d[k] = \beta e[k] + \alpha \sum_{\ell=0}^{k} e[\ell] \tag{40}$$

7) Update $\hat{\psi}[k]$.

$$\hat{\psi}[k+1] = \hat{\psi}[k] + d[k], \quad \text{alternatively } \hat{\psi}[k] = \hat{\psi}[k-1] + d[k-1]. \tag{41}$$

8) Wrap around $\psi[k+1]$. This is best described algorithmically as follows
   while($\psi[k+1] < -\pi$) $\psi[k+1] = \psi[k+1] + 2\pi$;
   while($\psi[k+1] > \pi$) $\psi[k+1] = \psi[k+1] - 2\pi$;
9) If $k <$ length $\{y[k]\}$, increment k and go to 4. Otherwise end the calculation.

$\hat{x}[k]$ will be your phase-corrected samples.

VI. Additional notes

A. Constants

Note that $\alpha < 1$ and $\beta < 1$. Typically we would want values of $\alpha \approx \beta/10$ (these are rules of thumb that come with experience). A good starting point is $\beta = 0.1$, which works well if you normalize your received signals to have rms value of 1.

Rules of thumb: If your system is too noisy, you should reduce the constants. If it has too much drift, you should reduce your constants.

B. 4-QAM/QPSK

Finally, this approach can work for 4-QAM (aka QPSK) signals with a different definition for the error signal $e[k]$, with everything else the same as in the BPSK system. The following is the expression for the error signal for 4-QAM

$$e[k] = -\frac{1}{\sqrt{2}}\left[\text{sign}\left\{\Re\{\hat{x}[k]\}\right\}\Im\{\hat{x}[k]\}\right] + \frac{1}{\sqrt{2}}\left[\text{sign}\left\{\Im\{\hat{x}[k]\}\right\}\Re\{\hat{x}[k]\}\right] \tag{42}$$

This works because of the following reasons.

Let

$$x[k] = x_I[k] + jx_Q[k]$$

Since we doing QPSK, then $x_I[k] = \pm\frac{1}{\sqrt{2}}$, $x_Q[k] = \pm\frac{1}{\sqrt{2}}$ (assuming $x[k]$ is normalized to unit amplitude). Assume that $n[k] \approx 0$, then by applying the Euler formula we have

$$\begin{aligned}\hat{x}[k] &= x[k]e^{j(\hat{\psi}[k]-\psi[k])}\\ &= (x_I[k] + jx_Q[k])\left(\cos(\hat{\psi}[k]-\psi[k]) + j\sin(\hat{\psi}[k]-\psi[k])\right)\end{aligned}$$

By expanding out the previous line and collecting real and imaginary terms we have,

$$\begin{aligned}\Re\{\hat{x}[k]\} &= x_I[k]\cos(\hat{\psi}[k]-\psi[k]) - x_Q[k]\sin(\hat{\psi}[k]-\psi[k])\\ \Im\{\hat{x}[k]\} &= x_I[k]\sin(\hat{\psi}[k]-\psi[k]) + x_Q[k]\cos(\hat{\psi}[k]-\psi[k])\end{aligned}$$

Next, we assume that $\hat{\psi}[k]$ is close to $\psi[k]$ and so we assume $\hat{\psi}[k]-\psi[k]$ is small. Since $\cos\theta \approx 1, \sin\theta \approx \theta$ for small θ, if we look at the expression for $\Re\{\hat{x}[k]\}$ above, we see that the term $x_I[k]\cos(\hat{\psi}[k]-\psi[k])$ dominates over $x_Q[k]\sin(\hat{\psi}[k]-\psi[k])$. Therefore,

$$\frac{1}{\sqrt{2}}\text{sign}\{\Re\{\hat{x}[k]\} = \text{sign}\{x_I[k]\} = x_I[k]$$

where the last equality is because of the fact that $x_I[k] = \pm\frac{1}{\sqrt{2}}$. Similarly,

$$\frac{1}{\sqrt{2}}\text{sign}\{\Im\{\hat{x}[k]\} = \text{sign}\{x_Q[k]\} = x_Q[k]$$

Therefore we have

$$\begin{aligned}\frac{1}{\sqrt{2}}\text{sign}\{\Re\{\hat{x}[k]\}\Im\{\hat{x}[k]\} &= x_I[k]\left(x_I[k]\sin(\hat{\psi}[k]-\psi[k]) + x_Q[k]\cos(\hat{\psi}[k]-\psi[k])\right)\\ &= (x_I[k])^2\sin(\hat{\psi}[k]-\psi[k]) + x_I[k]x_Q[k]\cos(\hat{\psi}[k]-\psi[k])\\ &= \frac{1}{2}\sin(\hat{\psi}[k]-\psi[k]) + x_I x_Q[k]\cos(\hat{\psi}[k]-\psi[k])\end{aligned}$$

Therefore we have

$$\begin{aligned}\frac{1}{\sqrt{2}}\text{sign}\{\Im\{\hat{x}[k]\}\Re\{\hat{x}[k]\} &= x_Q[k]\left(x_I[k]\cos(\hat{\psi}[k]-\psi[k]) - x_Q[k]\sin(\hat{\psi}[k]-\psi[k])\right)\\ &= x_I[k]x_Q[k]\cos(\hat{\psi}[k]-\psi[k]) - (x_Q[k])^2\sin(\hat{\psi}[k]-\psi[k])\\ &= x_I[k]x_Q[k]\cos(\hat{\psi}[k]-\psi[k]) - \frac{1}{2}\sin(\hat{\psi}[k]-\psi[k])\end{aligned}$$

Substituting into the expression for the error, yields

$$
\begin{aligned}
e[k] &= -\frac{1}{\sqrt{2}}\text{sign}\{\Re\{\hat{x}[k]\}\Im\{\hat{x}[k]\}\} + \frac{1}{\sqrt{2}}\text{sign}\{\Im\{\hat{x}[k]\}\Re\{\hat{x}[k]\}\} \\
&= -\frac{1}{2}\sin(\hat{\psi}[k]-\psi[k]) - x_I x_Q[k]\cos(\hat{\psi}[k]-\psi[k]) \\
&\qquad + x_I[k]x_Q[k]\cos(\hat{\psi}[k]-\psi[k]) - \frac{1}{2}\sin(\hat{\psi}[k]-\psi[k]) \quad = -\sin(\hat{\psi}[k]-\psi[k])
\end{aligned}
$$

Using the small angle approximation $\sin(\hat{\psi}[k]-\psi[k]) \approx \hat{\psi}[k]-\psi[k]$ yields

$$
e[k] \approx \psi[k] - \hat{\psi}[k]
$$