On economical timing-error detectors for QAM receivers

E.R. Pelet and J. Eric Salt

Abstract: This paper deals with timing jitter reduction in the timing recovery loop of a digital QAM receiver. The main contribution, which is derived analytically, is an economical prefilter to reduce the timing jitter in timing recovery loops containing either the early–late or the Gardner timing-error detector (TED). The proposed prefilter has the advantage of being an infinite impulse response filter that is placed inside the TED and runs at the symbol rate. For small roll off factors, it is shown with a computer simulation that a single-pole filter placed inside either the early–late or the Gardner detector is quite effective in reducing the timing jitter.

1 Introduction

Timing recovery is a critical component in a quadrature amplitude modulation (QAM) receiver. In a digital QAM receiver, the signal is typically sampled, digitised and down-converted to baseband before it is synchronised. In the synchronisation process, the signal is resampled [1] so that a sample point coincides with the decision time. A timing-error detector (TED) [2] is used to estimate the timing offset, which corresponds to the difference in time between the nearest sample point and the decision time.

The TED first renders the baseband signal to produce a timing signal, which holds the timing information. Normally, the TED squares the output of the matched filter to produce the timing signal [3, 4]. It is shown in [4] that the ensemble average of the squared signal has a constant term and a sinusoid with frequency equal to the symbol rate. It is also shown that the phase of the sinusoid depends solely on the timing offset. The TED then generates timing offset estimates from the timing signal.

Several algorithms exist to generate timing offset estimates. The Oerder and Meyr algorithm computes the Fourier coefficient for the component of the timing signal at frequency $1/T$, where $1/T$ denotes the symbol rate, to estimate the timing offset [5]. This algorithm, which uses four samples per symbol, is computationally more complex than algorithms that use two samples per symbol. Algorithms that take two samples per symbol are the objects of this paper.

The most common two-sample-per-symbol algorithms are the early–late [6, 7] and the Gardner [8] algorithms. Both algorithms operate on a timing signal produced by squaring.

The problem with the early–late and Gardner TEDs is that their output is quite noisy. The timing signal exhibits strong amplitude and phase jitters because of the random nature of the QAM data [9] and the channel noise. Noise at the output of the detector originating from the random nature of the data is often referred to as self-noise. This is the dominant noise. The self-noise introduces timing jitter in the timing recovery loop. Most of this jitter can be removed by operating the feedback loop with a small loop bandwidth (e.g. about 1% of the symbol rate). Further timing jitter reduction can be obtained by prefiltering the input to the timing detector [4]. This technique is illustrated in Fig. 1*a*. The prefilter is placed after the matched filter and before the resampler of the timing recovery loop containing the TED. The prefilter reshapes the QAM signal by introducing regular zero-crossings halfway between symbols. This greatly reduces the self-noise at the output of the detector; however, the QAM signal is distorted by this prefiltering operation. As shown in Fig. 1*a*, a second resampler is required to recover the data.

The prefilter [4] is a bandpass version of the matched filter which makes it a long finite impulse response (FIR) filter. It is assumed that the matched filter has a square-root raised cosine frequency response with arbitrary roll off factor, r. D'Andrea showed that the self-noise completely vanished for the Gardner + prefilter combination after timing had been acquired [10]. This result was obtained assuming no other source of noise in the timing recovery loop such as additive white Gaussian noise (AWGN). In a later publication [11], D'Andrea recognised that the prefilter was suboptimum in the presence of Gaussian noise and suggested a way to iteratively refine the coefficients of the filter to obtain better overall performance. D'Andrea also suggested in [11] that a shorter suboptimal prefilter that allowed some self-noise was more practical.

An approach different from D'Andrea's is taken to reveal an economical suboptimum filter that can be used to mitigate the self-noise produced by both the early–late and the Gardner detectors. Here, the term economical means less complex (i.e. less expensive to build). An analysis of the early–late algorithm reveals that two simple high-pass filters can be placed inside the early–late detector to significantly reduce the self-noise. The proposed structure is drawn in Fig. 1*b*. The high-pass filters are placed inside the TED in the timing recovery loop. Placing the filter inside the loop eliminates the need for a second resampler. The high-pass filters do not require a linear phase response and therefore can be infinite impulse response (IIR) filters.

doi:10.1049/iet-com:20060554

Paper first received 1st October 2006 and in revised form 9th April 2007

The authors are with the Department of Electrical Engineering, University of Saskatchewan, 57 Campus Drive, Saskatoon, Saskatchewan, Canada S7N 5A9

The authors are also with the TR*Labs*, 111 116 Research Dr., Saskatoon, Saskatchewan, Canada S7N 3R3

E-mail: erp588@mail.usask.ca

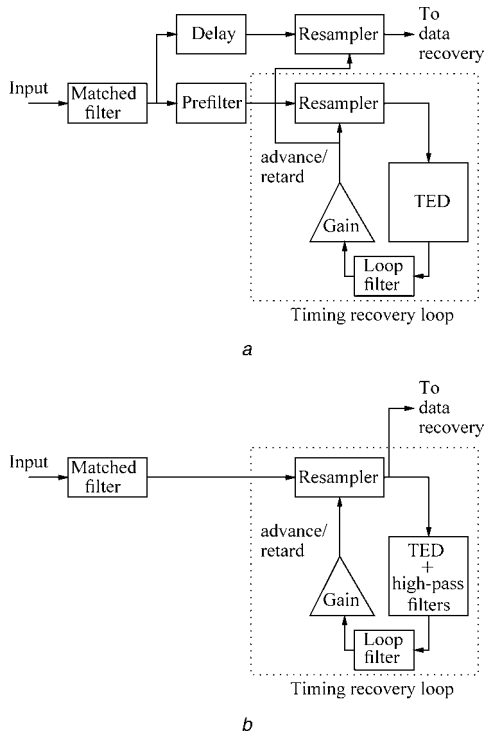

Fig. 1 *Timing recovery circuits for a digital QAM receiver*
a Structure that utilises a prefilter
b Structure that utilises the proposed high-pass filters

Furthermore, the filters run at the symbol rate, as opposed to the sampling rate.

The Gardner detector can be viewed as an early–late detector with a built-in prefilter [8]. The high-pass filters devised for the early–late detector can also be placed inside the Gardner detector to reduce the self-noise. Simulation results are given for the two structures shown in Fig. 1 with both the early–late and the Gardner detectors.

2 Early–late algorithm analysis

In the early–late algorithm, one sample is taken prior to the decision sample (early sample) and the other after the decision sample (late sample). The timing offset at symbol n is estimated with $|x(nT+D+\varepsilon T)|^2 - |x(nT-D+\varepsilon T)|^2$, where D specifies the difference in time between the decision sample and the early/late samples (e.g. $D=0.25T$ in a system sampled at four samples per symbol), εT is a real variable in the interval $[-0.5T, +0.5T]$ which denotes the timing offset, nT is the decision time for symbol n and $x(t)$ is the underlying baseband continuous-time QAM signal at the output of the matched filter. It is a complex signal equal to $x_I(t)+j\times x_Q(t)$, where $x_I(t)$ and $x_Q(t)$ are the respective in-phase and quadrature components of the baseband signal.

Critical to the operation of the TED is its input–output characteristic or S-curve [2]. In the case of the early–late detector, the S-curve is given by

$$E\{|x(nT+D+\varepsilon T)|^2\} - E\{|x(nT-D+\varepsilon T)|^2\} \quad (1)$$

where $E\{\cdot\}$ is the expectation operator.

A frequency-domain expression for $E\{|x(nT+\varepsilon T)|^2\}$ is derived. This expression reveals that there is a frequency band in the early and late signals that contains no timing information. This means that a filter can be used to remove the noise in that band without affecting the S-curve.

In the derivation that follows, εT is treated as a constant. In terms of Fig. 1, this mean that the feedback to the resampler is opened and the input controlling the resampler is held constant so that the timing offset is constant. Also, the early–late detector is quite insensitive to a small offset in the carrier frequency. This is a necessary quality as the fine tuning (correction of a small carrier offset) is often done after timing recovery. For the analysis, it is assumed that the carrier frequency offset has no significant effect on the timing detector.

Expressions for the digital signals, $x_I(nT+\varepsilon T)$ and $x_Q(nT+\varepsilon T)$, are required to determine an expression for $E\{|x(nT+\varepsilon T)|^2\}$. The continuous time signal, $x_I(t)$, can be expressed as [10]

$$x_I(t) = \sum_{i=-\infty}^{+\infty} d_I(i)h(t-iT;\, r) \quad (2)$$

where $h(t; r)$ is the impulse response of a raised cosine filter, $d_I(i)$ is the data sequence for the in-phase component of the QAM signal. A similar expression is obtained for $x_Q(t)$, where the data sequence is denoted by $d_Q(i)$. It is pointed out that $x_I(t)$ and $x_Q(t)$ are normally corrupted with AWGN. The derivation that follows only considers self-noise. The Gaussian noise has been neglected.

Using (2) with $t=(n+\varepsilon)T$ and a similar equation for $x_Q((n+\varepsilon)T)$, $E\{|x((n+\varepsilon)T)|^2\}$ can be written as

$$\begin{aligned} &E\{|x((n+\varepsilon)T)|^2\} \\ &\quad = E\left\{\left[\sum_{i=-\infty}^{+\infty} d_I(i)h((n-i+\varepsilon)T;\, r)\right]^2\right\} \\ &\qquad + E\left\{\left[\sum_{i=-\infty}^{+\infty} d_I(Q)h((n-i+\varepsilon)T;\, r)\right]^2\right\} \end{aligned} \quad (3)$$

For a constant ε, $x_I((n+\varepsilon)T)$ and $x_Q((n+\varepsilon)T)$ are stationary sequences [2]. Therefore the average power in these sequences can be expressed as an expectation for any value of n. Setting $n=0$ in (3) and expanding the sums yields

$$\begin{aligned} E\{|x((n+\varepsilon)T)|^2\} = \sum_{i=-\infty}^{+\infty}\sum_{l=-\infty}^{+\infty} &[E\{d_I(i)d_I(l)\} \\ &+ E\{d_Q(i)d_Q(l)\}] \\ &\times h((-i+\varepsilon)T;\, r)h((-l+\varepsilon)T;\, r) \end{aligned} \quad (4)$$

Assuming that $d_I(i)$ and $d_Q(i)$ are zero mean, independent sequences with the same variance, (4) simplifies to

$$E\{|x((n+\varepsilon)T)|^2\} = 2\sigma^2 \sum_{i=-\infty}^{+\infty} h^2((i+\varepsilon)T;\, r) \quad (5)$$

where σ^2 is the variance of $d_I(i)$ as well as $d_Q(i)$.

From Parseval's theorem [12], with the integration over 2π beginning at $-(1-r)\pi$

$$E\{|x((n+\varepsilon)T)|^2\} = \frac{2\sigma^2}{2\pi}\int_{-(1-r)\pi}^{(1+r)\pi} |H(\mathrm{e}^{j\omega};\, r,\, \varepsilon T)|^2\, \mathrm{d}\omega \quad (6)$$

where $H(\mathrm{e}^{j\omega}; r, \varepsilon T)$ is the discrete-time Fourier transform (DTFT) of $h((n+\varepsilon)T;\, r)$. A closed form expression for

$H(e^{j\omega}; r, \varepsilon T)$ is derived in the Appendix. Using (11) in the Appendix, the integrand in (6) can be expressed as

$$|H(e^{j\omega}; r, \varepsilon T)|^2 = \begin{cases} 1, & \omega \geq -(1-r)\pi \\ & \omega \leq (1-r)\pi \\ \cos^2(\pi\varepsilon) + \sin^2(\pi\varepsilon) & \\ \times \sin^2\left(\dfrac{\pi}{2r}\left(\dfrac{\omega}{\pi} - 1\right)\right), & \omega \geq (1-r)\pi \\ & \omega \leq (1+r)\pi \end{cases} \quad (7)$$

Clearly from (7), $|H(e^{j\omega}; r, \varepsilon T)|^2$ does not depend on T and only depends on ε in the interval $(1-r)\pi \leq \omega \leq (1+r)\pi$ rad/sample. The latter means that suppressing the early and late signals in the frequency band from $-(1-r)\pi$ to $(1-r)\pi$ rad/sample does not remove timing information. It can be shown using (1), (6) and (7) that the S-curve does not change if $|H(e^{j\omega}; r, \varepsilon T)|^2$ is set to 0 for $-(1-r)\pi \leq \omega \leq (1-r)\pi$ rad/sample.

From (7), the early and late signals, $x(nT - D + \varepsilon T)$ and $x(nT + D + \varepsilon T)$, can be filtered with a high-pass filter with passband corner at $(1-r)\pi$ rad/sample without removing timing information. Two high-pass filters are required, one for the early signal and one for the late signal. The high-pass filters do not require a linear phase response if they run at the sampling rate, since from (6), only the magnitude at the output of the high-pass filters is used in the early–late detection scheme.

A single-pole filter is an effective high-pass filter for small bandwidths. Since the required bandwidth is r, a single-pole filer is practical for small values of r. Two identical high-pass filters are required to process the early signal: one for the in-phase signal and one for the quadrature signal. Two more are required for the late signal. Therefore a total of four single-pole high-pass filters are needed. Fig. 2 shows where the filters are placed for both the early–late and Gardner detectors. Only the 'in-phase' circuits are shown.

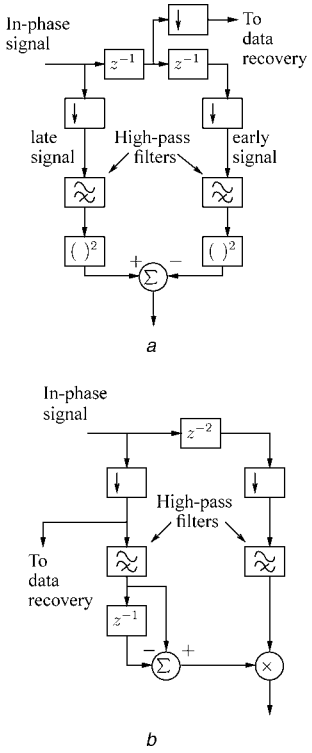

Fig. 2 *Early–late and Gardner detectors with high-pass filters*
a Early–lack detector
b Gardner detector

3 Performance

The performances of both the early–late and Gardner timing detectors were evaluated in three different configurations using Simulink/Matlab models. In the first configuration, the detectors were simulated in their conventional form (e.g. no prefilter and no high-pass filters). In the second configuration, the proposed single-pole high-pass filters were placed in the loop (see Fig. 2) in accordance with the structure shown in Fig. 1*b*. In the third configuration, the detectors were simulated with a FIR prefilter placed in front of the resampler but without high-pass filters in the loop. The structure of the third configuration is shown in Fig. 1*a*. The FIR for the prefilter was that of D'Andrea [10] and is referred to as D'Andrea's prefilter. D'Andrea's prefilter was simulated with an 81 tap FIR filter that was a least mean square fit to the frequency response of that filter given in [10].

The timing recovery loop containing the timing detector (see Fig. 1) was simulated with a critically damped phase-locked loop with a closed-loop bandwidth of 1.25% of the symbol rate. Such a loop bandwidth was obtained by placing the pole of the loop filter at $z = 0.82$ and setting the loop gain to $(1 - \sqrt{0.82})^2/g'(0)$, where $g'(0)$ is the slope of the S-curve detector when $\varepsilon T = 0$. The S-curve slope was determined for both the early–late and Gardner detectors with a computer simulation.

The operation of the timing recovery loop is as follows. The TED output is filtered by the loop filter and then used to control the resampler. The closed-loop system forces the resampler to sample at nearly the correct time. There is some fluctuation, referred to as timing jitter, because of self-noise produced by the TED.

The performance of the timing detector was evaluated by measuring the variance of the timing jitter in the resampled signal after the loop reached steady state. The resampler was an interpolation filter of length 30 samples, which was a truncated version of the impulse response [13], $h_r(t) = (\sin(\pi t/T))/(\pi t/T)$.

A digital 64QAM signal of length 16 000 symbols was generated with four samples per symbol. The roll off factor was $r = 0.1$. The 64QAM signal was corrupted with AWGN, the level of which was determined by the operating E_b/N_0, where E_b is the signal energy in a bit and N_0 is the one-sided spectral density constant of the noise. The ratio, E_b/N_0, was varied in 2 dB steps from 0 to 20 dB. The QAM signal was generated with no timing offset. A step function timing offset of $0.25T$ was introduced at the time of the 1000th symbol to strain the detector. The timing error of the resampled signal was recorded. Timing jitter was estimated from the data obtained after the loop reached steady state.

The timing jitter measurements were first performed with the 'bare' detectors in the loop (first configuration). The single-pole high-pass filters were then inserted in the loop to simulate the second configuration. The pole of the high-pass filters was placed at the same distance from the unit circle as the pole of the loop filter (i.e. at $z = -0.82$). The single high-pass filters are not ideal, which means that some timing information is lost in the filtering. This loss was compensated by increasing the loop gain in proportion to the decrease in the slope of the S-curve.

Curves of tracking mode timing jitter variances against E_b/N_0 are shown in Fig. 3 for the three configurations. Six curves are plotted: three are associated with the

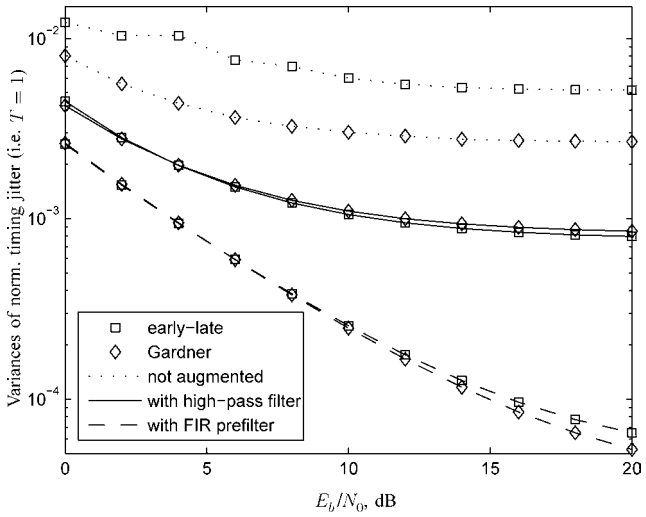

Fig. 3 *Variances of normalised timing jitter (i.e. $T = 1$) for a critically damped system with the early−late and the Gardner TEDs with and without augmentation in the form of a single-pole high-pass filter or D'Andrea's prefilter*

early−late detector, three are associated with the Gardner detector. The three curves associated with the early−late detector have data points shown with a square. The three curves associated with the Gardner detector have data points shown with a diamond. The dotted line marked with squares, which is the top curve, is the performance of the 'bare' early−late detector (first configuration). It is the noisiest circuit. At low signal-to-noise ratios, the 'bare' early−late detector is prone to symbol slips. A symbol slip did occur in one of the simulations, which caused the kink in the curve at $E_b/N_0 = 4$ dB. The solid line marked with squares shows the performance of the early−late detector augmented with the single-pole high-pass filters (second configuration). The dashed line marked with squares shows the performance when D'Andrea's prefilter is used (third configuration). The three curves for the Gardner detector illustrate its performance in the same three scenarios used for the early−late detector.

From these curves, it can be seen that the 'bare' Gardner detector outperforms the 'bare' early−late detector. With the high-pass filters in the detectors, the performances of both detectors significantly improve. In this case, the early−late detector has a very slight advantage. With D'Andrea's prefilter, the performances of both detectors are further improved. However, the performance improvement comes at the cost of a long FIR prefilter clocked at the sampling rate and an additional resampler for each of the in-phase and quadrature signal streams.

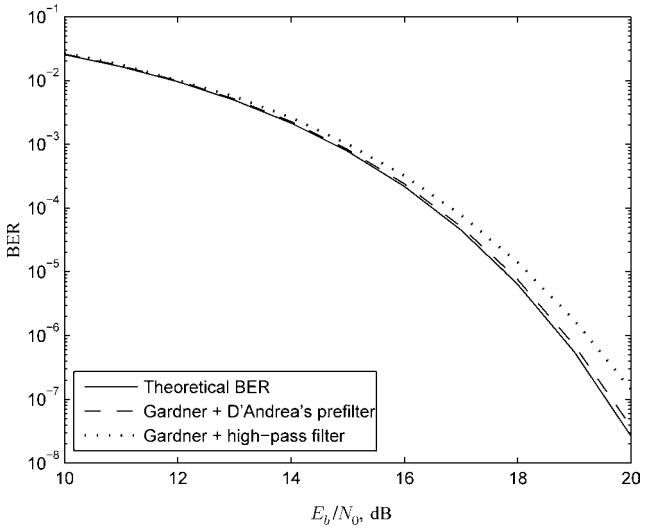

Fig. 4 *BER against E_b/N_0 for Gardner + high-pass filter and Gardner + D'Andrea's prefilter along with theoretical BER*

Bit-error-rate (BER) curves for 64QAM are plotted against E_b/N_0 in Fig. 4 for Gardner + high-pass filter and Gardner + D'Andrea's prefilter along with the theoretical BER curve. A closed-loop bandwidth of 0.25% of the symbol rate was used to generate the BER curves. To obtain this bandwidth, the pole of the loop filter was placed at $z = 0.96$. The pole of the high-pass filters was placed at $z = -0.96$. The performance degradation for the range shown in Fig. 4 is less than 0.1 dB for Gardner + D'Andrea's prefilter and less than 0.52 dB for Gardner + high-pass filter.

4 Conclusion

An analysis of the early−late algorithm revealed that the timing information in the early and late signals is in the frequency band $((1 - r)\pi, (1 + r)\pi)$ rad/sample. The frequency band $(-(1 - r)\pi, +(1 - r)\pi)$ rad/sample only contributes self-noise and can be suppressed with a high-pass IIR filter. For small values of r, a single-pole high-pass filter is effective. This low-cost suboptimum filter, which is placed inside the loop, does not affect the stability of the loop, but significantly reduces the timing jitter in steady-state operation. For a closed-loop bandwidth of 1.25% of a symbol and a signal-to-noise ratio, E_b/N_0, of 20 dB, the addition of the single-pole filters reduces the variance of the timing jitter by 8 dB.

Simulation also shows that the Gardner detector outperforms the early−late detector when they are used without enhancing filters, but looses its advantage when either the single-pole high-pass filter or D'Andrea's prefilter is used.

5 Acknowledgment

Financial support from the Natural Sciences and Engineering Research Council (NSERC) of Canada is gratefully acknowledged.

6 References

1 Gardner, F.M.: 'Interpolation in digital modems – part I: Fundamentals', *IEEE Trans. Commun.*, 1993, **41**, pp. 501–507
2 Meyr, H., Moeneclaey, M., and Fechtel, S.A.: 'Baseband communications', 'Digital communication receivers' (Wiley-Interscience, 1998, 1st edn.), pp. 80–94
3 Franks, L.E.: 'Carrier and bit synchronization in data communication – a tutorial review', *IEEE Trans. Commun.*, 1980, **28**, pp. 1107–1121
4 Franks, L.E., and Bubrouski, J.P.: 'Statistical properties of timing jitter in a PAM timing recovery scheme', *IEEE Trans. Commun.*, 1974, **22**, pp. 913–920
5 Oerder, M., and Meyr, H.: 'Digital filter and square timing recovery', *IEEE Trans. Commun.*, 1988, **36**, pp. 605–612
6 Stiffler, J.J.: 'Maximum-likehood symbol synchronization', 'Theory of synchronous communications' (Prentice-Hall, 1971, 1st edn.), pp. 221–231
7 Proakis, J.G.: 'Carrier and symbol synchronization', 'Digital communications' (McGraw-Hill, 1983, 2001, 4th edn.), pp. 359–365
8 Gardner, F.M.: 'A BPSK/QPSK timing-error detector for sampled receivers', *IEEE Trans. Commun.*, 1986, **34**, pp. 423–429
9 Gardner, F.M.: 'Self-noise in synchronizers', *IEEE Trans. Commun.*, 1980, **28**, pp. 1159–1163
10 D'Andrea, A.N., and Luise, M.: 'Design and analysis of a jitter-free clock recovery scheme for QAM systems', *IEEE Trans. Commun.*, 1993, **41**, pp. 1296–1299
11 D'Andrea, A.N., and Luise, M.: 'Optimization of symbol timing recovery for QAM data demodulators', *IEEE Trans. Commun.*, 1996, **44**, pp. 399–406

12 Couch II, L.W.: 'Signals and spectra', 'Digital and analog communication systems' (Prentice-Hall, 1983, 5th edn., 1997), pp. 46–50
13 Schafer, R.W., Oppenheim, A.V., and Buck, J.R.: 'Sampling of continuous-time signals', 'Discrete-time signal processing' (Prentice-Hall, 1989, 2nd edn., 1999), pp. 150–153
14 Proakis, J.G.: 'Signal design for band-limited channels', 'Digital communications' (McGraw-Hill, 1983, 2001, 4th edn.), pp. 554–561

7 Appendix

The Fourier transform of $h(t; r)$ is the raised cosine response denoted by $H_A(j\Omega; r)$. The Fourier transform of $h(t + \varepsilon T; r)$ is given by [14]

$$H_A(j\Omega;\, r,\, \varepsilon T) = H_A(j\Omega;\, r)\mathrm{e}^{j\Omega\varepsilon T} \tag{8}$$

$$= \begin{cases} T\mathrm{e}^{j\Omega\varepsilon T}, & |\Omega T| \le (1-r)\pi \\ \dfrac{T}{2}\left[1 + \cos\left(\dfrac{\pi}{2r}\left(\dfrac{|\Omega T|}{\pi}\right.\right.\right. & \\ \left.\left.\left. + r - 1\right)\right)\right]\mathrm{e}^{j\Omega\varepsilon T}, & |\Omega T| \ge (1-r)\pi \\ & |\Omega T| \le (1+r)\pi \\ 0, & |\Omega T| \ge (1+r)\pi \end{cases} \tag{9}$$

From $H_A(j\Omega; r, \varepsilon T)$ the Nyquist frequency is $(r+1)/2T$. Since r is greater than zero, the symbol rate, $1/T$, is less than the Nyquist rate, so sampling $h(t + \varepsilon T; r)$ at the symbol rate causes aliasing in the frequency domain. The DTFT of $h((k + \varepsilon)T; r)$ can be defined for any 2π interval of ω, where $\omega = \Omega T$. For $-(1-r)\pi < \omega \le (1+r)\pi$, it is given by

$$H(\mathrm{e}^{j\omega};\, r,\, \varepsilon T) = \frac{1}{T}\sum_{k=-\infty}^{+\infty} H_A\left(j\left(\frac{\omega - 2\pi k}{T}\right);\, r,\, \varepsilon T\right)$$

Since $H_A(j\Omega; r, \varepsilon T)$ is band limited to $|\Omega| < 2\pi/T$, then

$$H(\mathrm{e}^{j\omega};\, r,\, \varepsilon T) = \frac{1}{T}(H_A(\mathrm{e}^{j\omega/T};\, r,\, \varepsilon T) + H_A(\mathrm{e}^{j(\omega-2\pi)/T};\, r,\, \varepsilon T)) \tag{10}$$

for $-(1-r)\pi < \omega \le (1+r)\pi$. The functions $H_A(\mathrm{e}^{j\omega/T}; r, \varepsilon T)$ and $H_A(\mathrm{e}^{j(\omega-2\pi)/T}; r, \varepsilon T)$ overlap in the region $((1-r)\pi, (1+r)\pi)$. Using (9) and (10) and trigonometric identities $\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b)$ with $a = \pi/2r((\omega/\pi) - 1)$ and $b = \pi/(2r)$ $r = \pi/2$, $\mathrm{e}^{j\pi\varepsilon} + \mathrm{e}^{-j\pi\varepsilon} = 2\cos(\pi\varepsilon)$ and $(\mathrm{e}^{j\pi\varepsilon} - \mathrm{e}^{-j\pi\varepsilon})/2 = j\sin(\pi\varepsilon)$, the DTFT of $h((k + \varepsilon)T)$ can be written as

$$H(\mathrm{e}^{j\omega};\, r,\, \varepsilon T) = \begin{cases} \mathrm{e}^{j\omega\varepsilon}, & \omega > -(1-r)\pi \\ & \omega \le (1-r)\pi \\ \mathrm{e}^{j(\omega-\pi)\varepsilon}\big[\cos(\pi\varepsilon) & \\ -j\sin(\pi\varepsilon) & \\ \times \sin\left(\dfrac{\pi}{2r}\left(\dfrac{\omega}{\pi} - 1\right)\right)\big], & \omega \ge (1-r)\pi \\ & \omega \le (1+r)\pi \end{cases} \tag{11}$$