The True Performance of the Simplified Data-Transition Tracking Loop

M. K. Simon
Communications Systems and Research Section

The tracking (mean-squared timing-error) performance of a simplified data-transition tracking loop (SDTTL) suggested in the conference literature [1] about 10 years ago for symbol synchronization of a nonreturn-to-zero (NRZ) data stream is presented. In the original paper that suggested this simplification of the data-transition tracking loop (DTTL), which itself dates back about 30 years [2,3], the statement was made that [1, p. 34.8.1] "the SDTTL has exactly the same S-curve and noise performance as compared with the DTTL." Hence, the simplified version, which had the advantage of not requiring a transition detector in the in-phase arm and no delay in the quadrature-phase arm, would have represented the obvious choice for the system designer. Unfortunately, the authors arrived at these sweeping equivalence statements after examining only the noise-free *S-curves, which were found to be exactly alike for the two implementations. In this article, we show that while indeed the noise-free S-curves are exactly alike, the similarity in performance stops there; that is, the performances of the original DTTL and the SDTTL are quite different in the presence of noise. In fact, asymptotically for a large symbol signal-to-noise ratio (SNR), the mean-squared timing error (jitter) of the SDTTL is 3-dB worse than that of the DTTL, which considerably negates its implementation simplification. Even at a low SNR, where neither scheme would be motivated by maximum a posteriori (MAP) estimation of symbol synchronization considerations, the DTTL still outperforms the SDTTL.*

I. Introduction

Some 30 years ago, the data-transition tracking loop (DTTL) (see Fig. 1) was introduced as an efficient symbol synchronization means for tracking a nonreturn-to-zero (NRZ) data signal received in additive white Gaussian noise (AWGN). By using a particular approximation of the derivative of the rectangular pulse shape, which in reality is not differentiable at its edges, this closed-loop structure as originally proposed [2,3] can be shown for a large symbol signal-to-noise ratio (SNR) to be motivated by the maximum a posteriori (MAP) estimate of symbol timing based on an observation of, say, K symbols. A block diagram of the DTTL is presented in Fig. 1; the DTTL requires a transition detector in its in-phase (I) arm as well as a suitable delay element in the quadrature-phase (Q) arm to account for the time difference between the end of the integration across the data transition in this arm and the detection of this transition at the end of the second symbol in the in-phase arm. The I arm produces a signal representing the polarity of a data transition, and the Q-arm output is a signal whose absolute value is proportional to the timing error between the received signal epoch and the receiver's estimate of it. The

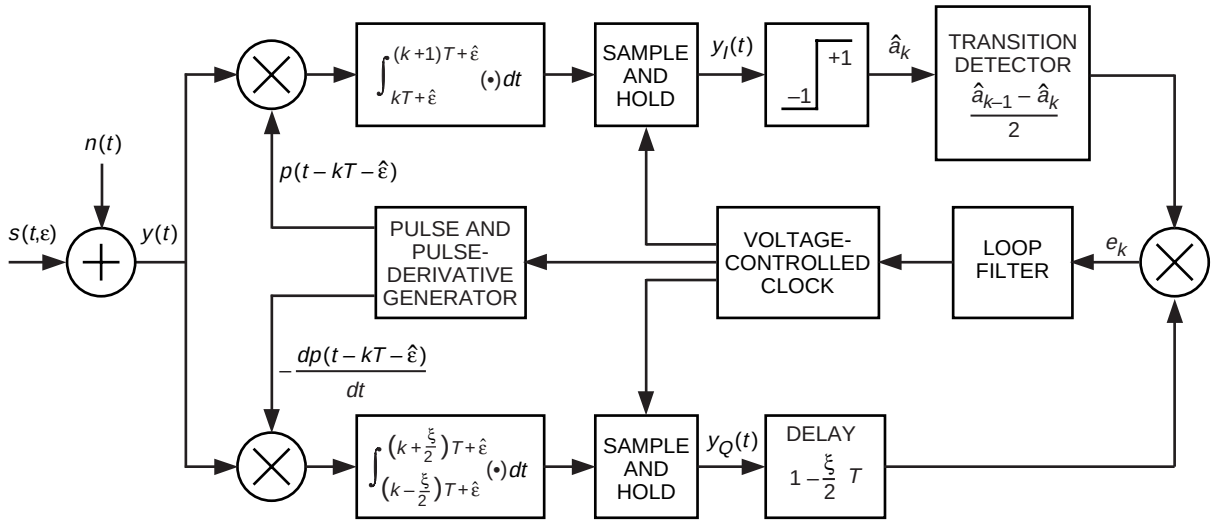

Fig. 1. The data-transition tracking loop (DTTL).

result of the product of the I and Q signals is an error signal that is proportional to this timing error, independent of the direction of the transition. The tracking performance of the DTTL as characterized by the mean-squared timing error has been well established in the past and is documented in [2,3].

About 10 years ago, Chiu and Lee [1] reported on a simplified data-transition tracking loop (SDTTL) which, by using an alternate approximation to the derivative of the rectangular pulse shape, suggested a structure that was able to eliminate the need for the in-phase-arm transition detector and quadrature-phase-arm delay element. Based on a limited analysis, i.e., illustration of the *noise-free* timing waveforms distributed throughout the loop and derivation of the associated *noise-free* S-curve, these authors claimed that the two structures had exactly equivalent S-curve and noise performances. If this were in fact true, then their "new approximation of the NRZ maximum-likelihood symbol synchronizer," as they referred to it [1, p. 34.8.1], would become the obvious choice for the symbol synchronization system designer. Unfortunately, a complete analysis of the SDTTL, as documented in this article, revealed that its performance in the presence of noise is quite different from that of the DTTL. In fact, as will be shown herein, the mean-squared timing error (jitter) of the SDTTL is always worse than that of the DTTL, the difference approaching 3 dB in the limit of a large symbol SNR. Also, the S-curves of the two schemes in the presence of noise are somewhat different—in particular, their slopes at the origins, which are influential in evaluating the mean-squared timing jitter. Interestingly enough, the slope of the S-curve at the origin for the SDTTL is, for any symbol SNR, larger than that of the DTTL, which represents an advantage of the former over the latter. Unfortunately, this is more than counteracted by the poorer noise performance of the SDTTL relative to the DTTL (an intuitive explanation for this behavior will be given later on), which accounts for the overall increase in the mean-squared timing jitter of the former relative to the latter.

II. SDTTL Versus DTTL Structure and Behavior

As mentioned in the introduction, the SDTTL differs from the DTTL in the manner in which the derivative of the rectangular NRZ pulse shape is approximated for use in determining the integrate-and-dump (I&D) interval in the loop's quadrature-phase arm. In particular, for the DTTL, the derivative of the rectangular pulse of width T seconds at either of its edges is approximated by a rectangle of width ξT $(0 \leq \xi \leq 1)$ symmetrically located around these edges [see Fig. 2(a)]. As such, the integration interval in the quadrature-phase arm as determined by the location of this derivative approximation is located *across* a potential data-transition point and as such ideally (for a zero timing offset) spans two successive

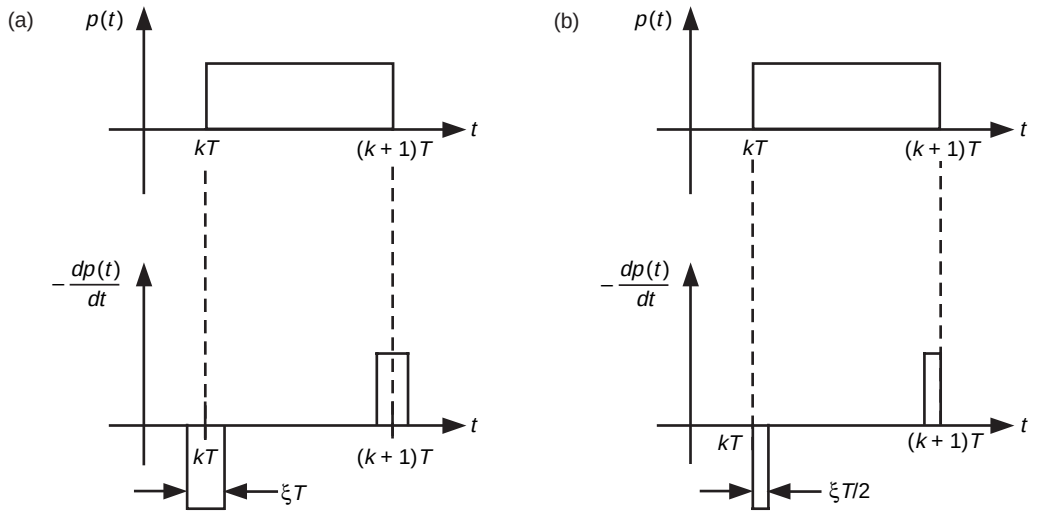

Fig. 2. Approximations of the derivative of the pulse shape: (a) DTTL and (b) SDTTL.

data symbols. The new approximation of the derivative of this pulse at either of its edges that leads to the SDTTL illustrated in Fig. 3 is a rectangle of width $\xi T/2$ occurring *within* a given data symbol [see Fig. 2(b)]. Thus, the integration in the quadrature-phase arm does not ideally span two successive data symbols but instead occurs at the beginning and end of a single symbol.

Because of the difference in the nature of the quadrature-phase-arm integration and the lack of a transition detector in the in-phase arm for the SDTTL, there is a fundamental difference in the manner in which the error signal (the output of the I–Q multiplier) behaves. For the DTTL, in the absence of noise, the upper-arm (I) output will be zero whenever no transition occurs in the data, regardless of the output of the lower (Q) arm, which typically provides the measure of timing offset. Even in the presence of noise, if no transition is detected, the upper-arm output will be zero and, hence, the error signal also is zero. Thus, for a random data sequence, on average the error signal will be zero half of the time and will have a value proportional to the timing offset the other half of the time (assuming good data detection). For the SDTTL, in the absence of noise, the upper-arm (I) output, which corresponds to the polarity of each data symbol, is never zero regardless of the transitions in the data. However, the lower-arm (Q) output will now be zero when no transition occurs, which again results in a zero error signal for that occurrence. In the presence of noise, the upper-arm output (which now corresponds to hard decisions on the data symbols) still is never equal to zero; however, the lower-arm output will now contain noise during intervals when no transition occurs and, hence, the error signal likewise will be unequal to zero during those same occurrences. Thus, for a random data sequence, on average the error signal will be noise only (multiplied by the polarity of the data symbol decision) half of the time and will have a value proportional to the timing offset the other half of the time (again assuming good data detection). It is the presence of noise in the error signal of the SDTTL during the no-transition intervals (half of the time for ideal detection) compared with a zero error signal of the DTTL in these same intervals that explains the asymptotic (in the limit of a large symbol SNR) 3-dB-worse performance of the former relative to the latter. The authors of [1] completely missed this all-important point in concluding that the two schemes had identical noise behavior.

In the following sections, we derive expressions for the S-curve and mean-squared timing error performances of the SDTTL and thereby justify the above qualitative description of its behavior relative to the DTTL.

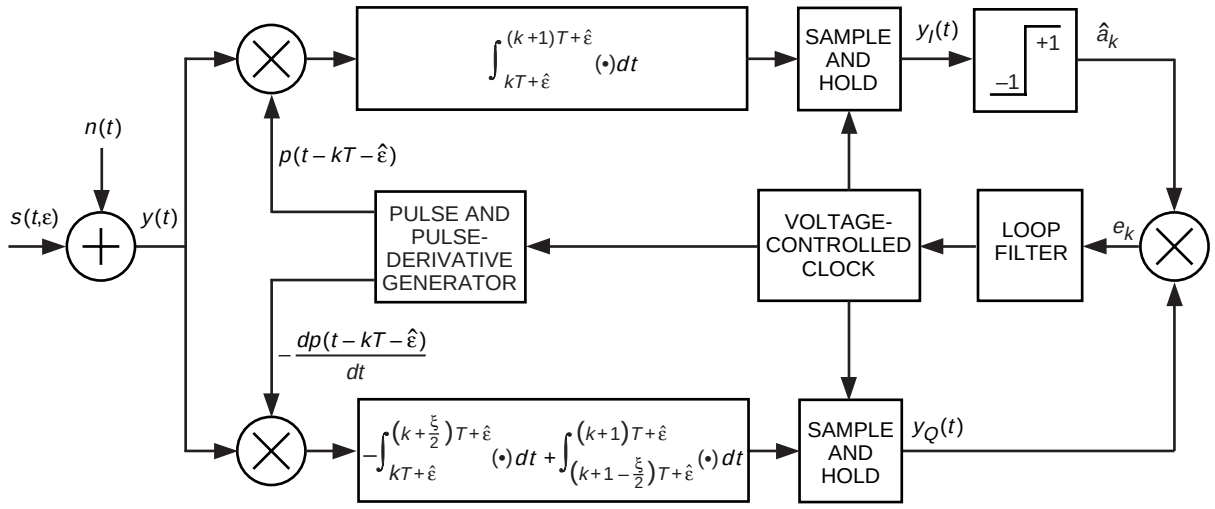

Fig. 3. The simplified data-transition tracking loop (SDTTL).

III. S-Curve Performance of the SDTTL

In the input to the loop there is assumed to be an NRZ data stream plus AWGN that has the mathematical description

$$\left.\begin{aligned} y(t) &= s(t,\varepsilon) + n(t) \\ s(t,\varepsilon) &= \sqrt{S}\sum_{n=-\infty}^{\infty} a_n p\left(t - nT - \varepsilon\right) \end{aligned}\right\} \tag{1}$$

where S is the data signal power, $\{a_n\}$ is an independent identically distributed (i.i.d.) ± 1 random data sequence, $p(t)$ is a unit amplitude rectangle in the interval $0 \le t \le T$, ε is the unknown symbol timing epoch assumed to be uniformly distributed in the interval $-(T/2) \le \varepsilon \le (T/2)$, and the AWGN $n(t)$ has single-sided power spectral density N_0 W/Hz. The output of the upper-arm I&D (following the sample and hold) in Fig. 3 is given by

$$\begin{aligned} y_I(t) &= \int_{kT+\hat{\varepsilon}}^{(k+1)T+\hat{\varepsilon}} s\left(t,\varepsilon\right) dt + \int_{kT+\hat{\varepsilon}}^{(k+1)T+\hat{\varepsilon}} n\left(t\right) dt \\ &= y_{Ik} \triangleq c_k + N_k, \qquad (k+1)\,T + \hat{\varepsilon} \le t \le (k+2)\,T + \hat{\varepsilon} \end{aligned} \tag{2}$$

where N_k is a zero-mean Gaussian random variable (r.v.) with variance $\sigma^2 = N_0 T/2$, and for the assumed NRZ signal of Eq. (1) we obtain

$$c_k = \sqrt{S}\left[a_{k-1}\lambda T + a_k\left(1-\lambda\right)T\right], \qquad 0 \le \lambda \le \frac{1}{2} \tag{3}$$

with $\lambda \triangleq \left(\varepsilon - \hat{\varepsilon}\right)/T$ denoting the normalized timing error. Similarly, the output of the lower-arm I&D (following the sample and hold) in Fig. 3 is given by

$$y_Q(t) = \int_{(k+1-(\xi/2))T+\hat{\varepsilon}}^{(k+1)T+\hat{\varepsilon}} s\left(t,\varepsilon\right) dt - \int_{kT+\hat{\varepsilon}}^{(k+(\xi/2))T+\hat{\varepsilon}} s\left(t,\varepsilon\right) dt + \int_{(k+1-(\xi/2))T+\hat{\varepsilon}}^{(k+1)T+\hat{\varepsilon}} n\left(t\right) dt$$

$$- \int_{kT+\hat{\varepsilon}}^{(k+(\xi/2))T+\hat{\varepsilon}} n\left(t\right) dt$$

$$= y_{Qk} \triangleq b_k + M_k, \qquad (k+1)\,T + \hat{\varepsilon} \le t \le (k+2)\,T + \hat{\varepsilon} \tag{4}$$

where M_k is a zero-mean Gaussian random variable (r.v.) with variance $\xi\sigma^2 = \xi N_0 T/2$, and for the assumed NRZ signal of Eq. (1) we obtain

$$b_k = \begin{cases} \sqrt{S}\left[-a_{k-1}\lambda T + a_k \lambda T\right], & 0 \le \lambda \le \dfrac{\xi}{2} \\ \sqrt{S}\left[-a_{k-1}\dfrac{\xi}{2}T + a_k\dfrac{\xi}{2}T\right], & \dfrac{\xi}{2} \le \lambda \le \dfrac{1}{2} \end{cases} \tag{5}$$

Clearly the noise sequences $\{M_k\}$ and $\{N_k\}$ are each i.i.d. and similarly independent of one another. The latter fact follows from the fact that the time intervals defining M_k and N_l are orthogonal even when $l = k$.

The error signal at the output of the I–Q multiplier in Fig. 3 is given by

$$e(t) = e_k = y_{Qk}\,\mathrm{sgn}\left(y_{Ik}\right) = \left(b_k + M_k\right)\mathrm{sgn}\left(c_k + N_k\right), \qquad (k+1)\,T + \hat{\varepsilon} \le t \le (k+2)\,T + \hat{\varepsilon} \tag{6}$$

The S-curve is by definition the statistical average of the error signal of Eq. (6) over the signal and noise probability distributions, i.e.,

$$g\left(\lambda\right) \triangleq E_{n,s}\left\{\left(b_k + M_k\right)\mathrm{sgn}\left(c_k + N_k\right)\right\} \tag{7}$$

Because of the independence of the I and Q noise components, Eq. (7) can be simplified to

$$g\left(\lambda\right) \triangleq E_s\left\{E_n\left\{b_k + M_k\right\} E_n\left\{\mathrm{sgn}\left(c_k + N_k\right)\right\}\right\} = E_s\left\{b_k \,\mathrm{erf}\left(\frac{c_k}{\sigma\sqrt{2}}\right)\right\} \tag{8}$$

where $\mathrm{erf}\,x$ denotes the error function with argument x. Substituting Eqs. (3) and (5) into Eq. (8) and performing the necessary averaging over the data symbols a_{k-1} and a_k gives the desired result, namely,

$$g\left(\lambda\right) = \begin{cases} \sqrt{S}T\lambda\,\mathrm{erf}\left(\sqrt{R_s}\left(1-2\lambda\right)\right), & 0 \le \lambda \le \dfrac{\xi}{2} \\ \sqrt{S}T\dfrac{\xi}{2}\,\mathrm{erf}\left(\sqrt{R_s}\left(1-2\lambda\right)\right), & \dfrac{\xi}{2} \le \lambda \le \dfrac{1}{2} \end{cases} \tag{9}$$

where $R_s \triangleq ST/N_0$ denotes the symbol SNR. In the limit of infinite symbol SNR, Eq. (9) becomes

$$g(\lambda) = \begin{cases} \sqrt{S}T\lambda, & 0 \le \lambda \le \dfrac{\xi}{2} \\ \sqrt{S}T\dfrac{\xi}{2}, & \dfrac{\xi}{2} \le \lambda \le \dfrac{1}{2} \end{cases} \tag{10}$$

The corresponding result to Eq. (9) for the DTTL is [2,3]

$$g(\lambda) = \begin{cases} \sqrt{S}T\left[\lambda \operatorname{erf}\left(\sqrt{R_s}(1-2\lambda)\right) - \dfrac{1}{8}(\xi - 2\lambda)\left[\operatorname{erf}\left(\sqrt{R_s}\right) - \operatorname{erf}\left(\sqrt{R_s}(1-2\lambda)\right)\right]\right], & 0 \le \lambda \le \dfrac{\xi}{2} \\ \sqrt{S}T\dfrac{\xi}{2}\operatorname{erf}\left(\sqrt{R_s}(1-2\lambda)\right), & \dfrac{\xi}{2} \le \lambda \le \dfrac{1}{2} \end{cases} \tag{11}$$

which has an asymptotic limit identical to Eq. (10), as previously confirmed by the authors of [1]. However, for a finite symbol SNR, clearly Eqs. (9) and (11) are different. Without belaboring the analysis, it is also straightforward to show that for $-(1/2) \le \lambda \le 0$, $g(\lambda) = -g(-\lambda)$, i.e., the S-curve is an odd function of the normalized timing error.

The slope of the S-curve at the origin ($\lambda = 0$) will be of interest in computing the mean-squared timing jitter performance. Taking the derivative of Eq. (9) with respect to λ and evaluating the result at $\lambda = 0$ gives, for the SDTTL,

$$K_g \triangleq \frac{dg(\lambda)}{d\lambda}\Big|_{\lambda=0} = \sqrt{S}T\operatorname{erf}\left(\sqrt{R_s}\right) \tag{12}$$

whereas the corresponding result for the DTTL, based on the derivative of Eq. (11), is

$$K_g \triangleq \frac{dg(\lambda)}{d\lambda}\Big|_{\lambda=0} = \sqrt{S}T\left[\operatorname{erf}\left(\sqrt{R_s}\right) - \frac{\xi}{2}\sqrt{\frac{R_s}{\pi}}\exp(-R_s)\right] \tag{13}$$

Since the second term in Eq. (13) is always positive, then comparing Eqs. (12) and (13), we see that the S-curve slope for the SDTTL is larger than that of the DTTL, which, as we shall see, would imply an advantage in mean-squared timing error performance.

IV. Noise Performance

The stochastic differential equation that characterizes the operation of either the DTTL or the SDTTL is [2,3]

$$\dot{\lambda} = -KF(p)\left[g(\lambda) + n_\lambda(t)\right] \tag{14}$$

where K is the total loop gain, $F(p)$ is the transfer function of the loop filter with p denoting the Heaviside operator, and $n_\lambda(t)$ is the equivalent additive noise that characterizes the variation of the loop error signal around its mean (the S-curve). Because of the I&D and sample-and-hold operations in the I and Q arms of the loops, $n_\lambda(t)$ is a piecewise (over intervals of T seconds) constant random process. In particular,

$$n_\lambda(t) = e_k - E_{n,s}\{e_k\} = e_k - g(\lambda), \qquad (k+1)T + \hat{\varepsilon} \le t \le (k+2)T + \hat{\varepsilon} \tag{15}$$

with a covariance function that is piecewise linear between the sample values

$$\begin{aligned} R_n(\tau)|_{\tau=mT} &= E\{n_\lambda(t)\, n_\lambda(t+\tau)\}|_{\tau=mT} \\ &= E_{n,s}\{(e_k - E_{n,s}\{e_k\})(e_{k+m} - E_{n,s}\{e_{k+m}\})\} \\ &= E_{n,s}\{e_k e_{k+m}\} - g^2(\lambda) \triangleq R(m,\lambda), \qquad m = 0, \pm 1, \pm 2, \cdots \end{aligned} \tag{16}$$

As is customary in the analysis of loops of this type, for loop bandwidths that are small compared with the reciprocal of the symbol time interval, $n_\lambda(t)$ can be approximated by a delta-correlated process with equivalent flat (with respect to frequency) power spectral density

$$N_0' \triangleq 2\int_{-\infty}^{\infty} R_n(\tau)\, d\tau = T\left[R(0,\lambda) + 2\sum_{m=1}^{\infty} R(m,\lambda)\right] \tag{17}$$

It is straightforward to show that

$$E_{n,s}\{e_k e_{k+m}\} = \begin{cases} E_s\{b_k^2\} + \xi\sigma^2, & m = 0 \\ g^2(\lambda), & m \neq 0 \end{cases} \tag{18}$$

Furthermore, from Eq. (5),

$$E_s\{b_k^2\} = \begin{cases} 2ST^2\lambda^2, & 0 \leq \lambda \leq \dfrac{\xi}{2} \\ 2ST^2\left(\dfrac{\xi}{2}\right)^2, & \dfrac{\xi}{2} \leq \lambda \leq \dfrac{1}{2} \end{cases} \tag{19}$$

Thus, combining Eqs. (16), (18), and (19), we have

$$R(0,\lambda) = \begin{cases} 2ST^2\lambda^2\left[1 - \operatorname{erf}^2\left(\sqrt{R_s}\,(1-2\lambda)\right)\right] + \xi\dfrac{N_0 T}{2}, & 0 \leq \lambda \leq \dfrac{\xi}{2} \\ 2ST^2\left(\dfrac{\xi}{2}\right)^2\left[1 - \operatorname{erf}^2\left(\sqrt{R_s}\,(1-2\lambda)\right)\right] + \xi\dfrac{N_0 T}{2}, & \dfrac{\xi}{2} \leq \lambda \leq \dfrac{1}{2} \end{cases} \tag{20}$$

$$R(m,\lambda) = 0, \qquad m \neq 0$$

from which the equivalent power spectral density is

$$N_0' = \begin{cases} 2ST^3\lambda^2\left[1 - \operatorname{erf}^2\left(\sqrt{R_s}\,(1-2\lambda)\right)\right] + \xi\dfrac{N_0 T^2}{2}, & 0 \leq \lambda \leq \dfrac{\xi}{2} \\ 2ST^3\left(\dfrac{\xi}{2}\right)^2\left[1 - \operatorname{erf}^2\left(\sqrt{R_s}\,(1-2\lambda)\right)\right] + \xi\dfrac{N_0 T^2}{2}, & \dfrac{\xi}{2} \leq \lambda \leq \dfrac{1}{2} \end{cases} \tag{21}$$

For large loop SNRs, it is customary to consider only the value of the equivalent power spectral density at $\lambda = 0$, namely,

$$N_0' = T\left[R(0,0)\right] = \xi\frac{N_0T^2}{2} \tag{22}$$

The equivalent quantity for the DTTL can be obtained from the results in [2,3] to be

$$N_0' = \xi\frac{N_0T^2}{4}\left[1 + \frac{\xi R_s}{2} - \frac{\xi}{2}\left[\frac{1}{\sqrt{\pi}}\exp\left(-R_s\right) + \sqrt{R_s}\,\mathrm{erf}\,\sqrt{R_s}\right]^2\right] \tag{23}$$

Note the 3-dB difference between Eqs. (22) and (23), as previously alluded to, in the limit as R_s approaches infinity. Comparing Eq. (22) with Eq. (23), we clearly see that the noise behavior of the SDTTL and the DTTL are quite different for any R_s.

V. Mean-Squared Timing-Error Performance

The mean-squared timing jitter σ_λ^2 of either the SDTTL or the DTTL is readily computed for a first-order loop filter ($F(p) = 1$) and large loop SNR conditions. In particular, linearizing the S-curve to $g(\lambda) = K_g\lambda$ and defining the single-sided loop bandwidth by B_L, we obtain

$$\sigma_\lambda^2 = \frac{N_0'B_L}{K_g^2} \tag{24}$$

where K_g is obtained from either Eq. (12) or Eq. (13) and N_0' from either Eq. (22) or Eq. (23). Making the appropriate substitutions in Eq. (24) gives the results

$$\sigma_\lambda^2 = \frac{\xi}{\rho\ \mathrm{erf}^2\sqrt{R_s}} \tag{25a}$$

for the SDTTL and

$$\sigma_\lambda^2 = \frac{\xi\left[1 + \frac{\xi R_s}{2} - \frac{\xi}{2}\left[\frac{1}{\sqrt{\pi}}\exp\left(-R_s\right) + \sqrt{R_s}\,\mathrm{erf}\,\sqrt{R_s}\right]^2\right]}{2\rho\left[\mathrm{erf}\left(\sqrt{R_s}\right) - \frac{\xi}{2}\sqrt{\frac{R_s}{\pi}}\exp\left(-R_s\right)\right]^2} \tag{25b}$$

for the DTTL, where $\rho \triangleq S/N_0B_L$ is the so-called phase-locked loop SNR. Figure 4 is a plot of the ratio of $\sigma_\lambda^2\,|_{SDTTL}$ to $\sigma_\lambda^2\,|_{DTTL}$, in dB, as a function of R_s, in dB, with quadrature-phase-arm normalized-window-width ξ as a parameter. The numerical results clearly illustrate the superior performance of the original DTTL for all symbol SNRs.

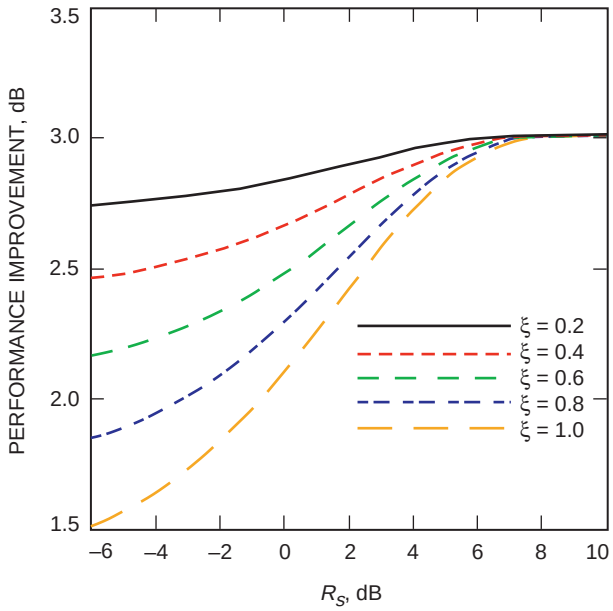

Fig. 4. Performance improvement of the DTTL relative to the SDTTL.

VI. Conclusion

The implementation of a closed-loop symbol synchronizer motivated by MAP estimation of timing error for NRZ signals is critically dependent on the manner in which the rectangular pulse shape derivative is approximated at its edges. Two such approximations that lead to the DTTL and the simplified DTTL have largely different performances in the presence of AWGN.

References

[1] J.-H. Chiu and L.-S. Lee, "The Simplified Data Transition Tracking Loop—A New Approximation of NRZ Maximum-Likelihood Symbol Synchronizer," *Globecom '87 Conference Record*, Tokyo, Japan, pp. 34.8.1–34.8.5, November 15–18, 1987.

[2] M. K. Simon, "An Analysis of the Steady-State Phase Noise Performance of a Digital Data-Transition Tracking Loop," *ICC '69 Conference Record*, Boulder, Colorado, pp. 20-9–20-15, June 1969. Also see JPL Technical Report 900-222, Jet Propulsion Laboratory, Pasadena, California, November 21, 1968.

[3] W. C. Lindsey and M. K. Simon, *Telecommunication Systems Engineering*, Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1973, reprinted by Dover Press, 1991.