Performance Analysis of Digital Tracking Loops for Telemetry Ranging Applications

Victor Vilnrotter,* Jon Hamkins,* Hua Xie,* and Shwan Ashrafi†

ABSTRACT. — In this article, we analyze mathematical models of digital loops used to track the phase and timing of communications and navigation signals. The limits on the accuracy of phase and timing estimates play a critical role in the accuracy achievable in telemetry ranging applications. We describe in detail a practical algorithm to compute the loop parameters for discrete update (DU) and continuous update (CU) loop formulations, and we show that a simple power-series approximation to the DU model is valid over a large range of time-bandwidth product ($B_L T$). Several numerical examples compare the estimation error variance of the DU and CU models to each other and to Cramér-Rao lower bounds. Finally, the results are applied to the problem of ranging, by evaluating the performance of a phase-locked loop designed to track a typical ambiguity-resolving pseudonoise (PN) code received and demodulated at the spacecraft on the uplink part of the two-way ranging link, and a data transition tracking loop (DTTL) on the downlink part.

I. Introduction

The goal of ranging is to accurately determine the distance between a spacecraft and a ground antenna as a function of time [1]. This is accomplished by measuring the time for an electromagnetic signal to travel between the spacecraft and the ground. By properly accounting for ground and spacecraft processing delays — carefully calibrated ahead of time — the round-trip light time can be determined. In the two-way ranging approach currently used by the Deep Space Network (DSN), a signal modulates the phase of an uplink carrier signal, which is then demodulated on the spacecraft and used to modulate the phase of a downlink carrier.

The basic timing relationship between the uplink and downlink signals used in the DSN for conventional two-way ranging is illustrated in Figure 1(a). The uplink signal is received at the spacecraft after a delay of τ_u seconds following transmission, turned around, and retransmitted towards the ground, where it is received after a further delay of τ_d seconds. In

* Communications Architectures and Research Section.
† University of Washington, Seattle.

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

the timing shown, the spacecraft has regenerated a clean copy of the acquired uplink ranging signal. The round-trip light time τ_{ud} can be determined by correlating the uplink transmission and received downlink ranging waveforms as seen on the ground. The round trip delay equals the sum of the one-way uplink and downlink light-times, i.e., $\tau_{ud} = \tau_u + \tau_d$, from which the range can be computed by multiplying by one-half the speed of light.

A new telemetry ranging method was presented in [2] — see Figure 1(b). As in the conventional two-way regenerative ranging, an uplink signal experiences a delay of τ_u seconds traveling from Earth to the spacecraft, and a downlink signal experiences a delay of τ_d seconds traveling from the spacecraft to Earth. The two-way delay is $\tau_{ud} = \tau_u + \tau_d$ seconds, as before. The difference with telemetry ranging is that the downlink signal is *not* a copy of the ranging signal. Instead, it is the regular science telemetry that the spacecraft wishes to send to the ground.

The two-way delay τ_{ud} is determined as follows. Since the uplink signal is a periodic waveform, we may define an associated "phase" that transitions from 0 through 2π over the entire signal, and repeats. Onboard the spacecraft, the uplink signal is acquired and tracked, and thus its phase is known at any instant of time. In particular, the phase is known at the instant that a telemetry codeword begins to be transmitted. The spacecraft latches this phase value, and inserts it into the telemetry stream to be transmitted in a subsequent codeword. On the ground, the arrival time of the codeword is determined from tracking loops, and the measured spacecraft phase at that ground-arrival time is compared with the uplink transmitted ranging signal to compute the two-way delay τ_{ud}.

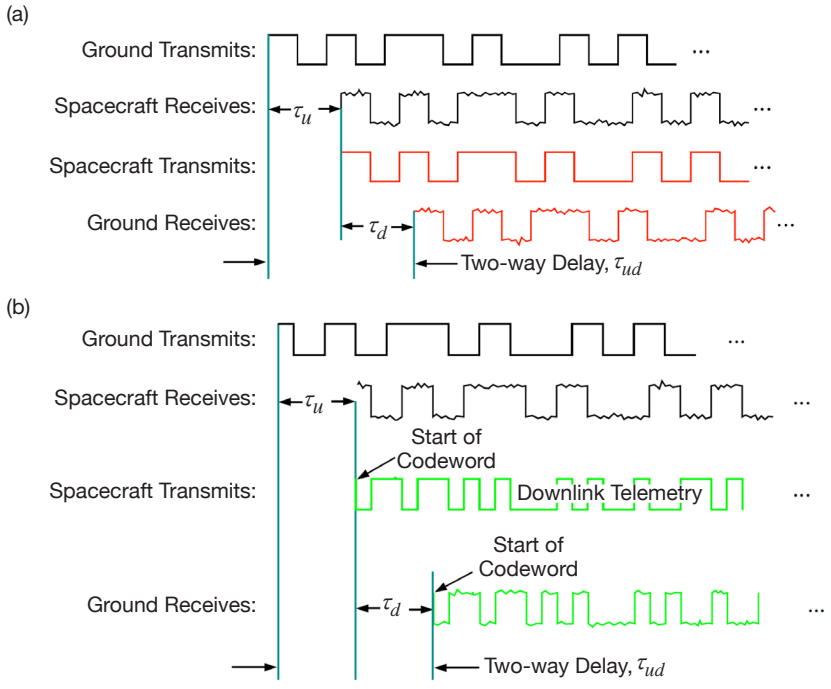

Figure 1. Timing of (a) conventional two-way regenerative ranging, and (b) telemetry ranging.

As can be seen from the description above, the ranging accuracy is limited by the accuracy of timing estimates both on the spacecraft and the ground. These estimates are made with digital tracking loops. In order to achieve the range accuracy of less than 1 m, each of these must be accurate to within 1 or 2 nanoseconds. This article explores the limits of range estimation accuracy by analyzing the theoretical limits of the digital tracking loops used.

With any modulation format, pseudonoise (PN) sequence, or telemetry data, it is first necessary to phase-lock to the carrier via a standard phase-lock loop for residual carrier signals or a Costas loop for suppressed carrier modulation. Following carrier lock, symbol synchronization is established to demodulate the received signals. For PN sequences, correlators can be used to acquire the signal to a fraction of a chip-duration, corresponding to tens of meters of range resolution for a 1-μs chip (corresponding to 1/2 MHz signal bandwidth), but much greater precision is required to achieve the goal of sub-meter resolution: this can be accomplished by data-transition tracking loops (DTTLs) on the ground, or for the case of ambiguity-resolving PN sequences that resemble a square wave (or sine wave, if shaped pulses are used), specially designed phase-locked loops on the spacecraft that lock to the demodulated signal. An example of such a phase-locked loop design is described in [3]. In all cases, the required precision can be attained by digital phase-locked loops; therefore, we concentrate on the analysis of digital phase-locked loop design applied to telemetry ranging.

An excellent systematic approach to digital phase-locked loop design that does not rely on conventional analog loop theory has been presented by Stephens and Thomas [4]. Their novel formulation enables the design engineer to specify the gain coefficients of the digital loop required to achieve a desired normalized loop bandwidth, $B_L T$, where T is the loop's update interval, which in turn determines the ability of the loop to track the phase of weak signals in a noisy environment, as typically encountered in deep-space communications and navigation applications.

When operating with weak signals in a noisy environment, it is often desirable to optimize the loop bandwidth to minimize the rms phase error. This optimization depends on the phase spectrum of the oscillator, especially on spacecraft where the reference oscillator is often an ultrastable oscillator (USO) with significant phase noise at low frequencies. As described in [5], phase-locked loops track out phase instabilities within the closed-loop bandwidth, suggesting that the loop bandwidth B_L should be made as large as possible in order to track out both low- and high-frequency phase noise contributions to the total phase error. However, additive noise entering the tracking loop also contributes to the total phase error, and this component increases with B_L. Therefore, an optimum value of loop bandwidth exists that minimizes the total phase error, by excluding noise while tracking the oscillator's phase fluctuations. The goal of digital loop design is to determine the gain coefficients to achieve the desired loop bandwidth, once the performance has been optimized.

The update interval T is a free design parameter that determines the rate at which the loop operates. Small update intervals lead to high update rates that require high-speed processing, leading to increased design complexity and cost; hence, low update rates often lead to more efficient design. The "continuous update" or CU model, valid for vanishingly small

update intervals as T approaches zero, yields tractable solutions for the gain coefficients in terms of the normalized loop bandwidth $B_L T$, provided that $B_L T \to 0$. However, for "discrete update" or DU loops, where efficiently designed loops operate with $B_L T \gg 0$, the gain coefficients obtained from the CU model are generally in error, leading to larger loop bandwidths than the CU model predicts; hence, suboptimum performance in terms of phase error.

Since the tracking-loop models do not depend on the method of extracting the error signal, these models are in principle applicable to any tracking loops, provided the error estimates and noise characteristics are suitably modified to take into account the preprocessing operations performed by different loops.

In this article, a useful mathematical model of DU loops originally described in [4] will be revisited and key concepts in the method of solution illustrated with simple examples. A general MATLAB computer program developed to compute the gain coefficients in real time for the DU model will be described, and a digital carrier tracking loop implemented in the simulation will be evaluated and compared to theoretical performance bounds. The impact of coefficient errors on residual carrier loop performance will be evaluated over a range of signal-to-noise ratios (SNRs) relevant for deep-space communication and navigation applications.

II. Signal and Noise Models for Telemetry Ranging

Telemetry ranging relies on the modulated ranging signals on the uplink, and telemetry or science data transmitted to the ground, to measure range. These signal formats are well established and will not be modified in any fundamental way to enable telemetry ranging, but some parameters of the received uplink signal will be measured at the spacecraft. Specifically, the nominally 7.2-GHz uplink carrier frequency will be tracked via a residual-carrier phase-locked loop and translated coherently to the nominally 8.4-GHz downlink frequency, the phase of the uplink PN ranging signal will be measured at the moment a telemetry codeword begins its transmission, and this measurement will be relayed to the ground as part of engineering data. We begin with a description of the uplink ranging signals and downlink telemetry with the current operational deep-space communications and navigation system.

A useful model for the modulated signal $s(t)$ used for both uplink and downlink signaling is in the form of a microwave carrier phase-modulated with subcarrier and binary data modulation. This signal is a real electromagnetic field of the form $s(t) = \sqrt{2}A\sin[\omega_0 t + \theta + m\,\mathrm{Sq}(\omega_{sc} t)d(t)]$, where ω_0 is the carrier frequency, θ is an unknown phase that must be determined and tracked as it changes, ω_{sc} is the subcarrier frequency, m is the modulation index $0 < m < \pi/2$, $d(t) = \pm 1$ is the binary data or PN code modulation, and Sq is a square-wave function analogous to sine with $\mathrm{Sq}(0) = 1$ (hence, direct modulation corresponds to $\omega_{sc} = 0$). The spectral components of the modulation are assumed to be far removed from the carrier, so that interference from the modulation can be ignored.

It is convenient to expand the sine of the sum into in-phase and quadrature components:

$$s(t) = \sqrt{2}A\Big(\sin\big[\omega_0 t + \theta\big]\cos\big[m\,\mathrm{Sq}(\omega_{sc}t)d(t)\big] + \cos\big[\omega_0 t + \theta\big]\sin\big[m\,\mathrm{Sq}(\omega_{sc}t)d(t)\big]\Big). \quad (1)$$

Recognizing that $\cos[m\,\mathrm{Sq}(\omega_{sc}t)\,d(t)] = \cos[m]$ and $\sin[m\,\mathrm{Sq}(\omega_{sc}t)\,d(t)] = \sin[m]\mathrm{Sq}(\omega_{sc}t)\,d(t)$, the received signal can be represented as

$$s(t) = \sqrt{2}A\cos\big[m\big]\sin\big[\omega_0 t + \theta\big] + \sqrt{2}A\sin\big[m\big]\cos\big[\omega_0 t + \theta\big]\mathrm{Sq}(\omega_{sc}t)d(t). \quad (2)$$

The first term in Equation (2) is the residual carrier with power $P_{rc} = P_t\cos^2[m]$, and the second term is the data-modulated subcarrier with power $P_{sc} = P_t\sin^2[m]$, where $P_t = A^2$ is the total power of the received signal. Power is conserved, since $P_{rc} + P_{sc} = P_t(\cos^2[m] + \sin^2[m]) = P_t$.

The signal is received in the presence of additive noise, which can be modeled as a continuous time noise process $n(t) = \sqrt{2}\,n_Q(t)\cos(\omega_0 t) - \sqrt{2}\,n_I(t)\sin(\omega_0 t)$, where $n_Q(t)$ and $n_I(t)$ are statistically independent zero-mean random Gaussian processes with variance σ^2. The received signal-plus-noise waveform $r(t)$ can therefore be represented as

$$r(t) = \sqrt{2}A\sin\Big[\omega_0 t + \theta + m\,\mathrm{Sq}(\omega_{sc}t)d(t)\Big] + \sqrt{2}\,n_Q(t)\cos(\omega_0 t) - \sqrt{2}\,n_I(t)\sin(\omega_0 t).$$

Before signal processing is applied, the received signal is downconverted to complex baseband, usually in two steps in a practical system: first, the signal is downconverted to an intermediate frequency (IF), then further downconverted from IF to complex baseband. The signal-processing algorithms are applied at baseband; hence, we can model downconversion as a single-step process without any loss in generality. With this model, the received signal is downconverted to complex baseband by multiplying with $\sqrt{2}\sin(\omega_{NCO}t + \hat{\theta})$ and $\sqrt{2}\cos(\omega_{NCO}t + \hat{\theta})$ and lowpass filtering, where ω_{NCO} is the radian frequency of the numerically controlled oscillator (NCO) driving the phase-locked loops either at the spacecraft or on the ground, and $\hat{\theta}$ is the loop's estimate of the received phase.

Since the received signal was modeled as a sine waveform, the "in-phase" or I component is obtained when the received signal is multiplied by $\sqrt{2}\sin(\omega_{NCO}t + \hat{\theta})$, and the "quadrature" or Q component obtained when multiplied by $\sqrt{2}\cos(\omega_{NCO}t + \hat{\theta})$. Letting $\omega_{NCO} = \omega_0$, using Equation (2), and carrying out the indicated operations, the I component becomes

$$\begin{aligned} I(t) &\equiv r(t)\sqrt{2}\sin\big(\omega_{NCO}t + \hat{\theta}\big)\big|_{LPF} \\ &= A\Big[\cos\big[m\big]\cos(\Delta\theta) - \sin\big[m\big]\sin(\Delta\theta)\mathrm{Sq}(\omega_{sc}t)d(t)\Big] + n_I(t) \\ &= A\cos(m)\cos(\Delta\theta) + n_I(t) \end{aligned}$$

where the double-frequency terms and the modulated subcarrier have been removed by lowpass filtering. Similarly, the Q component can be expressed as

$$\begin{aligned} Q(t) &\equiv r(t)\sqrt{2}\cos\big(\omega_{NCO}t + \hat{\theta}\big)\big|_{LPF} \\ &= A\Big[\cos(m)\sin(\Delta\theta) + \sin(m)\cos(\Delta\theta)Sq(\omega_{sc}t)d(t)\Big] + n_Q(t) \\ &= A\cos(m)\sin(\Delta\theta) + n_Q(t) \end{aligned}$$

where $\Delta\theta = \theta - \hat{\theta}$ is assumed to be small for a phase-locked loop, hence ignored in the noise components. The lowpass filter *LPF* eliminates the double-frequency terms, as well as the modulated subcarrier.

The downconverted *I* and *Q* components, which so far have been modeled as analog signals, are averaged via short-term integration and converted to digital samples. For our interest, the relevant quantities are the amplitude of the *I* and *Q* components and the SNR expressed in terms of P_t/N_0. For sampling intervals *T* shorter than the modulation, the signal samples can be expressed as

$$I_i = \frac{A}{T_s}\int_{iT_s}^{(i+1)T_s} A\cos[m]\cos[\Delta\theta]dt = A\cos[m]\cos[\Delta\theta]$$

and for the quadrature channel

$$Q_i = \frac{A}{T_s}\int_{iT_s}^{(i+1)T_s} A\cos[m]\sin[\Delta\theta]dt = A\cos[m]\sin[\Delta\theta].$$

Similarly,

$$n_{I,i} = \frac{A}{T_s}\int_{iT_s}^{(i+1)T_s} n_I(t)dt$$

and

$$n_{Q,i} = \frac{A}{T_s}\int_{iT_s}^{(i+1)T_s} n_Q(t)dt.$$

Since $\sin(\Delta\theta) \cong \Delta\theta$ for $\Delta\theta << 1$, the quadrature channel samples provide an error signal for the carrier tracking loops.

The variance of the noise components can be related to the two-sided noise spectral level $N_0/2$ by modeling the correlation function as

$$E[n(t_1)n(t_2)] \equiv \overline{n(t_1)n(t_2)} = \frac{N_0}{2}\delta(t_1,t_2)$$

where $n(t)$ applies to either the in-phase or quadrature components. The noise variance for either component can be expressed as

$$\sigma^2 = E\left\{\frac{1}{T_s^2}\int_{iT_s}^{(i+1)T_s} dt_1 \int_{iT_s}^{(i+1)T_s} dt_2 n(t_1)n(t_2)\right\} = \frac{1}{T_s^2}\int_{iT_s}^{(i+1)T_s} dt_1 \int_{iT_s}^{(i+1)T_s} dt_2 \overline{n(t_1)n(t_2)} = \frac{N_0}{2T_s},$$

valid for both the in-phase and quadrature noise components. The sample signal-to-noise ratio can now be expressed as the ratio of the total signal power to noise variance. When the phase error is close to zero, $\sin(\Delta\theta) \cong \Delta\theta \cong 0$, all of the residual carrier power is in the *I* component, yielding the following sample SNR for the *I* component:

$$\frac{A^2\cos^2(m)}{\sigma^2} = \frac{2T_sA^2\cos^2(m)}{N_0} = 2T_s\left(\frac{P_t}{N_0}\right)\cos^2(m).$$

Solving for P_t/N_0 yields $\frac{P_t}{N_0} = \frac{A^2}{2T_s\sigma^2}$, hence the noise variance can now be expressed as a function of P_t/N_0, in the following manner: $\sigma^2 = A^2/2T_s(P_t/N_0)$. These results will be used to adjust the variance of the additive noise in the simulation and discussed further in the numerical results section, where simulated loop performance will be evaluated.

III. Mathematical Model of Digital Tracking Loops

The simulation program designed to implement the telemetry ranging concept is based on current operational systems that rely heavily on digital tracking loops both on the spacecraft and on the ground, including carrier tracking loops, Costas loops, and DTTLs. Although differing on the method of obtaining the error signal needed for closing the loop and in some cases requiring different effective noise models due to squaring or other pre-processing operations, the closed-loop performance of the various tracking loops ultimately depends only on the gain coefficients, the effective noise environment, and the resulting loop bandwidth. Therefore, the gain coefficients are critical for establishing the desired transient characteristics and performance levels for both the CU and DU models, but the exact method of obtaining the required phase estimates does not impact loop design. Therefore, we can characterize loop behavior and establish performance criteria even with a simple loop structure, such as the residual carrier tracking loop: the same mathematical techniques can be applied to the other tracking loops as well, with similar conclusions regarding the impact of the CU and DU models for the gain coefficients on loop performance.

A. Digital Update Loop Model

The behavior of an *N*-th order digital loop with computational delay of n_c updates, $n_c = 0, 1, 2, \ldots$ can be described in terms of the following update equation in [4]:

$$\hat{\dot{\varphi}}_{n+1}T = K_1\tilde{\varphi}_{n-n_c} + K_2\sum_{i=0}^{n-n_c}\tilde{\varphi}_i + K_3\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\tilde{\varphi}_j + K_4\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\sum_{k=0}^{j}\tilde{\varphi}_k + \cdots + K_N\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\sum_{k=0}^{j}\cdots\sum_{m=0}^{l}\tilde{\varphi}_m \tag{3}$$

where the loop gain coefficients are $K_i \leq 1\ \forall i$, and $\tilde{\varphi}_n \equiv \varphi_n - \hat{\varphi}_n$ is the *n*-th residual phase term, defined in terms of the true phase φ_n and its estimate $\hat{\varphi}_n$. The update interval is T seconds, and the estimated rate of change of phase for update interval (*n*+1) on the left-hand side is given by $\hat{\dot{\varphi}}_{n+1} = (\hat{\varphi}_{n+1} - \hat{\varphi}_n)/T$. In the following development, we assume that observations start at *n* = 0, hence $\varphi_{-1} = \cdots\varphi_{-k}\cdots = 0$. Substituting $(\hat{\varphi}_{n+1} - \hat{\varphi}_n)/T$ for the rate of change of phase in Equation (3) yields

$$\hat{\varphi}_{n+1} - \hat{\varphi}_n = K_1\left(\varphi_{n-n_c} - \hat{\varphi}_{n-n_c}\right) + K_2\sum_{i=0}^{n-n_c}\left(\varphi_i - \hat{\varphi}_i\right) + K_3\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\left(\varphi_j - \hat{\varphi}_j\right) + \cdots + K_N\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\sum_{k=0}^{j}\cdots\sum_{m=0}^{j}\left(\varphi_m - \hat{\varphi}_m\right). \tag{4}$$

Separating the phase terms on the left-hand side and writing their estimates on the right-hand side yields the update equation for loops with phase and phase-rate feedback as

$$\hat{\varphi}_{n+1}-\hat{\varphi}_n+K_1\hat{\varphi}_{n-n_c}+K_2\sum_{i=0}^{n-n_c}\hat{\varphi}_i+K_3\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\hat{\varphi}_j+\cdots+K_N\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\sum_{k=0}^{j}\cdots\sum_{m=0}^{l}\hat{\varphi}_m$$
$$=K_1\varphi_{n-n_c}+K_2\sum_{i=0}^{n-n_c}\varphi_i+K_3\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\varphi_j+\cdots+K_N\sum_{i=0}^{n-n_c}\sum_{j=0}^{i}\sum_{k=0}^{j}\cdots\sum_{m=0}^{l}\varphi_m. \quad (5)$$

Analogous to the Heaviside differential operator $\frac{\partial}{\partial x}$, define the difference operator Δ as $\Delta x_n = x_n - x_{n-1}$. Apply the difference operator to Equation (3) N–1 times as in [4], and note that each sum collapses to a single term via the following argument:

$$\sum_{i=0}^{n}\Delta\varphi_i=(\varphi_n-\varphi_{n-1})+(\varphi_{n-1}-\varphi_{n-2})+\cdots+(\varphi_1-\varphi_0)+(\varphi_0-\varphi_{-1})+\cdots$$
$$=\varphi_n-\varphi_{-1}=\varphi_n \quad \text{because} \quad \varphi_{-1}=0.$$

This observation finally yields:

$$\Delta^N\hat{\varphi}_{n+1}+K_1\Delta^{N-1}\hat{\varphi}_{n-n_c}+K_2\Delta^{N-2}\hat{\varphi}_{n-n_c}+K_3\Delta^{N-3}\hat{\varphi}_{n-n_c}+\cdots+K_N\hat{\varphi}_{n-n_c}$$
$$=K_1\Delta^{N-1}\varphi_{n-n_c}+K_2\Delta^{N-2}\varphi_{n-n_c}+K_3\Delta^{N-3}\varphi_{n-n_c}+\cdots+K_N\varphi_{n-n_c}. \quad (6)$$

The difference Equation (5) specifies the phase response of the digital loop to the input signal phase, and can be solved for the loop's phase estimate in terms of the input phase by the method of z-transforms.

B. z-Transform Solution to the Difference Equation

The difference Equation (5) specifies the response of a second-order loop to the input phase, when the phase error is suitably small. Difference equations of this type are usually solved using the concept of the z-transform. Consider the single-sided z-transform of the sequence $\{x_n\}$, $Z(x_n)$, defined as $Z(x_n)=\sum_{n=0}^{\infty}x_n z^{-n}$, where $z=e^{sT}, s=-\beta(1\pm\eta)$ is the complex Laplace transform parameter that determines the loop type in the DU model. With $\hat{\Phi}_z$ and Φ_z denoting the z-transforms of $\hat{\varphi}_n$ and φ_n, respectively, and applying the z-transform relations $Z(x_{n+k})=z^kZ(x_n)$ and $Z(\Delta^N x_{n+k})=z^k(1-z^{-1})^NZ(x_n)$, the z-transform of Equation (5) can be expressed as [4]:

$$\frac{\hat{\Phi}(z)}{\Phi(z)}=\frac{K_1(z-1)^{N-1}+K_2z(z-1)^{N-2}+\cdots+z^{N-1}K_N}{z^{n_c}(z-1)^N+K_1(z-1)^{N-1}+K_2z(z-1)^{N-2}+\cdots+z^{N-1}K_N}\equiv H(z). \quad (7)$$

where $H(z)$ is the transfer function relating the phase estimate of the loop to the input phase and $(1-z^{-1})$ was replaced by $(z-1)/z$.

It is convenient to write the denominator of the transfer function as

$$D(z)=z^{n_c}(z-1)^N+K_1(z-1)^{N-1}+K_2z(z-1)^{N-2}+\cdots+z^{N-1}K_N, \quad (8)$$

yielding a simplified expression for the transfer function: $H(z)=[D(z)-z^{n_c}(z-1)^N]/D(z)$.

The above derivation is predicated on the assumption that both phase and phase-rate feedback were applied, so we could write the update equation as $\hat{\varphi}_{n+1} = \hat{\varphi}_n + \hat{\dot{\varphi}}_{n+1} T$. If only phase rate feedback is applied, then the update equation becomes $\hat{\varphi}_{n+1} = \hat{\varphi}_n + \frac{1}{2}(\hat{\dot{\varphi}}_{n+1} T + \hat{\dot{\varphi}}_n T)$, as described in [4]. For this case, the form of the transfer function remains the same, but the denominator becomes

$$D(z) = z^{n_c}(z-1)^N + \frac{1}{2}(1+z)\left[K_1(z-1)^{N-1} + K_2 z(z-1)^{N-2} + \cdots + z^{N-1}K_N\right].$$

The quantity of greatest interest to the design engineer is the single-sided loop noise bandwidth B_L, or equivalently the normalized loop bandwidth $B_L T$, which helps determine the performance of the loop in noisy environments. The normalized loop bandwidth depends on the closed-loop transfer function, which in turn depends on the loop gains. The normalized loop bandwidth can be determined as a contour integral over the unit circle, $2B_L T = \frac{1}{2\pi i}\oint H(z)H(z^{-1})z^{-1}dz$, where we made use of the fact that z^{-1} is the conjugate of z. This contour integral can be computed with the help of the residue theorem, according to which the value of the contour integral is the sum over the set of all poles $\{z_i; m_i\}$ of the integrand within the unit circle, where $m_i = 1, 2, \ldots$ is the multiplicity of the *i*-th pole. The general form of the residue theorem for poles with multiplicity m_i, applied to the problem of determining the normalized loop bandwidth, can be expressed as

$$2B_L T = \frac{1}{2\pi i}\oint H(z)H(z^{-1})z^{-1}dz = \sum_i \frac{1}{(m_i-1)!}\frac{d^{m_i-1}}{dz^{m_i-1}}(z-z_i)^{m_i}H(z)H(z^{-1})z^{-1}\big|_{z\to z_i} \tag{9}$$

where only roots inside the unit circle contribute to the sum. To evaluate this solution and determine the closed-loop bandwidth, it is necessary to first determine the roots of the transfer function and their multiplicity, namely the set $\{z_i; m_i\}$. The transfer function of an *N*th-order loop has *N* roots, some possibly with multiplicity greater than one. The roots can be determined directly by rewriting the denominator in root-factorized form as $D^*(z) = \prod_{i=1}^{N}(z - z_i)$. Expanding the product and equating the coefficients of like powers in Equation (8) yields *N* equations that can be solved for the roots of the transfer function in terms of the gain coefficients, $\{K_i\}$. These *N* equations are of the form [4]

$$\begin{aligned}
&\sum_{i=1}^{N} z_i = \binom{N}{1} - K_1 - K_2 - \cdots - K_N \\
&\sum_{i<j}^{N} z_i z_j = \binom{N}{2} - \binom{N-1}{1}K_1 - \binom{N-2}{1}K_2 - \cdots - K_{N-1}, \cdots, \\
&\prod_{i=1}^{N} z_i = 1 - K_1.
\end{aligned} \tag{10}$$

Once the roots have been determined in terms of the loop gains $\{K_i\}$, the loop bandwidth can be evaluated for the DU model in functional form. These results have been tabulated in [4] for the case of zero computational delay for loops of up to fourth order, part of which is reproduced in Table 1 for a "standard underdamped" loop design, as reference.

The following examples illustrate some of the above concepts, without introducing any unnecessary computational complexity.

Table 1. Normalized loop bandwidths in term of the *K* parameters.
(From [4], Table IV)

lst order $B_L T = \dfrac{K_1}{2(2-K_1)}$
2nd order $B_L T = \dfrac{2K_1^2 + 2K_2 + K_1 K_2}{2K_1(4-2K_1-K_2)}$
3rd order $B_L T = \dfrac{4K_1^2 K_2 - 4K_1 K_3 + 4K_2^2 + 2K_1 K_2^2 + 4K_1^2 K_3 + 4K_2 K_3 + 3K_1 K_2 K_3 + K_3^2 + K_1 K_3^2}{2(K_1 K_2 - K_3 + K_1 K_3)(8-4K_1-2K_2-K_3)}$

Example 1. The gain coefficients of a second-order loop with no computational delay are computed for the DU model first directly from Equation (8) by equating coefficients of like powers, then by simple substitution into the general solution of Equation (10). Letting $n_c = 0$ and $N = 2$, substituting into Equation (8) and expanding yields $D(z) \underset{N=2}{\equiv} (z-1)^2 + (z-1)K_1 + zK_2 = z^2 + (K_1 + K_2 - 2)z + (1-K_1)$. Writing the denominator in factorized form yields $D(z) \underset{N=2}{\equiv} (z-z_1)(z-z_2) = z^2 - (z_1+z_2)z + z_1 z_2$. Equating like coefficients and solving yields two equations in two unknowns, namely $z_1 + z_2 = 2 - K_1 - K_2$ and $K_1 = 1 - z_1 z_2$, yielding $K_2 = 1 + z_1 z_2 - (z_1 + z_2)$. Substituting into the general solution directly yields $z_1 + z_2 = 2 - K_1 - K_2$ and $z_1 z_2 = 1 - K_1$, which are seen to be the same as with the direct approach, verifying the validity of the general approach.

Example 2. For the sake of simplicity, consider a simple first-order loop with transfer function $H(z) = K_1/(z-(1-K_1)) = K_1/(z-z_1)$, where the second equality is the root-factorized form of a first-order loop. The single root is identified by inspection as $z_1 = 1 - K_1$, clearly inside the unit circle since $0 < K_1 < 1$. The complex conjugate of the transfer function is $H(z^{-1}) = K_1/(z^{-1}-(1-K_1)) = K_1/(z^{-1}-z_1)$. The squared magnitude of the transfer function with $i = 1$ and $m_i = 1$, $H_1(z)H_1(z^{-1}) = K_1^2/(z-z_1)(z^{-1}-z_1)$ yields

$$2B_L T = (z-z_1)H(z)H(z^{-1})z^{-1}\big|_{z\to z_1} = \frac{K_1^2 z^{-1}}{z^{-1}-z_1}\Big|_{z\to z_1} = \frac{K_1^2}{1-z_1 z}\Big|_{z\to z_1} = \frac{K_1^2}{1-(1-K_1)^2} = \frac{K_1}{(2-K_1)}$$

with the final result $B_L T = \frac{1}{2}K_1/(2-K_1)$.

The solutions presented so far determine the loop gain coefficients $\{K_i\}$ for the DU model, in principle for a loop of arbitrary type and degree. The loop bandwidth can be determined from these coefficients, as the above examples illustrate. However, when designing a tracking loop for deep-space applications, it is often more important to determine the gain coefficients for a given loop bandwidth, which may have been determined from loop bandwidth optimization or other external considerations. This is in effect the inverse problem, where the designer is given a loop bandwidth and asked to determine the gain coefficients for a tracking loop of a given order. The solution to this problem is not straightforward, and only approximate power series or iterative numerical solutions are known at this time.

For comparison, the results of the CU model presented in [4] are summarized in Table 2, which gives the gain parameters in terms of loop bandwidth. This is the desired form of the results, since the loop bandwidth is likely obtained from other considerations, hence the design engineer needs to determine the gain coefficients that realize the desired loop bandwidth.

Table 2. Loop gain parameters as functions of the normalized loop bandwidth, in the CU model for loops up to order 4. (From [4], Table III)

	Loop Constants			
	K_1	K_2	K_3	K_4
Supercritically damped: $\eta_i^2 = 0, \lambda_i = 1$, for all roots				
1st order	$4B_LT$	—	—	—
2nd order	$\frac{16}{5}B_LT$	$\frac{1}{4}K_1^2$	—	—
3rd order	$\frac{32}{11}B_LT$	$\frac{1}{3}K_1^2$	$\frac{1}{27}K_1^3$	—
4th order	$\frac{256}{93}B_LT$	$\frac{3}{8}K_1^2$	$\frac{1}{16}K_1^3$	$\frac{1}{256}K_1^4$
Standard underdamped: $\eta_i^2 = -1, \lambda_i = 1$, for all roots				
1st order	$4B_LT$	—	—	—
2nd order	$\frac{8}{3}B_LT$	$\frac{1}{2}K_1^2$	—	—
3rd order	$\frac{60}{23}B_LT$	$\frac{4}{9}K_1^2$	$\frac{2}{27}K_1^3$	—
4th order	$\frac{64}{27}B_LT$	$\frac{1}{2}K_1^2$	$\frac{1}{8}K_1^3$	$\frac{1}{64}K_1^4$

C. Power Series Expansion in Terms of Loop Bandwidth

Power series solutions for the loop gain coefficients in terms of the normalized closed-loop bandwidth B_LT were presented in [5] for loop orders up to four, for either supercritically damped or standard underdamped loop types, with both phase and phase-rate or phase-rate-only feedback and computational delay of up to one update interval, but no derivation was provided. For example, the gain coefficients for a supercritically damped second-order loop with zero computational delay are given by the following expressions in [5]:

$$K_1 \cong \tfrac{16}{5}B_LT - \tfrac{896}{125}(B_LT)^2 + \tfrac{46592}{3125}(B_LT)^3, \quad K_2/K_1^2 \cong \tfrac{1}{4} + \tfrac{2}{5}B_LT - \tfrac{12}{125}(B_LT)^2.$$

With a single update delay, the expressions are

$$K_1 \cong \tfrac{16}{5}B_LT - \tfrac{2368}{125}(B_LT)^2 + \tfrac{421888}{3125}(B_LT)^3, \quad K_2/K_1^2 \cong \tfrac{1}{4} + \tfrac{4}{5}B_LT - \tfrac{112}{125}(B_LT)^2.$$

The validity of these approximate formulas will be investigated in the section on numerical results.

D. Iterative Algorithm Converging on Specified Loop Bandwidth

An iterative solution for the loop gain coefficients $\{K_i\}$ has been implemented in MATLAB for any desired loop bandwidth, loop filter type, and feedback method. The input variables for this algorithm are the normalized closed-loop bandwidth $B_L T$, the filter order *N*, the computational delay n_c, the filter type (supercritically damped or standard underdamped), and the feedback type (either phase and phase-rate, or phase-rate only). The outputs of this program are the loop coefficients $\{K_i\}$.

In the loop-filter design tool, we define an *N* +1 dimensional symbolic vector in terms of the computational delay and *N* as

$$\mathbf{Z} = \left[z^{n_c}(z-1)^N, (z-1)^{N-1}, z(z-1)^{N-2}, \cdots, z^{N-1} \right]$$

for phase and phase-rate type feedback, and similarly as

$$\mathbf{Z} = \left[z^{n_c}(z-1)^N, \tfrac{1}{2}(1+z)(z-1)^{N-1}, \cdots, \tfrac{1}{2}(1+z)z^{N-1} \right]$$

for the phase-rate-only type feedback. Defining *N*+1 dimensional vector of gain coefficients as $\mathbf{Z} = (1, K_1, \cdots, K_N)$, the denominator of the transfer function can be computed as $D(z) = \mathbf{Z} \bullet \mathbf{K} := U$, where *U* is another name for $D(z)$ for convenience, but otherwise identical to it. Next, the normalized decay parameter $B_1 T$ is initialized, and the roots of $D(z)$ calculated as in [4, Equation (50)], and a symbolic parameter generated. We then equate the *N* equations obtained from taking up to the *N*–th derivative of *U* with that obtained from the corresponding derivatives of the polynomial of the roots, i.e., $D(z)$. The *N* equations in K_i obtained by letting z = 0 in each derivative can be solved using the MATLAB **solve** command. In case of nonzero computational delay, $n_c > 0$, $N + n_c$ linear equations are solved instead of *N* equations, by taking the first $N + n_c$ derivatives of the root polynomial.

In the following step, the normalized closed-loop bandwidth $B_L T$ is computed using the residue theorem: the multiplicity of each root is determined and the residue of the root z_i with multiplicity m_i is computed using the equation

$$\mathrm{Res}\{z_i\} = \frac{1}{(m_i - 1)!} \frac{d^{m_i - 1}}{dz^{m_i - 1}} (z - z_i)^{m_i} H(z).$$

Note that multiplication of $H(z)$ with $(z - z_i)^N$ should be computed as

$$\frac{D(z) - z^{n_c}(z-1)^N}{\prod_{j \neq i}(z - z_j)}$$

for root z_i; otherwise, pole cancellation can lead to very small numerical differences between the roots calculated initially and the roots of $D(z)$ in the denominator of $H(z)$.

An iterative process was implemented that converges on the *K*-vector for the desired value of $B_L T$. With $B_1 T$ set equal to $B_1 T$ + step_size, at each iteration the current value of $B_L T$ is compared to the desired value: when the algorithm approaches the desired value of $B_L T$, step_size is halved and the process continued, until the predetermined tolerance on $B_L T$

is achieved. Gain coefficients for loop orders of $N = 1, 2$, and 3 have been compiled using this algorithm, for a range of $B_L T$ values, with computational delays of $n_c = 0,1$, shown in Table 3.

IV. Numerical Results

A MATLAB program was developed to simulate the performance of the end-to-end telemetry ranging system from ground through spacecraft, where the uplink carrier is tracked and turned around coherently at the downlink frequency, and the arrival time of the binary modulation is measured, then relayed to the ground via engineering telemetry. Critical to implementing these functions are the residual carrier (RC) loop in the spacecraft, and Costas and DTTL loops on the ground. Although sample-level correlation is used to establish rough timing for the uplink PN sequences, this is only a preliminary estimate used primarily for initializing the tracking loops, which then provide the refined estimates required for range determination. We begin by evaluating the simulated performance of the RC loop on the spacecraft, both with conventional CU and the more accurate DU gain coefficients, and compare their performance under realistic operating conditions. First we highlight the differences in the various solutions to CU and DU loop gains, by comparing the loop gains generated by each of the three techniques considered in this article.

Three different estimates of the gain parameter K_1 for a second-order supercritically damped tracking loop, without computational delay, are shown in Figure 2(a). The CU model for the gain parameter K_1 is a simple linear function of normalized loop bandwidth, namely, $K_1 = \frac{16}{5} B_L T$, which approximates the accurate DU model for $B_L T < 0.5$. Based on this observation, it appears that the CU model can be used over a larger range of $B_L T$ then generally reported, namely ≤ 0.02. The three-term power-series approximation remains close to the DU model over an even greater range, namely $B_L T < 0.2$, hence this simple formula can be applied over this expanded range to approximate the first loop gain parameter. For larger values of $B_L T$, both CU and three-term power-series approximations deviate significantly from the DU model. For example, when $B_L T = 0.5$, both approximations overestimate the first gain parameter by more than a factor of two. These errors will also overestimate the loop bandwidth by roughly a corresponding amount, hence predict overly pessimistic phase-error performance. Somewhat surprisingly, the power series approximation for K_2 remains close to the true DU values over a much greater range, as shown in Figure 2(b), even though the CU approximation again diverges rapidly.

Next, we consider the differences in loop performance when using the CU and DU models to calculate the gain parameters for the simplest tracking loop structure in the telemetry ranging system model, namely, the residual-carrier tracking loop used on the spacecraft to lock onto the uplink carrier. Note that the loop equations apply to any closed-loop tracking system, including residual-carrier tracking loops, Costas loops, and DTTLs; therefore, the transient behavior, gain parameters, and normalized loop bandwidth are not affected by the details of the tracking loop system design. The main difference between the different loop structures is the method of obtaining the error signal, which does not impact the loop characteristics, but could impact the optimum value of the loop bandwidth. For example, Costas loops belong to the class of squaring loops because of the method of obtaining

Table 3. Gain coefficients generated by the iterative MATLAB tool for a broad range of B_LT covering both CU and DU models, and with computational delays of (a) zero and (b) 1.

	$N = 1$	$N = 2$		$N = 3$		
B_LT	K_1	K_1	K_2	K_1	K_2	K_3
(a) No computational delay						
0.001	3.99e-3	3.19e-3	2.55e-6	2.90e-3	2.81e-6	9.1e-10
0.005	1.98e-2	1.58e-2	6.31e-5	1.44e-2	6.94e-5	1.12e-7
0.01	3.92e-2	3.13e-2	2.49e-4	2.85e-2	2.73e-4	8.78e-7
0.02	7.69e-2	6.12e-2	9.76e-4	5.57e-2	1.06e-3	6.76e-6
0.03	1.11e-1	8.99e-2	2.12e-3	8.17e-2	2.31e-3	2.20e-5
0.05	1.82e-1	1.44e-1	5.58e-3	1.30e-1	6.03e-3	9.44e-5
0.075	2.61e-1	2.05e-1	1.18e-2	1.86e-1	1.27e-2	2.94e-4
0.1	3.33e-1	2.61e-1	1.96e-2	2.37e-1	2.10e-2	6.42e-4
0.15	4.62e-1	3.57e-1	3.93e-2	3.24e-1	4.14e-2	1.84e-3
0.2	5.71e-1	4.38e-1	6.26e-2	3.98e-1	6.52e-2	3.78e-3
0.25	6.67e-1	5.06e-1	8.84e-2	4.61e-1	9.09e-2	6.44e-3
0.3	7.50e-1	5.64e-1	1.16e-1	5.14e-1	1.18e-1	9.76e-3
0.35	8.24e-1	6.14e-1	1.46e-1	5.60e-1	1.44e-1	1.37e-2
0.4	8.89e-1	6.58e-1	1.72e-1	6.00e-1	1.71e-1	1.82e-2
0.45	9.47e-1	6.95e-1	2.01e-1	6.35e-1	1.98e-1	2.32e-2
0.5	1	7.28e-1	2.29e-1	6.66e-1	2.24e-1	2.86e-2
0.6	—	7.83e-1	2.85e-1	7.17e-1	2.73e-1	4.06e-2
0.7	—	8.26e-1	3.39e-1	7.59e-1	3.20e-1	5.37e-2
0.8	—	8.60e-1	3.92e-1	7.92e-1	3.63e-1	6.77e-2
0.9	—	8.87e-1	4.41e-1	8.20e-1	4.03e-1	8.23e-2
1	—	9.10e-1	4.89e-1	8.43e-1	4.40e-1	9.73e-2
1.2	—	9.42e-1	5.78e-1	8.78e-1	5.06e-1	1.28e-1
1.4	—	9.65e-1	6.59e-1	9.04e-1	5.64e-1	1.60e-1
1.6	—	9.79e-1	7.33e-1	9.24e-1	6.13e-1	1.91e-1
1.8	—	9.89e-1	8.01e-1	9.39e-1	6.56e-1	2.22e-1
2	—	9.95e-1	8.63e-1	9.50e-1	6.93e-1	2.52e-1
2.5	—	1	1	9.70e-1	7.69e-1	3.26e-1
3	—	—	—	9.81e-1	8.25e-1	3.95e-1
3.5	—	—	—	9.88e-1	8.67e-1	4.59e-1
4	—	—	—	9.92e-1	9.00e-1	5.19e-1
4.5	—	—	—	9.95e-1	9.25e-1	5.76e-1
5	—	—	—	9.97e-1	9.44e-1	6.29e-1
(b) 1 unit of computational delay						
0.001	3.97e-3	3.18e-2	2.54e-6	2.89e-3	2.80e-6	9.0e-10
0.005	1.94e-2	1.55e-2	6.14e-5	1.41e-2	6.75e-5	1.08e-7
0.01	3.78e-2	3.02e-2	2.36e-4	2.75e-2	2.59e-4	8.17e-7
0.02	7.18e-2	5.73e-2	8.74e-4	5.22e-2	9.55e-4	5.90e-6
0.03	1.03e-1	8.19e-2	1.83e-3	7.46e-2	1.99e-3	1.81e-5
0.05	1.57e-1	1.24e-1	4.48e-3	1.14e-1	4.82e-3	7.03e-5
0.075	2.15e-1	1.68e-1	8.74e-3	1.54e-1	9.28e-3	1.95e-4
0.1	—	2.05e-1	1.37e-2	1.87e-1	1.43e-2	3.90e-4
0.15	—	2.59e-1	2.49e-2	2.38e-1	2.51e-2	9.72e-4
0.2	—	—	—	2.74e-1	3.60e-2	1.78e-3
0.25	—	—	—	3.00e-1	4.62e-2	2.79e-3

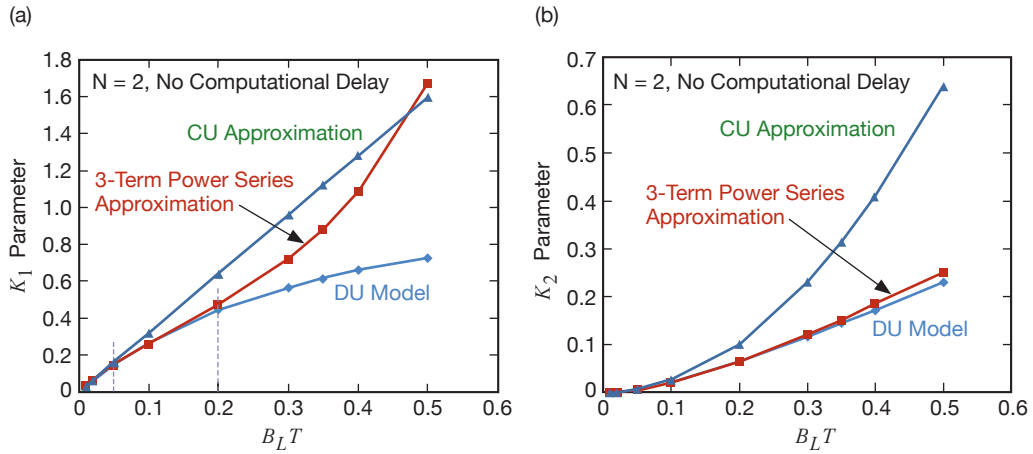

Figure 2. (a) Comparison of estimates of the loop parameter K_1 for a second-order supercritically damped loop, with no computational delay; (b) comparison of the K_2 parameters for the same conditions.

the error signal, hence the noise model changes due to the squaring operation, leading to "squaring loss" in loops of this type. Linear loops do not suffer from squaring loss, but may have other loss mechanisms that affect their performance, such as the "transition-detection" portion of the DTTL, which leads to additional errors at low SNRs. These differences will be examined in the following performance comparisons.

Before examining loop behavior, we need to describe the design of the roots to achieve the desired transient response. In the DU model developed in [3], the roots are parameterized pairwise as

$$\{z_1, z_2; z_3, z_4; \cdots\} = \{\exp[-\beta_1(1 \pm \eta_1)T; \exp[-\beta_1\lambda_1(1 \pm \eta_2)T; \cdots\}, \tag{11}$$

where η^2 is known as the damping parameter in analog loop design, and β is the decay-rate parameter. When $\eta^2 > 0$, there are two real roots and the loop is overdamped; when $\eta^2 = 0$, there are two equal roots and the loop is critically damped; when $\eta^2 < 0$, the roots form a complex conjugate pair, leading to underdamped and hence oscillatory behavior. The loop is called "supercritically damped" when $\eta_i^2 = 0$ and $\lambda_1 = 1 \quad \forall i$; with $\eta_i^2 = -1$ and $\lambda_1 = 1 \quad \forall i$, the loop is called "standard underdamped." These concepts are illustrated with the following example.

Example 3. Consider a supercritically damped loop with zero delay, so that $n_c = 0$, $\eta_i = 0, \lambda_i = 1$ and let $z_1 = z_2 = 0.9$, implying that $\beta_1 T = 0.10536$ via Equation (11). Substituting into the expression for the loop gains yields $K_1 = 0.19, K_2 = 0.01$. From the expression for normalized loop bandwidth in terms of the gain coefficients, we have

$$B_L T = \frac{2K_1^2 + 2K_2 + K_1K_2}{2K_1(4 - 2K_1 - K_2)} = 0.068596.$$

The power-series approximation for this value of $B_L T$ is $K_1 = 0.1906, K_2 = 0.009972$, not very different from the accurate DU values since this $B_L T$ is close to the region of agreement for these two models.

An example of the temporal response of a standard underdamped residual carrier phase-locked loop is shown in Figure 3, for a second-order loop with initial phase offset of π radians, indicated by the red line. Figures 3(a) and 3(b) show the loop's response, or NCO phase estimate, to the π radian phase offset and phase error over approximately 430,000 loop updates. The phase error is the difference between the loop's phase estimate and the true phase, hence it is a zero-mean process after the transient response has died down. Figures 3(c) and 3(d) illustrate the transient response in the beginning of the track for the first 250 samples. Statistics were obtained over the long observation interval after the transient response has died down, in order to characterize the loop's performance.

The variance of the phase error for a residual carrier tracking loop is shown in Figure 4 as a function of P_t/N_0. For this example, all of the power is in the carrier, hence there is no modulation, characterized by $m = 0$. The loop bandwidth is set to 100 Hz, which is typical for tracking loops on a spacecraft, and the update interval is $T = 0.0005$ s, yielding $B_L T = 0.05$. We can see in Figure 2 that for such small values of normalized loop band-

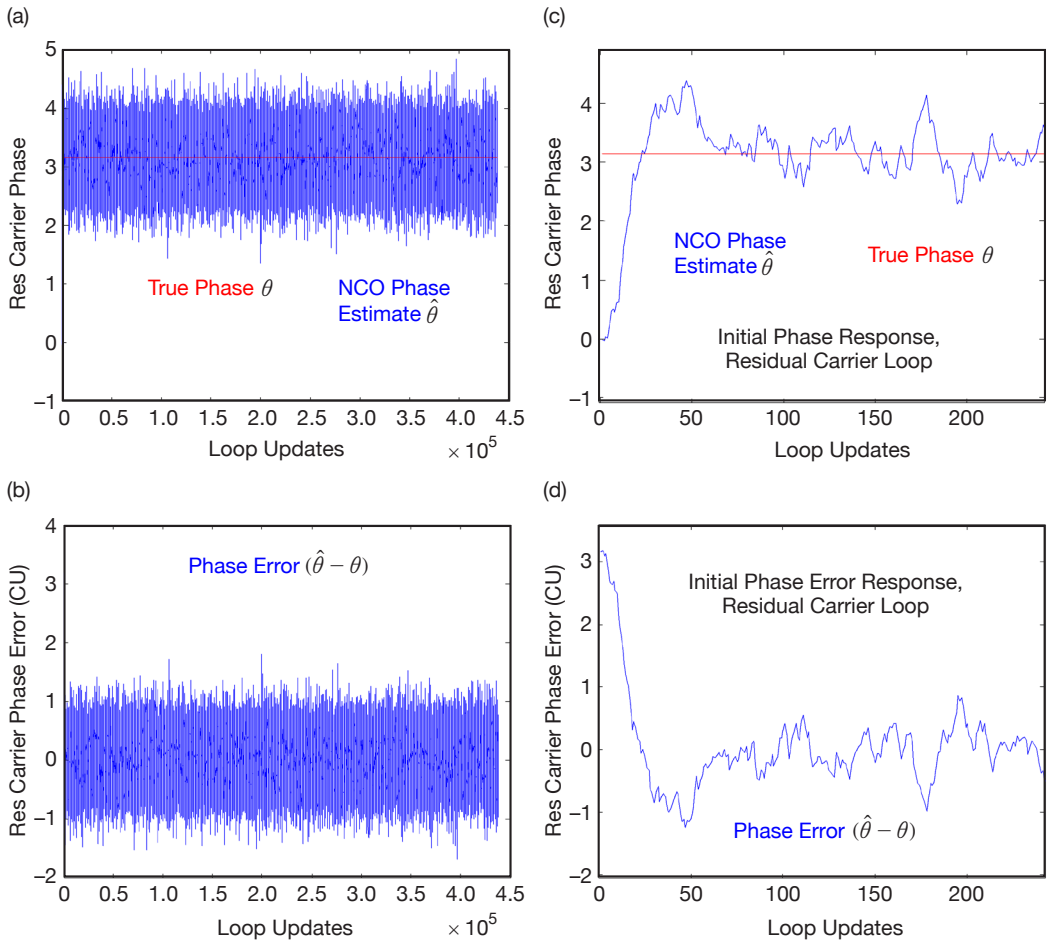

Figure 3. Phase (a) and phase-error (b) time-histories of the residual carrier tracking loop, with zero Doppler and initial phase offset of π radians; (c) and (d): transient response in the beginning of the track for the first 250 samples.

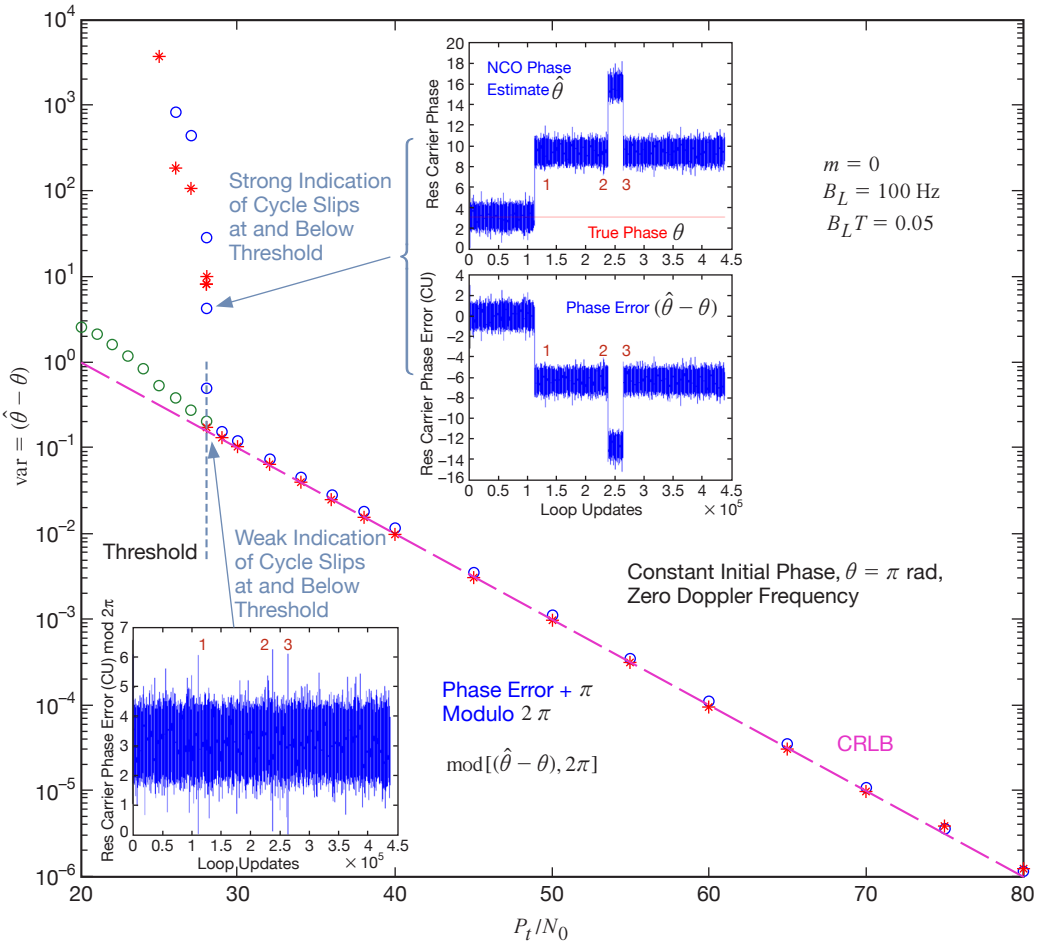

Figure 4. Residual carrier tracking loop performance as a function of P_t/N_0 (in dB), $m = 0$, $B_L T = 0.05$; top inset: loop phase response near threshold, showing three cycle slips of 2π radians each in the middle of the track for the loop using the approximate CU parameters (blue circles); bottom left inset: modulo 2π phase error, where the three large discontinuities have been replaced with narrow spikes, yielding weak indication of cycle slips. The dashed magenta line is the Cramér-Rao lower bound (CRLB) on residual carrier tracking loop performance, and the magenta asterisks represent simulation results with DU parameters.

width, the CU and DU models yield nearly identical results for both K_1 and K_2, hence we do not expect to see much difference in loop performance, measured in terms of the variance of phase error, $\mathrm{var}(\hat{\theta}-\theta)$. Note that even if the modulation index is greater than zero, meaning that part of the total power has been allocated to a modulated subcarrier, it is assumed that none of the modulation enters the tracking loop as interference because the subcarrier frequency is much greater than the loop bandwidth. Hence, the only impact on loop performance is the loss of residual carrier power when the modulation index is greater than zero.

A. The Cramér-Rao Lower Bound

The Cramér-Rao lower bound (CRLB) on estimator performance provides a convenient way to evaluate the tracking loop's performance, since it is a bound on the smallest average phase error that any phase estimator can achieve. For in-phase and quadrature sinu-

soidal signals with amplitude $A\cos(m)$ and unknown phase θ observed in the presence of zero-mean additive Gaussian noise with variance σ^2, the CRLB is given by the following expression:

$$\mathrm{var}(\hat{\theta}-\theta) \geq \frac{\sigma^2}{NA^2\cos^2(m)}$$

where N is the number of independent samples observed.

For a unit-amplitude signal such that $A = 1$ (as required for the phase-estimator in the phase-locked loops considered in this article), and expressing the variance of the noise in terms of the ratio of total signal power to noise spectral density P_t/N_0 as

$$\sigma^2 = \left(2T_s\frac{P_t}{N_0}\right)^{-1},$$

the CRLB becomes

$$\mathrm{var}(\hat{\theta}-\theta) \geq \frac{\sigma^2}{NA^2\cos^2(m)} = \left(2NT_s\cos^2(m)\frac{P_t}{N_0}\right)^{-1}.$$

A phase-locked loop with closed-loop bandwidth B_L can be viewed as a short-term integrator with effective integration time $T_L = 1/2B_L$, hence the effective number of T_s second samples ($T_s << T_L$) in a T_L second time interval is $N = T_L/T_s = 1/2B_LT_s$. Approximating N as the number of T_s second samples per loop integration time T_L, $N = T_L/T_s \cong (2B_LT_s)^{-1}$, we can express the CRLB as

$$\mathrm{var}(\hat{\theta}-\theta) \geq \underset{A=1}{=}\left(2NT_s\cos^2(m)(P_t/N_0)\right)^{-1} = \left(2\left(\frac{1}{2B_LT_s}\right)T_s\cos^2(m)(P_t/N_0)\right)^{-1} = \frac{B_L}{\cos^2(m)P_t/N_0}.$$

This form of the CRLB is most useful in evaluating simulated loop performance.

For loops that effectively employ a squaring operation to reconstruct the carrier from the modulated subcarrier (or from direct modulation), such as Costas loops or conventional DTTLs where the noisy transition estimate is multiplied by a noisy symbol delay estimate, the CRLB depends on $(T_sP_t/N_0)^2$, and is given by $\mathrm{var}(\hat{\theta}-\theta) \geq B_LT_s/(2\sin^2(m)(T_sP_t/N_0)^2)$. Therefore, we need to consider both a "high-SNR" and a "low-SNR" CRLB, when dealing with loops or parameter estimators that employ a squaring operation.

It is well known that for high SNRs, phase-locked loops are nearly optimal, closely approaching the CRLB in terms of phase-error variance. This behavior is illustrated in Figure 4, where it can be seen that for $P_t/N_0 > 27$ dB-Hz, the measured variance of the simulated phase error indeed follows the CRLB closely, for both the CU and DU models. However, for $P_t/N_0 < 27$ dB-Hz, there is a catastrophic increase in phase-error variance, due to cycle slips that contribute factors of 2π radians to the phase error each time they occur. Note that this threshold behavior starts at approximately 27 dB-Hz, where the variance of the phase error reaches ~ 0.3 rad^2, or a standard deviation of about 0.55 radians. With high-rms phase errors, occasional spikes of several radians can occur due to noise, leading to cycle slips that may eventually unlock the loop. In the insets in Figure 4, the three cycle-slip events are

labeled 1, 2, and 3, in order of occurrence. Measuring the number of cycle slips in a given time interval therefore provides a sharp indication of large errors, quickly warning the operator when the loop is not operating in its linear region, or has "lost lock."

Another method of determining the "in-lock" condition is by measuring the phase error modulo 2π, as shown in the inset on the lower left of Figure 4. However, with this approach, the phase error shows up only as a narrow spike of magnitude 2π radians for each discontinuity in phase, hence does not contribute much to the measured variance of the phase error. A very large number of cycle slips must occur before the measured variance changes significantly, hence this approach is not as decisive as the direct phase-error variance method. The modulo 2π phase error estimates are shown as green circles in Figure 4.

The impact of partitioning power between the residual carrier and the data modulation by changing the modulation index is illustrated in Figure 5, where two cases are shown: $m = 0$ and $m = 1.5$ radians. With $m = 0, \cos^2(0) = 1$, and all of the power is allocated to the carrier. Therefore, the loop benefits from the high power in the carrier, and achieves performance shown in Figures 4 and 5 (curves labeled $m = 0$).

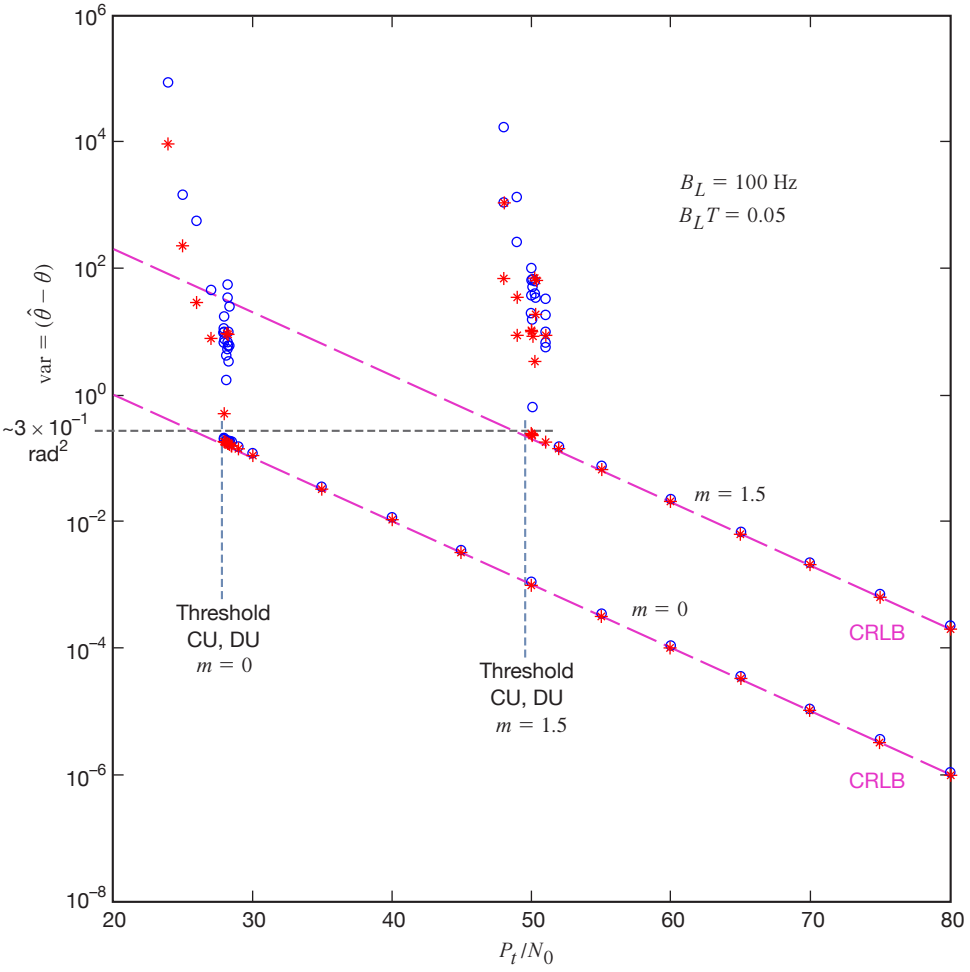

Figure 5. Residual carrier tracking loop performance in terms of phase-error variance, with two different (and extreme) values of modulation index, $m = 0$ and $m = 1.5$, as a function of P_t/N_0 (in dB).

When m = 1.5 radians, $\cos^2(1.5) = 5 \times 10^{-3}$, most of the power is allocated to the modulation and the power in the carrier has been reduced by 23 dB. Therefore, the entire performance curve is shifted to the right by 23 dB, leading to greatly reduced performance at any given value of P_t/N_0. The threshold again occurs at the point where the rms phase error begins to exceed 0.5 radians, but now this point is 23 dB higher, hence the loop will lose lock at a much higher value of P_t/N_0, rendering the entire link inoperable when it occurs. Therefore, care must be exercised in partitioning the power between the residual carrier and the modulation to ensure best end-to-end performance: with more power allocated to the modulation, detection performance improves provided the carrier-tracking loop remains in lock and tracks with acceptably small radio loss. But if the tracking loop begins to experience cycle slips or loses lock, then coherent data detection becomes impossible, and the link cannot be maintained.

The importance of using the correct gain parameters provided by the DU model for higher values of $B_L T$ is shown in Figure 6, selected for the purposes of illustration where the greatest divergence between the CU and DU models occurs in Figure 2, namely $B_L T = 0.5$. With B_L = 100 Hz, it follows that the loop update rate is T = 0.005 s, hence the loop is operating at 2-kHz update rate, which is reasonable for tracking loops on spacecraft.

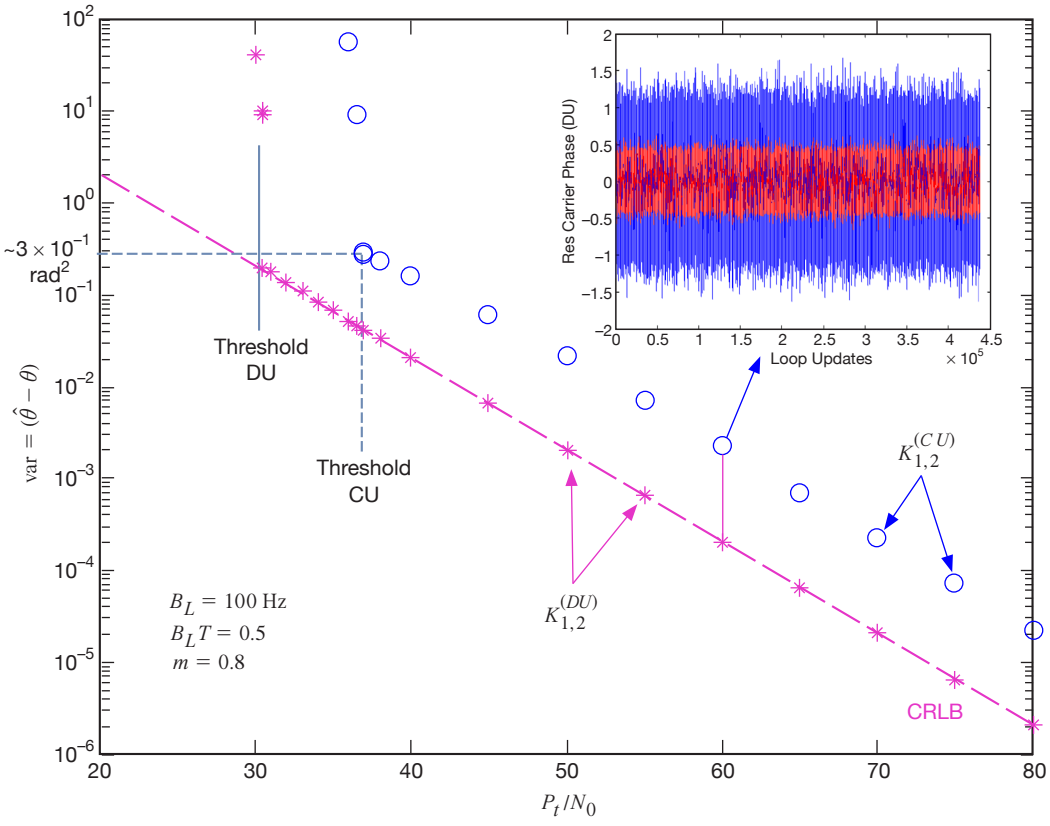

Figure 6. Comparison of residual carrier tracking loop performance using CU and DU gain parameter models, at a high value of normalized loop bandwidth: $B_L T \leq 0.5$.

As before, the CRLB is the dashed magenta line, representing the best theoretically attainable performance for any phase estimation strategy; the magenta asterisks are the measured phase-error variance obtained from the simulation when the DU gain parameters are used; and likewise the blue circles are the measured phase-error variance with the CU gain parameters. It is clear that loop performance is approximately 10 dB worse in terms of both phase-error variance and P_t/N_0 over the entire operating range where the loop is in lock. This conclusion is also supported by the inset, which is a time-history of the phase error for both the CU model (blue) and DU model (red) generated by the simulation: from visual inspection, the rms value of phase error produced by the CU parameters is approximately three times as great as with the DU parameters, confirming the 10 dB difference. These simulation results confirm the importance of using the more complex DU parameters instead of the readily available CU parameters for which straightforward closed-form expressions exist, but which lead to greatly degraded loop performance for large values of $B_L T$.

The simulation results presented in Figures 4–6 also confirm the analytic predictions of Figure 2, where it can be seen that for $B_L T \leq 0.05$, both CU and DU models yield similar estimates for the gain parameters, however the models diverge rapidly for larger normalized loop bandwidths. This highlights the importance of calculating the gain coefficients via the more complicated DU model, using algorithms similar to the MATLAB code described in this article, in system designs where large loop bandwidths and relatively slow update rates are required.

An example of loop behavior in the presence of squaring loss is shown in Figure 7, where the performance of the Costas loop used on the ground to recover the carrier with suppressed-carrier modulation ($m = \pi/2$) is evaluated. It is assumed in this example that the symbol boundaries are known, so that samples can be summed over a symbol duration to reduce the noise variance, before multiplying the I and Q components to generate an error signal. The CRLB for the high-SNR region is the same as for conventional phase-locked loops, that is, inversely proportional to P_t/N_0. However, in the low-SNR region (nominally below 60 dB-Hz), the CRLB is proportional to $(P_t/N_0)^{-2}$, yielding greater phase errors with decreasing SNR. The simulation points follow the envelope, or more precisely the greater of the low-SNR and high-SNR CRLB curves, until the variance of the phase error reaches approximately 0.3 rad^2 or, equivalently, standard deviation of 0.55 radians.

At this point, cycle slipping or loss of lock is likely to occur, as discussed earlier. With the steeper slope of the low-SNR CRLB, this critical point is reached at approximately 44 dB-Hz for the Costas loop under the stated conditions, causing catastrophic degradation in performance.

B. Range Resolution via Phase-Locked Loops

The phase-tracking ability of residual-carrier tracking loops can be applied to the problem of estimating range, when the received signal is an ambiguity-resolving compound PN code, the main component of which is a square wave, as is often the case [3]. An example of a unit-amplitude PN code received in the presence of noise and satisfying this condition is shown in Figure 8: this is a portion of the PN code used in the simulation. It can be seen

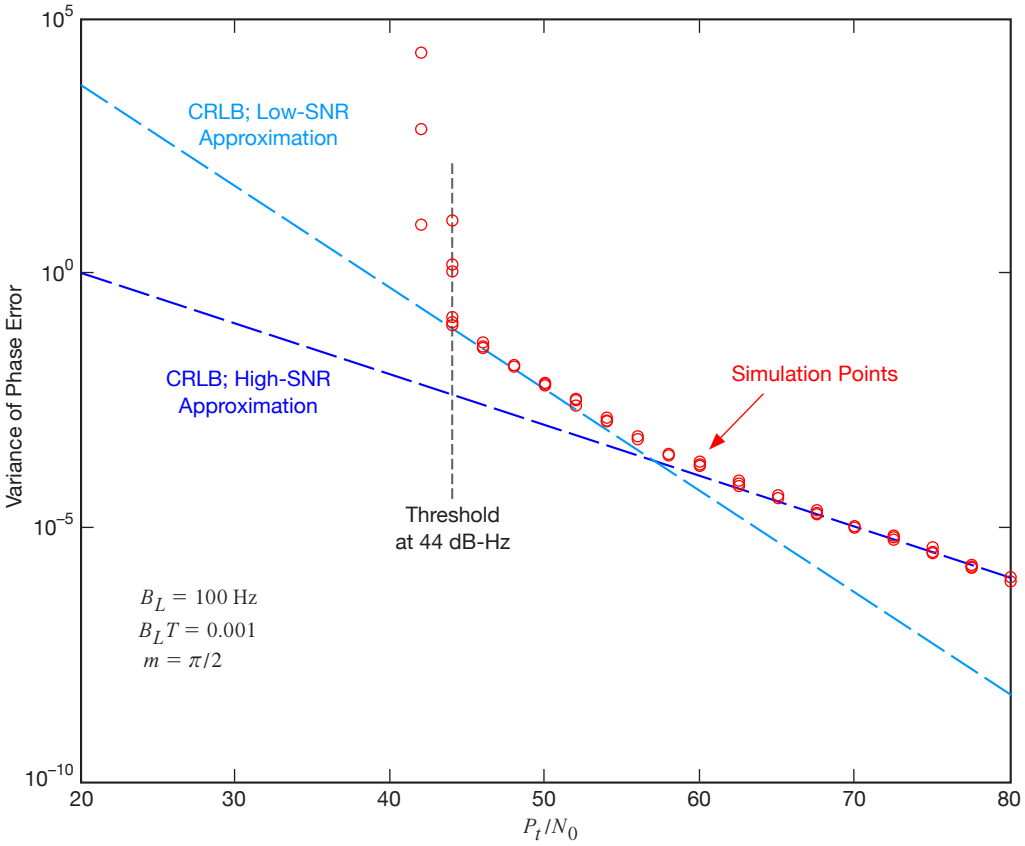

Figure 7. Simulated Costas loop performance as a function of P_t/N_0, showing the limiting cases of low-SNR and high-SNR CRLB, as well as the loss-of-lock threshold at 44 dB-Hz (10^6 chips/s, 10 samples/chip).

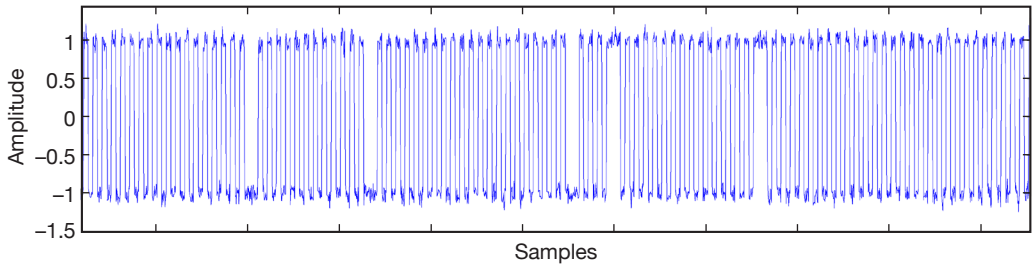

Figure 8. Example of a portion of the composite PN code used in the simulation, showing five full-cycle flips caused by higher-order components of the composite PN code.

that the received signal is dominated by a periodic square wave, with occasional sign flips over a chip, constructed in such a way that the phase continuity of the waveform after the sign flip is maintained.

This phase-locked loop concept for establishing chip-synchronization with ambiguity-resolving PN sequences has been simulated on the uplink portion of the link, where it is used to estimate the delay due to transit, thus equivalently estimating uplink range. After demodulation, and assuming that the carrier loop aboard the spacecraft is tracking with small phase error, the quadrature component is of the form $Q = A\sin[m]\cos(\Delta\theta)\mathrm{Sq}(\omega_{sc}t)d(t) + n_Q(t)$, where m is the modulation index and $d(t)$ describes the PN code modulation: for the simulation, we can approximate the received PN code as a square wave, but account for the loss in the error signal introduced by the sign flips caused by higher-order PN code components.

Since the generation of the error signal relies on a full cycle of the received waveform, multiplied by a sinusoid of the same frequency and essentially in phase with the received signal, a sign reversal over a chip cancels out the error signal over an entire cycle, introducing an effective loss. The fraction of gaps in typical PN codes used for ranging is about 10 percent, hence we assign a loss factor of 0.9 to the received signal power in the simulation.

For this simulation, we assume that the modulation index takes on its maximal value of $\pi/2$, implying that all of the power is in the modulation, even though in practice the power is typically divided optimally between the residual carrier and the modulation for best performance (alternately, a Costas loop could be used with a suppressed carrier to track the phase). With these approximations, and further assuming that the demodulated PN code (here approximated as a square wave) is first filtered to the fundamental component, resulting in a sinusoid of amplitude $4/\pi$ but with power a fraction $\frac{1}{2}(4/\pi)^2 = 0.81$ of the total received modulation power, and input directly to the second-order carrier tracking loop operating at the frequency of the approximating 1 μs chip square wave, or 0.5 MHz, updated at a rate high enough to ensure the validity of the CU model. An additional loss factor of 0.81 is assumed to account for mismatch between the approximate square wave input and the sinusoidal NCO waveform. Since the amplitude of the received signal must be unity in order to extract the correct phase, the standard deviation of the noise added to the signal was increased in proportion to the loss factor in the simulation.

In order to estimate range, it is necessary to translate the phase of the 0.5-MHz sinusoid (representing the square-wave component of the composite PN code with 1-μs chips) into units of range, namely meters. This can be accomplished by first expressing the phase-error results in terms of cycles instead of radians, then expressing the fractional cycle error in meters. Since a chip duration of 1 μs corresponds to 300 m of range, it follows that one radian error represents $300/\pi$ = 95.5 m of range (since a chip is half a cycle). The results of the simulation in terms of range resolution (meters) are shown in Figure 9, in terms of P_t/N_0 of the received signal, with a modulation index of $\pi/2$.

The CRLBs for both the lossy and lossless signal-power models are shown in Figure 9, translated to range in meters instead of carrier phase in radians. The CRLB for the lossless case,

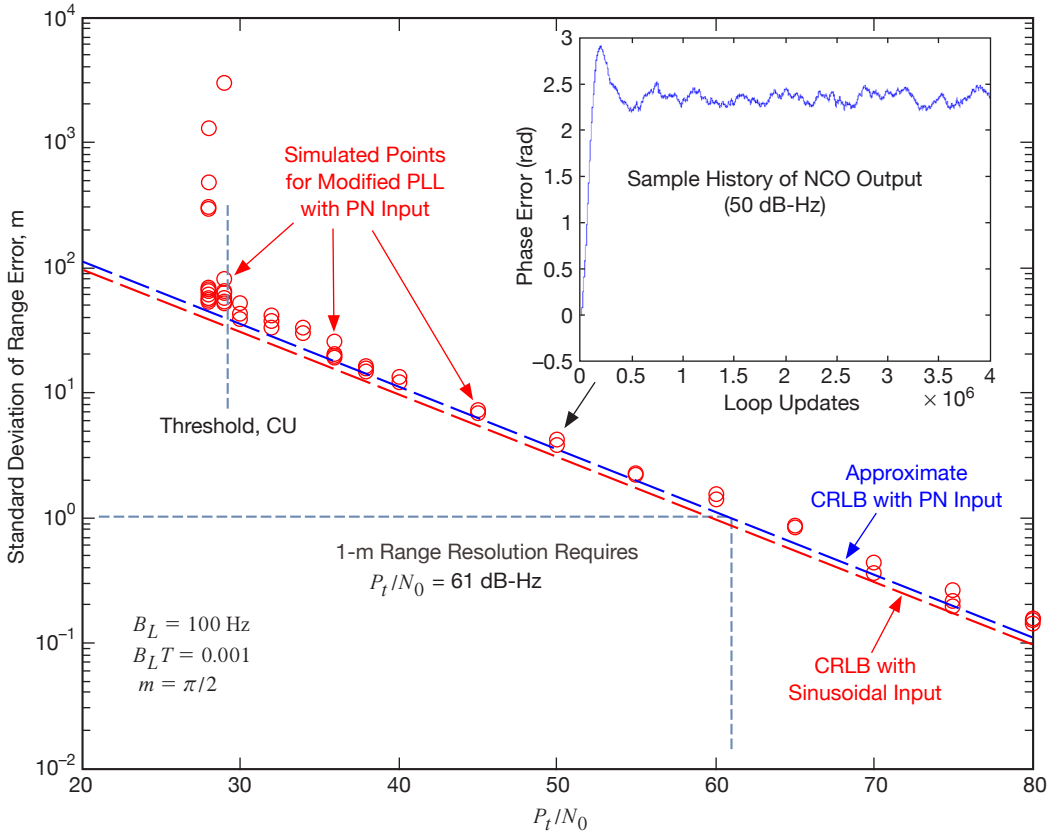

Figure 9. Range resolution performance of second-order carrier tracking loop applied to tracking ambiguity-resolving PN codes on the uplink part of the two-way link.

translated to meters of range, is shown by the red dashed line: the blue dashed line is the CRLB with losses included. The simulation follows the lossy CRLB accurately above threshold, verifying the validity of the model. Note that with total losses of $0.81 \times 0.9 = 0.729$, corresponding to –1.4 dB, the rms error in range does not degrade significantly, hence the system is not sensitive to small losses.

Symbol synchronization on the data-modulated downlink portion of the link is achieved with a DTTL. The basic idea is to process samples over two consecutive chips as shown in Figure 10. In the results presented here, each chip interval is sampled 10 times, although the sampling rate is a user adjustable parameter in the simulation.

In Figure 10, we assume that the chip boundaries are synchronized with the receiver clock, samples from each interval are summed, forming the "in phase" random variables I_{2k-1} and I_{2k}, and the sign of the random variables determined, assigning the value +1 if the sign of the in-phase sum is positive, and –1 otherwise. This operation constitutes a hard decision on the chips, viewed as binary symbols. Simultaneously, a "quadrature" sum Q_k is also obtained over the second half of the first interval and the first half of the second interval, again determined from the receiver clock. This sum can be used as an error signal for chip delay, since any offset between the receiver clock and the PN sequence generates an error signal when there is a symbol transition over the quadrature interval.

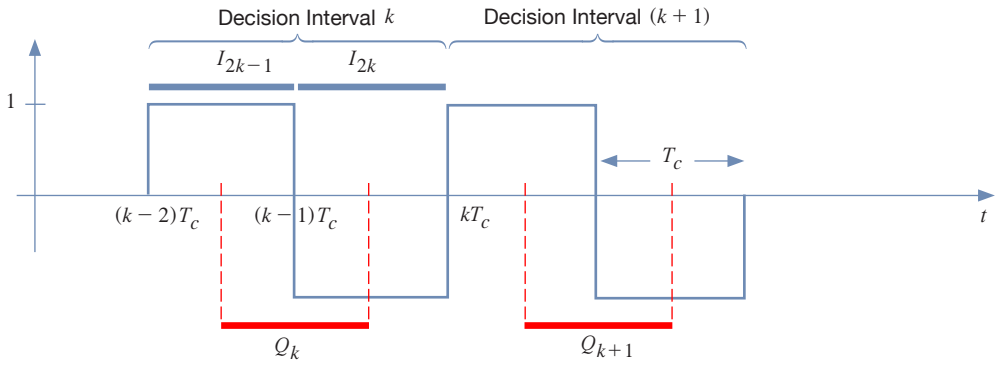

Figure 10. DTTL signal processing timing diagram, illustrating the in-phase and quadrature interval concepts (T_c refers to chip or symbol interval).

However, it is necessary to determine if a transition has occurred, as well as the direction of the transition, in order to correct the sign of the error signal when there is a transition, or exclude the quadrature data from the loop when there is no transition, since in that case only noise would be added to the loop, degrading its performance. This is accomplished by subtracting consecutive in-phase random variables, forming the decision variable $D_k = \frac{1}{2}(I_{2k-1} - I_{2k})$: if $D_k = +1$, then a transition from plus to minus was detected; if $D_k = -1$, then the transition was from minus to plus; if $D_k = 0$, there was no transition. The quadrature variable is multiplied by the decision variable, and the resulting delay error estimate input to the loop: $\delta_k = D_k Q_k$. Note that if the decision variable is zero, then there is no input to the loop. We observe that while this processing is sensible and works well at high SNR, it is not motivated by the maximum likelihood approach, where each pair of consecutive chips would be considered a distinct symbol that takes on one of four realizations, and would be estimated optimally by forming the weighted sum of samples over the two intervals and selecting the largest numerical value.

It is clear that the transition decision D_k is subject to error due to noise, especially at low SNR, and that the quadrature error signal Q_k is also a noisy estimate. Therefore, the product of these two noisy random variables constitutes a squaring operation, hence we expect to see squaring loss in the error variance of the DTTL, similar to that of the Costas loop described earlier.

The performance of the simulated DTTL is shown in Figure 11, as a function of P_t/N_0, along with approximate expressions for high- and low-SNR performance. The simulation used a half-chip window for the quadrature interval, to reduce the noise within the loop bandwidth: this suggested replacing P_t/N_0 with $2P_t/N_0$ in the CRLB expressions, leading to good agreement with the simulation. The output of the DTTL is an estimate of chip delay, hence the delay estimates are bounded by ±1. The delay estimates have been renormalized in Figure 11 to represent range, instead of fractional chip delay, by converting 1 μs chip duration to 300 m of range. Squaring loss is evident below 60 dB-Hz, as in the Costas loop, due to the inherent squaring operation performed in the DTTL.

Finally, the loss-of-lock threshold of the simulated DTTL was estimated by computing the probability of correct transition detection, assuming zero average delay in the received

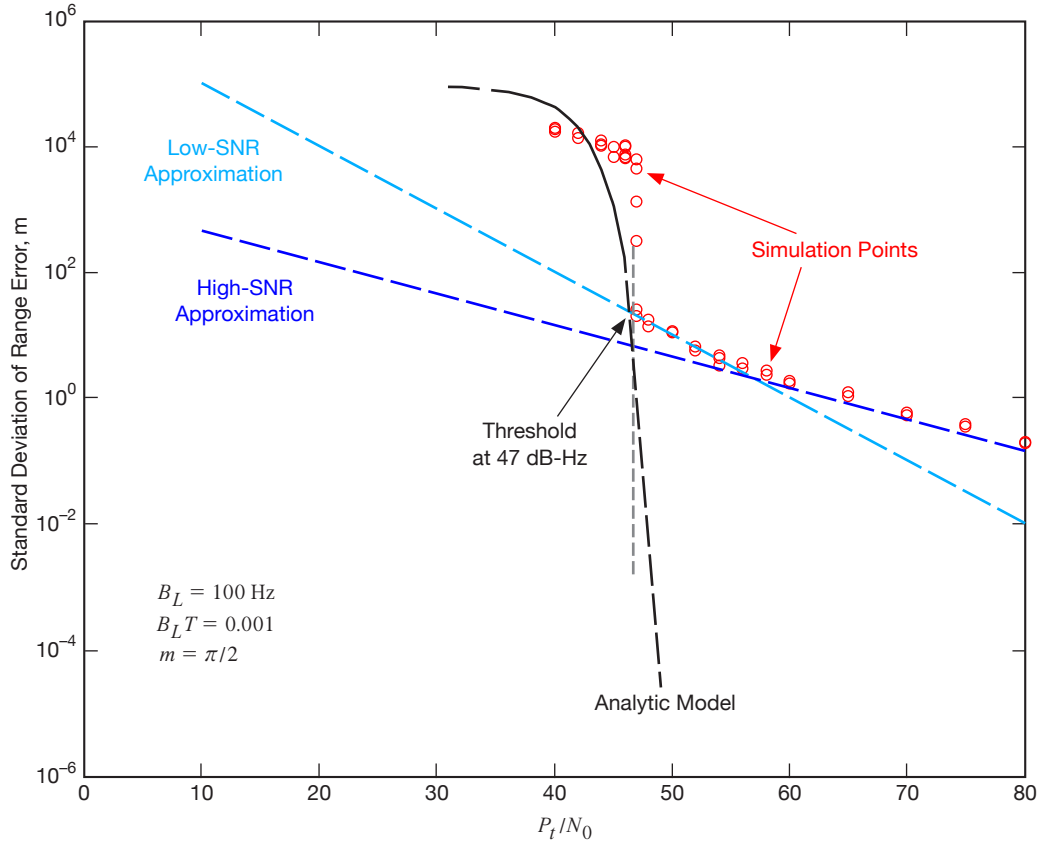

Figure 11. Performance of the simulated DTTL as a function of P_t/N_0, demonstrating agreement with the squaring loss analysis, and showing good agreement with the predicted threshold behavior.

symbols. The individual decisions were summed over an inverse loop bandwidth and multiplied by the gain factor K_1 (K_2 is so small in this case that it can be safely ignored). Since the loop bandwidth is narrow compared to the modulation, the probability of correct transition detection based on this weighted sum was determined, and scaled according to the maximum delay error observed in the simulation. As can be seen in Figure 11, this approximate analysis is consistent with the breakdown threshold of 47 dB-Hz in the simulation.

V. Summary

Telemetry ranging is a novel idea currently being explored at JPL. There are several potential advantages of telemetry ranging, including the ability to conduct simultaneous ranging and high-rate telemetry throughout the duration of a pass, without sacrificing power or bandwidth to a separate ranging channel. In addition, it has been shown in [2] that for low data rates the ranging resolution is as good as that of conventional ranging, but for data rates more than about 1 Mbps, it can be orders of magnitude better.

In order to achieve the range resolution required by deep-space links, extremely accurate range and range-rate determination is essential, which can be accomplished by the use of properly designed digital tracking loops. For the relatively high operating bandwidths employed by digital tracking loops on spacecraft to combat dynamics and pull-in require-

ments, it is important to use the correct digital gain coefficients to ensure that the design bandwidth is closely approximated in practice. The underlying theory needed to compute the correct digital gain coefficients as a function of design loop bandwidth has been reviewed, clarifying examples have been provided, and a MATLAB code designed to iteratively converge on the correct gain coefficients for any desired closed-loop bandwidth has been described. Since the gain coefficients and closed-loop bandwidth are essentially independent of the method of determining the error signal, a straightforward residual-carrier loop structure was simulated and analyzed to provide a clear description of the techniques, and performance compared to the Cramér–Rao lower bound. It was shown that range estimation errors of 1 m or less could be achieved with SNRs P_t/N_0 of 60 dB-Hz or greater, and that up to 10 dB improvement in phase-error performance could be achieved through the use of the correct digital gain coefficients when the normalized loop bandwidth is large, over a wide range of SNRs characteristic of deep-space telemetry and ranging applications.

References

[1] P. W. Kinman, "Pseudonoise and Regenerative Ranging," *DSN Telecommunications Link Design Handbook*, DSN No. 810-005, Module 214, Rev. E, Jet Propulsion Laboratory, Pasadena, California, March 31, 2004.
http://deepspace.jpl.nasa.gov/dsndocs/810-005/214/214-1.pdf

[2] K. Andrews, J. Hamkins, S. Shambayati, and V. Vilnrotter, "Telemetry-Based Ranging," *Proceedings of the 2010 IEEE Aerospace Conference*, Big Sky, Montana, March 2010.

[3] R. J. DeBolt, D. J. Duven, C. B. Haskins, C. C. DeBoy, and T. W. LeFevre, "A Regenerative Pseudonoise Range Tracking System for the New Horizons Spacecraft," 61st Annual Meeting of the Institute of Navigation, Fairfax, Virginia, Summer 2005.

[4] S. A. Stephens and J. B. Thomas, "Controlled-Root Formulation for Digital Phase-Locked Loops," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, no. 1, January 1995.

[5] V. A. Vilnrotter, W. J. Hurd, and D. H. Brown, "Optimized Tracking of RF Carriers with Phase Noise, Including Pioneer 10 Results," *The Telecommunications and Data Acquisition Progress Report*, vol. 42-91, Jet Propulsion Laboratory, Pasadena, California, pp. 1–17, November 15, 1987.
http://ipnpr.jpl.nasa.gov/progress_report/42-91/91O.PDF

[6] S. A. Stephens and J. B. Thomas, "Method of Implementing Digital Phase-Locked Loops," U. S. Patent 5,602,883, February 11, 1997.