Statistical Analysis of First-Order Bang-Bang Phase-Locked Loops Using Sign-Dependent Random Walk Theory

Stefan Tertinek, James P. Gleeson, and Orla Feely, *Fellow, IEEE*

***Abstract*—Bang-bang phase-locked loops (BBPLLs) are inherently nonlinear due to the hard nonlinearity introduced by the binary phase detector (BPD). This paper provides an exact statistical analysis of the steady-state timing jitter in a first-order BBPLL when the reference clock is subject to accumulative jitter. By elaborating on the analogy of viewing a first-order BBPLL as a single-integration delta modulator (DM) in the phase domain, we are able to relate hunting jitter and slew-rate limiting in a BBPLL to granular noise and slope overload in a DM. The stochastic timing-jitter behavior is modeled as a sign-dependent random walk, for which we obtain the asymptotic characteristic function and analytical expressions for the first four cumulants. These expressions are applied to the BBPLL to statistically analyze the static timing offset and the RMS timing jitter, including the effect of a frequency offset. The analysis shows that the RMS timing jitter is constant for small RMS clock jitter and grows quadratically with large RMS clock jitter, and that there exists an optimal bang-bang phase step for minimum RMS timing jitter. Computing the kurtosis reveals the effect of the BPD nonlinearity: the timing jitter is largely non-Gaussian.**

***Index Terms*—Bang-bang phase-locked loop, delta modulator, timing jitter, sign-dependent random walk, cumulants, kurtosis.**

I. Introduction

BANG-BANG phase-locked loops (BBPLLs) have been widely used for clock and data recovery (CDR) in serial data links [1], primarily due to their high-speed capabilities and inherent sampling phase alignment [2]. A typical implementation based on the charge-pump (CP) architecture is shown in Fig. 1(a). The distinct feature of BBPLLs is the binary phase detector (BPD) which binary quantizes the phase difference between input data and voltage-controlled oscillator (VCO) clock, generating only early/late phase-error information for the loop filter (LF). To suppress pattern-dependent jitter in a CDR application, the BPD is usually a tristate realization such as the Alexander topology [3]. BBPLLs have also been demonstrated for high-bandwidth digital frequency synthesis [4], [5]. A simple digital BBPLL (DBBPLL) implementation is shown in Fig. 1(b), which employs a D flip-flop as BPD [4]. Instead of the CP-based LF driving the VCO, a digital LF tunes the frequency of a digitally controlled oscillator (DCO) by a digital control word.

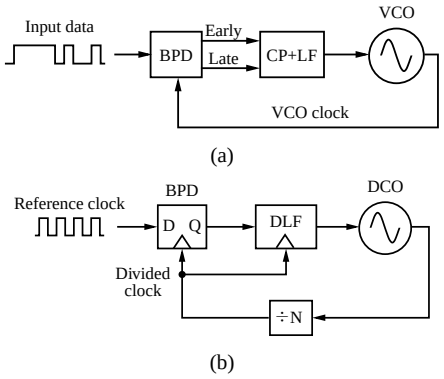

Fig. 1. BBPLL architecture for (a) CDR and (b) digital frequency synthesis.

Although the bang-bang loop principle has been successfully applied, a thorough understanding of the loop behavior is far from complete. The main reason is that the binary quantization introduced in the PD makes the loop inherently nonlinear, thus complicating its analysis. Traditionally, a wide class of PLLs can be linearized in steady state, allowing linear transfer functions to be applied in the analysis [6]. In BBPLLs, however, the hard nonlinearity introduced by the BPD causes limit cycles in steady state. In this case, a nonlinear analysis has been applied to derive conditions on the loop stability [7], and to investigate slew-rate limiting when the reference clock is frequency-modulated [8]. In practice, phase noise on the clock sources causes jitter on the clock edges, mainly in the form of non-accumulative jitter (white phase noise) and accumulative jitter (random-walk phase noise) [9], [10]. Since jitter effectively smoothes the binary PD characteristic [11], it is common to linearize the nonlinear loop [12] and apply linear transfer functions in the analysis [13]. Unlike a linear PD, the gain in a linearized BPD depends on the timing jitter at its input. For non-accumulative jitter, initial investigations considered the BPD as a stand-alone device [14], and only recently have the loop dynamics been taken into account to obtain a more accurate BPD gain expression [15].

Despite its widespread use, linearizing the binary loop precludes a thorough understanding of the nonlinear phase-jitter behavior. This is true particularly if clock jitter in the loop is small, in which case limit cycles and jitter interact

Manuscript received June 23, 2009; revised November 20, 2009. First published xxx, 2010; current version published xxx, 2010. S. Tertinek and O. Feely are funded by Science Foundation Ireland (SFI). J. P. Gleeson is funded by SFI and by MACSI. This paper was recommended by Associate Editor Sergio Callegari.

S. Tertinek and O. Feely are with the Circuits and Systems Group, Department of Electrical, Electronic and Mechanical Engineering, University College Dublin, Ireland (e-mail: stefan.tertinek@ucd.ie; orla.feely@ucd.ie).

J. P. Gleeson is with the Department of Mathematics and Statistics, University of Limerick, Ireland (e-mail: james.gleeson@ul.ie).

strongly. The limits of a linear BBPLL analysis were pointed out in [13], which motivates our nonlinear stochastic analysis.

A. Contribution of this Paper and Relation to Other Work

Recently, Markov theory has been applied to more accurately model the timing jitter in a first-order BBPLL [15]–[18], an approach known from early investigations into digital PLLs [19]. For a reference clock with non-accumulative jitter, Da Dalt [15] modeled the timing jitter as a Markov chain and derived a general BPD gain expression. By modeling the loop as a delayed Markov chain, Chun and Kennedy [16] gave an extension to a DBBPLL with nonzero loop delay and evaluated the timing-jitter performance. For a reference clock with accumulative jitter, our approach in [17] was to model the timing jitter as a discrete-time Markov process. A numerical solution of the Chapman-Kolmogorov equation allowed us to compute the steady-state timing jitter probability density function (PDF) and reveal its non-Gaussianity. An extension to a DBBPLL with nonzero loop delay was given in [18].

The aim of this paper is to provide an exact statistical analysis of the steady-state timing jitter in a first-order BBPLL when the reference clock is subject to accumulative jitter, thus analytically verifying our results in [17] and complementing previous work on the non-accumulative jitter case [15], [16].

The first contribution is to elaborate on the analogy between BBPLLs and delta modulation, which has appeared in the literature [1], [20], [21] but has never been fully exploited. In Sec. II a detailed comparison of the difference equations of both systems shows why and to what extent a first-order BBPLL can be viewed as a single-integration delta modulator (DM) in the phase domain. The analogy provides an intuitive description of the nonlinear loop behavior, and allows us to relate dither/hunting jitter and slew-rate limiting in a BBPLL to granular noise and slope overload in a DM.

The main contribution of this paper is the analysis of a sign-dependent random walk (SDRW) and its application to a first-order BBPLL. In Sec. III we formally define the SDRW as a RW whose step distribution depends on the sign of the RW's current position; a similar model was considered in [22], [23]. It will be shown that the SDRW is a suitable model for the loop's statistical timing-jitter behavior. The analogy with a DM enables us to apply and generalize existing theory on delta modulation. In particular, extending Fine's results in [24], we investigate the limiting behavior of the SDRW by deriving its asymptotic characteristic function, from which analytical expressions for the first four cumulants will be obtained. In Sec. IV the cumulant expressions will be applied to statistically analyze the timing jitter. We will show how varying the RMS clock jitter and the frequency offset influences the static timing offset; that the RMS timing jitter is constant for small RMS clock jitter and grows quadratically with large RMS clock jitter; and that there exists an optimal bang-bang phase step for minimum RMS timing jitter. Computing the kurtosis reveals that the timing jitter is largely non-Gaussian.

II. First-Order BBPLL as DM in the Phase Domain

This section elaborates on the analogy of viewing a first-order BBPLL as a single-integration DM in the phase domain.

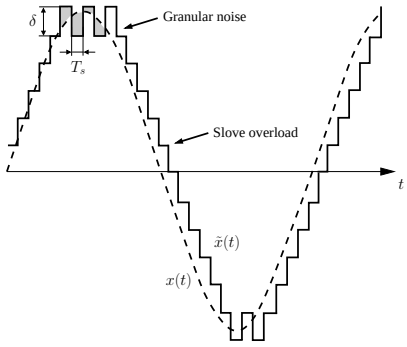

Fig. 2. A DM provides a staircase approximation of an analog signal [25].

We compare in detail the difference equations of both systems and use delta modulation terminology [25]–[27] to provide an intuitive description of the nonlinear BBPLL behavior.

A. Delta Modulation: Principle and Discrete-Time Model

A DM operates a 1-bit quantizer, a sampler and an integrator inside a feedback loop to provide a staircase approximation of an oversampled analog signal, as shown in Fig. 2 for a sinusoidal signal $x(t)$ [25]. At every sampling period T_s, the staircase signal $\tilde{x}(t)$ that approximates $x(t)$ is increased or decreased by the quantization step size δ. In this manner, a DM tracks an analog signal by changing the steps in the direction of the signal's slope, a behavior also exhibited by a BBPLL.

The performance of a DM is limited by two types of distortion [26]. Quantization distortion (granular noise) is caused by the granularity of the quantizer and occurs when $\tilde{x}(t)$ hunts around $x(t)$. Slope overload distortion is due to the DM's limited tracking speed and occurs when the slope of $x(t)$ exceeds the slope δ/T_s of $\tilde{x}(t)$. For a given sampling period T_s, a small δ will reduce granular noise but easily lead to slope overload; a large δ will allow the DM to track a fast varying signal but at the cost of increased granular noise. Minimizing both distortions results in a trade-off in selecting δ, a trade-off that also exists in a BBPLL.

Although implemented mostly with analog circuits and thus operated in continuous-time, a DM can be equally described at the sampling instants by the discrete-time model shown in Fig. 3 [27]. Given the sequence of input samples $x_n = x(nT_s)$ for $n \geq 0$, taken from $x(t)$ every T_s seconds (oversampling), a DM generates the staircase approximation recursively by

$$\tilde{x}_n = \tilde{x}_{n-1} + \delta \operatorname{sgn}(x_n - \tilde{x}_{n-1}) \tag{1}$$

where $\tilde{x}_n$ is the staircase value for $nT_s \leq t < (n+1)T_s$ and the quantizer is modeled by the signum function, defined as $\operatorname{sgn} x = 1$ for $x \geq 0$ and $\operatorname{sgn} x = -1$ for $x < 0$. The difference between the current input sample x_n and the previous staircase approximation $\tilde{x}_{n-1}$ is binary quantized into $\pm\delta$, and added to $\tilde{x}_{n-1}$ to form the approximation for the next input sample. The accumulation of the quantizer output values is represented by the discrete-time integrator with transfer function $z^{-1}/(1-z^{-1})$, the additional delay being part of any sampled feedback loop. Viewing the value $\tilde{x}_{n-1}$ that is subtracted from x_n as a prediction $\hat{x}_n = \tilde{x}_{n-1}$, we can define

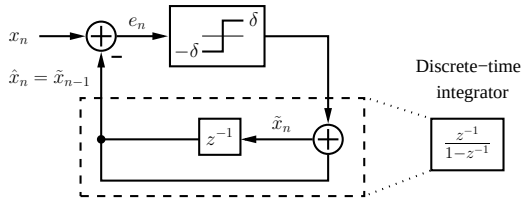

Fig. 3. Discrete-time equivalent model of a DM [27].

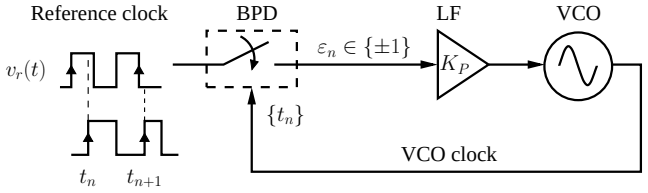

Fig. 4. Block diagram of a first-order BBPLL.

the prediction error

$$e_n = x_n - \hat{x}_n \tag{2}$$

which expresses the amount by which the input may be predicted exactly. Recursively, the sequence of predictions $\{\hat{x}_n\}$ is produced by

$$\hat{x}_{n+1} = \hat{x}_n + \delta \operatorname{sgn} e_n \tag{3}$$

and the corresponding prediction error sequence $\{e_n\}$ by

$$e_{n+1} = e_n + x_{n+1} - x_n - \delta \operatorname{sgn} e_n. \tag{4}$$

Equation (4) illustrates the dependence of the prediction error on the discrete-time derivative[1] $x_{n+1} - x_n$ and thus on the slope of the analog signal $x(t)$. In particular, slope overload is prevented if the no-overload condition [26]

$$\max \left| \frac{dx(t)}{dt} \right| < \frac{\delta}{T_s} \tag{5}$$

is satisfied, a condition that determines slew-rate limiting in the BBPLL.

B. Phase-Domain Model of a First-Order BBPLL

For a first-order BBPLL, both architectures in Fig. 1 may be represented by the block diagram in Fig. 4, assuming a 100% data transition density for the CDR loop in Fig. 1(a) [2]. We now rederive the difference equation governing the phase error in the phase domain [2], [7], following a derivation similar to the one in [19]. To emphasize its sampling nature, the BPD is represented as a sampler whose input signal (the reference clock) is sampled by the VCO clock. Since the loop is first order, the LF consists of a proportional path with gain coefficient K_P. The VCO is modeled as a linear block, with nominal frequency f_0 and frequency gain K_{F}.

The reference clock signal is a square wave of the form

$$v_r(t) = \operatorname{sgn}[\sin(\omega_0 t + \theta_r(t))] \tag{6}$$

which alternates between $+1$ and -1. We assume that its frequency be equal to the nominal VCO frequency $\omega_0 = 2\pi f_0$ (locked loop); any phase and frequency deviations will be incorporated into the excess phase $\theta_r(t)$, such as random phase noise and an actual frequency offset between both clocks.

The reference clock is sampled by the VCO clock; the sampling instants are the times of occurrence of the rising VCO clock edges. Denoting the time of the nth sampling instant by t_n, the sample value taken at t_n is the BPD output

$$\varepsilon_n = v_r(t_n) = \operatorname{sgn}[\sin(\omega_0 t_n + \theta_{r,n})] \tag{7}$$

where $\theta_{r,n} = \theta_r(t_n)$ is the sampled reference clock phase. The BPD output is either $+1$ or -1, indicating whether the reference clock has been sampled late or early with respect to its nth rising clock edge; no attempt is made to measure the actual time deviation from the clock edge, a feature significantly different from a linear PD [1]. The BPD output drives the VCO through the LF gain K_P, changing the VCO frequency so as to bring the sampling instants closer to the rising reference clock edges. Since the BPD values are binary, the VCO toggles between the two frequencies $f_0 + f_{\mathrm{bb}}$ and $f_0 - f_{\mathrm{bb}}$ which are set by the bang-bang frequency step $f_{\mathrm{bb}} = K_P K_{\mathrm{F}}$. During the nth update period—the time between the consecutive sampling instants t_n and t_{n+1}—the VCO operates at the frequency $f_0 + f_{\mathrm{bb}}\varepsilon_n$, and so produces the nonuniform sampling instants

$$t_{n+1} = t_n + \frac{1}{f_0 + f_{\mathrm{bb}}\varepsilon_n} = t_n + \frac{T_0}{1 + (f_{\mathrm{bb}}/f_0)\varepsilon_n} \tag{8}$$

for $n \geq 0$. In a practical application, the bang-bang frequency step f_{bb} is much smaller than the nominal VCO frequency f_0, in a CDR application typically around 0.1% [2]. Therefore, $f_{\mathrm{bb}}/f_0 \ll 1$, and with the approximation $1/(1+x) \cong 1 - x$ for x close to zero we can write (8) as

$$t_{n+1} \cong t_n + T_0 - \frac{\theta_{\mathrm{bb}}}{\omega_0}\varepsilon_n \tag{9}$$

where $\theta_{\mathrm{bb}} = 2\pi f_{\mathrm{bb}}/f_0$ is the bang-bang phase step of the loop. Writing (9) as a sum, assuming $t_0 = 0$, and plugging the result into (7) yields

$$\varepsilon_n = \operatorname{sgn}\left[\sin\left(\theta_{r,n} - \theta_{\mathrm{bb}} \sum_{i=0}^{n-1} \varepsilon_i\right)\right]. \tag{10}$$

The term involving the sum is the VCO phase $\theta_{v,n}$ during the nth update period. Writing it recursively as

$$\theta_{v,n+1} = \theta_{v,n} + \theta_{\mathrm{bb}}\varepsilon_n \tag{11}$$

shows that the BPD output ε_n causes the VCO phase to ramp up or down by θ_{bb} during every update period. Now define the phase error ϕ_n as

$$\phi_n = \theta_{r,n} - \theta_{v,n} \tag{12}$$

so that $\varepsilon_n = \operatorname{sgn}\phi_n$ for $-\pi \leq \phi_n \leq \pi$ from (10). Thus, with (11) and (12) we obtain the difference equation

$$\phi_{n+1} = \phi_n + \theta_{r,n+1} - \theta_{r,n} - \theta_{\mathrm{bb}} \operatorname{sgn}\phi_n \tag{13}$$

[1]More precisely, it is the backward-difference of the sampled input signal, which may be viewed as the discrete-time approximation to the continuous-time derivative of the input signal [26].

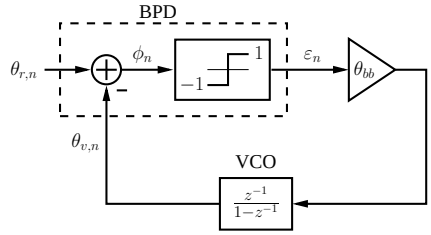

Fig. 5. Phase-domain model of a first-order BBPLL.

which governs the phase-error evolution as a function of the reference clock phase samples. The BBPLL operation in the phase domain, as expressed by (11)–(13), is summarized by the discrete-time model in Fig. 5.

C. First-Order BBPLL as DM in the Phase Domain

Comparing the models in Fig. 3 and Fig. 5 shows that for the given approximation, a first-order BBPLL can be viewed as performing single-integration delta modulation in the phase domain. In a DM, the sequence $\{\hat{x}_n\}$ produced by (3) forms a sequence of predictions for the input samples $\{x_n\}$. Changes in the prediction occur in steps of the step size δ, and the introduced error is the prediction error e_n. By analogy, in a first-order BBPLL, the sequence of VCO phases $\{\theta_{v,n}\}$ produced by (11) forms a sequence of predictions for the reference clock phase samples $\{\theta_{r,n}\}$. Changes in the prediction occur in steps of the bang-bang phase step $\theta_{\rm bb}$, and the phase error ϕ_n is the equivalent of the prediction error e_n in the DM. Furthermore, the sampling frequency f_s in the DM corresponds to the nominal VCO frequency f_0 in the BBPLL. Since a practical DM contains a zero-order hold circuit in its feedback loop [25], the staircase signal in Fig. 2 must be replaced by a signal with ramps, so that the behavior of both systems also correspond in continuous-time (compare Fig. 8 in [2]).

Referring to [28], Walker [2] provides an analogy between a second-order BBPLL and a sigma-delta modulator (SDM), showing that the proportional path performs first-order SD modulation on the frequency offset. The analogy with a DM has previously only been mentioned. Greshishchev et. al. [1] point out the similarity of a second-order binary PLL to a double-integration DM with prediction. In a slightly different context, an analogy with a DM has been established for the delay/phase-locked loop with BPD proposed in [20]. Interestingly, Muller and Leblebici [21] note that the formula for the onset of slew-rate limiting in a first-order BBPLL, which is known from [2], corresponds to the slope overload condition in a DM, which we will explicitly show below.

Although a SDM and a DM are intimately related [29], the argument to interpret the BBPLL in favor of a DM is the tracking behavior that is fundamental to both systems. Clearly, tracking performed by a DM in the voltage domain corresponds to tracking performed by a BBPLL in the phase domain. This implies that the distortions due to quantization described in Fig. 2 also occur in a BBPLL. Here, granular noise is commonly called self-generated hunting jitter [2], [21] or dither jitter[2] [14], [30], referring to the VCO phase hunting or dithering around the reference clock phase. Slope overload refers to slewing or slew-rate limiting [20], [31], usually in the context of tracking a modulated reference clock. From the analogy with a DM, slewing in a first-order BBPLL is prevented if the derivative of $\theta_r(t)$ satisfies condition (5):

$$\max\left|\frac{d\theta_r(t)}{dt}\right| < \frac{\theta_{\rm bb}}{T_0} = 2\pi f_{\rm bb}. \tag{14}$$

As an example, consider a sinusoidally phase-modulated reference clock with excess phase $\theta_r(t) = A\sin(2\pi f_m t)$, where A is the modulation amplitude and f_m is the modulation frequency. Using (14), slewing is prevented if $A < f_{\rm bb}/f_m$, which is the formula given in [2], [8], [21]. Similar to the trade-off in selecting the step size δ in a DM, there is a trade-off in selecting the bang-bang phase step $\theta_{\rm bb}$. A small $\theta_{\rm bb}$ will reduce hunting jitter, but it will also restrict the maximum modulation frequency in order to avoid additional jitter from slope overload; a large $\theta_{\rm bb}$ will allow a higher modulation frequency, but it will incur increased hunting jitter due to the coarser phase updates. In Sec. IV-D we will see that this trade-off also exists in a statistical sense—in choosing the optimal phase step to obtain the minimum RMS timing jitter.

D. Stochastic Difference Equation

Having elaborated on the analogy between a BBPLL and a DM, we now consider the stochastic difference equation that describes the stochastic phase-jitter process in the presence of accumulative reference clock jitter, complementing previous work on the non-accumulative jitter case [15], [16]. As in these references we assume ideal PLL building blocks, but we also consider a frequency offset Δf between reference clock and VCO clock, which almost always occurs in practice [6]. Hence, the excess phase of the reference clock is $\theta_r(t) = \Delta\omega t + \omega_0\alpha(t)$, where $\Delta\omega = 2\pi\Delta f$. The phase-noise term $\alpha(t)$ is a (nonstationary) Wiener process with linearly increasing variance, which gives rise to the Gaussian accumulative jitter [10]. Sampling $\theta_r(t)$ at the nonuniform time instants yields

$$\theta_{r,n+1} - \theta_{r,n} = \Delta\omega(t_{n+1} - t_n) + \omega_0(\alpha(t_{n+1}) - \alpha(t_n)) \tag{15}$$

where $\theta_{r,n} = \theta_r(t_n)$ as before. Although $\alpha(t)$ has stationary increments [10], the nonuniform sampling instants cause the increment process $\alpha(t_{n+1}) - \alpha(t_n)$ to be nonstationary. But since the deviations from the uniform sampling instants are small, we can use the uniform sampling approximation $t_{n+1} = t_n + T_0$ [2] to approximate the increment process by $\alpha(t_{n+1}) - \alpha(t_n) \cong \alpha((n+1)T_0) - \alpha(nT_0) = \xi_n$, where the jitter random variable (RV) ξ_n is Gaussian distributed with mean zero and variance σ^2 (the standard deviation σ will be called RMS clock jitter) and the sequence $\{\xi_n\}$ is independent, identically distributed (IID). Replacing the increment process in (15) by ξ_n, substituting (9) for $t_{n+1} - t_n$ and plugging the result into (13) gives the stochastic difference equation describing the phase jitter:

$$\phi_{n+1} = \phi_n + \Delta\omega T_0 - \theta'_{\rm bb}\,\mathrm{sgn}\,\phi_n + \omega_0\xi_n. \tag{16}$$

[2]We will use the term hunting jitter to avoid confusion with the term dither meaning an intentionally applied noise source.

Here $\theta'_{\rm bb} = (1 + \Delta f/f_0)\theta_{\rm bb}$ is the modified bang-bang phase step; neglected in [2], [7], it accounts for the phase update due to the frequency offset Δf and was considered by Hsieh and Yang [32] (see also [19] for the zero-crossing DPLL). Nonetheless, since Δf is typically much smaller than f_0 [6] we assume $\theta'_{\rm bb} \cong \theta_{\rm bb}$ in the following.

To be consistent with the aforementioned literature, the remainder of this paper treats the equivalent timing-jitter model. Since the timing jitter at the nth sampling instant is $\Delta t_n = \phi_n/\omega_0$, we obtain from (16) the stochastic difference equation

$$\Delta t_{n+1} = \Delta t_n + \Delta T - K \operatorname{sgn} \Delta t_n + \xi_n \tag{17}$$

with $\Delta T = \Delta f/f_0^2$ being the period deviation. The period quantization of the VCO clock is $K = K_P K_T$ (loop gain), where the VCO period gain $K_T = K_{\rm F}/f_0^2$ [7], [9]. For convenience we use the terminology introduced for the phase-domain model and call ΔT the frequency offset and K the bang-bang phase step.

Given the stability condition $|\Delta T| < K$ [7], we assume without loss of generality that ΔT is positive and irrational; the latter assumption is justified because the natural tolerances in an implementation make the exact frequency of the reference clock and the free-running VCO clock unknown, and with probability 1 they will be irrational [6], [7]. For simplicity we take $\Delta T \in [0, K)$ and let $\Delta T = 0$ denote the case when both clock frequencies are practically identical, i.e., when ΔT is a small irrational number.

Given the stochastic difference equation (17), the statistical analysis of the steady-state timing jitter will be based on the SDRW theory developed in the following section.

III. Sign-Dependent Random Walk Theory

The SDRW studied in this section is defined as a RW whose transition probability (or step distribution) depends on the sign of the RW's current position. To investigate the asymptotic behavior of the SDRW, we derive its limiting characteristic function (CF), from which analytical expressions for the first four cumulants will be obtained.

A. Sign-Dependent Random Walk Model

First, we recall the definition of a RW [33]. Let $\xi, \xi_1, \xi_2, \ldots$ be a sequence of IID RVs with distribution function F. Let $S_0 = 0$, and consider the nth partial sum of the RVs, recursively defined as

$$S_{n+1} = S_n + \xi_{n+1} \tag{18}$$

for $n \geq 0$. The sequence of RVs $\{S_n\}$ is called a RW starting at the origin, and ξ is called the step RV (or step) with the step distribution F.

Now let $\xi_\pm, \xi_1^\pm, \xi_2^\pm, \ldots$ be two sequences of IID RVs with respective distribution functions $F^\pm$. Let $U_0 = 0$, and define in accordance with (18) a sequence recursively as

$$U_{n+1} = \begin{cases} U_n + \xi_{n+1}^+, & U_n \geq 0 \\ U_n + \xi_{n+1}^-, & U_n < 0 \end{cases} \tag{19}$$

for $n \geq 0$. The sequence of RVs $\{U_n\}$ is called a sign-dependent RW starting at the origin because the RV realizing the next step depends on the sign[3] of U_n. Accordingly, ξ_+ and ξ_- are called the positive and negative step RVs and correspond to the steps taken in the positive and negative half-line, respectively. Throughout this section we assume that the distribution functions $F^\pm$

(i) are continuous, and satisfy $0 < F^\pm(0) < 1$ so that the steps taken in either half-line may be both positive and negative;
(ii) possess moments of sufficiently high order so that the cumulants derived later exist;
(iii) have mean values satisfying $\mu_+ < 0$ and $\mu_- > 0$ (drift conditions) so that the SDRW exhibits a convergent limiting behavior.

Although no proof for the statement in (iii) is given, the assumed drift conditions may be intuitively justified as follows. A RW has the property (see Theorem 1 in Appendix A) that if the mean of the step ξ is positive ($\mu > 0$), the RW will drift to ∞ with probability 1. Similarly, if the mean of the step is negative ($\mu < 0$), the RW will drift to $-\infty$ with probability 1. This behavior implies for the SDRW that if $\mu_+ < 0$ and $\mu_- > 0$, the SDRW will drift to $-\infty$ when in the positive half-line, and to ∞ when in the negative half-line, so that it will eventually fluctuate around the origin, exhibiting a convergent limiting behavior.

We mention that two similar models emphasizing the sign dependency have recently been treated in the literature. Carlsund [22] considers a random walk on the integers with sign-dependent transition probabilities. In a similar work [23], Lefebvre treats a Wiener process with infinitesimal parameters (mean and variance) depending on the sign of the process, a model he calls an asymmetric or sign-dependent Wiener process. Both models are different from the SDRW defined above, and only the first-passage time problem is studied.

Finally, to place the BBPLL in the context of the SDRW, consider the timing-jitter model (17) and split the signum function according to whether Δt_n is nonnegative or negative:

$$\Delta t_{n+1} = \begin{cases} \Delta t_n + \Delta T - K + \xi_n, & \Delta t_n \geq 0 \\ \Delta t_n + \Delta T + K + \xi_n, & \Delta t_n < 0. \end{cases} \tag{20}$$

Since the jitter RV ξ_n has zero mean, the model (20) can be obtained from the SDRW model (19) by setting $\mu_+ = \Delta T - K$, $\mu_- = \Delta T + K$ and $\sigma_\pm^2 = \sigma^2$, where the latter is the variance of the Gaussian reference clock jitter. Given these parameters, we may therefore view the timing-jitter process as performing a sign-dependent Gaussian random walk (SDGRW)—that is, a SDRW with Gaussian step distributions. This relation leads us to investigate the limiting behavior of the SDRW in the following, and apply the obtained results to the steady-state behavior of the timing jitter in Sec. IV.

[3] According to (19), the definition of the sign includes the equality in the positive argument; no complications arise when $F^\pm$ are continuous, as in our assumption (i).

B. Limiting CF

The limiting behavior of the SDRW is obtained from the convergence of the sequence of RVs $\{U_n\}$ to some limiting RV U as $n \to \infty$. Given the various modes of convergence, we will consider convergence in distribution. Denoting the distribution of U_n as F_{U_n}, the sequence $\{U_n\}$ is said to converge in distribution to U if the sequence of distributions $\{F_{U_n}\}$ converges to some limiting distribution F_U [33]. As in [24], it is more convenient to investigate the convergence of the sequence of CFs $\{\phi_{U_n}\}$ to some limiting CF ϕ_U, where the CF ϕ_{U_n} of the SDRW at epoch n is defined as

$$\phi_{U_n}(z) \equiv Ee^{izU_n} = \int_{-\infty}^{\infty} e^{izu}\, dF_{U_n}(u) \tag{21}$$

with z being a real variable and E denoting statistical expectation. (For notational simplicity, the argument z of a CF will be omitted throughout.) We now derive the limiting CF ϕ_U by extending Fine's result for a DM [24]. In the following, RVs with subscript $''+''$ ($''-''$) will refer to a RW with step RV ξ_+ (ξ_-). (For example, M_+^+ is the maximum of a RW with step RV ξ_+). To obtain a recursion for ϕ_{U_n}, write (19) as

$$U_{n+1} = U_n + \xi_{n+1}^+ \mathbf{1}_{[0,\infty)}(U_n) + \xi_{n+1}^- \mathbf{1}_{(-\infty,0)}(U_n) \tag{22}$$

where $\mathbf{1}_A$ is the indicator function of the set A. Using the decomposition $U_n = \hat{U}_n + \check{U}_n$, where $\hat{U}_n = \max\{0, U_n\}$ is the positive part of U_n with the CF

$$\phi_{U_n}^+ = \int_0^{\infty} e^{izu}\, dF_{U_n}(u) \tag{23}$$

and $\check{U}_n = \min\{0, U_n\}$ is the negative part of U_n with the CF

$$\phi_{U_n}^- = \int_{-\infty}^{0} e^{izu}\, dF_{U_n}(u) \tag{24}$$

we have

$$U_{n+1} = (\hat{U}_n + \xi_{n+1}^+)\mathbf{1}_{[0,\infty)}(U_n) + (\check{U}_n + \xi_{n+1}^-)\mathbf{1}_{(-\infty,0)}(U_n). \tag{25}$$

Multiplying both sides by iz, exponentiating and taking expectation gives

$$\phi_{U_{n+1}} = \phi_{\xi_{n+1}^+}\phi_{U_n}^+ + \phi_{\xi_{n+1}^-}\phi_{U_n}^- \tag{26}$$

a recursion for ϕ_{U_n}. Because we are interested in the limiting behavior of the sequence $\{U_n\}$, we let $n \to \infty$ in (26) to obtain the limiting CF ϕ_U. As shown in Appendix A, the limiting CF is given by

$$\phi_U = \phi_X \phi_{M_+^+} \phi_{M_-^-} \tag{27}$$

where

$$\phi_X = \frac{\phi_{\xi_+} - \phi_{\xi_-}}{iz(\mu_+ - \mu_-)}. \tag{28}$$

As in [24], because ϕ_U is a product of CFs, the limiting RV U can be decomposed into the sum of independent RVs

$$U = X + M_+^+ + M_-^- \tag{29}$$

where M_+^+ is distributed as the maximum of a RW with step RV ξ_+, and M_-^- is distributed as the minimum of a RW with step RV ξ_- (see Appendix A).

An illustrative interpretation of X can be given if $\xi_\pm$ have identical distribution functions up to a shift of mean. Then, by writing $\phi_{\xi_\pm} = \exp(iz\mu_\pm)\phi_\xi$, where the step RV ξ has zero mean, the CF ϕ_X in (28) becomes

$$\phi_X = \frac{e^{iz\mu_+} - e^{iz\mu_-}}{iz(\mu_+ - \mu_-)}\phi_\xi. \tag{30}$$

Since the fraction is the CF of a uniform distribution [34, p.1880], we have the decomposition $X = \eta + \xi$, where η is uniformly distributed on $[\mu_+, \mu_-]$, so that finally

$$U = \eta + \xi + M_+^+ + M_-^-. \tag{31}$$

If, in addition, we assume $\mu_+ = -K$ and $\mu_- = K$ where $K > 0$, then η is uniformly distributed on $[-K, K]$ and we obtain Fine's (31) (after subtracting ξ) [24]. Besides providing insight into the limiting behavior of the SDRW, the decomposition into a sum of independent RVs simplifies the computation of the cumulants in the sequel.

C. Cumulants of the SDRW

Moments and cumulants describe the statistical properties of a RV or, equivalently, its distribution. Given a RV X with CF ϕ_X, its kth moment EX^k is defined as

$$EX^k = i^{-k}\frac{d^k}{dz^k}\phi_X(z)\Big|_{z=0} \tag{32}$$

and its kth cumulant $c_{k,X}$ as

$$c_{k,X} = i^{-k}\frac{d^k}{dz^k}\log\phi_X(z)\Big|_{z=0} \tag{33}$$

where $\log\phi_X$ is the cumulant generating function [34, p.370]. Cumulants enjoy the additivity property: the kth cumulant of a sum of independent RVs equals the sum of the kth cumulants of the single RVs [34, p.371]. Applying this property to the decomposition (29) gives the simple relation

$$c_{k,U} = c_{k,X} + c_{k,M_+^+} + c_{k,M_-^-} \tag{34}$$

for which we now derive general expressions for $k = 1, ..., 4$.

The cumulants c_{k,M_+^+} and c_{k,M_-^-} can be directly obtained from their cumulant generating functions. Indeed, from Spitzer's identity (71) (see Theorem 2 in Appendix A), the cumulant generating function of the maximum RV M_+^+ is

$$\log\phi_{M_+^+} = \sum_{n=1}^{\infty}\frac{1}{n}(\phi_{\hat{S}_n^+} - 1) \tag{35}$$

and hence, by the cumulant definition (33),

$$c_{k,M_+^+} = \sum_{n=1}^{\infty}\frac{1}{n}E((\hat{S}_n^+)^k) = \sum_{n=1}^{\infty}\frac{1}{n}\int_0^{\infty} x^k\, dF_n^+(x). \tag{36}$$

Here, the last equality is due to the fact that $E((\hat{S}_n^+)^k)$ is the kth moment of $\hat{S}_n^+$ ($\{S_n^+\}$ being a RW with step RV ξ_+) whose distribution is concentrated on $[0, \infty)$. Similarly, from (72) the cumulant generating function of the minimum RV M_-^- is

$$\log\phi_{M_-^-} = \sum_{n=1}^{\infty}\frac{1}{n}(\phi_{\check{S}_n^-} - 1) \tag{37}$$

from which we find the cumulants

$$c_{k,M_-^-} = \sum_{n=1}^{\infty} \frac{1}{n} E((\check{S}_n^-)^k) = \sum_{n=1}^{\infty} \frac{1}{n} \int_{-\infty}^{0} x^k \, dF_n^-(x). \quad (38)$$

Here, the last equality is due to the fact that $E((\check{S}_n^-)^k)$ is the kth moment of $\check{S}_n^-$ ($\{S_n^-\}$ being a RW with step RV ξ_-) whose distribution is concentrated on $(-\infty, 0]$.

The cumulants $c_{k,X}$ will be obtained from the corresponding moments EX^k. In particular, evaluating the kth derivative of ϕ_X in (28) at zero and using the moment definition (32), we obtain

$$EX^k = \frac{E\xi_+^{k+1} - E\xi_-^{k+1}}{k(\mu_+ - \mu_-)}. \quad (39)$$

Note that the first k moments of X exist if the first $k+1$ moments of $\xi_\pm$ exist (see also assumption (ii)). Inserting (39) into the moment-cumulant transformations [34, p.371], and the result into (34), gives the following expressions for the first four cumulants of the SDRW:

- First cumulant (mean):

$$\mu_U = \frac{E\xi_+^2 - E\xi_-^2}{2(\mu_+ - \mu_-)} + c_{1,M_+^+} + c_{1,M_-^-}. \quad (40)$$

- Second cumulant (variance):

$$\sigma_U^2 = \frac{E\xi_+^3 - E\xi_-^3}{3(\mu_+ - \mu_-)} - \frac{1}{4}\left(\frac{E\xi_+^2 - E\xi_-^2}{\mu_+ - \mu_-}\right)^2 + c_{2,M_+^+} + c_{2,M_-^-}. \quad (41)$$

- Third cumulant (third central moment):

$$c_{3,U} = \frac{E\xi_+^4 - E\xi_-^4}{4(\mu_+ - \mu_-)} - \frac{(E\xi_+^2 - E\xi_-^2)(E\xi_+^3 - E\xi_-^3)}{2(\mu_+ - \mu_-)^2} + \frac{1}{4}\left(\frac{E\xi_+^2 - E\xi_-^2}{\mu_+ - \mu_-}\right)^3 + c_{3,M_+^+} + c_{3,M_-^-}. \quad (42)$$

- Fourth cumulant:

$$c_{4,U} = \frac{E\xi_+^5 - E\xi_-^5}{5(\mu_+ - \mu_-)} - \frac{3(E\xi_+^2 - E\xi_-^2)(E\xi_+^4 - E\xi_-^4) + 2(E\xi_+^3 - E\xi_-^3)^2}{6(\mu_+ - \mu_-)^2} + \frac{(E\xi_+^2 - E\xi_-^2)^2(E\xi_+^3 - E\xi_-^3)}{(\mu_+ - \mu_-)^3} - \frac{3}{8}\left(\frac{E\xi_+^2 - E\xi_-^2}{\mu_+ - \mu_-}\right)^4 + c_{4,M_+^+} + c_{4,M_-^-}. \quad (43)$$

Equations (40)–(43) express the first four cumulants in terms of the moments of the step RVs $\xi_\pm$ and the cumulants of the maximum and minimum RVs M_+^+ and M_-^-, which are given by (36) and (38).

D. Cumulants of the SDGRW

In general, given arbitrary step distributions $F^\pm$, it appears difficult to further simplify (40)–(43) because the n-fold convolution $F_n^\pm(x) = P(\xi_1^\pm + \ldots + \xi_n^\pm < x)$ in (36) and (38) may not be expressible in simple terms. For Gaussian step distributions, however, the convolution is easily performed, leading to the explicit cumulant expressions derived in the following. The Gaussianity of the step distributions also suggests the name SDGRW introduced above, in analogy with the widely used Gaussian random walk [33].

To derive the cumulants of the SDGRW, let $\xi_1, \xi_2, \ldots$ be a sequence of IID RVs with Gaussian distribution function F. Given the mean μ and the variance σ^2, the n-fold convolution $F_n(x) = P(\xi_1 + \ldots + \xi_n < x)$ is Gaussian with mean $n\mu$ and variance $n\sigma^2$, and so the integrals in (36) and (38) can be evaluated explicitly. Since F has the density

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)} \quad (44)$$

we have from (36) that

$$c_{k,M_+^+} = \sum_{n=1}^{\infty} \frac{1}{n} \int_0^{\infty} x^k f_n(x; \mu_+, \sigma_+^2) \, dx \quad (45)$$

and from (38), after the change of variable $x \to -x$, that

$$c_{k,M_-^-} = (-1)^k \sum_{n=1}^{\infty} \frac{1}{n} \int_0^{\infty} x^k f_n(x; -\mu_-, \sigma_-^2) \, dx \quad (46)$$

where $f_n(x; \mu, \sigma^2) \equiv f(x; n\mu, n\sigma^2)$. Formal integration shows that for $k = 1, ..., 4$

$$\sum_{n=1}^{\infty} \frac{1}{n} \int_0^{\infty} x^k f_n(x; \mu, \sigma^2) \, dx = \sigma^k G_k\left(-\frac{\mu}{\sigma}\right) \quad (47)$$

where

$$G_k(x) \equiv \sum_{n=1}^{\infty} g_k(n, x) \quad (48)$$

and

$$g_1(n, x) = \frac{1}{\sqrt{2\pi n}} e^{-nx^2/2} - \frac{1}{2} x \operatorname{erfc}\left(\sqrt{\frac{n}{2}} x\right) \quad (49)$$

$$g_2(n, x) = \frac{1}{2}(nx^2 + 1) \operatorname{erfc}\left(\sqrt{\frac{n}{2}} x\right) - \sqrt{\frac{n}{2\pi}} x e^{-nx^2/2} \quad (50)$$

$$g_3(n, x) = \sqrt{\frac{n}{2\pi}}(nx^2 + 2) e^{-nx^2/2} - \frac{1}{2} nx(nx^2 + 3) \operatorname{erfc}\left(\sqrt{\frac{n}{2}} x\right) \quad (51)$$

$$g_4(n, x) = \frac{1}{2} n(n^2 x^4 + 6nx^2 + 3) \operatorname{erfc}\left(\sqrt{\frac{n}{2}} x\right) - \sqrt{\frac{n}{2\pi}} nx(nx^2 + 5) e^{-nx^2/2} \quad (52)$$

with erfc being the complementary error function. The Gaussian distribution has the property that higher-order moments can be expressed in terms of the mean μ and the variance σ^2 since the latter two uniquely describe the distribution. In particular, the first five moments of the Gaussian distributed RV ξ are given by $E\xi^2 = \mu^2 + \sigma^2$, $E\xi^3 = \mu^3 + 3\mu\sigma^2$, $E\xi^4 = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4$ and $E\xi^5 = \mu^5 + 10\mu^3\sigma^2 + 15\mu\sigma^4$ [34, p.713]. Substitution of these moments together with (45)–(47) into (40)–(43) yields the following expressions for the first four cumulants of the SDGRW:

- First cumulant (mean):

$$\mu_U = \frac{\mu_+ + \mu_-}{2} + \frac{1}{2}\frac{\sigma_+^2 - \sigma_-^2}{\mu_+ - \mu_-} + \sigma_+ G_1\left(-\frac{\mu_+}{\sigma_+}\right) - \sigma_- G_1\left(\frac{\mu_-}{\sigma_-}\right). \quad (53)$$

- Second cumulant (variance):

$$\sigma_U^2 = \frac{(\mu_+ - \mu_-)^2}{12} + \frac{\sigma_+^2 + \sigma_-^2}{2} - \frac{1}{4}\left(\frac{\sigma_+^2 - \sigma_-^2}{\mu_+ - \mu_-}\right)^2 + \sigma_+^2 G_2\left(-\frac{\mu_+}{\sigma_+}\right) + \sigma_-^2 G_2\left(\frac{\mu_-}{\sigma_-}\right). \quad (54)$$

- Third cumulant (third central moment):

$$c_{3,U} = \frac{(\mu_+ - \mu_-)(\sigma_+^2 - \sigma_-^2)}{4} + \frac{1}{4}\left(\frac{\sigma_+^2 - \sigma_-^2}{\mu_+ - \mu_-}\right)^3 + \sigma_+^3 G_3\left(-\frac{\mu_+}{\sigma_+}\right) - \sigma_-^3 G_3\left(\frac{\mu_-}{\sigma_-}\right). \quad (55)$$

- Fourth cumulant:

$$c_{4,U} = -\frac{(\mu_+ - \mu_-)^4}{120} + \frac{(\sigma_+^2 - \sigma_-^2)^2}{4} - \frac{3}{8}\left(\frac{\sigma_+^2 - \sigma_-^2}{\mu_+ - \mu_-}\right)^4 + \sigma_+^4 G_4\left(-\frac{\mu_+}{\sigma_+}\right) + \sigma_-^4 G_4\left(\frac{\mu_-}{\sigma_-}\right). \quad (56)$$

These expressions generalize those in [24], [29], [35].

IV. Statistical Analysis of First-Order BBPLL Using Sign-Dependent Random Walk Theory

The SDRW theory developed in the previous section will now be applied to statistically analyze the steady-state timing jitter in a first-order BBPLL. As mentioned above, the timing-jitter model (20) can be recovered from the SDRW model (19) by assuming the step RVs $\xi_\pm$ to be Gaussian distributed with means $\mu_+ = \Delta T - K$ and $\mu_- = \Delta T + K$ and equal variances $\sigma_\pm^2 = \sigma^2$. Assuming ideal PLL blocks, the timing-jitter process may therefore be viewed as a SDGRW, and the requirements $\mu_+ < 0$ and $\mu_- > 0$ for the SDGRW to exhibit a convergent limiting behavior agree with the condition $|\Delta T| < K$ for the loop to be stable. Hence, we can immediately apply the cumulant expressions (53)–(56) in the analysis. In particular, the first cumulant (mean) and the second cumulant (variance) correspond to the static timing offset and the RMS timing jitter, respectively, and quantify the timing-jitter performance; the fourth cumulant is used in the definition of the kurtosis and quantifies the non-Gaussianity of the timing-jitter PDF.

The analytical cumulant expressions will be compared against Monte-Carlo simulation of (17), verifying also our numerical results in [17]. For the each simulation, 10^7 realizations of length 100 were generated and the statistics (histogram and cumulants) were computed for the last time instant. The mean, variance and kurtosis were computed using the corresponding MATLAB built-in functions. To illustrate the following discussion, Fig. 6 plots several timing-jitter PDFs obtained from the numerical method described in [17].

A. Decomposition of the Steady-State Timing Jitter Δt

We begin the analysis by recalling that (31) decomposes the limiting behavior of the SDRW into a sum of statistically independent contributions. Applying this decomposition to the steady-state timing jitter provides valuable insight into its nonlinear statistical behavior. More specifically, the steady-state timing jitter Δt can be decomposed as

$$\Delta t = \eta + \xi + o \quad (57)$$

where the RV η is uniformly distributed on $[\Delta T - K, \Delta T + K]$ and the jitter RV ξ is Gaussian distributed with mean zero and variance σ^2. The RV o is distributed (in the sense of equality in distribution) as the sum of the maximum and minimum of a RW as discussed above. Similar to the interpretation for a DM given in [24] and [29], the decomposition (57) suggests interpreting the timing jitter in terms of the Gaussian reference clock jitter and additionally introduced static and dynamic jitter components, as further discussed in Sec. II-C. In particular, due to its uniform distribution we refer to the RV η as self-generated hunting jitter (static component), a term which was used by Walker [2] but without giving a precise definition. Hunting jitter is introduced by the coarseness of the binary PD characteristic and is characterized by the VCO phase hunting randomly around the jittered reference clock phase. In line with the term overload noise used in DM theory [25], we refer to the RV o as overload jitter (dynamic component). Overload jitter is introduced by the inability of the VCO phase to track the reference clock phase faster than in steps of K, and is characterized by a sequence of equal VCO phase updates.

Note also that the hunting jitter η does not depend on ξ, whereas the overload jitter o does depend on ξ (via the RW maximum and minimum, both of which depend on ξ). This implies that the timing jitter Δt will be dominated by hunting jitter for small σ and by overload jitter for large σ. Moreover, η being independent of ξ also means that

$$\Delta t = \eta \qquad \text{for } \sigma = 0. \quad (58)$$

In other words, for a jitter-free reference clock, the timing jitter is uniformly distributed on $[\Delta T - K, \Delta T + K]$. In the following we will repeatedly refer to the decomposition (57) to attribute observed statistical properties to the underlying RVs.

B. Static Timing Offset Δt_{stat}

A PLL synchronizes, in both phase and frequency, the feedback clock with the reference clock by bringing their clock edges in close alignment. The average time (phase) difference between the two clocks in the locked state is the static timing offset (static phase error). A static timing offset is a significant problem in a BBPLL-based CDR circuit because it results in the incoming data no longer being sampled at the center of the data eye, thus increasing the bit error rate [3].

In our first-order loop, a static timing offset is caused by a frequency difference $\Delta T > 0$ between the reference clock and the VCO clock; the LF needs to have a nonzero output to tune the VCO to its correct average frequency. When an analog LF is used, a static timing offset may be caused by a

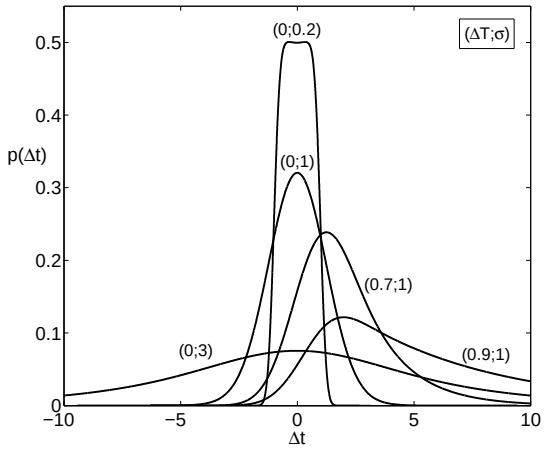

Fig. 6. Steady-state timing-jitter PDF for various ΔT and σ ($K = 1$) [17].

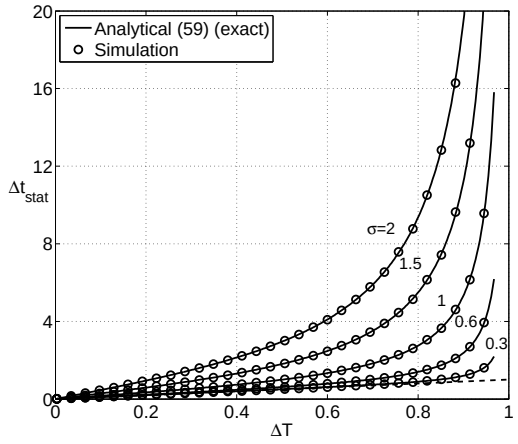

Fig. 7. Static timing offset $\Delta t_{\rm stat}$ versus frequency offset ΔT for $K = 1$.

mismatch between the two charge-pump currents, which can be considered as an additional frequency offset. Moreover, the presence of reference clock jitter causes any static timing offset to increase from its zero-jitter level [6]. We define the static timing offset $\Delta t_{\rm stat}$ as the statistical expectation of the timing jitter in steady state, for which we obtain from (53) the expression

$$\Delta t_{\rm stat} = \Delta T + \sigma G_1\left(\frac{K-\Delta T}{\sigma}\right) - \sigma G_1\left(\frac{K+\Delta T}{\sigma}\right) \quad (59)$$

where G_1 is given by (48) and (49). Figure 7 plots $\Delta t_{\rm stat}$ of (59) as a function of ΔT with parameter σ. The agreement between theory and simulation verifies the obtained analytical expression.

For $\Delta T = 0$, the static timing offset is zero independent of σ; the VCO clock samples the reference clock equally likely before and after the reference clock edges. The corresponding timing-jitter PDFs $(0;\sigma)$ in Fig. 6 are symmetric about zero and thus have zero mean.

For $\Delta T > 0$ and $\sigma = 0$, the static timing offset equals the frequency offset, as shown by the dashed line $\Delta t_{\rm stat} = \Delta T$; a larger frequency offset requires a larger average BPD output to tune the VCO to its correct average frequency. The result for $\Delta t_{\rm stat}$ also follows from (58) since the mean of a RV uniformly distributed on $[\Delta T - K, \Delta T + K]$ is ΔT.

For $\Delta T > 0$ and $\sigma > 0$, the combined effect of frequency offset and clock jitter causes the static timing offset to increase from its jitter-free level, as seen by the curves exceeding the dashed line. For small-enough σ, though, the static timing offset does not increase notably, even for a moderate frequency offset; this may be attributed to the dominant hunting jitter in this case. Note that since $K = 1$ in the figure, the loop is stable for $|\Delta T| < 1$, which explains the increase of the curves as ΔT tends to 1.

C. RMS timing jitter $\sigma_{\Delta t}$

For the loop design, a quantity of main interest is the timing-jitter variance $\sigma^2_{\Delta t}$; in practice, it is common to consider the standard deviation $\sigma_{\Delta t}$ and speak of the RMS timing jitter. Design questions to be addressed are how much jitter is transferred from the reference clock, how much jitter is generated by the loop itself and what is the minimum RMS timing jitter. An answer to these questions is given by (54), from which we obtain the expression

$$\sigma^2_{\Delta t} = \frac{K^2}{3} + \sigma^2 + \sigma^2 G_2\left(\frac{K-\Delta T}{\sigma}\right) + \sigma^2 G_2\left(\frac{K+\Delta T}{\sigma}\right) \quad (60)$$

where G_2 is given by (48) and (50). Although this formula is exact, its dependency on the parameters K, ΔT and σ via the infinite series G_2 complicates an intuitive explanation. An approximate formula providing a simple rule of thumb for the loop design can be derived by upper bounding the infinite series. As shown in Appendix B, the result is

$$\sigma^2_{\Delta t} \approx \frac{K^2}{3} + \sigma^2 + \frac{\sigma^4}{4}\left(\frac{1}{(K-\Delta T)^2} + \frac{1}{(K+\Delta T)^2}\right) \quad (61)$$

which yields

$$\sigma^2_{\Delta t} \approx \frac{K^2}{3} + \sigma^2 + \frac{\sigma^4}{2K^2} \qquad \text{for } \Delta T = 0. \quad (62)$$

A comparison of the exact expression (60) with the approximation (61) and the simulation results is shown in Fig. 8. The agreement between theory and simulation verifies the obtained analytical expressions, and suggests the use of (61) for the following explanations.

The timing-jitter behavior in different regions of the plot can be interpreted using the decomposition (57). For small σ, the hunting jitter dominates so that $\sigma_{\Delta t}$ is approximately constant. Indeed, as σ tends to zero, $\sigma_{\Delta t}$ approaches the small-σ asymptote $K/\sqrt{3}$ which corresponds to the standard deviation of the uniformly distributed RV in (58). Increasing σ causes $\sigma_{\Delta t}$ to rise, since the effect of the Gaussian clock jitter and the resulting overload jitter becomes gradually apparent. For large σ, overload jitter dominates, and $\sigma_{\Delta t}$ shows a linear increase with σ on the logarithmic scale. Because the last term in (61) is dominant in this case, we obtain the large-σ asymptote

$$\sigma^2_{\Delta t} \to \frac{\sigma^4}{4}\left(\frac{1}{(K-\Delta T)^2} + \frac{1}{(K+\Delta T)^2}\right) \quad (63)$$

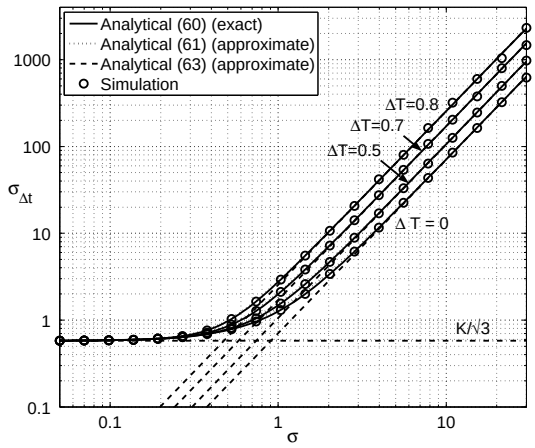

Fig. 8. RMS timing jitter $\sigma_{\Delta t}$ versus RMS clock jitter σ for $K = 1$.

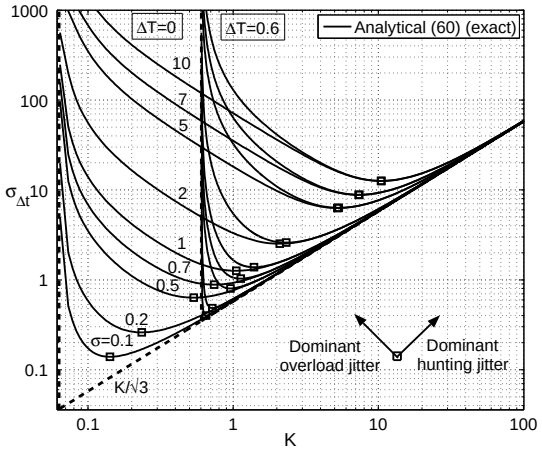

Fig. 9. Trade-off between hunting jitter and overload jitter leads to an optimal bang-bang phase step for minimum RMS timing jitter (symbol markers).

which is shown by the dashed lines. Interestingly, $\sigma_{\Delta t}$ increases proportional to the clock jitter variance σ^2, with the proportionality factor depending on K and ΔT. In particular, zero frequency offset gives the asymptote $\sigma^2/(\sqrt{2}K)$ reported in [18]. Contrasting this asymptote with the asymptote σ for the non-accumulative jitter case [16] shows that the nonlinear loop reacts differently to different clock jitter. In either jitter case, however, the RMS timing jitter $\sigma_{\Delta t}$ is independent of the loop delay in this large-σ regime (compare Fig. 7 in [16] with Fig. 6 in [18]). In other words, even in the presence of a loop delay, the SDRW model does capture the stochastic timing jitter dynamics in this regime, and the timing jitter performance can be predicted using (63).

D. Optimal BB Phase Step for Minimum RMS Timing Jitter

In Fig. 8 the RMS timing jitter $\sigma_{\Delta t}$ was plotted as function of σ for a normalized bang-bang phase step ($K = 1$). In practice, the RMS clock jitter σ is given, and the designer has to choose K such that the loop fulfills the timing-jitter specifications, possibly achieving minimal jitter. Formula (61) demonstrates the limits to the minimum RMS timing jitter: for any fixed σ and ΔT, there is a fundamental trade-off between hunting jitter and overload jitter. Clearly, a small K will yield small hunting jitter but large overload jitter, while a large K will yield small overload jitter but large hunting jitter.

This trade-off is illustrated in Fig. 9, which plots $\sigma_{\Delta t}$ of (60) as a function of K for zero and nonzero frequency offset. (Recall from Sec. II-D that a zero frequency offset means a small irrational number close to zero). It can be seen that for any fixed σ and ΔT, there exists an optimal bang-bang phase step $K_{\rm opt}$ that gives the minimum RMS timing jitter $\sigma_{\Delta t,\min}$ (symbol markers); the minima were obtained by numerically finding the minimum of (60). Predominance of hunting jitter and overload jitter can be ascribed to different regions of the plot. For $K > K_{\rm opt}$, the timing jitter is dominated by hunting jitter, the fundamental lower bound of which is $K/\sqrt{3}$ (dashed asymptote). For $K < K_{\rm opt}$, the timing jitter is dominated by overload jitter and increases rapidly as K approaches the stability boundary at ΔT (vertical dashed lines), obtained from the stability condition $|\Delta T| < K$. The figure also shows that for small σ, the minima are more sensitive to changes in K than for large σ; this may be crucial in a DBBPLL where changes in K due to quantization of the DLF gain K_P can degrade the timing-jitter performance. Notice finally that the minima become independent of ΔT with increasing σ, in which case $\sigma_{\Delta t,\min}$ is limited only by the clock jitter and not by a frequency offset.

To further illustrate the ultimate limits to the timing-jitter performance, we can derive an approximate closed-form formula for $K_{\rm opt}$ which allows us to analytically predict $\sigma_{\Delta t,\min}$. In particular, setting the derivative of (61) with respect to K to zero yields

$$K_{\rm opt}^2 = \Delta T^2 + \left(3\sigma^4\lambda\Delta T^2\right)^{1/3} + \left(\sigma^8/(24\lambda\Delta T^2)\right)^{1/3} \quad (64)$$

where $\lambda = 1+\sqrt{1-\sigma^4/(72\Delta T^4)}$. Substituting (64) into (61) gives an analytical prediction for $\sigma_{\Delta t,\min}$, which is not explicitly expressed here but which is plotted in Fig. 10 as a function of σ. The numerical (exact) results were again obtained by numerically solving for the minimum of (60), and show good agreement with the theory. The figure illustrates that as σ becomes smaller, $\sigma_{\Delta t,\min}$ attains the value $\Delta T/\sqrt{3}$ of the hunting jitter (dashed asymptotes). This value can be derived by setting $\sigma = 0$ in (64) and plugging the resulting optimal step size $K_{\rm opt} = \Delta T$ into the hunting jitter formula $K/\sqrt{3}$. This result also shows that, as a function of σ, the ultimate lower limit on the minimum RMS timing jitter can be obtained from the case $\Delta T = 0$. Setting the derivative of (62) with respect to K to zero gives $K_{\rm opt} \approx 1.107\sigma$, which means that for zero frequency offset, the optimal bang-bang phase step is approximately equal to the RMS clock jitter. Substituting this result into (62) gives the lower limit $\sigma_{\Delta t,\min} \approx 1.348\sigma$, which is just the line $\Delta T = 0$ in the figure.

E. Kurtosis of Δt

To provide further insight into the loop's nonlinear statistical behavior, we now compute the kurtosis of the timing jitter Δt. The kurtosis measures the degree by which a PDF deviates from Gaussianity, and is positive for a pointy PDF and negative for a flat-topped PDF. Since the clock jitter is assumed to be

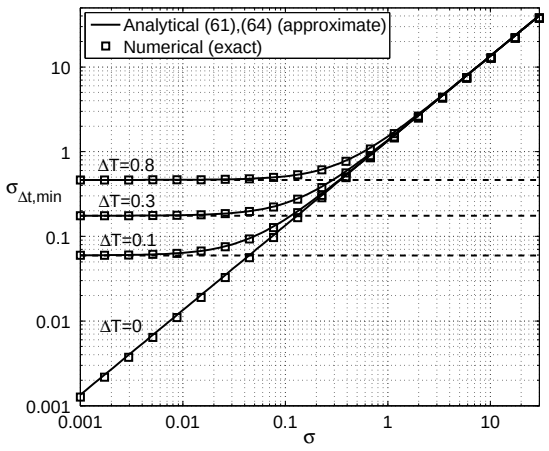

Fig. 10. Minimum RMS timing jitter $\sigma_{\Delta t,\min}$ versus RMS clock jitter σ.

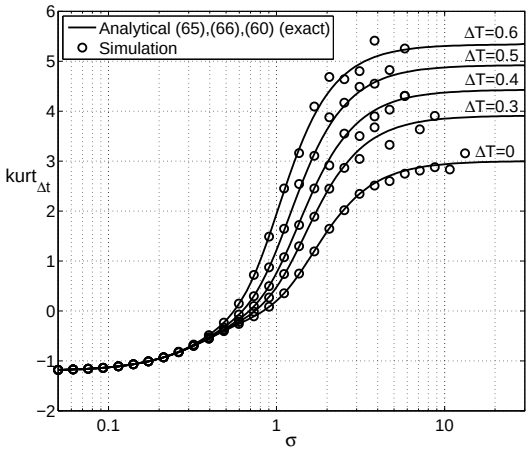

Fig. 11. Kurtosis of the timing jitter, $\mathrm{kurt}_{\Delta t}$, versus RMS clock jitter σ for $K = 1$.

Gaussian distributed, a nonzero kurtosis of Δt means a non-Gaussian timing-jitter PDF, and thus reveals the influence of the BPD nonlinearity.

For the timing jitter, the kurtosis is defined as [34, p.1009]

$$\mathrm{kurt}_{\Delta t} = \frac{c_{4,\Delta t}}{\sigma_{\Delta t}^4} \tag{65}$$

where the fourth cumulant $c_{4,\Delta t}$ is obtained from (56) as

$$c_{4,\Delta t} = -\frac{2K^4}{15} + \sigma^4 G_4\left(\frac{K-\Delta T}{\sigma}\right) + \sigma^4 G_4\left(\frac{K+\Delta T}{\sigma}\right) \tag{66}$$

with G_4 given by (48) and (52). The kurtosis $\mathrm{kurt}_{\Delta t}$ as a function of σ is plotted in Fig. 11, showing again good agreement with the simulation results.

For small σ, the kurtosis is negative and independent of ΔT; this is due to the uniform-like (flat-topped) PDF caused by the dominant hunting jitter. Indeed, as σ tends to zero, the kurtosis tends to the value -1.2 for a uniform distribution [34, p.1881]. The slight dependence of the kurtosis on σ is due to the Gaussian jitter smoothing the edges of the uniform-like PDF, as shown by the PDF $(0; 0.2)$ in Fig. 6.

For large σ, intuition suggests that the large reference clock jitter will linearize the binary PD characteristic so that the loop behaves linearly. The positive kurtosis in this region proves the opposite: the loop dynamics is still nonlinear, making the timing-jitter PDF non-Gaussian (see the PDF $(0; 3)$ in Fig. 6). Contrast this with the non-accumulative jitter case in [15], [16], for which it can be shown that the timing-jitter PDF is indeed Gaussian in this regime, leading to a linear[4] loop behavior.

Since the timing-jitter PDF is symmetric for $\Delta T = 0$, it also follows from the figure that for one particular value σ^*, the timing-jitter PDF will in fact be Gaussian, namely for the value of σ giving zero kurtosis (the zero crossing of the corresponding curve). In this case, the loop behaves effectively linearly, the Gaussian clock jitter with RMS σ^* being transformed into the Gaussian timing jitter with RMS $\sigma^*_{\Delta t}$. For $\Delta T = 0$, finding the zero of (66) numerically and plugging the result into (60) gives $\sigma^* \approx 0.83K$ and the RMS timing jitter $\sigma^*_{\Delta t} \approx 1.02K$. As a rule, the timing-jitter PDF is well approximated by a Gaussian PDF for σ close to σ^*, but it differs significantly for all other values of σ.

[4]Recall that a linear system transforms a Gaussian PDF into another Gaussian PDF.

V. Conclusions

Viewing a first-order BBPLL as a single-integration DM in the phase domain allows us to explain the bang-bang loop behavior using existing delta-modulation terminology and theory, linking hunting jitter and slew-rate limiting in a BBPLL to granular noise and slope overload in a DM.

When the reference clock is contaminated by accumulative jitter, the stochastic evolution of the timing jitter can be modeled as a SDRW. Analytical expressions for the first four cumulants of the SDRW have been obtained and applied to the BBPLL to statistically quantify the steady-state timing jitter. The main results can be summarized as follows:

- The steady-state timing jitter can be decomposed into three statistically independent components: accumulative jitter of the reference clock, hunting jitter and overload jitter. The latter two are static and dynamic components that are due to the binary phase-error quantization and that also occur in a DM [24].
- The static timing offset described by (59) equals the frequency offset for a jitter-free reference clock. Although the presence of clock jitter causes the timing offset to exceed its jitter-free value, the increased timing offset stays within limits, even for a moderate frequency offset, when the clock jitter is not too large; this behavior is due to the self-generated hunting jitter.
- Formula (61) explains intuitively how the RMS timing jitter depends on the bang-bang phase step, the RMS clock jitter and the frequency offset. For a fixed bang-bang phase step, the RMS timing jitter is constant for small RMS clock jitter and grows quadratically with large RMS clock jitter. For fixed RMS clock jitter, the hunting jitter is proportional to the bang-bang phase step, while the overload jitter is inversely proportional to it—a behavior also displayed by a DM [24].

- The opposing dependence of hunting jitter and overload jitter on the bang-bang phase step entails a trade-off in choosing the optimal phase step for minimum RMS timing jitter. For zero frequency offset, the optimal phase step is approximately equal to the RMS clock jitter.
- Computing the kurtosis has revealed the effect of the BPD nonlinearity: the timing-jitter PDF is Gaussian-like for σ close to some value σ^* (for which it is in fact Gaussian), but it is distinctively non-Gaussian for all other σ values.

The advantage of having treated the SDRW in the general form (19) is that non-ideal factors can be considered in the analysis. Mismatches between the two charge-pump currents or between the two output levels of the BPD can be accounted for by assuming appropriate mean values for the step RVs. Different reference clock jitter distributions can be taken into account by choosing appropriate step distributions. Flicker noise, which corresponds to accumulative jitter with correlation [9], can be modeled by assuming each sequence of step RVs to be correlated, but separate cumulant expressions need to be derived in this case. Non-zero loop delay cannot be fully modeled by the SDRW, but as was pointed in Sec. IV-C, the model does give a good prediction of the RMS timing jitter when the RMS reference clock jitter is large. To conclude, although the obtained analytical expressions enable an exact statistical loop analysis, the SDRW model is limited in that it only applies to a first-order loop subject to accumulative reference clock jitter.

The analysis of the present paper and of [15], [16], [18] may be extended into several directions. First, since a practical oscillator exhibits both non-accumulative and accumulative jitter [9], a more accurate loop analysis needs to consider the combined effect of both types of jitter, as well as the jitter from the VCO. Second, since BBPLLs are typically implemented as second-order loops [4], [5], the dynamics due to the LF integral path must be taken into account, particularly when the loop stability factor [2] is small. Further to this point, elaborating on the analogy between second-order BBPLLs and double-integration delta modulation with prediction [1] may turn out to be fruitful, in that existing delta-modulation theory [25] could also be applied to second-order loops. Despite the current restriction to first-order BBPLLs, the application of Markov theory has demonstrated a more accurate timing-jitter description by revealing statistical features that remain hidden from a linear analysis.

APPENDIX A
DERIVATION OF THE CHARACTERISTIC FUNCTION (27)

The CF (27) derived in the following generalizes the CF given by Fine in [24]. Because the derivation largely follows his approach, we present only those parts that lead to the generalization.

We begin by recalling some facts about the RW defined in (18) [33], [36]. To introduce the RVs used in the derivation, consider the sequence of points (n, S_n) for $n \geq 1$, and define the first ascending ladder point (T^+, H^+) as the first term in this sequence for which $S_n \geq 0$. Then, the first ascending ladder epoch $T^+ = \min\{n > 0 : S_n \geq 0\}$ marks the epoch at which the RW first enters the non-negative half-line, and the first ascending ladder height $H^+ = S_{T^+}$ marks the point of first entry. Similarly, define the first descending ladder epoch $T^- = \min\{n > 0 : S_n < 0\}$ and the first descending ladder height $H^- = S_{T^-}$, with the analogous interpretation of entry into the negative half-line. Because either half-line does not need to be entered at all, the above defined RVs are possibly defective, with the defects defined as

$$p^+ \equiv 1 - P(T^+ < \infty) = 1 - P(H^+ < \infty) \tag{67}$$

$$p^- \equiv 1 - P(T^- < \infty) = 1 - P(H^- < \infty). \tag{68}$$

Furthermore, let $M^+ = \max_{0 \leq n < \infty} S_n$ be the (possibly infinite) maximum of the RW, and $M^- = \min_{0 \leq n < \infty} S_n$ the (possibly infinite) minimum. The following two types of RWs are of interest to us.

Theorem 1. (Classification of RWs [33, p.379], [36, p.224])*: For a RW with step ξ and step distribution F, if $0 < E\xi < \infty$, then*

(i) *with probability 1, S_n drifts to ∞ and reaches a finite minimum M^-;*
(ii) *T^+ and H^+ are proper, T^- and H^- defective;*
(iii) *ET^+ and EH^+ are finite, and by Wald's identity*

$$ET^+ = \frac{1}{p^-} = \frac{EH^+}{E\xi}. \tag{69}$$

Similarly, if $-\infty < E\xi < 0$, then

(i) *with probability 1, S_n drifts to $-\infty$ and reaches a finite maximum M^+;*
(ii) *T^- and H^- are proper, T^+ and H^+ defective;*
(iii) *ET^- and EH^- are finite, and by Wald's identity*

$$ET^- = \frac{1}{p^+} = \frac{EH^-}{E\xi}. \tag{70}$$

Finally, two identities for the CF of the maximum M^+ and the minimum M^- will be required.

Theorem 2. (Spitzer's identity [36, p.230])*: Provided the maximum $M^+ < \infty$, we have*

$$\phi_{M^+} = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n}(\phi_{\hat{S}_n} - 1)\right\} \tag{71}$$

where $\hat{S}_n = \max\{0, S_n\}$. Similarly, provided the minimum $M^- > -\infty$, we have

$$\phi_{M^-} = \exp\left\{\sum_{n=1}^{\infty} \frac{1}{n}(\phi_{\check{S}_n} - 1)\right\} \tag{72}$$

where $\check{S}_n = \min\{0, S_n\}$.

To begin the derivation, consider the recursion for the CF in (26). Using $\phi_{U_{n+1}} = \phi^+_{U_{n+1}} + \phi^-_{U_{n+1}}$ and assuming that ϕ_U exists and that $\phi^+_{U_n} \to \phi^+_U$ and $\phi^-_{U_n} \to \phi^-_U$ as $n \to \infty$, we obtain

$$(1 - \phi_{\xi_+})\phi^+_U + (1 - \phi_{\xi_-})\phi^-_U = 0. \tag{73}$$

This equation reduces to Fine's (20) if $\phi_{\xi_+} = \exp(-izK)\phi_\xi$ and $\phi_{\xi_-} = \exp(izK)\phi_\xi$, corresponding to the model (20) with $\Delta T = 0$. Since (73) is a Wiener-Hopf equation, the general

procedure for its solution is the product factorization of the factors $1-\phi_{\xi_\pm}$. As in [24], by applying the Wiener-Hopf decomposition [33, p.571], these factors can be written as

$$1-\phi_{\xi_\pm} = \left(1-\phi_{H_\pm^+}\right)\left(1-\phi_{H_\pm^-}\right) \tag{74}$$

where $\phi_{H_\pm^+}$ and $\phi_{H_\pm^-}$ are the CFs of the first ascending and descending ladder heights of a RW with steps $\xi_\pm$, respectively. Defining the complex functions

$$X^+ = \frac{1-\phi_{H_-^+}}{z(1-\phi_{H_+^+})} \quad \text{and} \quad X^- = \frac{1-\phi_{H_+^-}}{z(1-\phi_{H_-^-})} \tag{75}$$

where $X^\pm \equiv X^\pm(z)$, the CF ϕ_U can be expressed as [24]

$$\phi_U = \frac{X^+ - X^-}{X^+(0) - X^-(0)} \tag{76}$$

where $X^+(0) = -iEH_-^+/p_+^+$ and $X^-(0) = -iEH_+^-/p_-^-$, with the defects p_+^+ and p_-^- defined in (67) and (68). For the denominator in (76), the latter two equations yield

$$X^+(0) - X^-(0) = \frac{i(p_+^+EH_+^- - p_-^-EH_-^+)}{p_+^+p_-^-}. \tag{77}$$

Consider now a RW with step RV ξ_+, for which $\mu_+ < 0$. By Theorem 1, the RW drifts to $-\infty$ and $p_+^+EH_+^- = E\xi_+ = \mu_+$ from (69). Similarly, a RW with step RV ξ_-, for which $\mu_- > 0$, drifts to ∞ and $p_-^-EH_-^+ = E\xi_- = \mu_-$ from (70). Thus, we can write (77) as

$$X^+(0) - X^-(0) = \frac{i(\mu_+ - \mu_-)}{p_+^+p_-^-}. \tag{78}$$

For the numerator in (76), substitute the decomposition (74) into (75) to obtain

$$X^+ - X^- = \frac{\phi_{\xi_+} - \phi_{\xi_-}}{z(1-\phi_{H_+^+})(1-\phi_{H_-^-})}. \tag{79}$$

Finally, plugging (78) and (79) into (76) gives the limiting CF of the SDRW in (27), which generalizes Fine's (28) [24]. The fact that the second and third term in (27) are the CFs of the RW maximum M_+^+ and minimum M_-^- follows from Theorem 2.

APPENDIX B
DERIVATION OF THE APPROXIMATION FOR THE RMS TIMING JITTER IN (61)

The idea to obtain the approximation (61) is to use an integral test.

Theorem 3. (Integral test [37, p.139])*: Let f be a continuous, positive and monotonically decreasing function on $[1,\infty)$. The series*

$$\sum_{n=1}^{\infty} f(n) \tag{80}$$

converges if and only if the improper integral

$$\int_1^\infty f(t)\,dt = L \tag{81}$$

is finite, in which case the series is upper bounded by $f(1)+L$.

Since G_2 in (48) is of the form (80), we can upper bound the variance $\sigma_{\Delta t}^2$ in (60) by applying the integral test as follows:

$$\sum_{n=1}^{\infty} g_2(n,x) \;<\; g_2(1,x) + \int_1^\infty g_2(t,x)\,dt \equiv H_2(x). \tag{82}$$

To show that $g_2(t,x)$ is monotonically decreasing in t on $[1,\infty)$ for every $x>0$, define $y=\sqrt{n/2}x > 0$ and write (50) as

$$g_2(y) = \left(y^2+\frac{1}{2}\right)\operatorname{erfc}(y) - \frac{y}{\sqrt{\pi}}e^{-y^2}. \tag{83}$$

Differentiating this equation with respect to y (denoted by prime) and using $\operatorname{erfc}'(y) = -2\exp(-y^2)/\sqrt{\pi}$ gives

$$g_2'(y) = 2y\left(\operatorname{erfc}(y) - \frac{y}{\sqrt{\pi}}e^{-y^2}\right). \tag{84}$$

The bound $\operatorname{erfc}(y) < y\exp(-y^2)/\sqrt{\pi}$ for $y>0$ [34, p.562] implies that $g_2'(y) < 0$ and thus that $g_2(t,x)$ is monotonically decreasing in t on $(0,\infty)$ for every $x>0$. Formal integration in (82) yields

$$H_2(x) = \frac{1}{4x^2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) + \frac{x^2}{4}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right) - \frac{x^2-1}{2\sqrt{2\pi}x}e^{-x^2/2} \tag{85}$$

so that the variance (60) can be upper bounded by

$$\sigma_{\Delta t}^2 < \frac{K^2}{3} + \sigma^2 + \sigma^2 H_2\left(\frac{K-\Delta T}{\sigma}\right) + \sigma^2 H_2\left(\frac{K+\Delta T}{\sigma}\right). \tag{86}$$

Simulations show that this upper bound is very tight; it is therefore convenient to simplify (86) further. The simple approximate formula for $\sigma_{\Delta t}^2$ is obtained by using the asymptotic expansion [34, p.562]

$$\operatorname{erfc}(x) \sim \frac{e^{-x^2}}{\sqrt{\pi}x}\left(1 - \frac{1}{2x^2} + \frac{3}{4x^4} - + \cdots\right). \tag{87}$$

Replacing the second erfc in (85) by the first two terms of this expansion shows that H_2 can be approximated by

$$H_2(x) \approx \frac{1}{4x^2}\operatorname{erfc}\left(\frac{x}{\sqrt{2}}\right). \tag{88}$$

Replacing H_2 in (86) by (88) and omitting erfc gives the approximation for the variance in (61).

REFERENCES

[1] Y. M. Greshishchev, P. Schvan, J. L. Showell, M.-L. Xu, J. J. Ojha, and J. E. Rogers, "A fully integrated SiGe receiver IC for 10-Gb/s data rate," *IEEE J. Solid-State Circuits*, vol. 35, no. 12, pp. 1949–1957, Dec. 2000.

[2] R. C. Walker, "Designing bang-bang PLLs for clock and data recovery in serial data transmission systems," in *Phase-Locking in High-Performance Systems - From Devices to Architectures*, B. Razavi, Ed. IEEE Press, 2003, pp. 34–45.

[3] B. Razavi, "Challenges in the design of high-speed clock and data recovery circuits," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 94–101, Aug. 2002.

[4] N. Da Dalt, E. Thaller, P. Gregorius, and L. Gazsi, "A compact triple-band low-jitter digital LC PLL with programmable coil in 130-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 40, no. 7, pp. 1482–1490, Jul. 2005.

[5] A. Rylyakov, J. Tierno, H. Ainspan, J.-O. Plouchart, J. Bulzacchelli, Z. T. Deniz, and D. Friedman, "Bang-bang digital PLLs at 11 and 20GHz with sub-200fs integrated jitter for high-speed serial communication applications," in *ISSCC Dig. Tech. Papers*,, San Francisco, Feb. 2009, pp. 94–95,95a.

[6] F. M. Gardner, *Phaselock Techniques*, 3rd ed. John Wiley & Sons, 2005.

[7] N. Da Dalt, "A design-oriented study of the nonlinear dynamics of digital bang-bang PLLs," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 1, pp. 21–31, Jan. 2005.

[8] A. Teplinsky, E. Condon, and O. Feely, "Driven interval shift dynamics in sigma-delta modulators and phase-locked loops," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 6, pp. 1224–1235, Jun. 2005.

[9] R. B. Staszewski, C. Fernando, and P. T. Balsara, "Event-driven simulation and modeling of phase noise of an RF oscillator," *IEEE Trans. Circuits Syst. I*, vol. 52, no. 4, pp. 723–733, Apr. 2005.

[10] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and $1/f$ noise," *IEEE Trans. Circuits Syst. I*, vol. 53, no. 9, pp. 1869–1884, Sep. 2006.

[11] J. Lee, K. S. Kundert, and B. Razavi, "Analysis and modeling of bang-bang clock and data recovery circuits," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1571–1580, Sep. 2004.

[12] F. A. Musa and A. Chan Carusone, "Modeling and design of multilevel bang-bang CDRs in the presence of ISI and noise," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 10, pp. 2137–2147, Oct. 2007.

[13] N. Da Dalt, "Linearized analysis of a digital bang-bang PLL and its validity limits applied to jitter transfer and jitter generation," *IEEE Trans. Circuits Syst. I*, vol. 55, no. 11, pp. 3663–3675, Dec. 2008.

[14] Y. Choi, D.-K. Jeong, and W. Kim, "Jitter transfer analysis of tracked oversampling techniques for multigigabit clock and data recovery," *IEEE Trans. Circuits Syst. II*, vol. 50, no. 11, pp. 775–783, Nov. 2003.

[15] N. Da Dalt, "Markov chains-based derivation of the phase detector gain in bang-bang PLLs," *IEEE Trans. Circuits Syst. II*, vol. 53, no. 11, pp. 1195–1199, Nov. 2006.

[16] B. Chun and M. P. Kennedy, "Statistical properties of first-order bang-bang PLL with nonzero loop delay," *IEEE Trans. Circuits Syst. II*, vol. 55, no. 10, pp. 1016–1020, Oct. 2008.

[17] S. Tertinek and O. Feely, "Investigation of first-order digital bang-bang phase-locked loops with reference clock jitter," in *Proc. NORCHIP'08*, Tallinn, Nov. 2008, pp. 217–222.

[18] ——, "Combined effect of loop delay and reference clock jitter in first-order digital bang-bang phase-locked loops," in *Proc. ISCAS'09*, Taipei, May 2009, pp. 2393–2396.

[19] A. Weinberg and B. Liu, "Discrete time analyses of nonuniform sampling first- and second-order digital phase lock loops," *IEEE Trans. Commun.*, vol. 22, no. 2, pp. 123–137, Feb. 1974.

[20] J. G. Kenney, D. Dalton, E. Evans, M. H. Eskiyerli, B. Hilton, D. Hitchcox, T. Kwok, D. Mulcahy, C. McQuilkin, V. Reddy, S. Selvanayagam, P. Shepherd, W. S. Titus, and L. DeVito, "A 9.95—11.3-Gb/s XFP transceiver in 0.13-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2901–2910, Dec. 2006.

[21] P. Muller and Y. Leblebici, *CMOS Multichannel Single-Chip Receivers for Multi-Gigabit Optical Data Communications*. Springer, 2007, p. 134.

[22] A. Carlsund, "First passage times for random walks and birth-and-death processes with sign depending transition probabilities," Ph.D. dissertation, Department of Mathematics, Royal Institute of Technology, Stockholm, 2003. [Online]. Available: http://www.math.kth.se/∼carlsund/

[23] M. Lefebvre, "First passage problems for asymmetric Wiener processes," *Journal of Applied Probability*, vol. 43, no. 1, pp. 175–184, 2006.

[24] T. L. Fine, "The response of a particular nonlinear system with feedback to each of two random processes," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 255–264, Mar. 1968.

[25] R. Steele, *Delta Modulation Systems*. John Wiley & Sons, 1975.

[26] S. Haykin, *An Introduction to Analog & Digital Communications*. John Wiley & Sons, 1989, chap. 5.7.

[27] R. M. Gray, *Source Coding Theory*. Kluwer Academic Publishers, 1990, chap. 6.7.

[28] I. Galton, "Analog-input digital phase-locked loops for precise frequency and phase demodulation," *IEEE Trans. Circuits Syst. II*, vol. 42, no. 10, pp. 621–630, Oct. 1995.

[29] T. Koski, "Statistics of the binary quantizer error in single-loop sigma-delta modulation with white Gaussian input," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 931–943, Jul. 1995.

[30] P. K. Hanumolu, G.-Y. Wei, U.-K. Moon, and K. Mayaram, "Digitally-enhanced phase-locking circuits," *Proc. CICC'07*, pp. 361–368, Sep. 2007.

[31] C. Kromer, G. Sialm, C. Menolfi, M. Schmatz, F. Ellinger, and H. Jäckel, "A 25-Gb/s CDR in 90-nm CMOS for high-density interconnects," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2921–2929, Dec. 2006.

[32] P.-H. Hsieh and C.-K. K. Yang, "Technique to reduce the resolution requirement of digitally controlled oscillators for digital PLLs," *IEEE Trans. Circuits Syst. II*, vol. 54,, no. 3, pp. 237–241, Mar. 2007.

[33] W. Feller, *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, 1966, vol. II.

[34] E. W. Weisstein, *CRC Concise Encyclopedia of Mathematics*. CRC Press, 1999.

[35] E. Masry and S. Cambanis, "Delta modulation of the Wiener process," *IEEE Trans. Commun.*, vol. 23, no. 11, pp. 1297–1300, Nov. 1975.

[36] S. Asmussen, *Applied Probability and Queues*, 2nd ed. Springer-Verlag New York, 2003.

[37] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill, 1976.