

DTC Linearization via Mismatch-Noise Cancellation for Digital Fractional- N PLLs

Eslam Helal^{1b}, Amr I. Eissa^{1b}, and Ian Galton^{2b}, *Fellow, IEEE*

Abstract—Digital-to-time converter (DTC) based quantization noise cancellation (QNC) has recently been shown to enable excellent fractional- N PLL performance, but it requires a highly-linear DTC. Known DTC linearization strategies include analog-domain techniques which involve performance tradeoffs and digital predistortion techniques which converge slowly relative to typical required PLL settling times. Alternatively, a DTC implemented as a cascade of 1-bit DTC stages can be made highly linear without special techniques, but such DTCs typically introduce excessive error from component mismatches which has so far hindered their use in low-jitter PLLs. This paper presents a background calibration technique that addresses this issue by adaptively canceling error from DTC component mismatches. The technique is entirely digital, is compatible with a large class of digital fractional- N PLLs, and has at least an order of magnitude lower convergence time than the above-mentioned predistortion techniques. The paper presents a rigorous theoretical analysis closely supported by simulation results which quantifies the calibration technique's convergence time and noise performance.

Index Terms—Fractional- N PLL, digital PLL, digital-to-time converter (DTC), dynamic element matching (DEM), background calibration, least-mean-square algorithm (LMS), mismatch-noise cancellation (MNC), quantization noise cancellation (QNC).

I. INTRODUCTION

THE signal processing performed within any fractional- N phase-locked loop (PLL) for frequency synthesis inevitably involves quantization. The resulting quantization error degrades the PLL's phase noise unless it is actively canceled prior to frequency modulation, a process known as quantization noise cancellation (QNC). An increasingly popular QNC method uses a digital-to-time converter (DTC) to cancel most of the quantization error prior to phase error measurement within the PLL. This prevents the quantization noise from being subjected to the inadvertent but inevitable nonlinearity of the phase error measurement circuitry, thereby avoiding fractional spurs which would otherwise be caused by nonlinearly distorting the quantization error [1], [2], [3], [4], [5].

However, for such DTC-based QNC to be effective, the DTC must be highly linear. Otherwise, it nonlinearly distorts

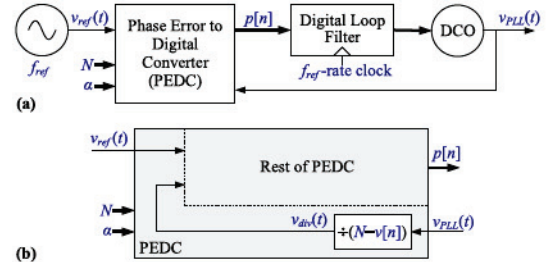


Fig. 1. a) General form of a digital fractional- N PLL driven by an f_{ref} -frequency reference oscillator, b) general form of a multi-modulus divider-based PEDC.

the quantization error directly, so it becomes a cause of fractional spurs in its own right. Several analog and digital techniques have been proposed to linearize DTC circuits to address this issue. The published analog techniques generally increase power consumption or circuit area considerably [6], [7], [8], [9], [10].

The published digital techniques perform predistortion via look-up tables (LUTs) to mitigate DTC nonlinearity, but the data with which the LUTs are populated must be measured in background via correlation algorithms which take considerably longer to converge than typical target PLL settling times [11], [12], [13]. As examples, the cold-start convergence times of the techniques presented in [11] and [12] are over 30,000 and 600,000 reference cycles, respectively.

Alternatively, a DTC implemented as a cascade of 1-bit DTC stages can be made highly linear without any special linearization techniques provided the stages are sufficiently buffered so that the state of each stage does not significantly affect the delays through the other stages. However, such DTCs typically introduce far more error from component mismatches than the more commonly used single-stage DTCs, which has so far stymied their application to DTC-based QNC in low-jitter PLLs. While dynamic element matching (DEM) can be applied to cause the DTC error arising from component mismatches, i.e., the DTC *mismatch noise*, to be free of nonlinear distortion and have a highpass spectral shape, the power of the mismatch noise nevertheless tends to be high enough that it significantly degrades the PLL's jitter.

This paper proposes an entirely digital DTC mismatch noise cancellation (MNC) technique that is applicable to a large class of digital fractional- N PLLs. The DTC-MNC technique adaptively measures and cancels DTC mismatch noise in background within the PLL prior to the PLL's digital loop filter, thereby making highly-linear DTCs comprised of 1-bit DTC stages practical for low-jitter digital fractional- N PLLs (although the technique is also applicable to single-stage DTCs). The DTC-MNC technique's convergence time is an order of magnitude faster than that of the fastest of the published predistortion techniques and results in significantly lower simulated jitter and spurious tones than the corresponding reported simulation and measurement results for the previously published predistortion techniques. The paper presents

Manuscript received 31 May 2022; revised 20 July 2022; accepted 8 August 2022. Date of publication 15 November 2022; date of current version 9 December 2022. This work was supported in part by the National Science Foundation under Award 1909678 and in part by Qualcomm, Inc. This article was recommended by Associate Editor H. Zhang. (Corresponding author: Eslam Helal.)

Eslam Helal and Amr I. Eissa are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92092 USA, and also with the Department of Electronics and Communications Engineering, Ain Shams University, Cairo 11566, Egypt (e-mail: ehelal@eng.ucsd.edu).

Ian Galton is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92092 USA.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TCSI.2022.3200047>.

Digital Object Identifier 10.1109/TCSI.2022.3200047

1549-8328 © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

a rigorous analysis closely supported by simulation results which quantifies the DTC-MNC technique's convergence time, and proves that the DTC-MNC technique has no convergence bias and is unconditionally stable.

II. DTC-BASED QUANTIZATION NOISE CANCELLATION

A. General Form of a Digital PLL With DTC-Based QNC

The general form of a digital fractional- N PLL driven by a reference oscillator of frequency f_{ref} is shown in Fig. 1a [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], and [26]. The PLL is comprised of a phase-error-to-digital converter (PEDC), a lowpass digital loop filter, and a digitally controlled oscillator (DCO). Its objective is to generate a low-noise oscillatory output signal, $v_{PLL}(t)$, with instantaneous frequency $f_{PLL} = (N + \alpha)f_{ref}$, where N is a positive integer and α is a fractional value bounded in magnitude by 1.

In many digital fractional- N PLLs, the PEDC incorporates a multi-modulus divider as shown in Fig. 1b. The multi-modulus divider is controlled such that its n th and $(n+1)$ th rising output edges are separated by $N - v[n]$ DCO cycles, where $v[n]$ is an integer-valued digital sequence generated within the PEDC. The rest of the PEDC digitizes the phase difference between $v_{ref}(t)$ and $v_{div}(t)$ to generate $p[n]$, and the PLL's feedback loop controls the DCO such that $p[n]$ stays bounded, thereby ensuring that the divider's average output frequency is f_{ref} . In some PLLs, $v[n]$ is generated by a digital delta-sigma ($\Delta\Sigma$) modulator such that its average value is $-\alpha$, and in other PLLs, $v[n]$ is generated within the PLL's feedback loop such that its average converges to $-\alpha$. In either case, the $N - v[n]$ division in conjunction with the feedback causing the divider's average output frequency to converge to f_{ref} causes the DCO's average output frequency to converge to $(N + \alpha)f_{ref}$.

The reason that $v[n]$ is restricted to integer values is that dividers are only capable of counting integer numbers of DCO cycles. Hence, in all such PLLs, $v[n]$ contains zero-mean quantization error which ultimately contributes to the PLL's overall phase error unless it is canceled prior to the DCO via QNC.

The quantization process with which the PEDC generates $v[n]$ happens in the digital domain so the quantization error is known to the system. One option is to perform QNC in the digital domain after the PEDC digitizes the phase difference between $v_{ref}(t)$ and $v_{div}(t)$. However, in most PLLs with divider-based PEDCs, quantization error is the dominant component in $v[n]$ and when the quantization error is subjected to the inevitable nonlinearity of the PEDC's phase error measurement and digitization circuitry, fractional spurs are induced which digital-domain QNC is unable to cancel. Therefore, it is desirable to perform QNC prior to phase error measurement and digitization if possible.

In principle, this can be done by inserting a DTC between the divider output and the rest of the PEDC. Ideally, the DTC would introduce a time delay of $T_D + \varepsilon[n]$ to the n th output edge of the divider, where $\varepsilon[n]$ represents the effect of the quantization error on the time of the n th rising output edge of the divider, and T_D is a constant that is large enough to ensure that $T_D + \varepsilon[n] > 0$ for DTC causality. Hence, the time of the n th rising edge of the DTC output is the ideal time of the n th rising edge of the divider output, i.e., the time that would have resulted had $v[n]$ not been quantized, aside from an additional fixed delay of T_D . The rest of the PEDC digitizes the phase difference between $v_{ref}(t)$ and $v_{DTC}(t)$, so the PLL's feedback controls the DCO such that the average value of this difference

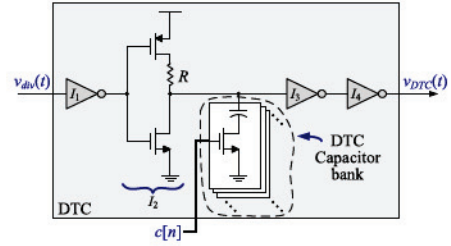


Fig. 2. A commonly-used DTC circuit.

converges to zero, thereby causing the average DCO frequency to converge to $(N + \alpha)f_{ref}$. The primary difference between this case and that of Fig. 1 is that QNC occurs prior to phase error measurement and digitization, which has the potential to significantly reduce spurious tones.

A commonly used DTC circuit is shown in Fig. 2 [7], [8], and [27]. It consists of inverters I_1 , I_2 , I_3 , and I_4 , and a bank of capacitors. The i th capacitor's top plate is connected to the output of inverter I_2 , and its bottom plate is connected to or disconnected from ground when the i th bit of the input codeword $c[n]$ is high or low, respectively. Hence, $c[n]$ controls the RC time constant at the output of inverter I_2 , and, consequently, the delay through the DTC.

Inverter I_2 differs from the other inverters, which are standard two-transistor inverters, in that it contains resistor R in series with the drain of the inverter's pMOS transistor. The pMOS transistor is chosen to be wide enough that its on-resistance is small compared to R . This makes the time constant at the output node of inverter I_2 relatively independent of the transistor's on-resistance when the inverter's output voltage transitions from low-to-high, thereby improving the DTC's linearity, i.e., the linearity of the delay between each rising edge of $v_{div}(t)$ and the corresponding rising edge of $v_{DTC}(t)$ as a function of $c[n]$. Another advantage of this design choice is that the large pMOS transistor size results in a relatively low flicker noise contribution from the transistor. As the PLL's timing information is only carried by the times of the rising edges of the divider and DTC outputs, it is not necessary for the nMOS transistor in inverter I_2 to be large or to include a resistor in series with its drain.

B. Nonideal DTC Behavior

A DTC's resolution specifies the number of different delays that the DTC is able to introduce. For example, if the capacitors in the DTC of Fig. 2 have values of $2^i C$ for $i = 0, 1, 2, \dots, b-1$, and the i th bit of $c[n]$ controls the transistor connected to the $2^i C$ capacitor, then the DTC is said to have b bits of resolution because it can introduce 2^b different delays.

In many applications, the minimum step-size of α is so small that it is not practical to implement a DTC with sufficient resolution to achieve delays of exactly $T_D + \varepsilon[n]$, so it is often necessary to have the DTC input be a quantized version of $\varepsilon[n]$. As described above, the purpose of the DTC is to cancel the effect of the quantization error in $v[n]$ prior to the PEDC's phase error measurement and digitization process, so quantizing $\varepsilon[n]$ prior to the DTC appears, at first glance, to defeat the purpose of the DTC. However, the error from quantizing $\varepsilon[n]$ prior to the DTC usually can be made much smaller than the quantization error in $v[n]$, so the quantization-noise-induced spurious tones it causes are much smaller than the those which would have occurred in the absence of the DTC. Furthermore, the quantization of $\varepsilon[n]$ is done in the digital domain, so the quantization error is available within the PEDC as digital sequence. Hence, if necessary, the small

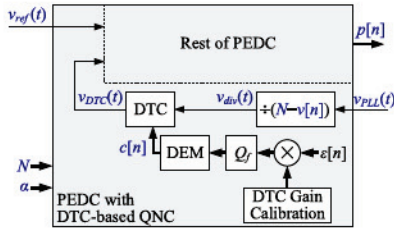


Fig. 3. The PEDC of Fig. 1b with a DEM encoder, quantizer Q_f , and DTC gain calibration.

amount of quantization error introduced by the quantization of $\varepsilon[n]$ can be canceled within the rest of the PEDC following the phase error measurement and digitization operation.

Another practical DTC limitation relates to component mismatches. In the DTC example described above, the 2^iC capacitor would typically be implemented as a parallel combination of 2^i unit capacitors of size C for each $i = 1, \dots, b - 1$. Mismatches among the different unit capacitors from fabrication errors and systematic layout asymmetries cause the b capacitors to deviate from their ideal values, which results in DTC nonlinearity.

If necessary, DEM can be applied to at least partially address this problem [28], [29]. Provided the number of DTC capacitors and their nominal values satisfy certain constraints, a digital DEM encoder can be used prior to the DTC to control which capacitors are connected and disconnected within the DTC during each reference period such that the error introduced by component mismatches is either white or highpass spectrally shaped noise instead of nonlinear distortion [30].

As described above, the DTC ideally introduces a delay to the n th rising edge of $v_{dio}(t)$ that well-approximates $T_D + \varepsilon[n]$. However, DTC gain error, which is inevitable in practice because of various types of nonideal circuit behavior, causes this delay to instead approximate $A_{DTC}(T_D + \varepsilon[n])$ where A_{DTC} is a constant that deviates from its ideal value of unity. Fortunately, background calibration techniques that adaptively measure A_{DTC} and compensate for it are well-known [4], [22].

Fig. 3 shows a DTC-enabled version of the PEDC of Fig. 1b in which DTC gain calibration, quantization, and DEM are applied to address the DTC's gain error, resolution limitation, and component mismatches, respectively. The details of the DEM encoder and the DTC gain calibration are not described in detail in this paper because they are well-known, established techniques that are described in detail in the cited references.

The remaining types of nonideal DTC behavior are circuit noise, and nonlinearity from sources other than component mismatches. Usually, for a given DTC topology, circuit noise can only be reduced at the expense of increased power consumption and/or area. Nevertheless, fractional- N PLLs with DTC-based QNC have been demonstrated with excellent phase noise performance and power efficiency, so the circuit noise issue has proven to be manageable [7], [8]. Unfortunately, DTC nonlinearity from sources other than component mismatches remains a significant issue, especially for DTCs with high dynamic range [6].

C. DTC Linearity Versus Component Mismatch Tradeoff

While careful sizing of the pMOS transistor and resistor in Inverter I_2 can reduce the nonlinearity of the type of DTC shown in Fig. 2 as described above, it is often not possible to reduce it sufficiently to prevent it from causing significant spurious tones. Consequently, low-jitter fractional- N PLLs with DTC-based QNC typically incorporate DTC linearization techniques as mentioned in the introduction.



Fig. 4. A DTC implemented as a cascade of M 1-bit DTCs.

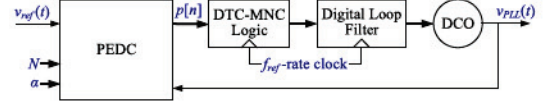


Fig. 5. General form of a digital fractional- N PLL with the MNC technique.

Alternatively, the DTC can be implemented as a cascade of 1-bit DTC stages as shown in Fig. 4. Ideally, the i th DTC stage introduces a delay of $\delta_i[n] = T_i + c_i[n]\Delta_i$, where T_i is a constant delay, $c_i[n]$ is the i th output bit of the DEM encoder preceding the DTC, and Δ_i is a constant which represents the DTC stage's delay *step-size*. For example, each 1-bit DTC stage in Fig. 4 can be implemented by the DTC shown in Fig. 2 except with a single capacitor and nMOS transistor in place of the full DTC capacitor bank. To the extent that the inverters at the input and output of each stage provide sufficient isolation that the i th stage's delay, $\delta_i[n]$, does not depend on $c_j[n]$ for any $j \neq i$, each DTC stage introduces one of only two possible delays to its input at any given time so each 1-bit stage is inherently linear (two points always lie on a straight line). However, component mismatches cause the two possible delays from each DTC stage to have static deviations from their ideal values, which, in the absence of DEM, would introduce overall DTC nonlinearity. Fortunately, by scrambling the usage pattern of the DTC stages, the DEM encoder causes error from component mismatches to introduce noise-like error instead of nonlinear distortion [30].

However, unlike the single-stage DTC of Fig. 2 wherein mismatch noise is dominated mainly by unit capacitor mismatches, every component within each stage of the multi-stage DTC contributes to the DTC's mismatch noise. As mentioned in the introduction, this typically causes the mismatch noise from the multi-stage DTC of Fig. 4 to be so high that its application to QNC in low-jitter PLLs has been problematic to date. The DTC-MNC technique presented in the next section addresses this problem.

III. ADAPTIVE DTC MISMATCH NOISE CANCELLATION

The PEDC of Fig. 3 generates an output sequence which can be written as

$$p[n] = r_{ideal}[n] + r_e[n], \quad (1)$$

where $r_{ideal}[n]$ is what $p[n]$ would have been had the DTC not introduced mismatch noise, and $r_e[n]$ is the component of $p[n]$ resulting from DTC mismatch noise. The purpose of the DTC-MNC technique is to adaptively measure and cancel $r_e[n]$. As explained shortly, this is accomplished by the block labeled DTC-MNC logic in the digital PLL shown in Fig. 5.

In general, DEM causes the DTC's mismatch noise, $e_{DTC}[n]$, to have the form

$$e_{DTC}[n] = \sum_{k=1}^L B_k S_k[n], \quad (2)$$

where L is a constant that depends on the details of the DEM encoder, each $S_k[n]$ is a white or spectrally shaped pseudo-random sequence that is known because it is generated within the DEM encoder, and each B_k is a constant that is unknown because it depends on the DTC's component mismatches [30], [32].

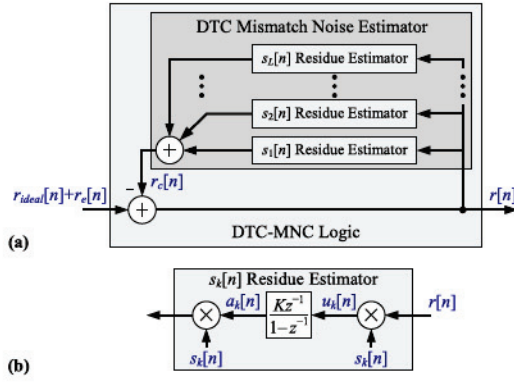


Fig. 6. DTC-MNC logic details; a) high-level view, b) details of the k th feedback loop's $s_k[n]$ residue amplifier.

The PEDCs in high-performance PLLs must be quite linear to avoid inducing large spurious tones, so by far the largest term in $p[n]$ resulting from $e_{DTC}[n]$ is a scaled and delayed version of $e_{DTC}[n]$. While PEDC nonlinearity causes $p[n]$ to also contain a nonlinearly distorted version of $e_{DTC}[n]$ which the DTC-MNC technique does not completely cancel, the power ratio of the linear to nonlinear terms is typically several tens of dB. Provided this ratio is larger than the desired level of cancellation of $e_{DTC}[n]$, which is usually only about 30 dB, then the nonlinearity of the PEDC can be neglected. Hence, (2) implies that $r_e[n]$ can be approximated as

$$r_e[n] = \sum_{k=1}^L b_k s_k[n], \quad (3)$$

where b_k is proportional to B_k , $s_k[n] = S_k[n - Q]$, and Q is a positive integer delay.

A. DTC Mismatch-Noise Cancellation Implementation

The details of the DTC-MNC logic block in Fig. 5 are shown in Fig. 6. The structure consists of L feedback loops, each of which contains the residue estimator block shown in Fig. 6b. The k th residue estimator accumulates $Kr[n]s_k[n]$ and multiplies the result by $s_k[n]$, where K is a constant called the DTC-MNC loop gain. For most types of DEM including those considered in this paper, each $s_k[n]$ sequence is limited to values of $-1, 0$, and 1 , so the multiplications are not hardware-intensive. As proven shortly, the k th feedback loop estimates and cancels the k th term of (3) in background, i.e., during normal operation of the PLL.

Although the DTC-MNC technique is applicable to any type of DEM, the analysis presented in this paper assumes that the DEM encoder has the general form of that presented in [31] and [32] and causes the DTC's mismatch noise to have either a white or first-order highpass shaped power spectral density (PSD). In all such cases, $s_k[n]$ for each k is a known, zero-mean pseudo-random sequence called a *switching sequence*, which takes on values of $-1, 0$, and 1 .

It is necessary in the analysis below to keep track of the values of n for which each switching sequence is non-zero, so the m th non-negative integer n for which $s_k[n] \neq 0$ is denoted as $J_{m,k}$. Therefore, $0 \leq J_{1,k} < J_{2,k} < J_{3,k} < \dots$, and $s_k[n] = 0$ if $n \neq J_{m,k}$ for any value of m .

The two most common options for the switching sequences in (3) are analyzed in this paper: *white switching sequences* and *first-order highpass shaped switching sequences*. The non-zero values of these sequences are given by

$$s_k[J_{r,k}] = w_k[r], \text{ and } s_k[J_{r,k}] = (-1)^{r-1} w_k \left[\left\lfloor \frac{r-1}{2} \right\rfloor \right], \quad (4)$$

respectively, where $r = 1, 2, 3, \dots$, the sequences $w_k[p]$, for all k and p , are independent zero-mean random variables, each of which is restricted to values of -1 and 1 , and $\lfloor x \rfloor$ for any real number x denotes the largest integer less than or equal to x .

The two switching sequence options correspond to the switching sequences generated by mismatch scrambling and first-order highpass mismatch shaping DEM encoders, respectively. For the latter case, (4) implies that each successive pair of non-zero $s_k[n]$ values is either $1, -1$, or $-1, 1$, where the choice between these two possibilities is made randomly with equal probability and independently from all other variables in the system.

B. DTC Mismatch-Noise Cancellation Analysis

The DTC-MNC logic is a special case of the multi-loop least-mean-square (LMS) like noise canceler analyzed in [33] for the case of white switching sequences, but not for the case of spectrally-shaped switching sequences. Furthermore, the DTC-MNC logic for both types of switching sequences differs in two ways from the noise canceler presented in [33], and these differences enable the significantly different and much more precise analysis presented in this paper. One difference is that the input to the DEM encoder that drives the DTC is not arbitrary; it is the sum of the quantization noise component of $v[n]$ and quantization noise from the Q_f quantizer in Fig. 3, each of which is the result of either dithered digital quantization or digital $\Delta\Sigma$ modulation in typical PLLs. The analysis presented in this paper relies on the properties of such DTC input sequences to accurately quantify the convergence speed of the DTC-MNC technique. The other difference is that the DTC-MNC logic is simpler than the noise canceler presented [33], which allows for much tighter error bounds than were derived in [33].

It follows from Fig. 6 and (3) that

$$a_k[n] = a_k[n-1] + K u_k[n-1] \quad (5)$$

for each $k = 1, 2, \dots, L$, where

$$u_k[n] = s_k[n] \left(r_{ideal}[n] + \sum_{l=1}^L s_l[n] (b_l - a_l[n]) \right). \quad (6)$$

The objective of the DTC-MNC logic is to cause $r_c[n] = r_e[n]$ such that $r[n] = r_{ideal}[n]$. Fig. 6 implies that

$$r_c[n] = \sum_{k=1}^L s_k[n] a_k[n], \quad (7)$$

with which (3) implies that this objective would be perfectly achieved if each $a_k[n]$ coefficient were equal to b_k . Therefore, the *convergence error* of each accumulator in Fig. 6b is defined as

$$z_k[n] = a_k[n] - b_k. \quad (8)$$

Combining (5), (6), and (8) with n replaced by $n+1$ gives

$$z_k[n+1] = z_k[n] + K s_k[n] r_{ideal}[n] - K s_k[n] \sum_{l=1}^L s_l[n] z_l[n] \quad (9)$$

for each $k = 1, 2, \dots, L$. Therefore, $z_k[n]$ for each $k = 1, 2, \dots, L$ is specified for all $n \geq 0$ by difference equations (9) with initial conditions

$$z_j[0] = a_j[0] - b_j \text{ for } j = 1, 2, \dots, L. \quad (10)$$

The system is considered to be “turned on” at time $n = 0$, so

$$z_j[n] = 0 \text{ for } n < 0 \text{ and } j = 1, 2, \dots, L. \quad (11)$$

The theorems presented below and discussed subsequently, which are proven in the appendix, apply to switching sequences given by (4) and system equations (9), (10), and (11). They quantify the convergence rate and noise performance of the DTC-MNC technique provided the switching sequences, which depend on the DEM encoder’s input sequence, and $r_{ideal}[m]$ satisfy the theorem hypotheses. Simulation results that closely support the theorems’ results are presented in the next subsection.

Theorem 1: For white switching sequences and $n \geq 0$, if neither $s_j[m]$ nor $r_{ideal}[m]$ depend on whether $s_k[n]$ is zero or nonzero for any j, k , and $n > m$, then

$$E\{z_k[n+1] | s_k[n]\} = \begin{cases} \bar{z}_k[n], & \text{if } s_k[n] = 0, \\ \bar{z}_k[n](1-K), & \text{if } s_k[n] \neq 0, \end{cases} \quad (12)$$

where $\bar{z}_k[n] = E\{z_k[n]\}$. If, in addition, $E\{s_k^2[n]\}$ does not depend on n , and $a_j[0] = 0$ for $j = 1, 2, \dots, L$, then

$$\bar{z}_k[n] = -b_k(1-c_kK)^n, \quad (13)$$

where $c_k = E\{s_k^2[m]\}$. $\square$

Theorem 2: For white switching sequences and $n \geq 0$, if $0 < K < 2c_{min}/(cL)$ neither $s_j[m]$ nor $r_{ideal}[m]$ depend on whether $s_k[n]$ is zero or nonzero for any j, k , and $n > m$, neither $E\{s_k^2[n]\}$ nor $E\{r_{ideal}^2[n]\}$ depend on n , and $E\{s_k^2[n]\} \neq 0$ for all k , then

$$\limsup_{n \rightarrow \infty} \{\sigma_z^2[n]\} \leq \frac{Kc\sigma_{r_{ideal}}^2}{2c_{min} - KcL}, \quad (14)$$

where

$$\sigma_z^2[n] = \frac{1}{L} \sum_{k=1}^L \bar{z}_k^2[n], \quad (15)$$

$\bar{z}_k^2[n] = E\{z_k^2[n]\}$, $\sigma_{r_{ideal}}^2 = E\{r_{ideal}^2[n]\}$, and c_{min} and c are the minimum and average values of $c_k = E\{s_k^2[n]\}$ over $k = 1, 2, \dots, L$, respectively. $\square$

Theorem 3: For first-order highpass shaped switching sequences and $n \geq 0$, if $0 < K < 1/L$, $E\{s_k^2[n]\}$ does not depend on n , neither $s_j[m]$ nor $r_{ideal}[m]$ depend on whether $s_k[n]$ is nonzero for any j, k , and $n > m$, and $a_j[0] = 0$ for $j = 1, 2, \dots, L$, then

$$|\bar{z}_k[n]| \leq |b_k| \left(1 - c_k K \left(\frac{1-LK}{1-K}\right)\right)^n, \quad (16)$$

where $c_k = E\{s_k^2[m]\}$. $\square$

Theorem 4: For first-order highpass shaped switching sequences and $n \geq 0$, if $0 < a < 1$, where

$$a = 1 - 2c_{min}K + K^2L \left(c_{max} \frac{2(1-K)L}{1-2K} + c \left(1 + L \frac{6+8KL+2K^2L^2}{1-2K}\right) \right), \quad (17)$$

$K < \min\{1/L, 1/2\}$, neither $s_j[m]$ nor $r_{ideal}[m]$ depend on whether $s_k[n]$ is zero or nonzero for any j, k , and $n > m$, and

neither $E\{s_k^2[n]\}$ nor $E\{r_{ideal}^2[n]\}$ depend on n , and $E\{s_k^2[n]\} \neq 0$ for all k , then

$$\limsup_{n \rightarrow \infty} \{\sigma_z^2[n]\} < \frac{cK^2}{1-a} \left(\frac{2+3KL+3K^2L^2+K^3L^3}{1-KL} \right) \sigma_{r_{ideal}}^2, \quad (18)$$

where $\sigma_z^2[n]$, $\sigma_{r_{ideal}}^2$, c_k , c_{min} , and c are as defined in the statement of Theorem 2, and c_{max} is the maximum value of $c_k = E\{s_k^2[n]\}$ over $k = 1, 2, \dots, L$. $\square$

Theorems 1 and 3 quantify the convergence rates of the DTC-MNC technique for white and first-order highpass switching sequences, respectively, in terms of the statistical means of $z_k[n]$, i.e., $\bar{z}_k[n]$, for all $n \geq 0$. Theorem 1 provides an exact expression for $\bar{z}_k[n]$ whereas Theorem 3 provides a tight upper bound on the magnitude of $\bar{z}_k[n]$. The theorems show that the convergence of the DTC-MNC technique is unbiased, which, with (8), implies that the mean values of $a_k[n]$ converge exactly to their ideal values, b_k , for all k . The theorems also show that each convergence rate is exponential with a convergence speed that increases with DTC-MNC loop gain K .

While Theorems 1 and 3 show that the means of $z_k[n]$ converge to their ideal values, they do not by themselves guarantee that the DTC-MNC logic is unconditionally stable, as they do not rule out the possibility that the variances of $z_k[n]$ could conceivably diverge. Theorems 2 and 4 address this issue by bounding the steady state variances of $z_k[n]$ for white and first-order highpass shaped switching sequences, respectively. They state conditions which ensure that the variances of $z_k[n]$ are bounded, thereby ensuring unconditional stability. The bounds they provide are in terms of K , the variance of $r_{ideal}[n]$, and how frequently the switching sequences are non-zero over time. The theorems imply that the maximum variances of $z_k[n]$ decrease with K and with the variance of $r_{ideal}[m]$. Together with Theorems 1 and 3, they quantify the convergence speed versus accuracy tradeoff associated with the choice of DTC-MNC loop gain K .

The theorems also provide insight into the tradeoffs between white and first-order highpass shaped switching sequences. Typically, K is small, e.g., less than 2^{-7} , so Theorems 1 and 3 imply that while the convergence rate is faster for white switching sequences than for first-order highpass shaped switching sequences, the difference is relatively small and decreases with K . However, Theorems 2 and 4 suggest that the variance of the convergence error is higher for first-order highpass shaped switching sequences than for white switching sequences. Nevertheless, first-order highpass shaped switching sequences suppress mismatch noise at low frequencies, so error from imperfect convergence introduced by DTC-MNC with these switching sequences tends to be suppressed at low frequencies. Consequently, the results suggest that white switching sequences become increasingly advantageous as the PLL’s bandwidth is increased whereas the opposite is true as the PLL’s bandwidth is decreased.

C. DTC Mismatch-Noise Cancellation Simulation Results

This subsection presents simulation results of a digital PLL enabled by the DTC-MNC technique, and compares them to the theoretical results presented above. The simulated PLL is based on that presented in [34] but with the modified PEDC shown in Fig. 7. The combination of the fixed divide-by-two and the multi-modulus divider can be viewed as a single

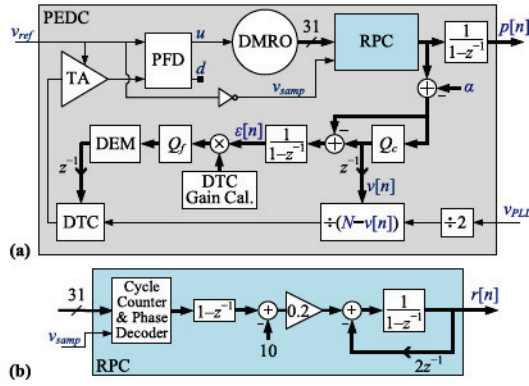


Fig. 7. Details of the simulated PLL a) high-level view of the PEDC b) digital ring phase calculator (RPC).

multi-modulus divider that divides by $2(N + \alpha)$, so the PEDC has the general form shown in Fig. 1b. The time amplifier (TA), PFD, cycle counter, phase decoder, and dividers are exactly as described in [34]. The dual-mode ring oscillator (DMRO), which has the same topology as that described in [34], has 31 delay elements and its output frequencies are approximately 3 GHz and 250 MHz when the u output of the PFD is high or low, respectively. Quantizers Q_f and Q_c are each implemented as 2nd-order $\Delta\Sigma$ modulators with LSB dither [35].

Like the original PLL presented in [34], the modified PLL has a reference frequency of $f_{ref} = 80$ MHz and its output frequency is tunable from 6 GHz to 7 GHz. All the PLL simulation results described in this section correspond to $f_{PLL} \cong 6.4$ GHz with $\alpha f_{ref} \cong 104$ kHz which is about a tenth of the PLL's 1 MHz bandwidth.

The DTC has 9 bits of resolution and has the form shown in Fig. 4 with $M = 20$ 1-bit DTC stages. It is driven by a segmented DEM encoder of the type presented in [32] with the option of either white or first-order highpass mismatch shaping, and the relative 1-bit DTC stage weights were chosen based on the tradeoffs presented in [31]. The i th 1-bit DTC stage has a nominal delay step-size of $\Delta_i = K_i \Delta$, where $\Delta = 1.4$ ps is the DTC's minimum delay step-size, $K_1, K_2, \dots, K_{12}$ equal 1, 1, 2, 2, 4, 4, ..., 32, 32, respectively, and $K_{13} = K_{14} = \dots = K_{20} = 64$. The delay between the DEM encoder and $p[n]$ is 2 reference periods, so the results presented in [32] imply that $r_e[n]$ is given by (3) with $L = 19$ and $s_k[n] = S_k[n - 2]$ where $S_k[n]$ is the DEM encoder's k th switching sequence. Hence, the DTC-MNC logic contains 19 residue estimators. The bus-width of each simulated residue estimator accumulator is 25 bits. The DTC gain calibration technique is as presented in [22].

The authors designed a transistor-level version of the DTC for the Global Foundries 22FDX process, wherein each of the 20 1-bit DTC stages has the form shown in Fig. 2 except with a single capacitor and transistor in place of the DTC capacitor bank. Circuit simulations predict that the DTC's mid-code phase noise floor relative to the 80 MHz reference frequency is -161 dBc/Hz and its power consumption is 1.8 mW, which is in line with state-of-the-art designs [6], [7]. Monte-Carlo simulations predict that the DTC's unit element delay mismatch has a standard deviation of 3.1%. For the specific case corresponding to the simulation results presented below, the resulting DTC integral nonlinearity (INL) ranges between -1.6 and 1.8 DTC LSBs across the 9-bit DTC input range. Circuit simulations further predict that the PLL's worst-case fractional spur resulting from imperfect isolation among the 1-bit DTC stages is lower than -70 dBc.

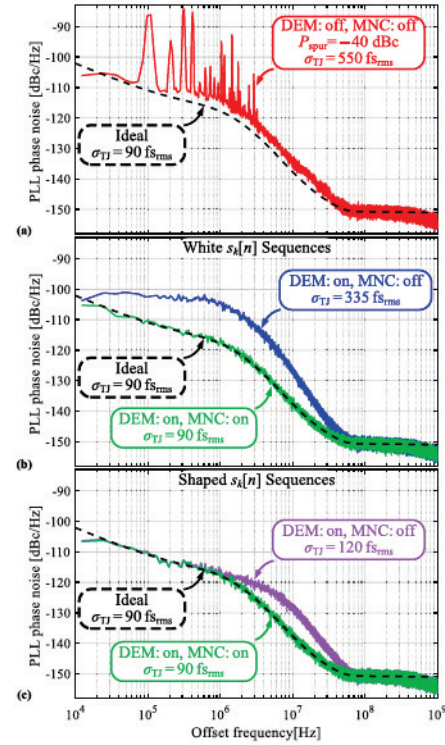


Fig. 8. PLL phase noise spectra from simulation (solid curves) and theoretical analysis (dashed curves).

The results presented in [31] ensure that the switching sequences satisfy (4), which is a requirement of the theorems presented above. By definition, $r_{ideal}[m]$ does not depend on $s_k[n]$, and it represents measured PLL phase error so it is reasonable to expect that $E\{r_{ideal}^2[n]\}$ does not depend on n once the PLL is locked. As quantified in [31], whether or not $s_k[n]$ is nonzero at time n is a complicated function of the DEM encoder's input code value at time n and some or all the values of $s_1[n], s_2[n], \dots, s_{k-1}[n]$ at time n . The DEM encoder's input sequence consists of quantization noise and accumulated quantization noise from the Q_f and Q_c $\Delta\Sigma$ modulators, respectively, and the LSB dither causes both quantization noise sequences to be asymptotically white and uniformly distributed prior to second-order noise shaping [35]. Consequently, it is reasonable to expect that $s_k[n]$ does not depend on whether future values of $s_j[n]$ are nonzero for any j and k , and that $E\{s_k^2[n]\}$ is nonzero and does not depend on n . These observations, which are further supported by simulation results performed by the authors, are consistent with the hypotheses of the four theorems presented above.

The authors used Cadence Spectre PNOISE circuit simulations to predict the phase noise of each PLL circuit block and Monte-Carlo simulations to determine component mismatches within the DTC. The results were back-annotated into a behavioral, event-driven C-language PLL simulator (along the lines of those presented in [36] and [37]) which generated all of the simulation results presented below.

Fig. 8 shows simulated PLL phase noise spectra which demonstrate the individual and combined effects of DTC mismatches, DEM, and DTC-MNC relative to the PLL's ideal phase noise spectrum. Without DEM or DTC-MNC (Fig. 8a), the DTC mismatches result in large spurious tones which degrade the PLL's RMS total jitter, σ_{TJ} , (integrated from 10 kHz to 80 MHz) to 550 fs from its ideal value of 90 fs which would have occurred in the absence of DTC

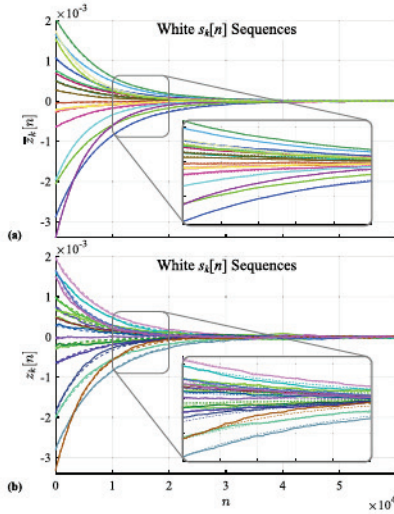


Fig. 9. Cold start trajectories of $\bar{z}_k[n]$ for white switching sequences predicted by Theorem 1 (dashed curves) with simulated trajectories (solid curves) of a) $\bar{z}_k[n]$ and b) $z_k[n]$.

mismatches. Enabling DEM without DTC-MNC causes the DTC mismatches to introduce noise rather than spurious tones, but with either white (Fig. 8b) or first-order highpass shaped (Fig. 8c) switching sequences, the noise significantly degrades the PLL's jitter. In both cases, enabling DTC-MNC cancels the noise as expected such that the simulated jitter differs insignificantly from its ideal value of 90 fs.

The results shown in Fig. 8 with DTC-MNC enabled correspond to a DTC-MNC loop gain of $K = 2^{-8}$. The theoretical results presented in Section III-B as well as the simulation results presented in Figures 9 and 10 imply that the corresponding DTC-MNC cold-start settling time — the time from when DTC-MNC technique is first enabled with uninitialized registers to the time at which the PLL's phase noise profile becomes visually indistinguishable from that which would have occurred in the absence of DTC mismatches — is less than 2000 reference periods, i.e., less than 25 μ s. As mentioned in the introduction, this is at least an order of magnitude faster than reported for the published DTC predistortion techniques [11], [12].

Fig. 9a shows simulated cold-start trajectories (solid curves) of the 19 $\bar{z}_k[n]$ sequences for DTC-MNC with $K = 2^{-12}$ and white switching sequences along with the corresponding theoretical trajectories (dashed curves) predicted by Theorem 1.

The simulated $\bar{z}_k[n]$ trajectories were obtained by averaging the $z_k[n]$ trajectories from ten separate simulation runs starting from the same initial state. As indicated in the figure, the simulated and calculated trajectories are extremely close, and the authors have verified that the simulated and corresponding theoretical trajectories become visually indistinguishable as the number of averages is increased. In principle, the averaging option is necessary because $\bar{z}_k[n]$ in Theorem 1 is the statistical mean of $z_k[n]$. Nevertheless, as shown in Fig. 9b, even without averaging, i.e., for only one simulation run, the simulated trajectories of $z_k[n]$ are very close to the trajectories of $\bar{z}_k[n]$ predicted by Theorem 1. Other values of K yield results similar to those shown in Fig. 9 aside from convergence-rate and noise variances differences.

Fig. 10 shows the simulated cold-start trajectories of $\sigma_z^2[n]$ for DTC-MNC with white switching sequences and various loop gains relative to the steady-state bounds predicted by Theorem 2. As expected, the simulated trajectories remain

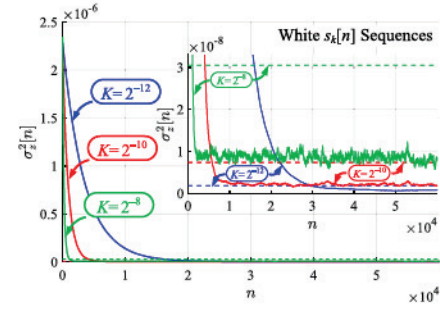


Fig. 10. Simulated cold start trajectories (solid curves) of $\sigma_z^2[n]$ for white switching sequences with the bounds predicted by Theorem 2 (dashed curves).

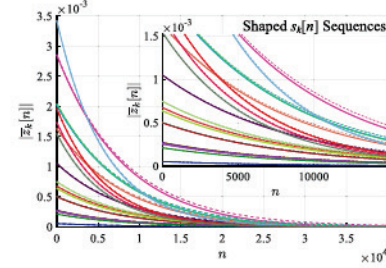


Fig. 11. Simulated cold start trajectories (solid curves) of the magnitudes of $\bar{z}_k[n]$ for first-order highpass shaped switching sequences with the bounds predicted by Theorem 3 (dashed curves).

below the bounds predicted by Theorem 2 after the initial settling transient. As can be seen in the figure, the bounds become tighter as K is decreased.

Fig. 11 shows results that correspond to those shown in Figures 9a, but for the case of first-order highpass shaped switching sequences. Given that Theorem 3 bounds the magnitude of $\bar{z}_k[n]$, Fig. 11 shows trajectories of the magnitudes of $\bar{z}_k[n]$, but otherwise the results including the convergence rates are very similar those shown in Fig. 9. Furthermore, as can be seen from Fig. 11, the bound provided by Theorem 3 is extremely tight.

Fig. 12 shows results that correspond to those shown in Fig. 10, but for the case of first-order highpass shaped sequences. As with the Fig. 10 results, Fig. 12 shows results for three values of DTC-MNC loop gain, K . The hypothesis of Theorem 4 for the parameters of this particular design example restricts K to be less than or equal to about 2^{-11} , so even though the simulation results suggest that $\sigma_z^2[n]$ has a steady-state bound and Theorem 3 ensures that $\bar{z}_k[n]$ converges to zero for all three cases, Theorem 4 only provides a bound for one of the three K values.

IV. CONCLUSION

An entirely digital background calibration technique has been presented that adaptively measures and cancels error resulting from DTC component mismatches that would otherwise degrade the phase noise of digital PLLs with DTC-based quantization noise cancellation. Aside from virtually eliminating DTC component mismatches as a source of phase noise in general, the technique indirectly addresses the well-known DTC nonlinearity problem because it facilitates the use of inherently-linear DTCs comprised of cascades of 1-bit DTC stages. Such DTCs tend to introduce excessive error from component mismatches, which has heretofore hindered their application to low-jitter PLLs. Published digital predistortion techniques provide an alternate means of mitigating DTC nonlinearity, but their convergence rates are at least an order of magnitude slower than that of the presented technique.

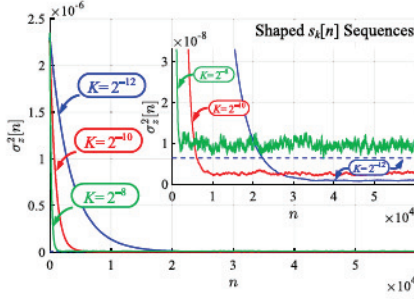


Fig. 12. Simulated cold start trajectories (solid curves) of $\sigma_z^2[n]$ for first-order highpass shaped switching sequences with the bound predicted by Theorem 4 (dashed curve).

A rigorous mathematical analysis has been presented that precisely quantifies the calibration technique's settling performance and provides conditions under which it is unconditionally stable. Closed-loop PLL simulations are notoriously time-consuming, so it is generally not practical to perform simulations over all possible PLL operating conditions. Hence, the results of the analysis are essential to ensure that the calibration technique is robust and works properly over all possible PLL operating conditions.

V. APPENDIX: PROOF OF THEOREMS 1-4

Proof of Theorem 1: Replacing n with $n-1$ in (9) gives

$$z_k[n] = z_k[n-1] + K s_k[n-1] r_{ideal}[n-1] - K s_k[n-1] \sum_{l=1}^L s_l[n-1] z_l[n-1] \quad (19)$$

for all n . Recursively substituting (19) into itself shows that $z_k[n]$ is a function, $f_{k,n}$, of only the variables $r_{ideal}[m]$ and $s_j[m]$ for $j = 1, 2, \dots, L$ and $m \leq n-1$, i.e.,

$$z_k[n] = f_{k,n}(r_{ideal}[m], s_j[m]; \quad j = 1, 2, \dots, L, m \leq n-1). \quad (20)$$

Given that $s_k[m]$ for all m is restricted to values of $-1, 0$, and 1 , it follows that $s_k^2[m] = 1$ when $s_k[m] \neq 0$. Hence, (9) can be written as

$$z_k[n+1] = z_k[n](1-K) + K s_k[n] r_{ideal}[n] - K s_k[n] \sum_{l=1, l \neq k}^L s_l[n] z_l[n] \quad (21)$$

whenever $s_k[n] \neq 0$.

The theorem hypothesis states that neither $s_j[m]$ nor $r_{ideal}[m]$ depend on whether $s_k[n]$ is zero or nonzero for any j, k , and $n > m$, so (20) implies that

$$E\{z_k[n] | s_k[n] \neq 0\} = E\{z_k[n]\} = \bar{z}_k[n], \quad (22)$$

and

$$E\{z_k[n] | s_k[n] = 0\} = E\{z_k[n]\} = \bar{z}_k[n]. \quad (23)$$

For white switching sequences, when $s_k[n] \neq 0$ it is independent of all other random variables in the system, and (20) implies that $z_k[n]$ is not a function of $s_k[n]$, so (21) and (22) imply

$$E\{z_k[n+1] | s_k[n] \neq 0\} = \bar{z}_k[n](1-K). \quad (24)$$

It follows from (9) that $z_k[n+1] = z_k[n]$ whenever $s_k[n] = 0$. This with (23) implies

$$E\{z_k[n+1] | s_k[n] = 0\} = \bar{z}_k[n]. \quad (25)$$

Combining (24) and (25) yields (12).

By definition, $s_k^2[n]$ is restricted to values of 1 and 0 , so the condition that $c_k = E\{s_k^2[n]\}$ is independent of n for $n \geq 0$ implies that

$$\Pr\{s_k[n] \neq 0\} = c_k \text{ and } \Pr\{s_k[n] = 0\} = 1 - c_k, \quad (26)$$

where $\Pr\{A\}$ denotes the probability of event A . The properties of conditional expectations imply

$$\begin{aligned} E\{z_k[n+1]\} &= E\{z_k[n+1] | s_k[n] = 0\} \Pr\{s_k[n] = 0\} \\ &\quad + E\{z_k[n+1] | s_k[n] \neq 0\} \Pr\{s_k[n] \neq 0\}, \end{aligned} \quad (27)$$

so it follows from (12) and (26), that

$$\bar{z}_k[n+1] = \bar{z}_k[n](1 - c_k K). \quad (28)$$

If (13) holds for any particular value of n , it follows from substituting (13) into (28) that it must hold for $n+1$. It follows from (8) that if $a_k[0] = 0$, then $\bar{z}_k[0] = -b_k$ so (13) holds for $n = 0$. Therefore, (13) must hold for all $n = 0, 1, 2, \dots$ by mathematical induction. $\square$

Proof of Theorem 2: The same reasoning that led to (22) and (23) implies

$$E\{z_k^2[n] | s_k[n] \neq 0\} = E\{z_k^2[n] | s_k[n] = 0\} = \bar{z}_k^2[n] \quad (29)$$

for all k and n . Given that $z_k[n+1] = z_k[n]$ whenever $s_k[n] = 0$, it follows from (29) that

$$E\{z_k^2[n+1] | s_k[n] = 0\} = \bar{z}_k^2[n]. \quad (30)$$

This with (26) implies

$$\bar{z}_k^2[n+1] = \bar{z}_k^2[n](1 - c_k) + E\{z_k^2[n+1] | s_k[n] \neq 0\} c_k. \quad (31)$$

Squaring both sides of (9) yields

$$\begin{aligned} z_k^2[n+1] &= z_k^2[n] + K^2 s_k^2[n] r_{ideal}^2[n] + 2K s_k[n] r_{ideal}[n] z_k[n] \\ &\quad + K^2 s_k^2[n] \sum_{l=1}^L \sum_{j=1}^L s_j[n] z_j[n] s_l[n] z_l[n] \\ &\quad - 2(z_k[n] + K s_k[n] r_{ideal}[n]) K s_k[n] \sum_{l=1, l \neq k}^L s_l[n] z_l[n]. \end{aligned} \quad (32)$$

As indicated by (20), $z_i[n]$ does not depend on $s_j[n]$ for any i, j , and n , and by the definition of white switching sequences, for each integer, i , $s_i[n]$ either equals zero or it has zero mean and is independent of all other random variables in the system. Therefore, taking the expectation of (32) conditioned on $s_k[n] \neq 0$ and applying (29) yields

$$\begin{aligned} E\{z_k^2[n+1] | s_k[n] \neq 0\} &= \bar{z}_k^2[n](1 - 2K) + K^2 \sigma_{r_{ideal}}^2 \\ &\quad + K^2 \sum_{l=1}^L \bar{z}_l^2[n] E\{s_l^2[n] | s_k[n] \neq 0\}. \end{aligned} \quad (33)$$

Given that $0 \leq s_l^2[n] \leq 1$, this implies

$$E\{z_k^2[n+1] | s_k[n] \neq 0\} \leq \bar{z}_k^2[n](1 - 2K) + K^2 \sigma_{r_{ideal}}^2$$

$$+K^2 \sum_{l=1}^L \overline{z_l^2}[n]. \quad (34)$$

It follows from (15), (26), (31), and (34) that

$$\overline{z_k^2}[n+1] \leq \overline{z_k^2}[n] (1 - 2c_k K) + c_k K^2 \sigma_{r_{ideal}}^2 + c_k K^2 L \sigma_z^2[n]. \quad (35)$$

This with (15) implies

$$\sigma_z^2[n+1] \leq \sigma_z^2[n] (1 - 2K c_{min} + cLK^2) + K^2 c \sigma_{r_{ideal}}^2. \quad (36)$$

Hence, $\sigma_z^2[n] \leq y[n]$, where $y[n]$ satisfies the constant-coefficient linear difference equation

$$y[n+1] = y[n] (1 - 2K c_{min} + cLK^2) + x[n], \quad (37)$$

with

$$x[n] = cK^2 \sigma_{r_{ideal}}^2 u[n], \quad (38)$$

and $u[n]$ is the unit step function. This implies that $y[n]$ can be viewed as the output of a linear time-invariant (LTI) system with input sequence $x[n]$, where $x[n]$ is a step function.

Solving the z -transform of (37) for the transfer function from $x[n]$ to $y[n]$ yields

$$B(z) = \frac{z^{-1}}{1 - z^{-1} (1 - 2K c_{min} + cLK^2)}, \quad (39)$$

where $B(z) = Y(z)/X(z)$, and $Y(z)$ and $X(z)$ are the z -transforms of $y[n]$ and $x[n]$, respectively. The condition $0 < K < 2c_{min}/(cL)$ implies that $|1 - 2K c_{min} + cLK^2| < 1$, so the LTI system is stable. As $x[n]$ is a step function, the properties of stable LTI systems imply that the limit of $y[n]$ as $n \rightarrow \infty$ is the zero-frequency gain of $B(z)$ times the amplitude of $x[n]$, i.e.,

$$\lim_{n \rightarrow \infty} \{y[n]\} = cK^2 \sigma_{r_{ideal}}^2 B(z^0) = \frac{K c \sigma_{r_{ideal}}^2}{2c_{min} - K c L}. \quad (40)$$

Given that $\sigma_z^2[n] \leq y[n]$ for all n , it follows that the upper bound of $\sigma_z^2[n]$ in the limit as $n \rightarrow \infty$, i.e., the limit supremum of $\sigma_z^2[n]$, is bounded by the right side of (40). $\square$

Lemma 1: For first-order highpass shaped switching sequences, if $l \neq k$, m is an even integer, and $0 < K < 1/L$, then

$$|E\{s_k[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| \leq K |E\{z_k[J_{m-1,k}]\}|. \quad (41)$$

Proof: Equation (41) holds by inspection if $s_l[J_{m,k}] = 0$. Therefore, it remains to show that (41) holds when $J_{m,k} = J_{p,l}$ for some integer p . By definition, $s_l[J_{p,l}]$ has zero mean and, if p is odd, is independent of all other contemporaneous and prior random variables in the system. It follows that (41) holds if p is odd. It remains to show that (41) holds when $J_{m,k} = J_{p,l}$ and p is even, so for the remainder of the proof suppose that $J_{m,k} = J_{p,l}$ and p is even.

As $z_l[n] = z_l[n-1]$ whenever $s_l[n-1] = 0$, it follows from (19) and the definition of $J_{p,l}$ that

$$z_l[J_{p,l}] = (1 - K)z_l[J_{p-1,l}] + K s_l[J_{p-1,l}]r_{ideal}[J_{p-1,l}] - K s_l[J_{p-1,l}] \sum_{\substack{j=1 \\ j \neq l}}^L s_j[J_{p-1,l}]z_j[J_{p-1,l}]. \quad (42)$$

By definition, $s_k[n]$ has zero mean for all n and k , the nonzero values of $s_k[n]$ are independent of $r_{ideal}[n']$ for all k and n' , and $s_l[J_{p,l}]s_l[J_{p-1,l}] = -1$. Hence, (42) implies

$$\begin{aligned} & E\{s_k[J_{m,k}]s_l[J_{p,l}]z_l[J_{p,l}]\} \\ &= (1 - K)E\{s_k[J_{m,k}]s_l[J_{p,l}]z_l[J_{p-1,l}]\} \\ &\quad + KE\{s_k[J_{m,k}]s_k[J_{p-1,l}]z_k[J_{p-1,l}]\} \\ &\quad + K \sum_{\substack{j=1 \\ j \neq l,k}}^L E\{s_k[J_{m,k}]s_j[J_{p-1,l}]z_j[J_{p-1,l}]\}. \end{aligned} \quad (43)$$

Equation (20) implies that $z_i[J_{p-1,l}]$ does not depend on $s_l[J_{p,l}]$ or $s_l[J_{p-1,l}]$ for any i , and, by definition, $s_l[J_{p,l}]$ has zero mean and is independent of all other contemporaneous and prior random variables in the system except $s_l[J_{p-1,l}]$. Hence, the first term on the right side of (43) is zero. Nearly identical reasoning implies that if $J_{m-1,k} > J_{p-1,l}$ then all the remaining terms on the right side of (43) are also zero, and if $J_{m-1,k} = J_{p-1,l}$ then all the remaining terms except the second term on the right side of (43) are zero. Furthermore, given that $J_{m,k} = J_{p,l}$ and $s_k[J_{p,l}] = -s_k[J_{p-1,l}]$, if $J_{m-1,k} = J_{p-1,l}$ then the second term on the right side of (43) is $-KE\{z_k[J_{m-1,k}]\}$ if $J_{m-1,k} = J_{p-1,l}$. Hence, the lemma holds when $J_{m-1,k} \geq J_{p-1,l}$.

For the remainder of the proof, suppose that $J_{m-1,k} < J_{p-1,l}$. As explained above, the first term on the right side of (43) is zero. The second term on the right side of (43) is also zero because $J_{m-1,k} \neq J_{p-1,l}$ implies that $s_k[J_{p-1,l}] = 0$. By the same reasoning followed in the first paragraph of the proof, the j th term in the summation on the right side of (43), i.e., is zero unless $J_{p-1,l} = J_{q,j}$ for some even integer q . If $J_{p-1,l} = J_{q,j}$ for some even integer q , then

$$\begin{aligned} & E\{s_k[J_{m,k}]s_j[J_{p-1,l}]z_j[J_{p-1,l}]\} \\ &= E\{s_k[J_{m,k}]s_j[J_{q,j}]z_j[J_{q,j}]\}. \end{aligned} \quad (44)$$

The right side of (44) has the same form as the left side of (41) but with different indices. Therefore, the results of the proof so far imply that the right side of (44) is either zero, $K|E\{z_k[J_{m-1,k}]\}|$, or

$$K \sum_{\substack{u=1 \\ u \neq j,k}}^L E\{s_k[J_{m,k}]s_u[J_{q-1,j}]z_u[J_{q-1,j}]\}. \quad (45)$$

As $0 < K < 1/L$, it follows that $K(L-2) < 1$. Therefore, (41) holds if the right side of (44) is either 0 or $-KE\{z_k[J_{m-1,k}]\}$ for each j not equal to l or k . The above reasoning can be applied recursively in each case where the right side of (44) is neither 0 nor $-KE\{z_k[J_{m-1,k}]\}$ for any j not equal to l or k . At each recursion step, each potentially non-zero term in the sum corresponding to (45) has the form of the right side of (44), but with a reduced value of $J_{q,j}$. After a finite number of conversion steps, $J_{q,j}$ for each term in the sum becomes small enough that $J_{m-1,k} \geq J_{q-1,j}$, in which case the term is either 0 or $-KE\{z_k[J_{m-1,k}]\}$. $\square$

Proof of Theorem 3: By definition, $s_k[n]$ has zero mean for all n and k , and the nonzero values of $s_k[n]$ are independent of $r_{ideal}[n']$ for all k and n' . Consequently, (21) implies

$$\begin{aligned} & E\{z_k[n+1]s_k[n] \neq 0\} \\ &= (1 - K)E\{z_k[n]s_k[n] \neq 0\} \end{aligned}$$

$$-K \sum_{\substack{l=1 \\ l \neq k}}^L E \{s_k[n] s_l[n] z_l[n] | s_k[n] \neq 0\}. \quad (46)$$

By definition, $s_k[n] \neq 0$ for $n \geq 0$ if and only $n = J_{m,k}$ for some integer m . Hence, (46) can be written as

$$E\{z_k[J_{m,k} + 1]\} = (1 - K) E\{z_k[J_{m,k}]\} - K \sum_{\substack{l=1 \\ l \neq k}}^L E\{s_k[J_{m,k}] s_l[J_{m,k}] z_l[J_{m,k}]\}. \quad (47)$$

The right-most equation in (4) implies that for each odd value of m , $s_k[J_{m,k}]$ is independent of all other contemporaneous and prior random variables in the system, so (47) implies

$$E\{z_k[J_{m,k} + 1]\} = (1 - K) E\{z_k[J_{m,k}]\} \quad (48)$$

for odd values of m . For even values of m , (48) does not hold as $z_l[J_{m,k}]$ depends on $s_k[J_{m-1,k}]$ and $s_k[J_{m,k}] = -s_k[J_{m-1,k}]$ if m is even.

Applying Lemma 1 and the triangle inequality to (47) gives

$$|E\{z_k[J_{m,k} + 1]\}| \leq |E\{z_k[J_{m,k}]\}| (1 - K) + (L - 1) K^2 |E\{z_k[J_{m-1,k}]\}|. \quad (49)$$

Given that $m - 1$ is odd when m is even, it follows from (48) that when m is even $|E\{z_j[J_{m-1,k}]\}| \leq |E\{z_j[J_{m,k}]\}|/(1 - K)$. Therefore, (49) can be written as

$$|E\{z_k[J_{m,k} + 1]\}| \leq |E\{z_k[J_{m,k}]\}| \left(1 - K \left(\frac{1 - LK}{1 - K}\right)\right). \quad (50)$$

This inequality holds for even values of m , but given that $0 < K < 1/L$ and, by definition, $L \geq 1$, it is a less restrictive inequality than (48), so it must also hold for odd values of m .

As mentioned above, $s_k[n] \neq 0$ for $n \geq 0$ if and only if $n = J_{m,k}$ for some value of $m \geq 1$. Hence, (50) implies

$$|E\{z_k[n + 1] | s_k[n] \neq 0\}| \leq |E\{z_k[n] | s_k[n] \neq 0\}| \left(1 - K \left(\frac{1 - LK}{1 - K}\right)\right). \quad (51)$$

This with (22), (23), (25) and (26) implies

$$|E\{z_k[n + 1] | s_k[n] \neq 0\}| c_k + |E\{z_k[n + 1] | s_k[n] = 0\}| (1 - c_k) \leq |\bar{z}_k[n]| \left(1 - c_k K \left(\frac{1 - LK}{1 - K}\right)\right). \quad (52)$$

Given that c_k and $1 - c_k$ are both non-negative, it follows from the definition of $\bar{z}_k[n]$, (27), the triangle inequality, and (52) that

$$|\bar{z}_k[n + 1]| \leq |\bar{z}_k[n]| \left(1 - c_k K \left(\frac{1 - LK}{1 - K}\right)\right). \quad (53)$$

If (16) holds for any particular value of n , it follows from substituting (16) into (53) that it must hold for $n + 1$. It follows from (8) that if $a_k[0] = 0$, then $\bar{z}_k[0] = -b_k$ so (16) holds for $n = 0$. Therefore, (16) must hold for all $n = 0, 1, 2, \dots$ by mathematical induction. $\square$

Lemma 2: For first-order highpass shaped switching sequences, if $K < 1/L$ and $E\{r_{ideal}^2[n']\}$ does not depend on n , then for any n'

$$|E\{r_{ideal}[n'] s_i[n] z_i[n]\}| < \frac{K \sigma_{ideal}^2}{1 - KL}, \quad (54)$$

when $n = J_{m,i}$ and m is a non-negative even integer, and

$$E\{r_{ideal}[n'] s_i[n] z_i[n]\} = 0, \quad (55)$$

otherwise.

Proof: If $s_i[n] = 0$, then (55) holds by inspection. Otherwise, $n = J_{m,i}$ for some non-negative integer m . If m is odd, then by the definition of the first order highpass shaped switching sequences, $s_i[J_{m,i}]$ is independent of all other contemporaneous and prior random variables in the system and it has zero mean, so (55) holds in case too. Therefore, (55) holds unless $n = J_{m,i}$ and m is a non-negative even integer, so to show that (54) holds, it is sufficient to evaluate $E\{r_{ideal}[n'] s_i[J_{m,i}] z_i[J_{m,i}]\}$ for the case where m is even.

For the remainder of the proof, suppose that m is even. Equation (42) holds with p replaced by m and l replaced by i , and the definition of the first-order highpass shaped switching sequences implies that $s_i[J_{m,i}] s_i[J_{m-1,i}] = -1$ when m is even, so

$$\begin{aligned} s_i[J_{m,i}] z_i[J_{m,i}] &= s_i[J_{m,i}] z_i[J_{m-1,i}] (1 - K) - K r_{ideal}[J_{m-1,i}] \\ &\quad + K \sum_{\substack{l=1 \\ l \neq i}}^L s_l[J_{m-1,i}] z_l[J_{m-1,i}]. \end{aligned} \quad (56)$$

By definition, $s_i[J_{m,i}]$ has zero mean and is independent of all other contemporaneous and prior random variables in the system except $s_i[J_{m-1,i}]$, and (20) implies that $z_i[J_{m-1,i}]$ does not depend on either $s_i[J_{m,i}]$ or $s_i[J_{m-1,i}]$, so it follows from (56) that

$$\begin{aligned} E\{r_{ideal}[n'] s_i[J_{m,i}] z_i[J_{m,i}]\} &= -K E\{r_{ideal}[n'] r_{ideal}[J_{m-1,i}]\} \\ &\quad + K \sum_{\substack{l=1 \\ l \neq i}}^L E\{r_{ideal}[n'] s_l[J_{m-1,i}] z_l[J_{m-1,i}]\}. \end{aligned} \quad (57)$$

The Cauchy-Schwarz inequality for random variables implies that

$$|E\{r_{ideal}[p] r_{ideal}[q]\}| \leq \sigma_{ideal}^2 \quad (58)$$

for any integers p and q [38]. This with (57) and the triangle inequality implies that

$$\begin{aligned} |E\{r_{ideal}[n'] s_i[J_{m,i}] z_i[J_{m,i}]\}| &\leq K \sigma_{ideal}^2 + K \sum_{\substack{l=1 \\ l \neq i}}^L |E\{r_{ideal}[n'] s_l[J_{m-1,i}] z_l[J_{m-1,i}]\}|. \end{aligned} \quad (59)$$

Hence,

$$|E\{r_{ideal}[n'] s_i[J_{m,i}] z_i[J_{m,i}]\}| \leq K \sigma_{ideal}^2 + K (L - 1) A_1. \quad (60)$$

where

$$A_1 = \max_{i, l \neq i} \{ |E\{r_{ideal}[n'] s_l[J_{m-1,i}] z_l[J_{m-1,i}]\}| \}. \quad (61)$$

As A_1 has the same form as the expectation in the lemma statement, the results of the proof so far apply to it. Substituting (59) with a change of variables into (61) and substituting the result into (60) yields

$$\begin{aligned} |E\{r_{ideal}[n'] s_i[J_{m,i}] z_i[J_{m,i}]\}| &\leq K \sigma_{ideal}^2 + K^2 (L - 1) \sigma_{ideal}^2 \\ &\quad + K^2 (L - 1)^2 A_2, \end{aligned} \quad (62)$$

where A_2 , like A_1 , has the same form as the expectation in the lemma statement. Recursively repeating this process $N - 2$ more times yields

$$|E\{r_{ideal}[n']s_i[J_{m,i}]z_i[J_{m,i}]\}| \leq \sigma_{ideal}^2 K \sum_{r=0}^{N-1} (K(L-1))^r + K^N (L-1)^N A_N, \quad (63)$$

where A_N has the same form as the expectation in the lemma statement. For a sufficiently large value of N , $A_N = |E\{r_{ideal}[n']s_u[n'']z_u[n'']\}|$ where $n'' < 0$ in which case it follows from (11) that $A_N = 0$. Therefore,

$$|E\{r_{ideal}[n']s_i[J_{m,i}]z_i[J_{m,i}]\}| < \sigma_{ideal}^2 K \sum_{r=0}^{\infty} (K(L-1))^r, \quad (64)$$

which implies (54) (in which L has been used in place of $L - 1$ to simplify the expression at the expense of a slightly looser bound). $\square$

Lemma 3: For first-order highpass shaped switching sequences, if p is even and $K < \frac{1}{2}$ then

$$E\{z_i^2[J_{p-1,i}]\} \leq \frac{1}{1-2K} E\{z_i^2[J_{p,i}]\}. \quad (65)$$

Proof: Equation (42) can be rewritten as

$$z_i[J_{p,i}] = z_i[J_{p-1,i}] + A, \quad (66)$$

where

$$A = K s_i[J_{p-1,i}] \left(r_{ideal}[J_{p-1,i}] - \sum_{j=1}^L s_j[J_{p-1,i}] z_j[J_{p-1,i}] \right). \quad (67)$$

Therefore,

$$z_i^2[J_{p,i}] = z_i^2[J_{p-1,i}] + A^2 + 2A z_i[J_{p-1,i}], \quad (68)$$

so

$$E\{z_i^2[J_{p,i}]\} \geq E\{z_i^2[J_{p-1,i}]\} + 2E\{A z_i[J_{p-1,i}]\}. \quad (69)$$

Given that p is even, $p - 1$ is odd, so $s_i[J_{p-1,i}]$, which has zero mean, is independent of all other contemporaneous and prior random variables in the system. Consequently,

$$2E\{A z_i[J_{p-1,i}]\} = -2KE\{z_i^2[J_{p-1,i}]\}. \quad (70)$$

This with (69) and $K < \frac{1}{2}$ implies (65). $\square$

Lemma 4: For first-order highpass shaped switching sequences and any integers $j \neq l$ and $m \geq 1$, if $K < \min\{1/L, \frac{1}{2}\}$ and $E\{r_{ideal}^2[n]\}$ does not depend on n , then

$$\begin{aligned} & |E\{s_j[J_{m,k}]z_j[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| \\ & < K^2 \left(\frac{1+KL}{1-KL} \right) \sigma_{ideal}^2 \\ & + \frac{K(1-K)L}{1-2K} \left(E\{z_j^2[J_{m,k}]\} + E\{z_l^2[J_{m,k}]\} \right) \\ & + \frac{2K(1+KL)}{1-2K} \sum_{r=1}^L E\{z_r^2[J_{m,k}]\}. \end{aligned} \quad (71)$$

Proof: By definition, $s_i[J_{r,i}]$ for each nonnegative odd value of r has zero mean and is independent of all other contemporaneous and prior random variables in the system, and (20) implies that $z_{r'}[J_{r,i}]$ does not depend on $s_i[J_{r,i}]$ for any i and r' . Therefore,

$$E\{s_j[J_{m,k}]z_j[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\} = 0 \quad (72)$$

unless $J_{m,k} = J_{p,j}$ and $J_{m,k} = J_{q,l}$, where p and q are even integers. For the remainder of the proof, suppose $J_{m,k} = J_{p,j}$ and $J_{m,k} = J_{q,l}$, where p and q are even integers, so

$$\begin{aligned} & E\{s_j[J_{m,k}]z_j[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\} \\ & = E\{s_j[J_{p,j}]z_j[J_{p,j}]s_l[J_{q,l}]z_l[J_{q,l}]\}, \end{aligned} \quad (73)$$

and (56) with a change of indices implies

$$\begin{aligned} & s_j[J_{p,j}]z_j[J_{p,j}] \\ & = s_j[J_{p,j}]z_j[J_{p-1,j}](1-K) \\ & - K \left(r_{ideal}[J_{p-1,j}] - \sum_{\substack{r'=1 \\ r' \neq j}}^L s_{r'}[J_{p-1,j}]z_{r'}[J_{p-1,j}] \right) \end{aligned} \quad (74)$$

and

$$\begin{aligned} & s_l[J_{q,l}]z_l[J_{q,l}] \\ & = s_l[J_{q,l}]z_l[J_{q-1,l}](1-K) \\ & - K \left(r_{ideal}[J_{q-1,l}] - \sum_{\substack{r'=1 \\ r' \neq l}}^L s_{r'}[J_{q-1,l}]z_{r'}[J_{q-1,l}] \right). \end{aligned} \quad (75)$$

The remainder of the proof bounds the expectation of the product of (74) and (75) by bounding the magnitudes of the expectations of the various product terms individually and applying the triangle inequality.

The expectations of several of the product terms contain r_{ideal} . Given that $J_{p,j} = J_{q,l}$ and $J_{q,l} > J_{q-1,l}$, the definition of the switching sequences and (20) imply that $s_j[J_{p,j}]$ is independent of $z_j[J_{p-1,j}]r_{ideal}[J_{q-1,l}]$, so

$$E\{K(1-K)s_j[J_{p,j}]z_l[J_{p-1,j}]r_{ideal}[J_{q-1,l}]\} = 0. \quad (76)$$

The same reasoning further implies that

$$E\{K(1-K)s_l[J_{q,l}]z_l[J_{q-1,l}]r_{ideal}[J_{p-1,j}]\} = 0. \quad (77)$$

Lemma 2, inequality (58), and the triangle inequality imply that

$$|R| \leq K^2 \left(1 + \frac{2K(L-1)}{1-KL} \right) \sigma_{ideal}^2 < K^2 \left(\frac{1+KL}{1-KL} \right) \sigma_{ideal}^2, \quad (78)$$

where R is the expectation of the remaining product terms that contain r_{ideal} .

Equation (20) implies that $z_i[n]$ is not a function of $s_k[n']$ for any i, k , and $n' > n-1$. Given that $j \neq l$, if $J_{p-1,j} > J_{q-1,l}$ then $s_j[J_{p,j}]$ is independent of $z_j[J_{p-1,j}]s_l[J_{q,l}]z_l[J_{q-1,l}]$. Otherwise, $s_l[J_{q,l}]$ is independent of $z_l[J_{q-1,l}]s_j[J_{p,j}]z_j[J_{p-1,j}]$. Hence,

$$E\{s_j[J_{p,j}]z_j[J_{p-1,j}]s_l[J_{q,l}]z_l[J_{q-1,l}]\} = 0. \quad (79)$$

Given that the switching sequences are bounded in magnitude by 1, each of the terms of the product of (74) and (75) that have yet to be considered satisfy

$$K(1-K) |E\{s_j[J_{p,j}]z_j[J_{p-1,j}]s_r[J_{q-1,l}]z_r[J_{q-1,l}]\}| \leq K(1-K) E\{|z_j[J_{p-1,j}]z_r[J_{q-1,l}]\}|, \quad (80)$$

$$K(1-K) |E\{s_l[J_{q,l}]z_l[J_{q-1,l}]s_{r'}[J_{p-1,j}]z_{r'}[J_{p-1,j}]\}| \leq K(1-K) E\{|z_l[J_{q-1,j}]z_{r'}[J_{p-1,j}]\}|, \quad (81)$$

or

$$K^2 |E\{s_{r'}[J_{p-1,j}]z_{r'}[J_{p-1,j}]s_r[J_{q-1,l}]z_r[J_{q-1,l}]\}| \leq K^2 E\{|z_{r'}[J_{p-1,j}]z_r[J_{q-1,l}]\}|. \quad (82)$$

The Cauchy-Schwarz inequality implies that

$$E\{|z_u[J_{p-1,j}]z_w[J_{q-1,l}]\}| \leq \sqrt{E\{z_u^2[J_{p-1,j}]\}} \sqrt{E\{z_w^2[J_{q-1,l}]\}}, \quad (83)$$

for any u and w . This with Lemma 3 implies

$$E\{|z_u[J_{p-1,j}]z_w[J_{q-1,l}]\}| \leq \frac{1}{1-2K} \sqrt{E\{z_u^2[J_{p,j}]\}} \sqrt{E\{z_w^2[J_{q,l}]\}}. \quad (84)$$

For any non-negative real numbers a and b ,

$$\sqrt{a}\sqrt{b} \leq \left(\max\{\sqrt{a}, \sqrt{b}\}\right)^2 = \max\{a, b\} \leq a + b, \quad (85)$$

so (84) implies

$$E\{|z_u[J_{p-1,j}]z_w[J_{q-1,l}]\}| \leq \frac{1}{1-2K} \left(E\{z_u^2[J_{p,j}]\} + E\{z_w^2[J_{q,l}]\}\right). \quad (86)$$

Given that $J_{m,k} = J_{p,j}$ and $J_{m,k} = J_{q,l}$, it follows from (74)-(82), (86), and the triangle inequality that

$$\begin{aligned} & |E\{s_j[J_{m,k}]z_j[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| \\ & < K^2 \left(\frac{1+KL}{1-KL}\right) \sigma_{r_{ideal}}^2 \\ & + \frac{K(1-K)}{1-2K} \sum_{\substack{r=1 \\ r \neq l}}^L \left(E\{z_j^2[J_{m,k}]\} + E\{z_r^2[J_{m,k}]\}\right) \\ & + \frac{K(1-K)}{1-2K} \sum_{\substack{r'=1 \\ r' \neq j}}^L \left(E\{z_l^2[J_{m,k}]\} + E\{z_{r'}^2[J_{m,k}]\}\right) \\ & + \frac{K^2}{1-2K} \sum_{\substack{r=1 \\ r \neq l}}^L \sum_{\substack{r'=1 \\ r' \neq j}}^L \left(E\{z_{r'}^2[J_{m,k}]\} + E\{z_r^2[J_{m,k}]\}\right), \end{aligned} \quad (87)$$

which implies (71) (wherein L has been used in place of $L-1$ for simplicity at the expense of a slightly looser bound). $\square$

Proof of Theorem 4: Equations (26) hold by the same argument made in the proof of Theorem 1, and (29), (30), (31), and (32) hold by the same arguments made in the proof of Theorem 2.

By definition, $n = J_{m,k}$ for some integer m if and only if $s_k[n] \neq 0$, so

$$\Pr\{s_k[n] \neq 0\} = d_{n,k} + \Pr\{n = J_{m,k} \text{ for some odd } m\}, \quad (88)$$

where

$$d_{n,k} = \Pr\{n = J_{m,k} \text{ for some even } m\}. \quad (89)$$

This and (26) imply that

$$0 \leq d_{n,k} \leq c_k, \quad (90)$$

and

$$\Pr\{n = J_{m,k} \text{ for some odd } m\} = c_k - d_{n,k}. \quad (91)$$

Therefore, it follows from (31) that

$$\begin{aligned} & \overline{z_k^2}[n+1] \\ & = E\left\{z_k^2[n+1] \middle| n = J_{m,k} \text{ for some even } m\right\} d_{n,k} \\ & + E\left\{z_k^2[n+1] \middle| n = J_{m,k} \text{ for some odd } m\right\} (c_k - d_{n,k}) \\ & + \overline{z_k^2}[n] (1 - c_k). \end{aligned} \quad (92)$$

Equation (32) implies that $z_k^2[n+1] = z_k^2[n]$ whenever $s_k[n] = 0$, so it follows from (32) and the definition of $J_{m,k}$ that

$$\begin{aligned} & z_k^2[J_{m,k}+1] \\ & = z_k^2[J_{m,k}] + K^2 r_{ideal}^2[J_{m,k}] \\ & + 2K r_{ideal}[J_{m,k}] s_k[J_{m,k}] z_k[J_{m,k}] \\ & + K^2 \sum_{l=1}^L \sum_{j=1}^L s_j[J_{m,k}] z_j[J_{m,k}] s_l[J_{m,k}] z_l[J_{m,k}] \\ & - 2(z_k[J_{m,k}] s_k[J_{m,k}] + K r_{ideal}[J_{m,k}]) \\ & K \sum_{l=1}^L s_l[J_{m,k}] z_l[J_{m,k}], \end{aligned} \quad (93)$$

which can be rearranged as

$$\begin{aligned} & z_k^2[J_{m,k}+1] \\ & = z_k^2[J_{m,k}] (1-2K) \\ & + K^2 r_{ideal}^2[J_{m,k}] + 2K(1-K) s_k[J_{m,k}] r_{ideal}[J_{m,k}] z_k[J_{m,k}] \\ & + K^2 \sum_{l=1}^L z_l^2[J_{m,k}] \\ & + K^2 \sum_{l=1}^L \sum_{\substack{j=1 \\ j \neq l}}^L s_j[J_{m,k}] z_j[J_{m,k}] s_l[J_{m,k}] z_l[J_{m,k}] \\ & - 2K(z_k[J_{m,k}] s_k[J_{m,k}] + K r_{ideal}[J_{m,k}]) \\ & \sum_{\substack{l=1 \\ l \neq k}}^L s_l[J_{m,k}] z_l[J_{m,k}]. \end{aligned} \quad (94)$$

The next steps in the proof apply the triangle inequality to bound $E\{z_k^2[J_{m,k}+1]\}$ by summing bounds on the magnitudes of the expectations of the individual terms in (94).

Lemmas 2 and 4 (using L in place of $L-1$ for simplicity at the expense of slightly looser bounds) can be applied to bound magnitudes of the expectations of several of these terms. Lemma 2 implies that

$$\begin{aligned} & 2K(1-K) |E\{s_k[J_{m,k}] r_{ideal}[J_{m,k}] z_k[J_{m,k}]\}| \\ & \leq \begin{cases} \frac{2K^2(1-K)}{1-KL} \sigma_{r_{ideal}}^2, & \text{if } m \text{ is even,} \\ 0, & \text{if } m \text{ is odd,} \end{cases} \end{aligned} \quad (95)$$

and

$$2K^2 \sum_{\substack{l=1 \\ l \neq k}}^L |E \{r_{ideal}[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| < \frac{2K^3 L \sigma_{r_{ideal}}^2}{1 - KL}. \quad (96)$$

Lemma 4 with (15) and (29) imply

$$\begin{aligned} & K^2 \sum_{l=1}^L \sum_{j=1}^L |E \{s_j[J_{m,k}]z_j[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| \\ & < 2K^3 L^3 \left(\frac{2+KL}{1-2K} \right) \sigma_z^2[J_{m,k}] + K^4 L^2 \left(\frac{1+KL}{1-KL} \right) \sigma_{r_{ideal}}^2, \end{aligned} \quad (97)$$

and

$$\begin{aligned} & 2K \sum_{\substack{l=1 \\ l \neq k}}^L |E \{z_k[J_{m,k}]s_k[J_{m,k}]s_l[J_{m,k}]z_l[J_{m,k}]\}| \\ & < \frac{2K^2(1-K)L^2}{1-2K} \overline{z_k^2}[J_{m,k}] + 2K^2 L^2 \left(\frac{3+2KL}{1-2K} \right) \sigma_z^2[J_{m,k}] \\ & + 2K^3 L \left(\frac{1+KL}{1-KL} \right) \sigma_{r_{ideal}}^2. \end{aligned} \quad (98)$$

It follows from (15), (29), (30), (92), (94), (95), (96), (97), (98), the triangle inequality, and again using L in place of $L-1$ that

$$\begin{aligned} \overline{z_k^2}[n+1] & < \overline{z_k^2}[n] (1 - 2c_k K) + c_k \frac{2K^2(1-K)L^2}{1-2K} \overline{z_k^2}[n] \\ & + c_k K^2 L \left(1 + L \frac{6+8KL+2K^2L^2}{1-2K} \right) \sigma_z^2[n] \\ & + c_k K^2 \left(1 + KL \frac{4+3KL+K^2L^2}{1-KL} \right) \sigma_{r_{ideal}}^2 \\ & + d_{n,k} \frac{2K^2(1-K)}{1-KL} \sigma_{r_{ideal}}^2. \end{aligned} \quad (99)$$

Summing both sides of (99) over k , dividing the result by L , and applying (15) yields

$$\sigma_z^2[n+1] < a \sigma_z^2[n] + x[n] \quad (100)$$

where a is given by (17),

$$\begin{aligned} x[n] & = u[n] K^2 \left[c \left(1 + KL \frac{4+3KL+K^2L^2}{1-KL} \right) \right. \\ & \quad \left. + d_n \frac{2(1-K)}{1-KL} \right] \sigma_{r_{ideal}}^2, \end{aligned} \quad (101)$$

$$d_n = \frac{1}{L} \sum_{k=1}^L d_{n,k}, \quad (102)$$

and $u[n]$ is the unit step function (because the system is “turned on” at time $n = 0$). Hence, $\sigma_z^2[n] < y[n]$, where $y[n]$ satisfies the constant coefficient, linear difference equation

$$y[n+1] = ay[n] + x[n]. \quad (103)$$

The definitions of the first-order switching sequences, $J_{m,k}$, and c_k imply

$$\mathbf{p}_{n+1,k} = \mathbf{M}_k \mathbf{p}_{n,k}, \text{ where } \mathbf{M}_k = \begin{bmatrix} 1-c_k & c_k \\ c_k & 1-c_k \end{bmatrix}, \quad (104)$$

and

$$\mathbf{p}_{n,k} = \begin{bmatrix} \Pr \{ \text{largest } J_{m,k} \leq n \text{ has odd } m \} \\ \Pr \{ \text{largest } J_{m,k} \leq n \text{ has even } m \} \end{bmatrix}. \quad (105)$$

The matrix $\mathbf{M}_k$ is a Markov matrix and $\boldsymbol{\pi} = [\frac{1}{2} \ \frac{1}{2}]$ is a probability vector that satisfies $\boldsymbol{\pi} \mathbf{M}_k = \boldsymbol{\pi}$, so it follows that all elements of $\mathbf{M}_k^n$, and, hence, of $\mathbf{p}_{n,k}$ converge to $\frac{1}{2}$ as $n \rightarrow \infty$ [38]. Given that $d_{n,k} = c \Pr \{ \text{largest } J_{m,k} \leq n \text{ has odd } m \}$, it follows from (102) that

$$\lim_{n \rightarrow \infty} d_n = \frac{c}{2} \quad (106)$$

Solving the z -transform of (103) for the transfer function from $x[n]$ to $y[n]$, yields

$$B(z) = \frac{z^{-1}}{1 - az^{-1}}, \quad (107)$$

where $B(z) = Y(z)/X(z)$, and $Y(z)$ and $X(z)$ are the z -transforms of $y[n]$ and $x[n]$, respectively. The properties of stable LTI systems imply that

$$\lim_{n \rightarrow \infty} \{y[n]\} = B(z^0) \lim_{n \rightarrow \infty} x[n]. \quad (108)$$

The limit supremum of $\sigma_z^2[n]$ as $n \rightarrow \infty$ is less than or equal to (108) because $\sigma_z^2[n] < y[n]$, so (18) follows from (17), (101), (106), (107) and (108). $\square$

REFERENCES

- [1] A. Swaminathan, A. Panigada, E. Masry, and I. Galton, “A digital requantizer with shaped requantization noise that remains well behaved after nonlinear distortion,” *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5382–5394, Nov. 2007.
- [2] E. Familier and I. Galton, “A fundamental limitation of DC-free quantization noise with respect to nonlinearity-induced spurious tones,” *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4172–4180, Aug. 2013.
- [3] M. Zanuso, S. Levantino, C. Samori, and A. Lacaita, “A 3 MHz-BW 3.6 GHz digital fractional- N PLL with sub-gate-delay TDC, phase-interpolation divider, and digital mismatch cancellation,” in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2010, pp. 476–477.
- [4] D. Tasca, M. Zanuso, G. Marzin, S. Levantino, C. Samori, and A. Lacaita, “A 2.9–4.0-GHz fractional- N digital PLL with bang-bang phase detector and 560-fs_{rms} integrated jitter at 4.5-mW power,” *IEEE J. Solid-State Circuits*, vol. 46, no. 12, pp. 2745–2758, Dec. 2011.
- [5] N. Pavlovic and J. Bergervoet, “A 5.3 GHz digital-to-time-converter-based fractional- N all-digital PLL,” in *Proc. IEEE Int. Solid-State Circuits Conf.*, Feb. 2011, pp. 54–56.
- [6] W. Wu et al., “A 14-nm ultra-low jitter fractional- N PLL using a DTC range reduction technique and a reconfigurable dual-core VCO,” *IEEE J. Solid-State Circuits*, vol. 56, no. 12, pp. 3756–3767, Dec. 2021.
- [7] W. Wu et al., “A 28-nm 75-fs_{rms} analog fractional- N sampling PLL with a highly linear DTC incorporating background DTC gain calibration and reference clock duty cycle correction,” *IEEE J. Solid-State Circuits*, vol. 54, no. 5, pp. 1254–1265, May 2019.
- [8] A. Santiccioli et al., “A 66-fs-rms jitter 12.8-to-15.2-GHz fractional- N bang-bang PLL with digital frequency-error recovery for fast locking,” *IEEE J. Solid-State Circuits*, vol. 55, no. 12, pp. 3349–3361, Dec. 2020.
- [9] S. Levantino, G. Marucci, G. Marzin, A. Fenaroli, C. Samori, and A. L. Lacaita, “A 1.7 GHz fractional- N frequency synthesizer based on a multiplying delay-locked loop,” *IEEE J. Solid-State Circuits*, vol. 50, no. 11, pp. 2678–2691, Nov. 2015.
- [10] A. T. Narayanan et al., “A fractional- N sub-sampling PLL using a pipelined phase-interpolator with an FOM of -250 dB,” *IEEE J. Solid-State Circuits*, vol. 51, no. 7, pp. 1630–1640, Jul. 2016.
- [11] S. Levantino, G. Marzin, and C. Samori, “An adaptive pre-distortion technique to mitigate the DTC nonlinearity in digital PLLs,” *IEEE J. Solid-State Circuits*, vol. 49, no. 8, pp. 1762–1772, Aug. 2014.

- [12] N. Markulic et al., "A DTC-based subsampling PLL capable of self-calibrated fractional synthesis and two-point modulation," *IEEE J. Solid-State Circuits*, vol. 51, no. 12, pp. 3078–3092, Dec. 2016.
- [13] B. Liu et al., "A fully synthesizable fractional- N MDLL with zero-order interpolation-based DTC nonlinearity calibration and two-step hybrid phase offset calibration," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 2, pp. 603–616, Feb. 2021.
- [14] R. B. Staszewski et al., "All-digital PLL and transmitter for mobile phones," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2469–2482, Dec. 2005.
- [15] R. Tonietto, E. Zuffetti, R. Castello, and I. Bietti, "A 3 MHz bandwidth low noise RF all digital PLL with 12 ps resolution time to digital converter," in *Proc. 32nd Eur. Solid-State Circuits Conf.*, Sep. 2006, pp. 150–153.
- [16] C. Weltin-Wu, E. Temporiti, D. Baldi, and F. Svelto, "A 3 GHz fractional- N all-digital PLL with precise time-to-digital converter calibration and mismatch correction," in *Proc. IEEE Int. Solid-State Circuits Conf.-Dig. Tech. Papers*, Feb. 2008, pp. 344–618.
- [17] C. Hsu, M. Z. Straayer, and M. H. Perrott, "A low-noise wide-BW 3.6-GHz digital $\Delta\Sigma$ fractional- N frequency synthesizer with a noise-shaping time-to-digital converter and quantization noise cancellation," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2776–2786, Dec. 2008.
- [18] V. Kratyuk, P. K. Hanumolu, K. Ok, U.-K. Moon, and K. Mayaram, "A digital PLL with a stochastic time-to-digital converter," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 8, pp. 1612–1621, Aug. 2009.
- [19] E. Temporiti, C. Weltin-Wu, D. Baldi, M. Cusmai, and F. Svelto, "A 3.5 GHz wideband ADPLL with fractional spur suppression through TDC dithering and feedforward compensation," *IEEE J. Solid-State Circuits*, vol. 45, no. 12, pp. 2723–2736, Dec. 2010.
- [20] C. Venerus and I. Galton, "Delta-sigma FDC based fractional- N PLLs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 5, pp. 1274–1285, May 2013.
- [21] G. Marucci, S. Levantino, P. Maffezzoni, and C. Samori, "Analysis and design of low-jitter digital bang-bang phase-locked loops," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 61, no. 1, pp. 26–36, Jan. 2014.
- [22] A. Elkholy, T. Anand, W.-S. Choi, A. Elshazly, and P. K. Hanumolu, "A 3.7 mW low-noise wide-bandwidth 4.5 GHz digital fractional- N PLL using time amplifier-based TDC," *IEEE J. Solid-State Circuits*, vol. 50, no. 4, pp. 867–881, Apr. 2015.
- [23] C. Weltin-Wu, E. Familier, and I. Galton, "A linearized model for the design of fractional- N digital PLLs based on dual-mode ring oscillator FDCs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 62, no. 8, pp. 2013–2023, Aug. 2015.
- [24] F.-W. Kuo et al., "An all-digital PLL for cellular mobile phones in 28-nm CMOS with -55 dBc fractional and -91 dBc reference spurs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 11, pp. 3756–3768, Nov. 2018.
- [25] E. Alvarez-Fontecilla, A. I. Eissa, E. Helal, C. Weltin-Wu, and I. Galton, "Delta-sigma FDC enhancements for FDC-based digital fractional- N PLLs," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 3, pp. 965–974, Mar. 2021.
- [26] E. Alvarez-Fontecilla, E. Helal, A. I. Eissa, and I. Galton, "Spectral breathing and its mitigation in digital fractional- N PLLs," *IEEE J. Solid-State Circuits*, vol. 56, no. 10, pp. 3191–3201, Oct. 2021.
- [27] N. Markulic, K. Raczkowski, P. Wambacq, and J. Craninckx, "A 10-bit, 550-fs step digital-to-time converter in 28 nm CMOS," in *Proc. 40th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2014, pp. 79–82.
- [28] W. S. Chang, P. C. Huang, and T. C. Lee, "A fractional- N divider-less phase-locked loop with a subsampling phase detector," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 2964–2975, Dec. 2014.
- [29] X. Lin and G. Zhang, "A 12 bit 100 fs resolution multi-stage digital-to-time converter with dynamic element matching," in *Proc. IEEE Int. Conf. Integr. Circuits, Technol. Appl. (ICTA)*, Nov. 2021, pp. 31–32.
- [30] I. Galton, "Why dynamic-element-matching DACs work," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 2, pp. 69–74, Feb. 2010.
- [31] K. L. Chan, N. Rakuljic, and I. Galton, "Segmented dynamic element matching for high-resolution digital-to-analog conversion," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 11, pp. 3383–3392, Dec. 2008.
- [32] K. L. Chan, J. Zhu, and I. Galton, "Dynamic element matching to prevent nonlinear distortion from pulse-shape mismatches in high-resolution DACs," *IEEE J. Solid-State Circuits*, vol. 43, no. 9, pp. 2067–2078, Sep. 2008.
- [33] D. Kong and I. Galton, "MSE analysis of a multi-loop LMS pseudo-random noise canceler for mixed-signal circuit calibration," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 67, no. 9, pp. 3084–3098, Sep. 2020.
- [34] E. Helal, E. Alvarez-Fontecilla, A. I. Eissa, and I. Galton, "A time amplifier assisted frequency-to-digital converter based digital fractional- N PLL," *IEEE J. Solid-State Circuits*, vol. 56, no. 9, pp. 2711–2723, Sep. 2021.
- [35] S. Pamarti, J. Welz, and I. Galton, "Statistics of the quantization noise in 1-bit dithered single-quantizer digital delta-sigma modulators," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 54, no. 3, pp. 492–503, Mar. 2007.
- [36] M. Cassia, P. Shah, and E. Bruun, "Analytical model and behavioral simulation approach for a $\Sigma\Delta$ fractional- N synthesizer employing a sample-and-hold element," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 11, pp. 850–859, Nov. 2003.
- [37] I. L. Syllaos, R. B. Staszewski, and P. T. Balsara, "Time-domain modeling of an RF all-digital PLL," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 6, pp. 601–605, Jun. 2008.
- [38] P. Billingsley, *Probability and Measure*. New York, NY, USA: Wiley, 1986.



Eslam Helal received the B.Sc. (Hons.) and M.Sc. degrees in electrical engineering from Ain Shams University (ASU), Cairo, Egypt, in 2014 and 2018, respectively. He is currently pursuing the Ph.D. degree with the University of California at San Diego, La Jolla, CA, USA.

From 2014 to 2018, he was an Analog/RF Design Engineer at Si-Ware Systems (now Goodix Technology), Cairo. He was also a Teaching and Research Assistant with the Department of Electronics and Communications Engineering, ASU. His research

interests include analog/mixed-signal integrated circuits and systems, frequency synthesizers, and data converters.



Amr I. Eissa received the B.Sc. (Hons.) and M.Sc. degrees in electrical engineering from Ain Shams University (ASU), Cairo, Egypt, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree in electronic circuits and systems with the University of California at San Diego, La Jolla, CA, USA.

From 2013 to 2016, he was a Teaching and Research Assistant with the Department of Electronics and Communications Engineering, ASU. His research interests include analysis and design of

analog/mixed-signal integrated circuits and systems.



Ian Galton (Fellow, IEEE) received the B.Sc. degree in electrical engineering from Brown University, Providence, RI, USA, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology in 1989 and 1992, respectively.

Since 1996, he has been a Professor of electrical engineering with the University of California, San Diego, La Jolla, CA, USA, where he teaches and conducts research in the field of mixed-signal integrated circuits and systems for communications.

His research involves invention, analysis, and integrated circuit implementation of critical communication system blocks, such as data converters and phase-locked loops.