A Time Amplifier Assisted Frequency-to-Digital Converter Based Digital Fractional-N PLL

Eslam Helal, Enrique Alvarez-Fontecilla, Amr I. Eissa, and Ian Galton, Fellow, IEEE

Abstract—This article presents a wide input-range delay chain based time amplifier (TA) and its application to a 6.5-GHz digital fractional-N phase-locked loop (PLL). The TA includes a delay-averaging linearity enhancement technique and the PLL is based on an improved dual-mode ring oscillator (DMRO) delta-sigma (ΔΣ) frequency-to-digital converter (FDC). The TA mitigates contributions to the PLL’s phase noise from DMRO flicker noise, which would otherwise degrade the PLL’s in-band phase noise, and from ΔΣ FDC quantization error, which would otherwise degrade the PLL’s phase noise at high bandwidth settings. This paper also presents a delay-free asynchronous DMRO phase sampling scheme, and the first experimental demonstration of a recently-proposed ΔΣ FDC digital gain calibration technique. The TA-assisted PLL achieves a random jitter of 145 fs rms, a total jitter that ranges from 151 to 270 fs rms as a result of fractional spur, and a worst-case fractional spur of −49 dBc without requiring nonlinearity calibration.

Index Terms—Averaging resistors, delta-sigma (ΔΣ) modulation, digital phase-locked loop (PLL), dual-mode ring oscillator (DMRO), frequency synthesizer, frequency-to-digital converter (FDC), gain calibration, jitter, phase sampling, time amplifier (TA).

I. INTRODUCTION

MANY types of phase-locked loops (PLLs) use a phase-frequency detector (PFD) with subsequent circuitry to measure the time differences between corresponding edges of the reference signal and a divided-down version of the PLL output signal. In such PLLs, using a time amplifier (TA) to amplify the edge time differences prior to the PFD and subsequently dividing the measured time differences by the gain of the TA attenuates the noise introduced by the measurement process without otherwise changing the loop dynamics.

Several TAs have been proposed over the last two decades [1]–[6], yet most suffer from significant drawbacks such as narrow input range [1]–[4], gain and input range dependency on technology parameters [1]–[4], high nonlinearity [1]–[5], a tradeoff between gain and input range [1]–[3], and a tradeoff between linear input range and noise [4]. The TA presented in [6] avoids most of these issues, but its relatively complicated implementation limits its noise performance which reduces its suitability for high-performance PLLs.

A low-noise inverter based delay chain TA with an analog delay-averaging nonlinearity mitigation technique is presented in this paper. The gain of the TA is nearly constant across a wide input range and is relatively insensitive to process, voltage, and temperature (PVT) variations, as it depends on a ratio of inverter delays. The TA’s principle of operation is similar to that of the TA presented in [6], but its implementation is simpler and it achieves better noise performance.

The proposed TA is demonstrated in the context of a 6.5 GHz digital fractional-N PLL based on a dual-mode ring oscillator (DMRO) delta-sigma (ΔΣ) frequency-to-digital converter (FDC) [7]–[9]. As demonstrated in [7], this type of PLL can achieve good fractional spur performance, but the DMRO’s $1/f^3$ phase noise component degrades the PLL’s in-band phase noise, and ΔΣ FDC quantization error limits the PLL’s performance at high bandwidth settings. The PLL presented in this paper applies the proposed TA to overcome these issues by attenuating both noise sources by approximately...
16 dB. Additionally, it incorporates and is the first experimental demonstration of several ΔΣ FDC improvements presented in [9]. These improvements include an all-digital background gain calibration technique that simplifies the DMRO design, and various architecture changes that relax the ΔΣ FDC’s timing constraints. A modified delay-free asynchronous DMRO is the reference frequency, \( v_{\text{ref}} \), where \( \alpha \) is a positive integer, and \( N \) is a fractional frequency offset that ranges from \( -1/2 \) to \( 1/2 \).

The PLL consists of a TA-assisted DMRO, a digital loop filter (DLF), and a divide-by-2 block with output \( v_{\text{div2}}(t) \). The ΔΣ FDC quantization error \( e(q[n]) \) is used to cancel most of the contribution of \( e_{\text{eq}}(n) \) prior to the DLF [7]–[12]. The sequence \( e_{\text{eq}}[n] \) is an estimate of \( e(q[n]) \), and it is used to cancel most of the contribution of \( e(q[n]) \) prior to the digital loop filter (DLF) [7]–[12].

Fig. 1(a) and (1) imply that the DLF input, \( p[n] \), is a measure of the PLL average phase error over the \( n \)th reference period and \( e(q[n]) \) is a ΔΣ FDC quantization error [8], [9]. The sequence \( e_{\text{eq}}[n] \) is an estimate of \( e(q[n]) \), and it is used to cancel most of the contribution of \( e(q[n]) \) prior to the digital loop filter (DLF) [7]–[12].

Fig. 1(b) shows a simplified block diagram of the TA-assisted DMRO ΔΣ FDC. It consists of a PFD, a DMRO, a digital ring phase calculator (RPC), a multi-modulus divider (MMD), and a TA. The signal \( v_{\text{ samp}}(t) \), which is an inverted version of \( v_{\text{ ref}}(t) \), is used within the RPC to sample the DMRO phase each reference period. The signal processing details of the RPC including the gain calibration technique are shown in Fig. 1(c) [9].

**B. TA-Assisted DMRO ΔΣ FDC Behavior**

An analysis similar to that presented in [8] but modified to include the TA and the improvements presented in [9] yields the ΔΣ FDC behavioral model shown in Fig. 2. In this model, \( J_{\text{TA}}(n) \) is the TA’s output jitter during the \( n \)th reference period, \( \theta_{\text{PLL}}(t), \theta_{\text{ref}}(t), \) and \( \theta_{\text{DMRO}}(t) \) are the respective phase errors in cycles of \( v_{\text{PLL}}(t), v_{\text{ref}}(t), \) and the DMRO, \( \tau_n, i_n, \rho_n, \) and \( \gamma_n \), for \( n = 0, 1, 2, \ldots \), are the respective times of the nth rising edges of \( v_{\text{ PLL}}(t), v_{\text{ ref}}(t), v_{\text{ TA}}(t), \) and \( v_{\text{ samp}}(t) \), \( T_{\text{ ref}} = 1/f_{\text{ ref}} \), and \( T_{\text{ PLL}} = 1/f_{\text{ PLL}} \). The behavioral model in Fig. 2 does not include error sources corresponding to the PFD, the MMD, or the divide-by-2 block. Simulations performed by the authors indicate that these blocks do not significantly affect the PLL’s phase noise, so they are omitted in the figure for simplicity.

The MMD is identical to those in analog PLLs, so, as illustrated in Fig. 2, \( \tau_n \) is an accumulated version of \( 2T_{\text{ PLL}}(N - v[n - 1]) \) plus noise, where \( 2T_{\text{ PLL}} \) is the divide-by-2 block’s output period and \( N - v[n - 1] \) is the MMD modulus.

As explained in Section III-B, the TA is implemented as a chain of \( N_{\text{ TA}} \) nominally identical delay cells. The propagation delay of each delay cell is \( \tau_{\text{ fast}} \) when \( v_{\text{ ref}}(t) \) is low and \( \tau_{\text{ slow}} \) otherwise, where \( \tau_{\text{ fast}} < \tau_{\text{ slow}} \). As also explained in Section III-B, the TA delays the rising edges of \( v_{\text{ div2}}(t) \) such that the pulse-width of the PFD output, \( u(t) \), during the \( n \)th reference period, i.e., \( u_n = \rho_n - i_n \), is given by

\[
u_n = -A_{\text{ TA}}(t_n - \tau_n) + N_{\text{ TA}}\tau_{\text{ slow}} + J_{\text{ TA}}[n],
\]

where

\[A_{\text{ TA}} = \tau_{\text{ slow}}/\tau_{\text{ fast}}\]

is the TA gain and \( N_{\text{ TA}}\tau_{\text{ slow}} \) is a constant offset term introduced as a byproduct of the TA’s operation. Thus, the combined behavior of the TA and the PFD is equivalent to that of an inverting amplifier with input \( t_n - \tau_n \) and additive noise and offset terms.

The DMRO is a ring of \( N_{\text{ FDC}} \) nominally identical delay cells. Ideally, its frequency is \( f_{\text{ high}} \) when \( u(t) \) is high and \( f_{\text{ low}} \) otherwise, where \( f_{\text{ high}} > f_{\text{ low}} \). As explained in [8] and

---

1By definition, \( \theta_{\text{ ref}}(t) \) is the phase error at time \( t \) of \( v_{\text{ ref}}(t) \) in units of cycles of \( v_{\text{ ref}}(t) \). Accordingly, \( T_{\text{ ref}}\theta_{\text{ ref}}(t) \) has units of seconds and it represents the reference oscillator’s absolute jitter. Similarly, \( \theta_{\text{ PLL}}(t) \) represents the phase error at time \( t \) of \( v_{\text{ PLL}}(t) \) in units of cycles of \( v_{\text{ PLL}}(t) \), so \( T_{\text{ PLL}}\theta_{\text{ PLL}}(t) \) has units of seconds and it represents the PLL output signal’s absolute jitter.
illustrated in Fig. 2, the behavior of the DMRO is that of an accumulator with gain

\[ A_{DMRO} = f_{high} - f_{low} \]  

followed by an additive noise source, an additive \( n f_{low}T_{ref} \) term, and a quantizer, \( Q_r \), with quantization step-size \( \Delta_r = (2N_R)^{-1} \).

As explained in [9], the RPC extracts the information encoded in the sampled and quantized DMRO phase and computes a fixed-point measure of \(-\alpha - e_{PLL}[n]\) each reference period. This measure is quantized to the nearest integer to compute \( y[n] \), and the resulting quantization error, \( \hat{e}_q[n] \), is used within the DLC to perform QNC. This coarse quantization operation is represented by a unity step-size quantizer, \( Q_c \), in Fig. 2.

An analysis similar to that presented in [9] shows that the DMRO locks to an average frequency of \( M_{f_{ref}} \) and the average \( u(t) \) pulsewidth, \( T_u \), is

\[ T_u = A_{DMRO}^{-1}(M - T_{ref}f_{low}). \]  

The parameter \( M \) is chosen so that the falling edges of \( u(t) \) occur between rising edges of \( v_{ref}(t) \) and \( v_{samp}(t) \), i.e., \( t_u < \rho_n < \gamma_n \) for all \( n \). This with the TA operation described in Section III-B causes the rising edges of \( v_{ref}(t) \) to precede the rising edges of \( v_{DMRO}(t) \), i.e., \( \tau_n < t_u \), for all \( n \), so that

\[ \gamma_n - 1 < t_u < \tau_n < \rho_n < \gamma_n \]  

when the PLL is locked.

A simplified version of the TA-assisted \( \Delta \Sigma \) FDC behavioral model that is valid for constant \( g_n \) is shown within the dashed contour in Fig. 2, wherein all noise components are input-referred and lumped into \( e_{PLL}[n] \), the offset components are omitted, and the quantizers \( Q_r \) and \( Q_c \) are replaced by their respective additive error sequences, \( e_{eq}[n] \) and \( \hat{e}_q[n] \). The model implies that the behavior of the \( \Delta \Sigma \) FDC when

\[ g_n = \frac{1}{2T_{PLL}A_{TA}A_{DMRO}} \]  

is identical to that of a second-order \( \Delta \Sigma \) modulator, the output of which is given by (1) with

\[ e_q[n] = g_ne_{eq}[n] + \hat{e}_q[n]. \]  

As explained in [9], when (7) is not satisfied, \( \hat{e}_q[n] \) is imperfectly canceled by QNC so it leaks into the DLC input, thereby degrading the PLL’s phase noise. The gain calibration technique shown in Fig. 1(c) causes \( g_n \) to converge to the right side of (7), which effectively circumvents this problem [9].

It follows from Fig. 2 and (7) that the power contribution of the DMRO’s phase noise to \( y[n] \), and, hence, to the PLL’s phase noise, is proportional to both \( A_{TA}^2 \) and \( A_{DMRO}^2 \). The original DMRO \( \Delta \Sigma \) FDC PLL presented in [7] does not incorporate a TA, so it corresponds to the case of \( A_{TA} = 1 \) in (7), and its in-band phase noise is dominated by the DMRO’s \( 1/f^3 \) phase noise component. In the absence of a TA, modifying the DMRO to increase \( A_{DMRO} \) and/or reduce the DMRO’s \( 1/f^3 \) phase noise component are the only options that would have mitigated this problem.

Unfortunately, these options are not attractive. In principle, increasing the widths of the transistors that make up the DMRO’s delay cells increases \( A_{DMRO} \) via (4) and decreases the DMRO’s \( 1/f^3 \) phase noise component by reducing transistor flicker noise, but in practice \( A_{DMRO} \) increases only up to a point beyond which parasitic capacitances and supply resistance cause \( A_{DMRO} \) to decrease with further transistor width increases. After this point, \( A_{DMRO} \) can only be increased further by reducing the number of DMRO delay cells, \( N_R \). Given that \( e_{eq}[n] \) is proportional to \( \Delta_r = (2N_R)^{-1} \), this would reduce the effectiveness of QNC, which would require the PLL bandwidth to be reduced to compensate for the increase in quantization noise power. Interestingly, increasing the number of DMRO delay cells does provide a modest net benefit. For example, doubling \( N_R \) reduces \( A_{DMRO}^2 \) by 6 dB, but as shown in [13] it decreases the power of the DMRO’s \( 1/f^3 \) phase noise component by 9 dB. Hence, each doubling of \( N_R \) reduces the power contribution of the DMRO’s \( 1/f^3 \) phase noise component to the PLL’s phase noise by 3 dB. Unfortunately, achieving large phase noise reductions in this manner typically requires impractically large numbers of DMRO delay cells.

These tradeoffs are avoided in this work because the TA provides amplification prior to the DMRO. As described above, the DMRO’s contribution to the PLL’s phase noise is proportional to \( A_{TA}^2 \), so each doubling of \( A_{TA} \) reduces the power of the DMRO’s contribution to the PLL’s phase noise by 6 dB.

Both the TA and the DMRO are made up of dual-delay inverter based delay cells, but the TA is an open-loop chain and the DMRO is a ring, so transistor flicker noise gives rise to \( 1/f \) noise in \( J_{TA}[n] \) and \( 1/f^3 \) noise in \( \theta_{DMRO}[\gamma_n] \). Nevertheless, as implied by Fig. 2, the contributions of \( J_{TA}[n] \) and \( \theta_{DMRO}[\gamma_n] \) to \( y[n] \) are first-order and second-order highpass shaped, respectively, so flicker noise injected by each TA transistor has a similar contribution to the PLL’s phase noise as that injected by each DMRO transistor.

Yet, it is not the case that using the TA simply transfers the problem of reducing the effect of flicker noise from the DMRO to the TA. As implied by (3), \( A_{TA} \) depends on a ratio of inverter delays, so the TA’s flicker noise can be reduced by increasing transistor widths without significantly decreasing \( A_{TA} \) or incurring other side effects similar to those mentioned above that come with reducing the DMRO’s \( 1/f^3 \) phase noise component. Furthermore, the TA is only active for a fraction of each reference period, whereas the DMRO operates continuously. As the power of the noise introduced by a chain of inverters grows at least proportionally to the number of inverters that transition as explained in [13], it follows that the TA noise contribution can be made small compared to that of the DMRO.

For instance, in the implemented PLL, 100 TA delay cells transition each reference period, whereas 660 DMRO delay cells transition each reference period on average. Moreover, in contrast to the TA, each delay cell within the DMRO transitions four to six times each reference period. Given that flicker noise changes slowly relative to \( T_{ref} \), having an edge propagating four to six times through the same delay cell effectively increases the power of the DMRO’s \( 1/f^3 \) phase

\[ 2T_{PLL}A_{TA}A_{DMRO} \]
noise component by approximately 3-5 dB compared to the case where the edge propagates through four to six different delay cells. These features made it possible for the TA to suppress the DMRO’s contribution to the PLL’s phase noise without the TA’s noise being a limitation.

III. IMPLEMENTATION DETAILS

The implemented PLL is shown in Fig. 3. It has four power supply domains, which correspond to the dashed boxes in Fig. 3. The place-and-route (PNR) digital block is clocked at a rate of \( f_{PLL}/8 \) by \( v_{clk}(t) \) and contains the DLC, the DCO control logic, the \( \Delta \Sigma \) FDC’s \( z^{-1} \) register, and all RPC components except the cycle counter and phase-sampling flip-flops.

As shown in Fig. 4, the PNR digital block comprises three sub-blocks, FDC digital, DLC, and DCO digital, that are clocked sequentially by gated versions of \( v_{clk}(t) \). The signal \( v_{ref}(t) \) is timed such that it goes high once each reference period when the DMRO phase information is ready to be processed by the PNR digital block. The \( clk_{FDC} \), \( clk_{DLC} \), and \( clk_{DCO} \) clock signals are generated by the flip-flop chain driven by \( v_{clk}(t) \), and the numbers of flip-flops between adjacent clock signals are such that enough time is allocated for each digital sub-block to meet digital timing constraints across PVT variations for an input clock frequency of 1 GHz.

The details of the sub-blocks within the PNR digital block are similar to those presented in [7]. Most of the differences are in the \( \Delta \Sigma \) FDC’s digital sub-block to incorporate the improvements proposed in [9] which include the gain calibration technique shown in Fig. 1(c). As explained in [9], the \( f_{ref} \) rate multiplier prior to the RPC’s accumulator in Fig. 1(c) represents most of the gain calibration technique’s added complexity. Its inputs have respective bit-widths of 12 and 14 bits, and its output has a bit-width of 25 bits. The RPC’s accumulator would have required 24 bits in the absence of the multiplier, so the inclusion of the gain calibration technique negligibly increases the power consumption and circuit area of the RPC accumulator and subsequent digital sub-blocks. Furthermore, the relaxed timing of the implemented \( \Delta \Sigma \) FDC architecture relative to that presented in [7] causes the power consumption and circuit area of the multiplier to be negligible relative to those of the overall digital block.

A. Timing

As shown in the timing diagram of Fig. 5, the MMD loads its inputs, \( mod_{d}[n] \) and \( mod_{p}[n] \), \( 30T_{PLL} \) after the rising edge of \( v_{div}(t) \). Fig. 5 implies that for the earliest possible \( v_{ref}(t) \) rising edge, the \( \Delta \Sigma \) FDC portion of the PNR digital block has a time budget of about \( 51T_{PLL} \) to generate its outputs after \( v_{dly}(t) \) goes high. This constraint is easy to satisfy in practice as the \( \Delta \Sigma \) FDC logic does not require more than two periods of \( v_{clk}(t) \) (i.e., \( 16T_{PLL} \)) to compute the MMD inputs. Accordingly, the implemented PLL has no significant timing bottle-necks, which makes its implementation much simpler than that of prior \( \Delta \Sigma \) FDCs [7], [8], [12], [14], [15].

B. TA

Fig. 6(a) shows a conceptual block diagram of the proposed TA. It consists of \( N_{TA} \) nominally identical inverter based delay cells, where \( N_{TA} \) is an even number. The delay of each delay cell, \( \tau_{delay} \), takes on one of two values: \( \tau_{fast} \) when \( v_{rel}(t) \) is low and \( \tau_{slow} \) when \( v_{rel}(t) \) is high.

It follows from (6) that during the \( nth \) reference period, the time, \( \tau_n \), of the rising edge of \( v_{ref}(t) \) occurs after the time, \( \tau_n \), of the corresponding rising edge of \( v_{div}(t) \), but before the time, \( \rho_n \), at which the rising edge of \( v_{div}(t) \) finishes propagating through the TA. Therefore, at time \( \tau_n \), when the rising edge of \( v_{div}(t) \) starts propagating through the TA, the...
delay cells have a delay of $\tau_{fast}$. When $v_{ref}(t)$ goes high at time $t_n$, the rising edge of $v_{div}(t)$ has already propagated through $[(t_n - t_n)/\tau_{fast}]$ delay cells and a fraction, given by $[(t_n - t_n)/\tau_{fast} - [(t_n - t_n)/\tau_{fast}]],$ of a delay cell. Thus, at time $t_n$, the rising edge of $v_{div}(t)$ has propagated through an equivalent of $(t_n - t_n)/\tau_{fast}$ delay cells, including both integer and fractional parts. At this time, the TA's delay cells are switched to have a delay of $\tau_{slow}$, so the remaining TA delay cells through which the edge must propagate contribute a combined delay of $(N_{TA} - (t_n - t_n)/\tau_{fast})\tau_{slow}$. Consequently, the time, $\rho_n$, at which $v_{TA}(t)$ goes high is given by

$$\rho_n = t_n + (t_n - t_n)/\tau_{fast} + (N_{TA} - (t_n - t_n)/\tau_{fast})\tau_{slow}. \quad (9)$$

This implies that the pulse-width of $u(t)$ during the $n$th reference period, $u_n = \rho_n - t_n$, is given by (2) with $A_{TA}$ given by (3), where the jitter term, $J_{TA}[n]$, represents the combined effect of all transistor noise sources within the TA.

It follows from the explanation above that for the TA to provide time-difference amplification it is necessary to ensure:

$$0 < t_n - t_n < N_{TA}\tau_{fast}. \quad (10)$$

Otherwise, the TA would only introduce a fixed delay between $v_{div}(t)$ and $v_{TA}(t)$. Fig. 5 implies that the time at which the MMD loads its inputs also imposes a constraint on the maximum value of $t_n - t_n$. Specifically, the MMD must load its inputs at the time of the rising edge of $clk_{FDC}$ at the earliest, which can occur up to $37T_{PLL}$ after the falling edge of $v_{ref}(t)$. Therefore, $t_n - t_n$ must satisfy

$$t_n - t_n < 1/2T_{ref} - 7T_{PLL}. \quad (11)$$

in addition to (10). Moreover, for the $\Delta \Sigma$ FDC to work properly, $u(t)$ must go low before the DMRO phase is sampled at time $\gamma_n$, which requires

$$0 < u_n < 1/2T_{ref} + 10T_{PLL}. \quad (12)$$

Equations (10)-(12) impose design constraints on the TA parameters $N_{TA}$, $\tau_{slow}$, $\tau_{fast}$, and $A_{TA}$.

As shown in Fig. 6(b), each of the TA’s dual-delay inverters consists of a standard inverter in parallel with a larger tri-state inverter. When $v_{ref}(t)$ goes high, the tri-state inverter is disabled by disconnecting its ground and power supply terminals from the supply rails, thereby increasing $\tau_{delay}$ from $\tau_{fast}$ to $\tau_{slow}$.

Ideally, $\tau_{delay}$ changes instantaneously from $\tau_{fast}$ to $\tau_{slow}$ when $v_{ref}(t)$ goes high, in which case the TA performs linear amplification. Unfortunately, the $\tau_{fast}$-to-$\tau_{slow}$ transitions are...
non-instantaneous in practice, which causes TA nonlinearity. Moreover, as illustrated in Fig. 6(c), this transition also depends on whether the cell’s input, $d_{n-1}(t)$, goes from low to high or vice versa.

The TA topology shown in Fig. 7(a) is proposed to reduce such nonlinearity. It consists of two nominally identical delay chains in parallel, where the input of one delay chain is an inverted version of that of the other delay chain, both delay chains are controlled by $v_{\text{ref}}(t)$, and each pair of parallel delay cells are cross-connected with averaging resistors. As shown in Fig. 7(a) for the top and bottom delay chains in isolation, the odd-indexed and even-indexed delay cells have inputs that transition in opposite directions, so they have different $\tau_{\text{fast}}$-to-$\tau_{\text{slow}}$ transitions. This causes a quasi-periodic artifact in the input-output characteristics of the delay chains. Driving the bottom delay chain by an inverted version of $v_{\text{div}}(t)$ causes its input-output characteristic to be shifted with respect to that of the top delay chain such that, when averaged via the cross-coupled resistor network, the nonlinearity of the cross-coupled delay chains is considerably smaller than that of either delay chain in isolation. Behavioral simulations of the PLL in which the TA’s nonlinear behavior is considered and all other spur-generation mechanisms are neglected suggest that the power of the PLL’s worst-case fractional spur decreases by 7 dB when the proposed nonlinearity mitigation technique is used.

In addition to having improved linearity, the proposed TA topology’s pseudo-differential nature can be exploited to implement a TA power-saving (PS) mode. Without the PS mode, the falling edge of $v_{\text{div}}(t)$ propagates through the TA each reference period. This resets the delay cells’ states for the next rising edge of $v_{\text{div}}(t)$, but the power consumed by the resulting delay cell transitions represents a significant portion of the TA’s total power consumption. The idea behind the PS mode is to swap the differential inputs and swap the differential outputs of the TA each reference period to obviate the need to reset the delay cells, so the falling edge of $v_{\text{div}}(t)$ can be prevented from propagating through the TA to save power.

The implemented TA, which includes the nonlinearity mitigation technique and PS mode option as described above, is shown in Fig. 7(b). It comprises the TA core shown in Fig. 7(a) as well as input and output swapping circuitry used when the PS mode is enabled. The transistor-level details of the TA’s delay cells are shown in Fig. 7(c). The TA core was designed to maximize the value of $A_{\text{TA}}$ while satisfying the constraints in (10)-(12). Specifically, $N_{\text{TA}} = 100$, $\tau_{\text{fast}} = 10$ ps, $\tau_{\text{slow}} = 70$ ps, and $A_{\text{TA}} = 7$. Simulation results predict that the TA’s gain varies by $\pm$7% across process corners, $\pm$10% across process corners and temperature variations (0°C to 85°C), and $\pm$14% across process corners, temperature variations, and supply voltage variations ($\pm$10%).

The PS mode is enabled and disabled via the PSen signal. When enabled, the $\phi_1(t)$ and $\phi_2(t)$ signals are used to implement the input and output swapping operations. The signal $\phi_1(t)$, which is derived from $v_{\text{div}}(t)$, is used to swap the inputs and the outputs of the TA core each reference
period, whereas \( \phi_2(t) \) is used to control the input and output latches. As illustrated in the timing diagram shown in Fig. 7(b), these latches prevent the falling edges of \( \zeta(t) \) from propagating through the TA core, and also prevent the output swapping circuitry from disturbing \( v_{TA}(t) \) while the swapping occurs.

The TA was laid out such that systematic mismatch among its unit cells is negligible, and the unit cells are sized such that the power of the PLL’s worst-case fractional spur caused by random mismatches among the TA’s delay cells is approximately \(-50\) dBc. This was determined by performing a Monte Carlo simulation in Cadence to obtain 90 different TA input-output characteristics, and the results were imported into a bit-exact, event-driven, custom behavioral PLL simulator. Fig. 7(d) shows a histogram of the simulated PLL’s worst-case fractional spur power. As shown in Fig. 7(d), the worst-case fractional spur power’s expected value is \(-51.7\) dBc, and its standard deviation is 2.5 dBc.

As mentioned in Section I, the proposed TA achieves better noise performance than a comparably configured TA of the type presented in [6]. One reason for this difference is that the TA in [6] incorporates two ring oscillators that both contribute noise to the output whereas the proposed TA incorporates a single delay-chain that contributes noise to the output. Another reason is that the TA presented in [6] requires NAND gate based delay cells instead of inverter based delay cells which each introduce more phase noise than comparable inverter based delay elements.

C. DMRO and Phase Sampling Scheme

The DMRO, which is shown in Fig. 8, consists of \( N_R = 127 \) inverter delay cells and has \( \text{ADMRO} = 670 \) MHz \((\text{f}_{\text{high}} = 730 \) MHz and \( \text{f}_{\text{low}} = 60 \) MHz\). Each DMRO delay cell contains a dual-delay inverter that is similar to that used in the TA. It includes a standard \( \times 1 \) inverter in parallel with a \( \times 16 \) tri-state inverter, and the tri-state inverter’s power and ground lines are connected to or disconnected from the supply rails when \( u(t) \) is high or low, respectively. This modulates each delay cell’s propagation delay such that the DMRO frequency is \( \text{f}_{\text{high}} \) when \( u(t) \) is high and \( \text{f}_{\text{low}} \) when \( u(t) \) is low. In both cases, the DMRO outputs swing from rail to rail, which allow the DMRO outputs to drive standard digital logic without the need for level-shifting. The \( \times 2 \) inverter shown within the dashed box in Fig. 8 is used to buffer the delay cell’s input to reduce the disturbance to the DMRO when its phase is sampled.

As explained in Section II-B, the TA causes the PLL phase noise contributed by the DMRO to be attenuated in power by a factor of \( A^2_{\text{TA}} \). Additionally, the DMRO’s 1/\( f^3 \) phase noise component is further mitigated by using a large number of stages [13]. This comes at the expense of higher digital complexity and higher power consumption, primarily due to the charging and discharging of the gates controlled by \( u(t) \).

To prevent the DMRO from running with multiple stages transitioning simultaneously, even for a brief period of time, the first delay cell includes a switch between the ground terminal of the \( \times 1 \) inverter and the ground rail. At startup, both \( u(t) \) and the enable signal are set low. This opens the ring so that any transition propagating through it eventually reaches the first stage and stops propagating. The switch is subsequently closed after which the DMRO operates normally.

The DMRO phase sampling scheme is shown in Fig. 8. As explained below, it addresses the issue that the sampling clock, \( v_{\text{samp}}(t) \), and the DMRO are asynchronous yet avoids the delay incurred by the DMRO sampling scheme in [7]. It consists of a cycle counter followed by sampling flip-flops and a phase decoder. The principle behind the sampling of the cycle counter’s outputs is based on that of the asynchronous sampling schemes presented in [16] and [17]. To the knowledge of the authors, the proposed phase decoder implementation described below is introduced for the first time in this work.

The cycle counter consists of two 4-bit counters that are clocked, respectively, by the rising and falling edges of the DMRO delay cell with output \( d_1(t) \). On each rising edge of the ref-rate signal \( v_{\text{samp}}(t) \), the counter outputs \( c_{\text{pos}}(t) \) and \( c_{\text{neg}}(t) \) are sampled to generate \( c_{\text{pos}}[n] \) and \( c_{\text{neg}}[n] \), and the DMRO outputs \( d_1(t), d_2(t), \ldots, d_{127}(t) \) are sampled to generate \( d_1[n], d_2[n], \ldots, d_{127}[n] \). The phase decoder consists of a lookup table (LUT) that quantizes the sampled DMRO outputs to a 10-bit sequence, \( t_k[n] \), which represents the fractional part of the sampled DMRO phase, and logic that computes \( c_k[n] \), which represents the integer part of the sampled DMRO phase.

The number of bits of \( t_k[n] \) was chosen to ensure that the
contribution to the PLL’s phase noise from the error introduced by the LUT’s quantization operation is negligible compared to those of the other error sources.

The top and bottom counters in the cycle counter are clocked when \( t_R[n] \) is between 63\( \Delta_r \) and 189\( \Delta_r \), and \( c_R[n] \) is set to \( c_{pos}[n] \). Ideally, \( c_R[n] \) should be set to \( c_{neg}[n] \) when \( t_R[n] \) is between 190\( \Delta_r \) and 253\( \Delta_r \), and to \( c_{neg}[n] + 1 \) when \( t_R[n] \) is between 0 and 62\( \Delta_r \), so as to account for the bottom counter being clocked half a DMRO cycle after the top counter is clocked. Yet to work correctly this would require \( c_{pos}(0) = c_{neg}(0) \) and the initial DMRO fractional phase to be such that the top counter is clocked before the bottom counter after startup, which are hard to ensure in practice.

These requirements are avoided via the \( c_{corr}[n] \) correction logic shown in Fig. 8. As both sampled counter outputs are reliable when \( t_R[n] \) is around 63\( \Delta_r \) and 190\( \Delta_r \), the \( c_{corr}[n] \) logic block in Fig. 8 computes

\[
c_{corr}[n] = \begin{cases} 
    c_{pos}[n] - c_{neg}[n] - 1, & \text{if } t_R[n] \in [53\Delta_r, 73\Delta_r], \\
    c_{pos}[n] - c_{neg}[n], & \text{if } t_R[n] \in [180\Delta_r, 200\Delta_r], \\
    c_{corr}[n] - 1, & \text{otherwise}, 
\end{cases}
\]

and \( c_R[n] \) is set to \( c_{neg}[n] + c_{corr}[n] \) when \( t_R[n] \) is between 190\( \Delta_r \) and 253\( \Delta_r \), to \( c_{neg}[n] + c_{corr}[n] + 1 \) when \( t_R[n] \) is between 0 and 62\( \Delta_r \), and to \( c_{pos}[n] \) otherwise.

D. MMD

As shown in Fig. 9, the MMD consists of a finite-state machine (FSM), a 4/5 prescaler, an edge-select flip-flop, and a resynchronization flip-flop. As explained below, the MMD causes the rising edges of \( v_{div2}(t) \) during the \( n \)th and \( (n + 1) \)th reference periods to be separated by \( N - v[n] \) periods of \( v_{div2}(t) \).

When the FSM’s \( p_{sel}(t) \) output bit is low, the prescaler divides by 4. Otherwise, it divides by 5. At the beginning of each MMD cycle, the FSM sets \( p_{sel}(t) \) low for five periods of \( v_{pres}(t) \), so the first five periods of \( v_{pres}(t) \) each have a duration of four \( v_{div2}(t) \) periods. Then, the FSM sets \( p_{sel}(t) \) so that \( mod_d[n] \) counts to 4 followed by \( mod_d[n] \) counts to 5 occur, where

\[
mod_d[n] = N - v[n] - 20 - 4\lfloor (N - v[n] - 20)/4 \rfloor \quad \text{and} \quad mod_d[n] = \lfloor (N - v[n] - 20)/4 \rfloor - mod_d[n],
\]

after which \( N - v[n] \) periods of \( v_{div2}(t) \) will have occurred.

As illustrated in the timing diagram shown in Fig. 9 for the example case of \( mod_d[n] = 1 \), the FSM’s \( p_{pass}(t) \) output goes high at the start of the last full \( v_{pres}(t) \) period prior to the next rising edge of \( v_{div}(t) \), which causes the edge-select flip-flop’s output to go high on the next rising edge of \( v_{pres}(t) \). The resynchronization flip-flop samples the edge-select flip-flop output on the next rising edge of \( v_{div2}(t) \) to prevent the MMD output edge from being corrupted by noise and modulus-dependent delay error that originated in the prior MMD components.

All MMD blocks were built using standard cells, with the exception of the resynchronization flip-flop which was custom-designed to minimize its contribution to the PLL’s phase noise.

E. DCO

The DCO is similar to that presented in [7]. It consists of a single-turn center-tapped inductor, a cross-coupled pair

---

**TABLE I**

<table>
<thead>
<tr>
<th>Block</th>
<th>Area (mm²)</th>
<th>Power (mW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PNR Digital</td>
<td>0.0242</td>
<td>4.76</td>
</tr>
<tr>
<td>Reference Buffers</td>
<td>0.00158</td>
<td>0.165</td>
</tr>
<tr>
<td>( \Delta ) FDC</td>
<td>0.00713</td>
<td>9.44/8(1)</td>
</tr>
<tr>
<td>DCO</td>
<td>0.137</td>
<td>8.75</td>
</tr>
<tr>
<td>Total Area</td>
<td>0.6321</td>
<td>23.15/21.7(1)</td>
</tr>
<tr>
<td>Active Area(1)</td>
<td>0.1683</td>
<td>23.15/21.7(1)</td>
</tr>
</tbody>
</table>

1 Without and with TA PS mode enabled, respectively.
2 Without decoupling capacitors.
of nMOS transistors, a tail resonant tank of the type proposed in [18], a triode MOS transistor tail source, an integer frequency control element (FCE) bank driven by $c_I[p]$, and a fractional FCE bank driven by $c_F[p]$. The implemented FCEs are of the type presented in [15], and the minimum-size FCE has an equivalent frequency step of $\Delta_{\text{min}} = 160 \text{ kHz}$ at 6.5 GHz. The DCO’s 16-bit input sequence, $d[n]$, is split into integer and fractional parts. The integer part is encoded to drive the integer FCE bank, which comprises eight $\Delta_{\text{min}}$ FCEs and five pairs of $16\Delta_{\text{min}}, 8\Delta_{\text{min}}, 4\Delta_{\text{min}}, 2\Delta_{\text{min}},$ and $\Delta_{\text{min}}$ FCEs. The fractional part is up-sampled and re-quantized by a second-order $\Delta \Sigma$ modulator that generates a five-level output sequence. This output sequence is scrambled by a dynamic element matching (DEM) encoder, the outputs of which drive four $\Delta_{\text{min}}$ FCEs within the fractional FCE bank. The PLL controls the DCO over a range of 41 MHz with a minimum step size of 625 Hz.

The DCO also contains a binary-weighted capacitor array controlled via a serial peripheral interface (SPI), which is in parallel with the integer and fractional FCE banks. The capacitor array has 7 bits of tuning over a frequency range of 5.6-6.6 GHz.

IV. MEASUREMENT RESULTS

The prototype IC contains the PLL in Fig. 3 as well as an SPI port and test circuitry to measure internal signals during testing. It was fabricated in the GlobalFoundries 22-nm CMOS 22FDX technology. A die photograph is shown in Fig. 10, and area and power breakdowns are presented in Table I. The IC is packaged in a QFN28 package with a ground paddle and was tested with an Ironwood SG-MLF-7003 compression elastomer socket. Except where noted otherwise, all of the measurements presented below were taken with a common set of PLL parameters set via the SPI.

Unfortunately, the DCO tank’s quality factor is severely degraded by a layout issue to the point that the DCO as-fabricated does not even oscillate, and the problem was not flagged by simulations prior to fabrication because of a post-layout extraction tool flaw. Removing metal near the DCO’s main inductor via focused ion beam (FIB) surgery made the DCO functional, but even with its maximum current setting and its supply set to 0.9 V, its oscillation amplitude is extremely low. Consequently, the DCO’s power consumption is that of a high-performance DCO, yet it achieves relatively poor phase noise performance (e.g., 10 dB worse at a 1-MHz offset than expected\textsuperscript{2}) and its low oscillation amplitude makes it highly sensitive to interference from other circuit blocks. While the PLL’s overall measured performance is nevertheless in line with the current state of the art, these issues limited its performance as quantified later in this section. The IC’s measured output power is around −34 dBm, so an amplifier module was used to boost the output power to around −2 dBm.

Fig. 11 shows the measured phase noise of the PLL at $f_{\text{PLL}} = 6.5 \text{ GHz}$ with and without the TA enabled for PLL bandwidths of 1 and 4.5 MHz. The integrated random jitter (i.e., the jitter omitting spurious tones), $\sigma_{\text{RJ}}$, is also reported in Fig. 11, where the integration band extends from 10 kHz to 80 MHz. To estimate the expected noise reduction when the TA is enabled, $A_{\text{TA}}$ was calculated indirectly from (7) using measured values of $g_a$ read through the SPI. It was found that $g_a$ converged to about 0.758 and 4.832 with and without the TA enabled, respectively, with which two equations based on (7) were solved to find $A_{\text{TA}} = 6.37$. This suggests that the TA reduces the power of the portions of the PLL’s phase noise contributed by both the DMRO’s circuit noise and its quantization noise by 16 dB.

In the case of Fig. 11(a), the in-band spot phase noise at a 100-kHz offset frequency decreases from $−99$ to $−107 \text{ dBc/Hz}$ when the TA is enabled, whereas in the case of Fig. 11(b), the in-band spot phase noise at a 1-MHz offset frequency decreases from $−100$ to $−112 \text{ dBc/Hz}$ when the TA is enabled. In the former case, the PLL’s in-band phase noise has comparable contributions from the DMRO, reference

\textsuperscript{2}The spot phase noises of the DCO after the FIB surgery when tuned to 6.5 GHz are $−59$, $−117$, and $−148 \text{ dBc/Hz}$ at offset frequencies of 10 kHz, 1 MHz, and 100 MHz, respectively.
signal, and DCO, whereas in the latter case, the in-band phase noise is mostly dominated by the DMRO phase noise. Accordingly, as the TA suppresses the DMRO’s contribution to the PLL’s phase noise, the PLL’s in-band spot phase noise reduction is more significant in Fig. 11(b). Nonetheless, as shown in Fig. 11(a), the spot phase noise at a 1-MHz offset frequency decreases from $-100$ to $-112$ dBc/Hz when the TA is enabled, which occurs because the PLL’s phase noise is dominated by DMRO quantization error around that offset frequency.

Fig. 12 shows the PLL’s measured phase noise with $\alpha_{f_{\text{ref}}}$ set to 18 MHz, the PLL bandwidth set to 1 MHz, and the TA enabled. In this case, the integrated total jitter (i.e., the jitter including spurious tones), $\sigma_{TJ}$, was 151 fs$_{\text{rms}}$. This represents the best-case total jitter because it corresponds to a case where the spurious tones are well outside the PLL bandwidth.

The largest measured fractional spur and $\sigma_{TJ}$ versus $\alpha_{f_{\text{ref}}}$ are shown in Fig. 13(a) and Fig. 13(b), respectively, for a PLL bandwidth of 1 MHz. The fractional frequency offset, $\alpha$, was swept such that $\alpha_{f_{\text{ref}}}$ ranges from 1 kHz to 40 MHz with 20 equally-spaced values per decade on a log scale. The integration band of the jitter extends from 10 kHz to 80 MHz to include all significant spurs. The spur powers were measured with the spectrum analyzer’s averaging option disabled, and for each value of $\alpha$, the instrument was configured to ensure that five negative and positive fractional spur harmonics were always visible. In each case, the largest fractional spur was one of the first three harmonics of $\alpha_{f_{\text{ref}}}$, and was no higher than $-49$ dBc. The measured worst-case spurious tone powers are in line with those predicted by simulation results that include random mismatches among the TA delay cells.

For some values of $\alpha_{f_{\text{ref}}}$ > 5 MHz, spurs with power lower than $-60$ dBc and frequencies that are not multiples of $\alpha_{f_{\text{ref}}}$ were measured. The authors have not definitively determined the origin of these spurs, but suspect they are from external interference that is parasitically coupled into the DCO and their effect is exacerbated by the DCO’s abnormally low amplitude. These interference spurs are not reported in Fig. 13(a), although their contribution to $\sigma_{TJ}$ is taken into account in Fig. 13(b), which is why $\sigma_{TJ}$ increases somewhat for $\alpha_{f_{\text{ref}}}$ > 5 MHz.

As shown in Fig. 14, the measured reference spur power is lower than $-80$ dBc. As mentioned above, the authors believe
that the DCO’s low oscillation amplitude makes it extremely sensitive to external interference. This theory is supported by the observation that increasing the DCO supply, which increases its oscillation amplitude somewhat, tends to reduce the measured spurs. For example, measurements taken with the DCO supply set to 1.1 V yields a reference spur of \(-85\) dBc.

Accordingly, the reported reference spur power in Fig. 14 is a worst-case bound on the reference spur performance of the PLL, as the power of this spur is expected to decrease when the DCO problem mentioned above is fixed in a future version of the PLL.

Fig. 15 shows the measured phase noise of the PLL with and without the gain calibration technique enabled for a PLL bandwidth of 4.5 MHz. The results demonstrate the effect of non-ideal Δ Σ FDC forward path gain, i.e., the effect of \(g_n\) not satisfying (7), on the PLL’s performance at high bandwidth settings. As indicated in Fig. 15, the spot phase noise at a 20-MHz offset frequency decreases by 32 dB when enabling the gain calibration technique, which causes \(\sigma_{TJ}\) to decrease from 2.7 ps\(_{\text{rms}}\) to 248 fs\(_{\text{rms}}\).

Measurements indicate that enabling the TA PS mode has several effects: 1) it decreases the PLL’s power consumption by 1.45 mW, which corresponds to 37% of the TA power consumption when the PS mode is disabled, 2) it increases the best-case \(\sigma_{TJ}\) by 20 fs because the swapping circuitry shown in Fig. 7(b) introduces noise into the reference path, 3) it decreases the worst-case \(\sigma_{TJ}\) by 30 fs due to the slightly better fractional spur performance, and 4) it increases the reference spur power by 14 dB. The authors believe that the reference spur power increase is related to coupling from the analog domain to the DCO, again because of the DCO’s low oscillation amplitude.

Table II summarizes the performance of the PLL with and without the TA PS mode enabled, along with that of the best digital PLLs published to date [7], [19]–[27]. As shown in Table II, the PLL achieves one of the best in-band spot phase noises, and its spurious tone performance is comparable to that of other state-of-the-art digital PLLs, even though no dedicated spur mitigation technique is used. In contrast, automatic time-to-digital converter (TDC) gain tracking is used to reduce the fractional spur from \(-35\) to \(-55\) dBc in [22], a TDC calibration technique is used to reduce the fractional spur power from \(-43\) to below \(-74\) dBc in [21], and a phase interpolation nonlinearity calibration technique is used to reduce the fractional spur from \(-24.58\) to \(-53.1\) dBc in [20]. Similarly, digital-to-time converter (DTC) range reduction techniques are used in [19], [24], and [25] to improve fractional spur performance.

The PLL’s best-case \(\sigma_{TJ}\) is lower than most of the other PLLs in Table II, but its power consumption is higher than those of the other PLLs. As previously mentioned, the implemented DCO consumes the power of a DCO with much better phase noise. Simulations run by the authors suggest that for a properly designed DCO with similar phase noise to that of the implemented DCO, the power consumption should be around 4 mW instead of 8.75 mW. Alternatively, if the DCO had performed as expected, the PLL’s best-case \(\sigma_{TJ}\) would have been 115 fs\(_{\text{rms}}\) instead of 151 fs\(_{\text{rms}}\). Furthermore, as mentioned in Section III, the PNR digital was overdesigned to be clocked at 1 GHz, which is supported by measurements...
given that the digital domain power supply can be reduced from 0.8 to 0.55 V without affecting the PLL’s performance. In this case, the power consumption of the PNR digital goes down from 4.76 to 2.66 mW. Therefore, the implemented PLL’s power consumption is higher than necessary, and it could potentially be lowered by approximately 6.85 mW.

Nonetheless, as shown in Table II, even with the higher-than-necessary digital power consumption and worst-than-expected DCO performance, the PLL achieves a GaAs figure of merit (FoM) comparable to or better than prior-art digital PLLs [28]. Had the DCO performed as expected, i.e., with performance comparable to that of the DCO presented in [7], the PLL’s best-case FoM would have been −245.1 dB and −245.4 dB with and without the TA PS mode disabled, respectively. Alternatively, had the PLL’s power consumption been 6.85 mW lower as explained above, the PLL’s best-case FoM could potentially be lowered by approximately 6.85 mW.

In this case, the power consumption of the PNR digital goes down lower [13].

ACKNOWLEDGMENT

The authors are grateful to Colin Weltin-Wu, Yiwu Tang, and Dongmin Park for helpful advice, Raghavendra Haresamdrum for his constant support with different software tools, Julian Puscar and Mahmoud Abbad for digital-flow advice, and Dongmin Park for helpful advice. Raghavendra Haresamdrum for his constant support with different software tools, Julian Puscar and Mahmoud Abbad for digital-flow advice, and Dongmin Park for helpful advice.

REFERENCES


Enrique Alvarez-Fontecilla received the B.Sc. and M.Sc. degrees in electrical engineering from the Universidad Católica de Chile (PUC), Santiago, Chile, in 2011 and 2013, respectively. He is currently pursuing the Ph.D. degree with the University of California at San Diego, San Diego, CA, USA. From 2012 to 2015, he was an Adjunct Assistant Professor with the School of Engineering and also with the Institute of Philosophy, PUC.

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, San Diego, CA, USA. From 2013 to 2016, he was a Teaching and Research Assistant with Electronics and Communications Engineering Department, ASU. His research interests include the analysis and design of analog/mixed-signal integrated circuits and systems.

Ian Galton (Fellow, IEEE) received the B.Sc. degree from Brown University, Providence, RI, USA, in 1984, and the M.S. and Ph.D. degrees from the California Institute of Technology, Pasadena, CA, USA, in 1989 and 1992, respectively, all in electrical engineering. Since 1996, he has been a Professor of Electrical Engineering with the University of California at San Diego, San Diego, CA, where he teaches and conducts research in the field of mixed-signal integrated circuits and systems for communications. His research involves the invention, analysis, and integrated circuit implementation of critical communication system blocks such as data converters and phase-locked loops.