On-Chip Jitter Measurement Using Jitter Injection in a 28 Gb/s PI-Based CDR

Joshua Liang, Ali Sheikholeslami, Senior Member, IEEE, Hirotaka Tamura, Fellow, IEEE, and Hisakatsu Yamaguchi

Abstract—We present a technique to measure random jitter in a phase interpolator (PI)-based clock and data recovery (CDR) circuit by injecting a controlled amount of square-wave jitter into its edge clock and monitoring its effect on the autocorrelation function of the CDR’s bang-bang phase detector output. Jitter is injected by adjusting the code of the edge PI while the autocorrelation function is measured by on-chip counters. Since the injected jitter only affects the edge clock, the CDR remains operational during jitter injection. Using this technique, the rms relative jitter between the clock and data at the CDR input can be estimated with sub-picosecond accuracy as demonstrated in measurements of a 28 Gb/s half-rate digital CDR fabricated in 28 nm CMOS.

Index Terms—Bang-bang phase detector (BB-PD), built-in self-test (BIST), clock and data recovery (CDR), jitter measurement, phase interpolator (PI).

I. INTRODUCTION

As data rates increase and the unit interval (UI) shrinks, it becomes increasingly important to minimize the jitter of clock and data recovery (CDR) circuits. This can be difficult, however, as jitter in a CDR is contributed by numerous sources such as the transmitter, intersymbol interference (ISI), as well as various receiver components such as the phase detector (PD), phase interpolator (PI), and phase-locked loop (PLL). Jitter is also contributed by secondary sources such as power supply noise. As a result, it is difficult to accurately predict and budget for each jitter contribution when designing a CDR. Furthermore, even if a design is optimized, process voltage and temperature variation and ageing can still degrade performance by pushing the design away from its optimal operating point.

On-chip jitter measurement can help address these challenges, allowing a CDR to be adaptively tuned based on the observed jitter. For example, a jitter measurement circuit can be added to the CDR as shown in Fig. 1, and used to measure the relative jitter between the CDR’s input data and recovered clock. Gradient descent [1] or other optimization techniques can then be used to minimize the relative jitter, and help achieve the lowest bit error rate (BER). Unfortunately, adding jitter measurement circuits is generally costly in terms of power, area, and the extra circuits that need to be designed.

In this paper, we present a technique to measure the rms value of the relative jitter in PI-based CDRs with reduced circuit overhead. By injecting a known amount of square-wave jitter using the CDR’s PI and monitoring the autocorrelation function of the CDR’s bang-bang PD (BB-PD) output [2], the rms relative jitter of the CDR is estimated with sub-picosecond accuracy without requiring any dedicated analog circuits such as eye monitors. The scheme, shown in Fig. 2, is demonstrated in a 28 Gb/s half-rate CDR implemented in 28 nm CMOS. Measurements results also show how this measurement scheme can be used to select the optimal loop filter and equalizer settings of the CDR.

The remainder of this paper is organized as follows. First, Section II provides background on the most common on-chip
II. BACKGROUND

To be of practical use in high data rate applications, an on-chip jitter measurement circuit must have sufficiently fine measurement resolution. At 25 Gb/s and beyond, the random jitter (RJ) of reference clocks, usually generated by high-performance PLLs, is tightly constrained to be hundreds of femtoseconds or less [3], [4]. This is because such RJ is unbounded and scales with the target BER. Optimizing these PLLs using on-chip jitter measurement would require femtosecond resolution, which may not be practical with existing techniques. PLLs can instead be optimized using other methods [5]. On the other hand, deterministic jitter sources in the CDR such as dithering, PI nonlinearity, ISI, and power supply-induced jitter are bounded, and can have much higher rms values, pushing the total rms jitter of CDRs beyond 1 ps [6], [7]. A measurement technique with sub-picosecond resolution can therefore be valuable in optimizing a CDR’s performance.

Jitter can be measured on-chip in a CDR by adding a dedicated analog jitter measurement circuit or eye monitor. Examples of jitter measurement circuits include PD-based circuits [8], which use a linear PD to convert jitter into analog signals, or time-to-digital converters (TDCs), which convert timing information to digital codes using some form of delay line [9], [10]. Achieving sub-picosecond resolution from either of these circuits is challenging and may require the use of calibration [11]. The sampling rates of published TDCs such as [9]-[11] are also limited to several gigahertz. Another class of measurement circuits estimate jitter statistically by sampling a signal asynchronously [12]. In [13], a dedicated jittery voltage-controlled oscillator generates an asynchronous clock, which is then used to estimate the rms jitter of another clock by sampling it along with a reference signal. These approaches all require the design of custom analog circuits and often also require additional supporting circuitry such as dedicated reference clocks that are sometimes supplied from off-chip.

In CDRs, jitter can also be measured using an eye monitor [14]-[18]. As shown in Fig. 3, an eye monitor consists of a dedicated sampler with an adjustable clock phase that can be generated by a PI. The cumulative distribution function of the jitter can be measured by comparing the eye sampler output $S_n$ to the adjacent data samples $D_n$ and $D_{n+1}$ and counting the number of transitions occurring before and after $S_n$, as the phase of the eye monitor clock $C_{EYE}$ is swept across the UI. If the eye sampler has an adjustable voltage threshold, it can also be used to measure the receiver’s vertical eye opening, which can help drive adaptation [14].

As with other measurement circuits, eye monitors require extra analog circuits, which consume additional power, area, and design effort. More importantly, these circuits also increase loading on the critical clock and high-speed data paths of the analog front end, potentially degrading CDR performance. In this paper, we avoid this additional analog overhead by using the CDR’s existing BB-PD to measure jitter.

III. PROPOSED TECHNIQUE

Although a BB-PD only outputs the sign of the relative jitter between the clock and data ($\psi_{ER}$), it can also be modeled as a linear gain $K_{PD}$ with quantization noise $\psi_{PD}$ added at its output [19] as shown in Fig. 4. As such, the BB-PD output $\Gamma$ can be used to measure $\psi_{ER}$. However, $K_{PD}$ itself also depends on the probability density function (PDF) and standard deviation of $\psi_{ER} (\sigma_{ER})$. An eye monitor can be used to estimate $\sigma_{ER}$ and, therefore, $K_{PD}$ [18], but that would require additional analog hardware. Instead in this paper, a controlled amount of square-wave jitter ($\psi_{INJ}$) is injected into the CDR by adding a signal to the PI code of the CDR’s edge clock $C_{EDGE}$, as shown in Fig. 2. Since the injected jitter amplitude is known, it can be used as a reference to extract $K_{PD}$ and $\sigma_{ER}$. When injecting jitter, $K_{PD}$ can be estimated by either correlating the PD output $\Gamma$ with the injected signal, or by estimating the autocorrelation function $R(n)$ of the PD output as shown in Fig. 5(a). The second method has the additional benefit of allowing the power spectral density (PSD) of the PD output and, therefore, jitter, to be estimated [20] and is, therefore, used in this paper.
To extract $K_{PD}$, we inject square-wave jitter with amplitude $A$ into the recovered clock. Then to estimate $R(n)$, we correlate the PD output with itself delayed by $n$, and take the average as shown in Fig. 5(a). (Note that since $\psi_{INJ}$ is periodic, $\psi_{ER}$ is actually cyclostationary, making $R(n)$ a periodic function of $k$. Here for simplicity, $R(n)$ refers to the time-averaged autocorrelation function, which is averaged over $k$ [18].)

Fig. 5(b) shows how $R(n)$ is affected by the injected jitter. When $\psi_{INJ} = 0$, and if $\psi_{ER}$ is largely white, $R(n)$ is dominated by a delta function. When $\psi_{INJ}$ is a square wave and the other jitter sources are zero, $R(n)$ becomes a triangular wave, with a delta function caused by $\psi_{PD}$, which we also assume to be white. When $\psi_{INJ}$ is added to the existing jitter, $R(n)$ is again a delta function with a triangular wave superimposed on top of it. By measuring the amplitude of the triangular wave ($\Delta$) compared to the rest of $R(n)$, we can extract $\sigma_{ER}$, the rms value of the CDR’s relative jitter. With jitter injection, the total CDR input jitter ($\psi_{EFF}$) is equal to $\psi_{ER} + \psi_{INJ}$ and has a variance equal to $\sigma_{EFF}^2 = \sigma_{ER}^2 + \sigma_{INJ}^2$. This gives

$$R(0) = \alpha_T = K_{PD,EFF}^2 [\sigma_{ER}^2 + \sigma_{INJ}^2 + \sigma_{PD}^2]$$

(6)

where $K_{PD,EFF}$ is the PD gain in the presence of jitter injection. Since $\psi_{INJ}$ is uncorrelated with $\psi_{ER}$, $\Delta$ corresponds to the power of $\psi_{INJ}$ at the PD output

$$\Delta \approx K_{PD,EFF}^2 \sigma_{INJ}^2.$$  

(7)

Combining (7) and (6) and setting $\sigma_{INJ} = A$ gives

$$\sigma_{ER} \approx A \sqrt{\frac{\alpha_T - \sigma_{PD}^2}{\Delta} - 1.}$$

(8)

Assuming that $\psi_{ER}$ is Gaussian and $A$ is small compared to $\psi_{EFF}$, $\sigma_{EFF}$ can still be approximated as Gaussian. $\sigma_{PD}$ can then be substituted with (5), giving $\sigma_{ER}$ as a function of $\Delta$

$$\sigma_{ER} \approx A \sqrt{\frac{2}{\pi} \frac{a_T^2}{\Delta} - 1.}$$

(9)

Since the BB-PD output is binary, $R(n)$ can be measured with simple logic and digital counters that count the number of correlated outputs [18]. $\Delta$ can then be measured from the resulting waveform. Because a PD output of zero provides no useful information, we can also calculate a normalized $R(n)$ ($R'(n)$) only using the non-zero PD output samples

$$R'(n) = E[\Gamma(k)\Gamma(k-n)|\Gamma(k) \neq 0, \Gamma(k-n) \neq 0].$$

(10)

Calculating $R(n)$ in this way removes the effect of transition density. $\alpha_T$ then becomes 1 in (9), making $\sigma_{ER}$ only a function of $\Delta$. Although the relationship is nonlinear, in a digital implementation, it could be stored in a simple lookup table, whose number of addresses would depend on the number of bits with which $\Delta$ is estimated, and whose bit width would be determined by the accuracy with which the response is to be fitted. In addition, in some applications, the above calculations may not all be required. For example, to minimize jitter in an adaptive loop, $\Delta$ itself can be used to drive the adaptive loop, since the largest value of $\Delta$ corresponds to the smallest value of $\sigma_{ER}$. In this case, a precise estimate of $\sigma_{ER}$ is not needed and since measuring $\Delta$ provides enough information, the implementation can be simplified.

A. Analysis Using Linear Model

Fig. 5(a) shows the linear model of a PI-based CDR with BB-PD. Here $\psi_{ER}$ has contributions from the data ($\psi_{DATA}$), reference clock ($\psi_{REF}$), PI ($\psi_{PI}$), and BB-PD ($\psi_{PD}$). $R(n)$ is defined as

$$R(n) = E[\Gamma(k)\Gamma(k-n)]$$

$$= E[K_{PD}^2 \psi_{ER}(k)\psi_{ER}(k-n) + \psi_{PD}(k)\psi_{PD}(k-n)]$$

(1)

where we have assumed that $\psi_{ER}$ and $\psi_{PD}$ are uncorrelated and that all of the jitter is stationary so that $R(n)$ is not a function of time $k$. The variance of the PD output ($\sigma_{PD}^2$) is found by evaluating $R(n)$ at $n = 0$, which is proportional to the variances of $\psi_{ER}$ and $\psi_{PD}$

$$R(0) = E[\Gamma^2] = \sigma_{PD}^2 = K_{PD}^2 \sigma_{ER}^2 + \sigma_{PD}^2.$$  

(2)

Since a BB-PD only outputs +1, −1, or 0 when there is no transition, the variance of the PD output is also equal to the transition density of the input data $\alpha_T$. We can, therefore, solve for $\sigma_{ER}$ by setting (3) equal to $\alpha_T$. This gives

$$\sigma_{ER} = \sqrt{\alpha_T - \sigma_{PD}^2} / K_{PD}.$$  

(3)

The value of $\sigma_{PD}$ depends on the distribution of $\psi_{ER}$. For example, if $\psi_{ER}$ is Gaussian, $\sigma_{PD}$ can be derived [19] as

$$\sigma_{PD} = \sqrt{\alpha_T - \frac{2}{\pi} a_T^2}.$$  

(4)

The value of $\sigma_{PD}$ depends on the distribution of $\psi_{ER}$. For example, if $\psi_{ER}$ is Gaussian, $\sigma_{PD}$ can be derived [19] as

$$\sigma_{PD} = \sqrt{\alpha_T - \frac{2}{\pi} a_T^2}.$$  

(5)
Fig. 6. MATLAB simulation results with Gaussian \( \psi_{ER} \), comparing rms jitter estimated using (9) compared to actual jitter.

B. Simulation Results

The proposed technique was simulated in MATLAB by adding a square wave to Gaussian RJ and modeling the BB-PD as a sign operation. \( R(n) \) was estimated and the amplitude of the triangular wave \( \Delta \) was found by taking the average difference between the peaks and valleys of \( R(n) \). Fig. 6 shows the actual and estimated rms jitter \( (\sigma_{ER}) \), found using (9) and normalized to \( \sigma_{INJ} \). The plot shows good agreement between the two curves when \( \sigma_{ER} / \sigma_{INJ} \) is greater than approximately 2. However, the curves diverge when \( \sigma_{ER} \) becomes too small, since the derivations of Section III-A assumed that \( \sigma_{INJ} \ll \sigma_{ER} \), and also used the linearized model of the PD, which is only an approximation. To improve the accuracy of the jitter estimates when \( \sigma_{ER} \) is small, the effect of the injected jitter can be analyzed in terms of the PDF of \( \psi_{ER} \). Next, we provide this analysis for the case when \( \psi_{ER} \) is Gaussian.

C. Analysis Using Jitter PDF

Without jitter injection, the BB-PD simply outputs the sign of \( \psi_{ER} \). We can, therefore, write \( R(n) \) as

\[
R(n) = E[ \text{sgn}(\psi_{ER}(k)) \text{sgn}(\psi_{ER}(n-k))] \tag{11}
\]

Normalizing \( R(n) \) as in (10) allows us to ignore the effect of transition density. \( R'(n) \) can, then, be written as

\[
R'(n) = P[\text{sgn}(\psi_{ER}(k)) = \text{sgn}(\psi_{ER}(n-k))] - P[\text{sgn}(\psi_{ER}(n-k)) \neq \text{sgn}(\psi_{ER}(n-k))]. \tag{12}
\]

Assuming that the PSD of \( \psi_{ER} \) does not have large low-frequency content, then samples of \( \psi_{ER} \) separated by a long time \( n \) should be uncorrelated with each other. This means that the signs of \( \psi_{ER}(k) \) and \( \psi_{ER}(n-k) \), which are also assumed to have zero mean, are also uncorrelated. In other words, without jitter injection \( R'(n) \approx 0 \) for large \( n \).

When \( \psi_{INJ} \) is added, peaks and valleys are introduced into \( R'(n) \). \( R'(n) \) will reach its peak when \( \psi_{INJ}(k) \) and \( \psi_{INJ}(n-k) \) have the same sign and phase and reach a minimum when they have the opposite sign. Taking the case when \( \psi_{INJ}(k) = \psi_{INJ}(n-k) = A \), the peak of \( R'(n) \) can be written as

\[
R'_{\text{max}} = P[\text{sgn}(\psi_{ER}(k) + A) = \text{sgn}(\psi_{ER}(n-k) + A)] - P[\text{sgn}(\psi_{ER}(k) + A) \neq \text{sgn}(\psi_{ER}(n-k) + A)]. \tag{13}
\]

Assuming that \( \psi_{ER} \) is white and stationary, then any samples \( \psi_{ER}(n) \) and \( \psi_{ER}(n-k) \) of \( \psi_{ER} \) are also independent and identically distributed \( (\text{iid}) \) random variables. \( R'_{\text{max}} \) can then be evaluated in terms of the PDF of \( \psi_{ER} \)

\[
R'_{\text{max}} = P[\psi_{ER} > -A] + P[\psi_{ER} < -A] - 2P[\psi_{ER} > -A]P[\psi_{ER} < -A] \tag{14}
\]

\[
= [2P[\psi_{ER} > -A] - 1]^2 \tag{15}
\]

\[
= 2\left(\int_{-A}^{\infty} f_{\psi_{ER}}(\psi_{ER})d\psi_{ER}\right) - 1. \tag{16}
\]

If \( \psi_{ER} \) has a zero-mean Gaussian distribution, the integral in (16) can be replaced with the \( Q \)-function

\[
R'_{\text{max}} = 2Q\left(\frac{-A}{\sigma_{ER}}\right) - 1. \tag{17}
\]

Since \( R'_{\text{max}} \approx \Delta \), \( \sigma_{ER} \) can then be estimated from \( \Delta \) as

\[
\sigma_{ER} \approx \frac{-A}{Q^{-1}\left(1+\frac{\sqrt{2}}{2}\right)}. \tag{18}
\]

Fig. 7 shows \( \sigma_{ER} \) normalized to \( \sigma_{INJ} \) comparing the values as estimated using (9) and (18). As shown in Fig. 7, the two equations diverge when \( \sigma_{ER} / \sigma_{INJ} \) becomes less than approximately 2, which is also where (9) starts to become less accurate in Fig. 6. Fig. 8 re-plots the results of Fig. 6 comparing jitter as estimated using (9) from the linear analysis of Section III-A, compared to (18). As shown in Fig. 8, as \( \sigma_{ER} \) becomes small compared to \( \sigma_{INJ} \), the jitter estimate based on the Gaussian PDF becomes more accurate. Therefore in cases where the shape of the jitter PDF is known, or can be assumed to be Gaussian, an accurate jitter estimate can be obtained even when \( \sigma_{ER} \) is much smaller than \( \sigma_{INJ} \). When a PI is used for jitter injection, the smallest value of \( \sigma_{INJ} \) (which is also equal to \( A \)) is equal to the PI’s step size. In this paper, where the smallest PI step size is approximately 0.56 ps, it is therefore theoretically possible to estimate jitter in the hundreds of femtoseconds. However, several sources of error will limit the achieved performance.

D. Limitations and Sources of Error

The accuracy of the proposed technique depends on the precision of the injected jitter’s amplitude and on the accuracy
with which we can measure and model the effect of the injected jitter on \( R(n) \).

The amplitude accuracy of the jitter injection is primarily limited by nonlinearity in the PI, which can cause the injected jitter amplitude to differ from the expected value \( A \). Such inaccuracy will cause gain error as seen from the expressions in Section III-A. However, this error is mitigated whenever frequency offset is present, as in this paper. With frequency offset, the CDR continuously rotates through all PI codes. Any error in the PI step size will, therefore, be averaged across all PI codes, greatly reducing the impact of any nonlinearity on jitter measurement. However, when no frequency offset is present, the CDR edge PI will only alternate between two codes, making PI nonlinearity more important. To address this, the PI step size could be calibrated using similar techniques to those used in TDCs [11], [21].

Accurate modeling of \( R(n) \) also depends on making correct assumptions about the jitter’s PDF and spectral content. For example, if the jitter being measured has a uniform, rather than Gaussian distribution, (9) and (18) will become inaccurate, and have a gain error of approximately 38%. This is because uniformly distributed jitter leads to a lower \( K_{PD} \) value compared to the Gaussian case [22], changing the expression for \( \sigma_{PD} \) in (8), and the PDF used in (16).

Last, we assumed in Section III-C, that \( R(n) \) is zero for large \( n \), which makes \( \Delta \) easy to extract. This assumption will not be valid if the jitter has large low-frequency, or periodic content. Large low-frequency content will give \( R(n) \) a slowly decaying response as a function of \( n \), so that \( R(n) \) will not be zero until \( n \) is much larger than the time constant of the jitter. In the case of periodic jitter, which could include any excessive CDR dithering, a periodic waveform will also exist in \( R(n) \). If the periodic jitter amplitude becomes comparable to that of the injected jitter, this waveform will start to reduce the accuracy with which \( \Delta \) is estimated. However, both of these cases will be observable from the \( R(n) \) waveform and could be addressed by using large values of \( n \) and more detailed fitting or analysis of the \( R(n) \) waveform.

IV. IMPLEMENTATION

The proposed measurement technique was applied to the 28 Gb/s half-rate PI-based digital CDR shown in Fig. 9.

The input data are first equalized by a continuous time linear equalizer (CTLE) before being sampled by the half-rate front end. Half-rate clocks are generated by an injection-locked oscillator, which takes an external 14 GHz reference clock and generates quadrature clocks for the 7-bit CMOS inverter-based PIs shown in Fig. 10.

In each PI, the upper 2 bits of the PI code select the polarities of the two input clocks while the lower 5 bits control their respective weights by enabling and disabling inverter slices for each input. The polarity of each clock input is selected using a mux, which passes either the clock or its inverted version. A source-switched mux is used, which achieves higher bandwidth at the cost of increased signal feed-through. The 5-bit inverter weight is decoded to control 16 thermometer-coded slices. To save area, the LSB is implemented by adding a half-size LSB slice. Fig. 11 shows the control signals for \( CK_{OUT} \) in Fig. 10. As shown in Fig. 11, at each quadrant boundary, the polarity of either \( CK_I \) or \( CK_Q \) is switched to achieve full 2\( \pi \) phase rotation. Because the polarity signals \( POL_I \) and \( POL_Q \) switch when the weights of their corresponding clocks are zero, the effect of the mux switching is reduced.
Duty cycle distortion (DCD) in each PI was corrected from off-chip by programming an on-chip current-mode digital-to-analog converter (IDAC) shown in Fig. 10. To calibrate the DCD of $C_{K_{\text{EDGE}}}$, the IDAC code is adjusted to minimize the difference between the average of the even and odd PD outputs when the CDR is locked. This is because when locked, DCD in $C_{K_{\text{EDGE}}}$, which is aligned to the edges of the input data, will cause the average even and odd PD outputs to differ, with one becoming biased early and the other being biased late. To calibrate the DCD of $C_{K_{\text{DATA}}}$, the polarity of the CDR’s PD is inverted in a calibration mode, so that the CDR aligns $C_{K_{\text{DATA}}}$ to the edges of the input data instead of the center. DCD can, then, be corrected as before by comparing the even and odd PD outputs. This scheme assumes of course that the input data does not have significant DCD and can be used as a reference. To avoid this reliance on the duty cycle of the input data, an analog DCD calibration circuit could also be included on-chip [23].

Data from the analog front end is demuxed by 32 and sent to the digital core clocked at 875 MHz. In the digital core, the edge and data samples pass through PD logic, generating 32 parallel BB-PD outputs. Majority voting (MV) reduces this parallel output to a single binary output which drives the digital loop filter and $R(n)$ measurement blocks.

Each value of $R(n)$ is measured using a FIFO, which first delays the MV output by the desired delay $n$, and two counters that measure the correlation between the delayed and current MV outputs. One counter counts the total number of samples, while the second counts the number of correlated outputs over the same interval. In this paper, approximately twenty thousand samples of the MV output are used to measure each point of $R(n)$. This corresponds to a measurement time of roughly 23 $\mu$s per point. If incorporated in an adaptive loop, this could lead to adaptation times on the order of milliseconds, depending on the adaptation step size, the number of samples used to estimate each value of $R(n)$ (corresponding to the amount of averaging used), and the number of points of $R(n)$ being measured. While $R(n)$ is measured on-chip, in this paper $\Delta$ is extracted off-chip. In this implementation, measuring $R(n)$ from the downsampled PD output and after MV greatly simplifies the $R(n)$ measurement block. However, MV also introduces a lowpass filtering effect that must be considered.

A. Effect of Majority Voting

MV consists of a moving average lowpass filter (LPF) followed by a slicer as shown in Fig. 12(a). The slicer can be modeled identically to a BB-PD, as a linear gain $K_{\text{MV}}$ with a quantization noise source $\psi_{\text{MV}}$ added to its output as shown in Fig. 12(b). The values of both $K_{\text{MV}}$ and $\psi_{\text{MV}}$ depend on the PDF of the output of the moving average filter $\psi_A$. If the BB-PD output is a white, stationary random process with iid samples, then if the demux ratio $N$ is large, by the central limit theorem [24], the output of the moving average filter $\psi_A$, which is the sum of $N$ consecutive PD outputs is a random process that can be approximated as having a zero-mean Gaussian distribution. This gives a lower bound for $\sigma_A$

$$\sigma_A \geq \sqrt{N}\sigma_\Gamma.$$  \hspace{1cm} (19)

Note that $\sigma_A$ could be higher if the BB-PD outputs are correlated with each other. Analogous to the analysis in [19] for a BB-PD, and since there is no effect from transition density, the slicer gain $K_{\text{MV}}$ and the standard deviation of of $\psi_{\text{MV}} (\sigma_{\text{MV}})$ can, then, be approximated by

$$K_{\text{MV}} \approx \frac{1}{\sqrt{\pi/2} \sigma_A},$$  \hspace{1cm} (20)

$$\sigma_{\text{MV}} \approx \sqrt{1 - \frac{2}{\pi}} \approx 0.6.$$  \hspace{1cm} (21)

Because of its lowpass filtering effect, MV also heavily suppresses the quantization noise of the BB-PD $\psi_{\text{PD}}$ that precedes it, although in doing so, it also increases quantization noise [25]. Ignoring $\psi_{\text{PD}}$ for simplicity, $K_{\text{PD}}$ and $K_{\text{MV}}$ can be combined into an effective gain $K_{\text{EFF}}$ as shown in Fig. 12(c). To check the validity of this simplification, a CDR was...
Fig. 13. Simulated spectrum of (a) relative jitter ($\psi_{ER}$) and of (b) output of MV for CDR using nonlinear and linearized models of BB-PD and MV for two different loop gain settings.

Fig. 14. Normalized frequency response of MV.

Simulated in Simulink, replacing the nonlinear BB-PD and MV blocks of Fig. 12(a) with the simplified model in Fig. 12(c). Fig. 13 shows the spectrum of the CDR’s relative jitter $\psi_{ER}$ and output from the MV block with the CDR configured with two different loop gain settings. As shown in Fig. 13, the results of the nonlinear and linearized models match quite closely. The linear model does, however, fail to predict some spurs, due to its lack of nonlinearity.

Note that the model in Fig. 12(c) is identical to that of a BB-PD preceded by a LPF. The analysis of Section III therefore still applies, except that the LPF response of MV, which is shown in Fig. 14 for this design, also filters the jitter that can be observed from the PD output. This suppresses wideband jitter such as ISI jitter, and limits the effective measurement bandwidth to roughly 387 MHz in this implementation. Since the measurement bandwidth is inversely proportional to $N$, if wider measurement bandwidth is required, voting could be performed over fewer samples (by reducing $N$), or voting could be omitted entirely at the cost of higher complexity in the $R(n)$ measurement block.

**B. Jitter Injection**

In the half-rate architecture, jitter can be injected into the CDR’s edge clock without affecting the data sampling phase, by passing the PI code of the edge clock through a jitter injection stage which alternately adds or subtracts an offset from the PI code. As a result, square-wave jitter is injected at 437.5 MHz (half the clock rate of the digital core). Since this is well beyond the CDR’s loop bandwidth of approximately 10 MHz, it gets heavily suppressed by the CDR’s loop filter, minimizing its impact on the CDR’s operation.

Although the injected jitter frequency exceeds the 3 dB frequency of the MV filter response, as long as the injected square wave is aligned to the window over which voting is performed, it will not be attenuated. If, however, the injected jitter is misaligned, as can occur due to the delay of the analog front end, the square wave will be attenuated. This means that $\sigma_{INJ}$ will be smaller than is assumed by (9) and (18) and the estimated jitter will have a gain error. Fig. 15 shows this effect and plots the simulated gain of the estimated jitter as a function of the phase skew $\Phi$ of the injected jitter compared to the voting window. In the extreme case, if the injected jitter is skewed by $\pi/2$ compared to the voting window (corresponding to 16UI in this paper), $\psi_{INJ}$ can be completely averaged out, and jitter cannot be measured.

In this paper, we were able to measure the misalignment of the injected jitter by examining the subsampled PD output at the input to the MV stage, using a test mux. By comparing the first PD output within each voting window, with the other bits within the same voting window, we found the bit index where the average PD output changed sign. Because this bit position corresponds to the UI where the injected jitter changes sign, we determined that the injected jitter was misaligned by approximately 11UI. Based on Fig. 15(b), this corresponds to a measurement gain of approximately 2, meaning that the measured results had to be divided by this factor to correct for the gain error.

**V. MEASUREMENT RESULTS**

The complete CDR was fabricated in 28 nm CMOS and consumes 106.6 mW while operating at 28 Gb/s. Digital logic accounts for 32% of the total power consumption. Fig. 16 shows the die photograph of the fabricated chip. The digital core occupies relatively large area for two reasons. First,
the density of the synthesized logic was very low. Second, a large amount of test registers were included to monitor internal signals at up to 437.5 MHz. Although a digital power breakdown is not available, $R(n)$ measurement circuits only account for roughly 12.5% of the digital core area.

Fig. 17 shows the test setup. The chip was tested by wire bonding it to a custom printed circuit board (PCB). The CDR’s 14 GHz reference clock was supplied by a Rhode and Schwarz SMB100A signal generator. PRBS data were generated by an Agilent N4951B pattern generator, clocked by an N4960A serial BERT controller having an rms jitter of approximately 421 fs as measured using an Agilent EXA 9010A spectrum analyzer and shown in Fig. 18(a). Fig. 18(b) shows the eye diagram of the pattern generator output, which has 1.03 ps rms jitter when observed through short SMA cables using an Agilent DCA-86100D oscilloscope with precision timebase module.

Fig. 18(c) shows the phase noise of the CDR’s recovered clock when locked to PRBS31 data with 50 ppm frequency offset. The recovered clock has an rms jitter of approximately 950 fs. At low frequencies, the CDR’s phase noise tracks that of the input while mid-band phase noise is determined by the performance of the CDR and includes contributions from PI nonlinearity.

A. PI Linearity

Fig. 19(a) and (b) show the differential nonlinearity (DNL) and integral nonlinearity (INL) of the PI, measured by observing $CK_{EDGE}$ through a test mux, followed by a CMOS buffer chain and output driver. The highest DNL occurs when the PI changes quadrants as the PI’s muxes switch the polarity of one of the PI’s input clocks. Even though the weight of the clock whose polarity is switched is zero at the time the quadrant changes, simulations show that capacitive coupling between the output of the mux and output of the PI’s inverter banks contributes to a discontinuity in the PI’s response as the mux output changes. Mismatch between the PI inverter slices and
the half slice also causes even/odd variation in the PI response. In addition, it was observed that the DCD of \( \text{CK} \) the half slice also causes even/odd variation in the PI response.

**C. Jitter Measurement Results**

To assess the accuracy of our measurement technique, we measure the rms jitter of the CDR’s input data (\( \sigma_{\text{DAT}, \text{Meas}} \)) and recovered clock (\( \sigma_{\text{CK}, \text{Meas}} \)) at the input and output of the CDR using the Agilent DCA-86100D sampling scope, as shown in Fig. 17. All jitter measurement results were performed with the CDR reference clock having a frequency offset of 50–100 ppm with respect to the received data, while the CDR is locked to 28 Gb/s PRBS31 data. Since our technique estimates the relative rms jitter at the CDR input, we must compare our results to \( \sigma_{\text{ER}, \text{Meas}} \). Because simulations show that the CDR’s output jitter is dominated by PI and reference clock jitter and not data jitter, \( \psi_{\text{DAT}, \text{Meas}} \) and \( \psi_{\text{CK}, \text{Meas}} \) are essentially uncorrelated, meaning that \( \sigma_{\text{ER}, \text{Meas}} \) can be approximated as

\[
\sigma_{\text{ER}, \text{Meas}} \approx \sqrt{\sigma_{\text{CK}, \text{Meas}}^2 + \sigma_{\text{DAT}, \text{Meas}}^2}. \tag{22}
\]

In the following measurements, the PRBS generator output was applied to the DUT through low loss cables. As such, the CTLE only had to compensate for ISI induced by the short PCB trace and wirebond. Unless otherwise noted, the CTLE was always configured to achieve the lowest jitter, as confirmed by jitter tolerance measurements. Nonetheless, since \( \sigma_{\text{DAT}, \text{Meas}} \) is measured at the input of the PCB, it does not contain any residual ISI that might still be left at the CTLE output. However, as discussed earlier, MV also suppresses wideband jitter such as ISI in our implemented measurement scheme. As such, our estimated jitter can be fairly compared against \( \sigma_{\text{ER}, \text{Meas}} \).

The accuracy of the proposed technique was tested in several conditions, as the CDR’s input data jitter, loop filter settings and CTLE settings were varied.

**D. Measuring the Effect of Data Jitter**

First, \( \sigma_{\text{ER}} \) was estimated as different amounts of RJ were added to the CDR’s input data, by driving the external jitter input of the Agilent N4960A serial BERT controller and clock synthesizer with a NoiseCom NC6110 noise generator. Using this setup, RJ could be injected up to a bandwidth of 350 MHz.

Since this falls within our system’s measurement bandwidth, the filtering effect of MV does not have a large impact on these jitter estimates. (If RJ is injected over a much wider bandwidth, it could be filtered by MV, leading to offset and/or gain error in the estimated jitter.)

Fig. 22 compares \( \sigma_{\text{ER}} \) as estimated with our technique compared to \( \sigma_{\text{ER}, \text{Meas}} \), derived from oscilloscope measurements. As shown in Fig. 22, the results show good agreement, with results matching well within 1 ps.

**E. Measuring the Effect of Loop Gain Setting**

Next, we examine the ability of this technique to measure the effect of different loop gain settings on \( \psi_{\text{ER}} \). Fig. 23 shows...
the estimated and measured jitter as the CDR’s loop gain setting is swept for two test cases. In each case, the CDR is driven with reference clocks having different phase noise profiles and frequency offsets.

The results are shown in Fig. 23 and again show good agreement between the estimated and measured jitter. This example demonstrates how this technique can potentially be used in adapting a CDR’s loop gain to minimize jitter.

**F. Measuring the Effect of CTLE Setting**

Jitter was also estimated as the CTLE setting was swept. Since ISI-induced jitter is suppressed in this implementation, the proposed technique can only be used to observe the impact of the CTLE on the CDR’s recovered clock jitter. Since the oscilloscope measurements also do not include the effect of CTLE under or over-equalization, the estimated and measured jitter can be fairly compared. Given this, Fig. 24(a) again shows good matching between the jitter as estimated using our technique and based on oscilloscope measurements. Both techniques measure CTLE code 11 as having the lowest jitter.

To test whether this is optimal, the high-frequency jitter tolerance of the CDR was also measured at 100 MHz for the corresponding CTLE codes with jitter injection disabled. The results shown in Fig. 24(b) confirm that the highest jitter tolerance is achieved with code 11 as expected. This demonstrates that even though the proposed technique cannot fully characterize the effect of the CTLE on ISI, the jitter estimates can still be used to help optimize a CDR’s equalizer settings.

**G. Effect of Jitter Injection on CDR Operation**

Last, we also examine the effect of the injected jitter on jitter tolerance. Since only the edge clock is modulated, we expect the data samples to be relatively unaffected. To verify this, we perform jitter tolerance measurements with jitter injection enabled. As shown in the jitter tolerance results plotted in Fig. 25, although some degradation is seen compared to the normal case, the CDR is able to maintain error-free operation with $\text{BER} < 10^{-12}$, making it possible to perform measurements while the CDR remains operational. Note that jitter tolerance was measured with an adaptive loop gain...
algorithm [20] enabled, which automatically compensates for any changes in the CDR’s loop gain caused by variations in jitter. If the injected jitter is too large, it could affect the gain of the BB-PD, and consequently, the loop gain of the CDR. If jitter injection is used for on-chip adaptation, this technique can be used to characterize the effects of varying data jitter, CDR loop gain, and CTLE settings on CDR jitter performance. When combined with adaptation logic, this technique can be used for on-chip adaptation.

ACKNOWLEDGMENT

The authors would like to thank CMC microsystems for providing CAD tools and test equipment.

REFERENCES


TABLE I

<table>
<thead>
<tr>
<th>Reference</th>
<th>Analog Circuit Overhead</th>
<th>Digital Circuit Overhead</th>
<th>Off-Chip Post-Processing</th>
<th>Ext. Ref. Ck?</th>
<th>Freq./Data Rate</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>[8]</td>
<td>Analog PD</td>
<td>-</td>
<td>N/A</td>
<td>Yes</td>
<td>2.5 Gb/s</td>
<td>1.56 ps</td>
</tr>
<tr>
<td>[9]</td>
<td>TDC</td>
<td>Counters</td>
<td>Yes</td>
<td>No</td>
<td>3.36 GHz</td>
<td>&lt;0.7 ps</td>
</tr>
<tr>
<td>[10]</td>
<td>TDC</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>250 MHz</td>
<td>&lt;1 ps</td>
</tr>
<tr>
<td>[11]</td>
<td>TDC</td>
<td>-</td>
<td>Yes</td>
<td>Yes</td>
<td>1 GHz</td>
<td>0.1 ps</td>
</tr>
<tr>
<td>[12]</td>
<td>Sampler, Edge Detect</td>
<td>FIFO</td>
<td>Yes</td>
<td>Yes</td>
<td>5 Gb/s</td>
<td></td>
</tr>
<tr>
<td>[13]</td>
<td>Sampler, VCO</td>
<td>Counters</td>
<td>No</td>
<td>No</td>
<td>6 Gb/s</td>
<td>&lt;0.3 ps</td>
</tr>
<tr>
<td>[16]</td>
<td>Eye-Mon.</td>
<td>-</td>
<td>Yes</td>
<td>No</td>
<td>40 Gb/s</td>
<td></td>
</tr>
<tr>
<td>This work</td>
<td>None</td>
<td>Counters</td>
<td>Yes</td>
<td>Yes</td>
<td>28 Gb/s</td>
<td>+0.33 ps</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.25 ps</td>
</tr>
</tbody>
</table>

VI. CONCLUSION

A jitter measurement technique has been proposed which makes use of existing circuits in a half-rate PI-based CDR, allowing jitter to be measured without having to add additional analog circuits. Table I compares this work to prior measurement techniques. Similar to most prior techniques, off-chip post-processing is used to recover the estimated jitter, although such processing could easily be incorporated on-chip.

Measurements demonstrate how this technique can be used to characterize the effects of varying data jitter, CDR loop gain, and CTLE settings on CDR jitter performance. When combined with adaptation logic, this technique can be used for on-chip adaptation.
Joshua Liang received the B.A.Sc. degree in engineering science and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2007, 2009, and 2017, respectively.

From 2009 to 2011, he was an Analog Designer with Zarlink Semiconductor (now Microsemi), Ottawa, ON, Canada, where he worked on circuits for low-jitter clock synthesis. Since 2012, he has been involved in research on wireline transceivers. In 2017, he joined Huawei, Markham, ON, Canada, where he focused on the research and development of high-speed interfaces.

Dr. Liang was a recipient of the Ontario Graduate Scholarship, the Ontario Queen Elizabeth II Graduate Scholarship in Science and Technology, and the NSERC Postgraduate Scholarships. He received the Analog Devices’, Toronto, ON, Canada, Outstanding Student Designer Award in 2016.

Ali Sheikholeslami (S’98–M’99–SM’02) received the B.Sc. degree in electrical engineering from Shiraz University, Shiraz, Iran, in 1990, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1994 and 1999, respectively.

In 1999, he joined the Department of Electrical and Computer Engineering, University of Toronto, where he is currently a Professor, an Associate Chair, and a Researcher. He was on research sabbatical with the Fujitsu Laboratories, Kawasaki, Japan, in 2005–2006, and with Analog Devices, Toronto, ON, Canada, in 2012–2013. He has co-authored over 70 journal and conference papers and a graduate-level textbook entitled Understanding Jitter and Phase Noise, and holds 10 patents. His current research interests include analog and digital integrated circuits, high-speed signaling, and very large-scale integration memory design.

Dr. Sheikholeslami has received numerous teaching awards including the 2005–2006 Early Career Teaching Award and the 2010 Faculty Teaching Award, both from the Faculty of Applied Science and Engineering at the University of Toronto. He served on the Memory, Technology Directions, and Wireline Subcommittees of the ISSCC in 2001–2004, 2002–2005, and 2007–2013, respectively. He currently serves as the Education Chair for both ISSCC and SSCS, and as a member of the SSCS Administration Committee. Being the SSCS Education Chair, he oversees the SSCS Distinguished Lecturer Program, Webinars, Circuit Contests, and other educational activities.

He was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS I during 2010–2012, and the Program Chair of the 2004 IEEE ISMVL. He is an Associate Editor of IEEE Solid-State Circuits Magazine, in which he has a regular column entitled Circuit Intuitions. He is currently a Registered Professional Engineer at Ontario.

Hirotaka Tamura (M’02–SM’10–F’13) received the B.S., M.S., and Ph.D. degrees in electronic engineering from Tokyo University, Tokyo, Japan, in 1977, 1979, and 1982, respectively.

In 1982, he joined the Fujitsu Laboratories, Kawasaki, Japan. He was involved in the development of different exploratory devices such as Josephson junction devices and high-temperature superconductor devices. In 1996, he focused on CMOS high-speed signaling and was involved in the development of a multichannel high-speed I/O for server interconnects. Since 1996, he has been involved in the area of architecture- and transistor-level design for CMOS high-speed signaling circuits. Since 2014, he has been expanding his area to cover devices, circuits, and architectures for post-Moore-era computing.

Hisakatsu Yamaguchi received the B.S. degree in electrical engineering from the Tokyo University of Science, Chiba, Japan, in 1994, and the M.S. degree in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1996.

In 1996, he joined the Fujitsu Laboratories Ltd., Kawasaki, Japan, where he was involved in research on DRAMs with high-speed I/Os and was responsible for developing MPEG4 Codec ICs. He is currently developing high-speed I/Os for high-end servers and super-computers.

Mr. Yamaguchi served on the Technical Program Committees of ISSCC from 2012 to 2016.