A 20-Gb/s Full-Rate Linear Clock and Data Recovery Circuit With Automatic Frequency Acquisition

Jri Lee, *Member, IEEE*, and Ke-Chung Wu

Abstract—A 20-Gb/s full-rate clock and data recovery circuit employing a mixer-type linear phase detector and automatic frequency locking technique is described. The phase detector achieves high-speed operation by mixing the clock with the data-transition pulses, providing output proportional to the phase error. The frequency acquisition loop utilizes the data phases rather than the clock phases to distill the frequency difference, and no external reference is used in this design. Fabricated in 90-nm CMOS technology, this circuit reveals rms and peak-to-peak jitter of 480 fs and 4.22 ps in response to a $2^{31}-1$ PRBS on the recovered clock while consuming 154 mW from a 1.5-V supply.

Index Terms—Bit error rate (BER), clock and data recovery (CDR), frequency detector (FD), jitter generation, jitter tolerance, linear phase detector (PD).

I. INTRODUCTION

CLOCK and data recovery (CDR) circuits have found extensive usage in modern communication systems. Fig. 1(a) summarizes the data rate of representative CMOS circuits published since 1995. It can be clearly seen that the speed of the circuits improves even faster than that of the devices (the device transit frequency f_T improves only 2.3 times from 0.18-μm to 90-nm nodes [1]). This is largely due to the continuous progression of CDR architectures and circuit techniques. In addition, the power efficiency also improves steadily approximately 1.4 times per year, which is primarily attributed to scaling. Unfortunately, the scaling of analog circuits has less influence on speed because the passive devices and interconnects do not scale. Overall, novel architectures, new broadband circuit techniques, and advanced processes are essential for modern CDR design.

Based on the operation of phase detectors (PDs), CDR circuits are traditionally classified into two categories [2]: binary (bang-bang or Alexander) [3] and linear (Hogge) [4]. The jitter transfer and jitter tolerance of the former vary with the input (phase) magnitude. Linear PDs, on the other hand, present a straightforward design flow, since a typical linear PLL model can fit the requirement very well. However, in contrast to their binary counterparts, linear CDR circuits thus far confront a speed limitation at around 10 Gb/s. This is primarily because the linear operation usually involves pulse generation and
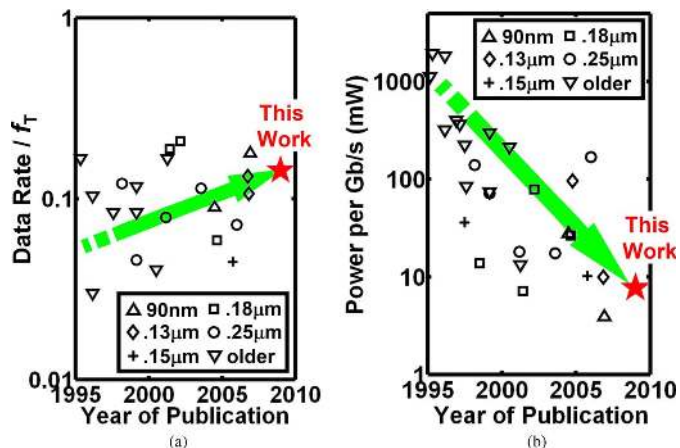

Fig. 1. Evolution of CMOS CDR circuits in terms of (a) data rate and (b) power efficiency. Technology information obtained from [1].

pulsewidth comparison, both of which are very challenging at high speed. Let us take a Hogge PD [Fig. 2(a)] as an example. In 90-nm CMOS, since the rise and fall times of a fan-out-of-4 inverter are as large as 33 ps, it is impossible to build a digital PD running beyond 10 Gb/s. Current-mode logic (CML) can speed up the operation to some extent, but the fundamental issue still remains. Fig. 2(b) shows the simulated linear range of standard full-rate Hogge PDs designed with CML topologies in 90-nm and 0.13-μm CMOS technologies. Even though the circuits are optimized (e.g., adding proper delays to compensate for the skew), the operation range still drops dramatically after 1 Gb/s. This is because at high speed, the finite transition times compress the width of the reference pulses, making the PD characteristic imbalanced. Other limitations of the Hogge type PDs were analyzed in [2]. CK-to-Q delay of the flip-flops in a Hogge PD could never be fully compensated over process, supply voltage, and temperature (PVT) variations. Half-rate [5] or even quarter-rate operation relaxes the stringent speed requirement. In fact, parallelism has been widely adopted in modern CDR designs [6], [7]. However, it also causes other design issues. For example, the phase error between the multiple clock phases could cause significant skew and jitter, potentially requiring calibration circuitry that could increase design complexity as well as power consumption. The large area would be another negative concern if we were to develop a multi-channel receiver [8].

Another design consideration is the implementation of a frequency acquisition loop. The Pottbacker frequency detector (FD) [9] and other similar approaches [10], [11] can determine
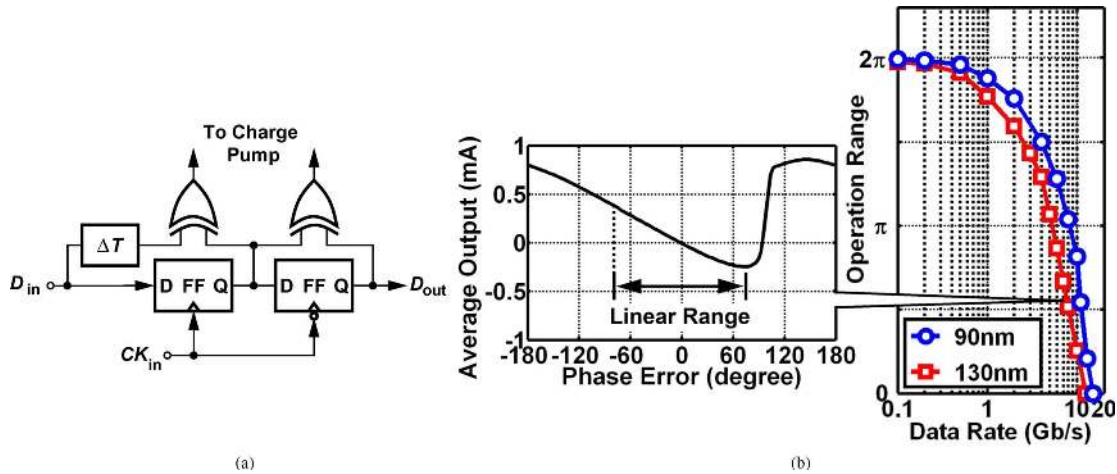
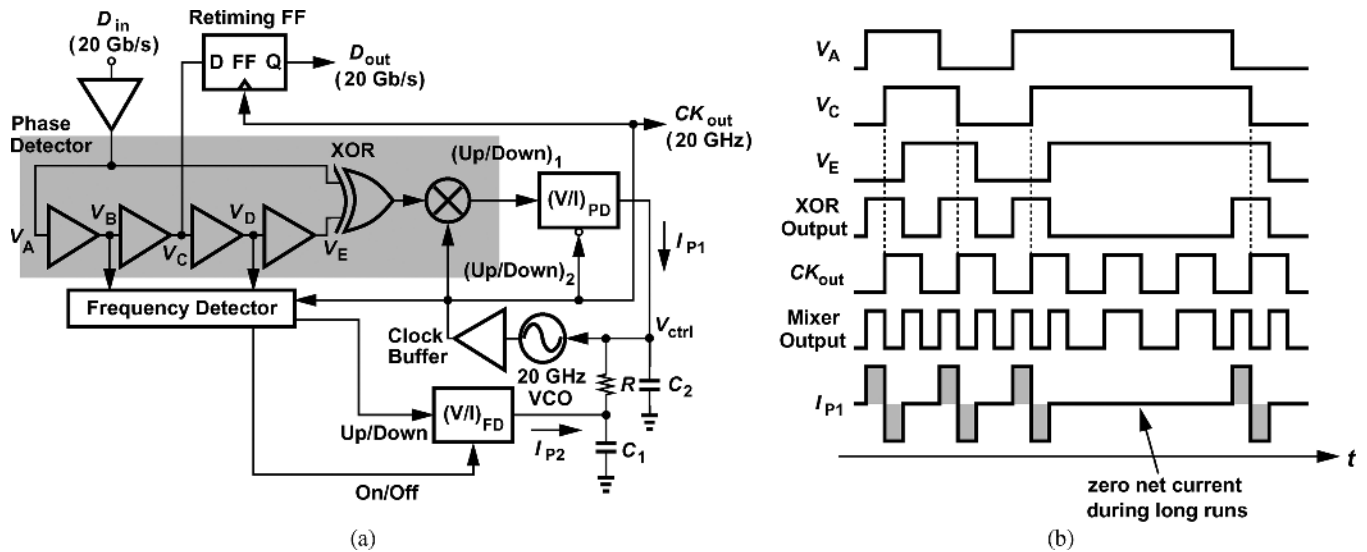Fig. 2. (a) Conventional full-rate Hogge PD. (b) Simulated operation ranges.

Fig. 3. (a) CDR architecture, (b) "conceptual" PD operation upon lock.

the frequency error without an external reference, but they require quadrature clocks to do so. For sub-rate architectures, this type of frequency detection needs to further break down the clock phases, making the clock generation and distribution more difficult. Creating multiphase clocks with *LC* oscillators would lead to higher phase noise as well, since the oscillators may operate at frequencies away from the resonance of the tanks [11]. In this paper, we propose an automatic frequency detection loop that activates itself as the loop is out of lock, and turns off when the frequency acquisition is accomplished. Minimizing the hardware and power consumption, this technique also requires no external reference nor lock detector. We discuss the realization of the FD in Section II.

This paper presents the design and analysis of a 20-Gb/s CDR circuit in 90-nm CMOS technology. Using a mixer-type full-rate linear PD structure, this work completely eliminates the above issues. The reference-free frequency detector makes use of the by-product of the PD, necessitating no quadrature clocks and saving significant power. The automatic turn-off

function considerably simplifies the design, because the digital blocks for frequency comparison and lock detection are eliminated. The prototype achieves output clock jitter of 480 fs,rms and 4.22 ps,pp in response to a $2^{31}-1$ PRBS while consuming 154 mW from a 1.5-V supply.

The paper is organized as follows. Section II describes the CDR topology, describing design issues and considerations. Section III presents the building blocks of the circuits. Transistor-level analysis of the blocks is also included. Section IV summarizes the measurement results.

II. CDR ARCHITECTURE

The proposed CDR circuit is shown in Fig. 3(a). It incorporates a full-rate voltage-controlled oscillator (VCO), a linear PD based on a transition detector and mixer, an automatic frequency detector, the corresponding V-to-I converters, and a retiming flip-flop. We describe the operation and requirements of the PD and FD in the following subsections.

A. Phase Detection

As described in Section I, the pulse generation and comparison involved in Hogge PD limit the speed because of the long rise and fall signal edges of the XOR gates and finite CK-to-Q delay in the flip-flop. These issues can be alleviated by using a mixer-based phase detector. As shown in Fig. 3(a), the input data passes through a chain of delay cells, providing a total delay (from V_A to V_E) approximately equal to half a bit (25 ps). An XOR gate examines this fixed phase difference, creating a pulse nominally equal to 25 ps upon occurrence of data transitions. Acting as a reference for phase detection, this pulse sequence is mixed with the clock from the VCO. To create a rough sketch on the PD operation, we plot conceptual waveforms of important nodes under locked condition in Fig. 3(b).[1] When a data edge is present, the mixer produces an output pulse whose width is proportional to the phase difference between the XOR output and the clock. This result can be used for phase alignment. During consecutive bits, on the other hand, the mixer generates a periodic signal which is in phase with the clock. This signal has a zero average, given that the duty cycle of the clock is 50%. In other words, for random data the mixer provides an average output voltage proportional to the phase error between the two inputs. A V-to-I converter then translates the voltage into current and injects it into the loop filter. As a result, the center tap V_C always aligns with the clock, and data sampling can be accomplished in the retiming flip-flop using the falling clock edges.

What happens if the clock duty cycle deviates from 50%? If $(V/I)_{PD}$ were solely driven by the mixer, the distortion would lead to finite residue current and modulate the control voltage. Fortunately, we can apply the complement of the clock ($\overline{CK}$) into $(V/I)_{PD}$ to overcome this difficulty. Since the clock and the mixer's output are in phase, it completely cancels out the periodic disturbance for consecutive bits. The $(V/I)_{PD}$ design is presented in Section III-C. As illustrated in Fig. 3(b), I_{P1} [the output current of $(V/I)_{PD}$] reveals pure zero output during long runs.

In real operation with finite bandwidth, however, sharp transitions at 20 Gb/s cannot be created. Fortunately, unlike the Hogge PD, this work need not generate narrow pulses at all. At such a high speed, the clock and the XOR outputs become rounded because the higher order harmonics are suppressed. As a result, the phase detection is nothing more than mixing two sinusoidal signals. As illustrated in Fig. 4(a), if the delay is exactly 25 ps, the XOR gate and the clock outputs can be simply modeled as

$$V_{XOR} = \begin{cases} A\cos(\omega t + \pi), & \text{for data transitions} \\ -A, & \text{for long runs} \end{cases} \quad (1)$$

$$CK_{out} = B\cos(\omega t + \theta) \quad (2)$$

[1]For simple illustration, we temporarily assume infinite bandwidth and therefore obtain sharp transitions, which is not true in reality. Fig. 4(a) presents more realistic waveforms.
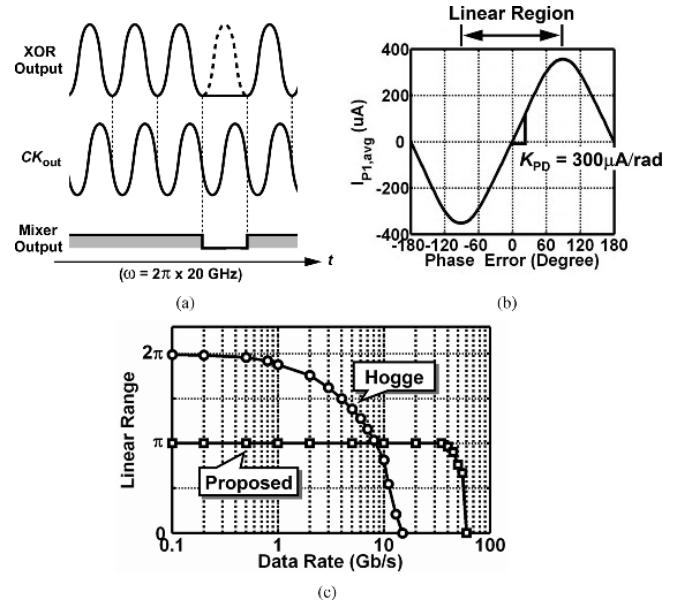

Fig. 4. (a) Proposed PD operation. (b) Average output of $(V/I)_{PD}$ as a function of phase error, (c) comparison of operation bandwidth.

where $\omega = 2\pi \times 20$ GHz, A and B denote the magnitudes of the two signals, respectively, and θ the phase error. The mixer's output thus becomes

$$V_{mixer} = \begin{cases} \frac{AB}{2}[\cos(\theta - \pi) \\ \quad + \cos(2\omega t + \theta + \pi)], & \text{for data transitions} \\ -AB\cos(\omega t + \theta), & \text{for long runs.} \end{cases} \quad (3)$$

In other words, when the data edge is presented, the phase difference is obtained as a near-dc output $(AB/2)\cos(\theta - \pi)$. The second-order term is filtered out by means of intrinsic parasitics. Simulation suggests a suppression of 34 dB. The fundamental modulation during consecutive bits is eliminated by $\overline{CK}$ as described in the previous paragraph. Simulated at the transistor level in 90-nm CMOS, Fig. 4(b) depicts the average output current as a function of phase error. As expected, it presents a sinusoidal characteristic with a PD gain [together with $(V/I)_{PD}$ of 300 μA/rad in the vicinity of origin. A linear operation region of about 180° is obtained. To compare with the Hogge PD, we also plot the simulated linear ranges as shown in Fig. 4(c). The proposed structure achieves a large operation bandwidth all the way from dc to 40 Gb/s. Note that the sharp pulses of I_{P1} in Fig. 3(b) do not exist in reality either. Upon phase locking, the output current I_{P1} would only be modulated by less than 30 μA at a rate of 20 GHz because of the low-pass filtering effect. Nonetheless, it is equivalent to a 20-GHz periodic phase modulation on the input and is rejected by the limited loop bandwidth of the CDR (≈ 15 MHz).

It is noteworthy that under locked condition, the clock edges always align with the center of the generated pulses, whether or not the delay from V_A to V_E ($\Delta T_{A\rightarrow E}$) is exactly 25 ps. Fig. 5 reveals two cases where the delay is longer and shorter than half bit period. Obviously, V_C still coincides with the clock, keeping an optimal phase for data retiming. Note that the buffered clock
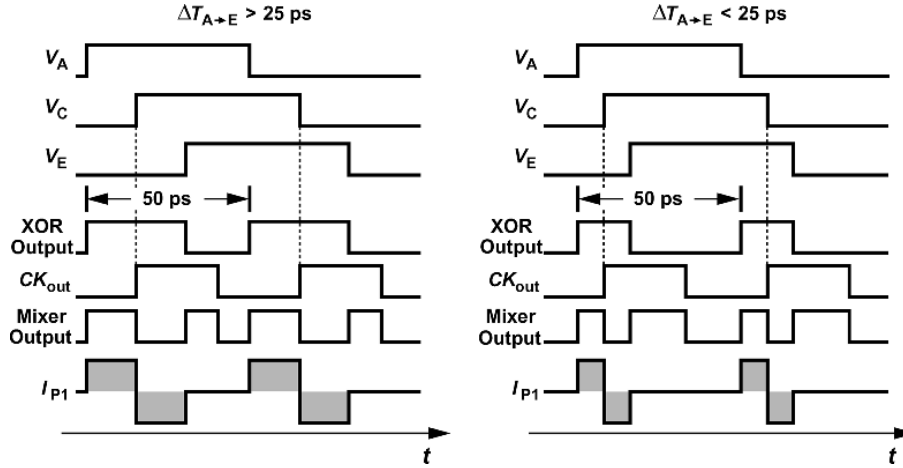
Fig. 5. Waveforms of important nodes as $\Delta T_{A \rightarrow E}$ deviates from 25 ps.
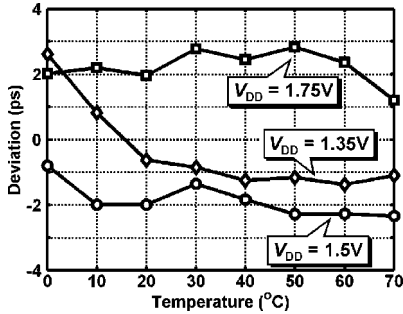

Fig. 6. Simulated deviation of sampling points.

CK_{out} directly drives the mixer, $(V/I)_{\text{PD}}$, and the retiming flip-flop simultaneously, so no phase error is expected. The only possible source of alignment degradation is the delay of the XOR gate. Here, we incorporate inductive peaking in the XOR gate to minimize this effect. The XOR gate delay is estimated to be less than 5 ps. Such an intrinsic alignment proves robust over PVT variations. To verify it, we examine the sampling points of the retiming flip-flop under locked condition, and plot the simulated deviation of the sampling points over different conditions in Fig. 6. The clock edge deviates from the data center of V_C by only 57 mUI for the worst case. The delay actually requires no manual tuning at all.

The reader may also wonder how large the delay error it can tolerate. Using the more accurate model with round (sinusoidal) pulses in Fig. 4(a), we assume the delay deviation is $\Delta\%$ with respect to a clock period. The imbalanced XOR pulses thus yield a corresponding factor $\sin[\pi(1/2 + \Delta\%)]$ for the fundamental term of its Fourier series. The resulting near-dc term now becomes

$$V_{\text{mixer}} = \frac{AB}{2} \sin\left[\pi\left(\frac{1}{2} + \Delta\%\right)\right] \cos(\theta - \pi). \quad (4)$$

That is, the input-output characteristic keeps the sinusoidal shape but with lower PD gain. Fig. 7(a) illustrates the average output of PD for different delay deviations. It is obvious that

for $\Delta\% = \pm 25\%$ (deviation $= \pm 12.5$ ps), the PD gain in the vicinity of origin gets reduced by 30%, just as predicted by (4). Simulation suggests the maximum $\Delta\%$ to be less than 7.1% with variations of $\pm 10\%$ supply, $0 \sim 100^{\circ}$C temperature, and different corners [Fig. 7(b)]. In other words, the delay needs no additional tracking or control, since it is irrelevant to the clock.

B. Frequency Acquisition

Due to the limited lock range, the correct frequency must be acquired before phase locking. Conventional dual-loop architecture [12], [13] requires a local frequency reference. Some referenceless approaches such as Pottbacker frequency detector can avoid the need for a reference by using quadrature clocks to distill the frequency information. We propose here a more compact solution for frequency acquisition.

In this CDR design, the nominal 25-ps delay from V_A to V_E implies a 12.5-ps delay from V_B to V_D, which allows us to extract the frequency difference. Indeed, the 12.5-ps data delay corresponds to a 90° phase shift[2] of a 20-GHz clock, making it possible to realize a rotational frequency detector without using quadrature clocks. The proposed FD is shown in Fig. 8(a). Here, the clock is sampled by using the PD's by-product V_B and V_D, producing two outputs Q_1 and Q_2, respectively. Whether Q_1 is leading or lagging Q_2 depends on the polarity of the frequency error [Fig. 8(b)]. Similar to the Pottbacker FD in [9], Q_1 is further sampled by Q_2 through another flip-flop. The polarity of frequency error Q_3 is therefore obtained. The up/down signal is subsequently applied to a second V-to-I converter $(V/I)_{\text{FD}}$, which injects a current into the loop filter and corrects the VCO frequency accordingly. Like the Pottbacker FD, the $V_B \rightarrow V_D$ delay need not be exactly 12.5 ps. Simulation shows that a range of $\pm 27\%$ on the delay variation is tolerable for the FD to function properly.

To minimize the disturbance to the VCO, the frequency acquisition should be turned off upon lock and re-enabled if

---

[2]As a matter of fact, a precise 90° separation on adjacent phases is not mandatory. A looser condition (such as 80° or 100°) would still allow an FD to achieve similar performance, given that the initial frequency deviation stays within a certain range.
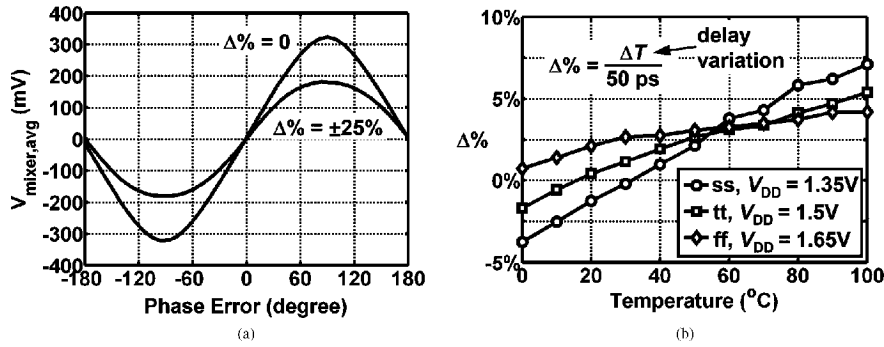
Fig. 7. (a) Simulated PD characteristic under delay variation. (b) $\Delta\%$ for different PVT variations.
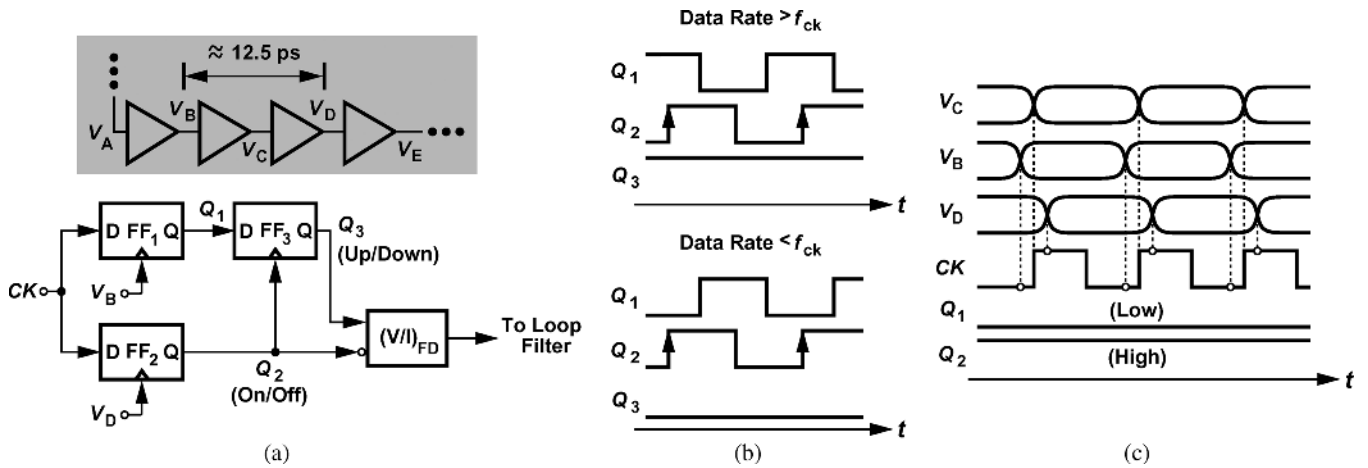
Fig. 8. (a) Frequency detector. (b) FD operation. (c) Waveforms under phase locking.
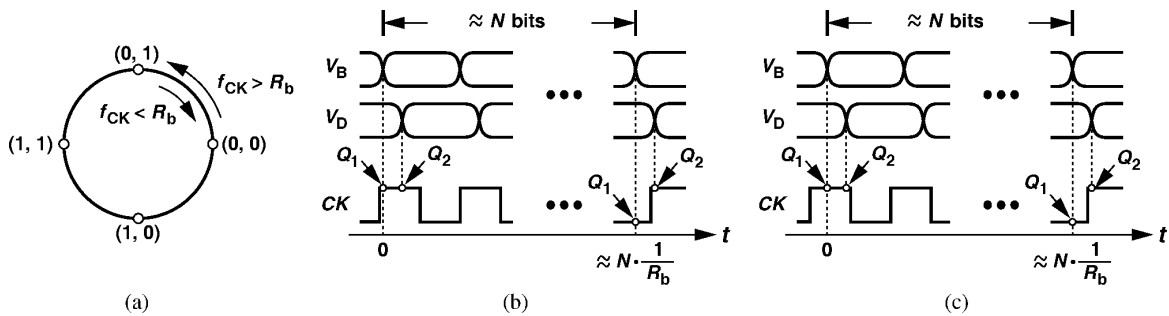
Fig. 9. (a) States of (Q_1, Q_2). (b) The worst case and (c) the best case of states changing.

necessary. A simple modification can achieve such a tristate operation. As illustrated in Fig. 8(c), when the phase lock is accomplished, Q_1 and Q_2 would stay low and high, respectively. Following the design of [15], we apply $\overline{Q_2}$ (the reverse of Q_2) to $(V/I)_{FD}$ to fulfill the automatic switching off when the frequency acquisition is completed. As compared with typical realizations that usually involve significant peripheral control circuits, this work achieves a compact yet powerful design. The spontaneous operation suppresses undesirable interference onto the loop filter and saves considerable power and area for lock detector, logic controller, and other auxiliary circuits. Note that other FD designs such as [16] using quadrature clocks can

achieve graceful shut down function as well, but the FD in [16] requires at least four flip-flops and four NAND gates.

It is instructive to examine the FD operation in detail and quantize the capture range. The states of Q_1 and Q_2 can be characterized in Fig. 9(a), where the rotating direction indicates the sign of the beat frequency. For example, a clockwise rotation suggests the clock frequency (f_{CK}) is less than the data rate (R_b). Of course, the rotation rate represents the beat frequency. For such an FD to make a right decision on every sampling, we must require the states of Q_1 and Q_2 to jump no more than one step at a time. That is, the average output current $I_{P2,avg}$ remains fixed (either positive or negative) for low frequency error,

forming a binary characteristic. This situation continues until the above condition is violated. To determine the points where $I_{P2,avg}$ begins to drop, we study one worst case as illustrated in Fig. 9(b). Here, without loss of generality, we assume f_{CK} is less than R_b and the transition of V_B (and thus Q_1) is already very close to the clock edge. Starting from $(1, 1)$, the state either stays at $(1, 1)$ or moves to $(0, 1)$ in the next sampling. As we know, for a PRBS of $2^N - 1$, the longest run length between transitions is N bits. Since the longest run accumulates the most error, we can determine the largest beat frequency at which the average output current begins to degrade. That is, after N bits, the sampled Q_2 remains high. The boundary condition gives

$$N \cdot \left| \frac{1}{f_{CK}} - \frac{1}{R_b} \right| = \frac{1}{4f_{CK}}. \tag{5}$$

It follows that the deviation is given by

$$\Delta f_1 \triangleq |f_{CK} - R_b| = \frac{R_b}{4N}. \tag{6}$$

If $N = 7$, for example, the binary range is equal to $\pm 3.6\%$. It can be easily proven that Δf is symmetric with respect to the origin. Strictly speaking, the use of N bits as the longest period of error accumulation is not exactly correct because the flip-flops in the FD are single-edge triggered. The actual accumulation time would be longer than $N \cdot (1/R_b)$. For example, the longest distance between two adjacent rising edges in a $2^7 - 1$ PRBS is 13 bits, so the binary characteristic begins to roll off at around $R_b/(4 \cdot 13)$.

The above analysis is based on the worst-case scenario. In practice, Q_1 may stay far away from the clock edge before the N-bit long run. The best-case scenario can be shown in Fig. 9(c), where the phase error accumulated over N bits must be less than a half rather than a quarter of a clock cycle in order to maintain a saturated $I_{P2,avg}$. Thus, the widest binary range would be twice as large as that in (6):

$$\Delta f_2 = \frac{R_b}{2N}. \tag{7}$$

Depending on the initial phase relationship, the binary range in reality lies between the two extremes Δf_1 and Δf_2.

The FD performance begins to degrade beyond the binary range, as the sequence of Q_1 and Q_2 becomes chaotic and the erroneous samplings occur in FF_3. It is expected to see the average output eventually approaching zero as the sequential states of Q_1 and Q_2 become totally random, i.e., no reliable average on Q_3 can be obtained. The vanishing point can be roughly estimated as follows. For random data, the expected interval between two adjacent transitions is two bits. Since FF_1 and FF_2 are single-edge triggered, V_B and V_D on average sample the clock every four bits. Now, if the frequency error is so significant that (Q_1, Q_2) steps more than one state in each sampling, the beat-frequency sequences become totally corrupted and the FD has no way to judge the polarity. Under such a circumstance, we have

$$4 \cdot \left| \frac{1}{f_{CK}} - \frac{1}{R_b} \right| \geq \frac{1}{2f_{CK}}. \tag{8}$$

Fig. 10. Simulated FD characteristic.

$M_{1,2}$	$M_{3,4}$	R_{SS}	C_{SS}	L	I_B
$\frac{12}{0.1}$	$\frac{2}{0.5}$	75 Ω	2 pF	700 pH	4 mA

Fig. 11. VCO design.

It follows that

$$f_{CK,max} = \frac{9}{8}R_b \quad \text{and} \quad f_{CK,min} = \frac{7}{8}R_b. \tag{9}$$

In other words, the capture of the FD is about $\pm 12.5\%$. In fact, the vanishing point is slightly larger than the prediction of (9) because of the finite rising and falling times. Fig. 10 reveals the simulated FD characteristic for a $2^7 - 1$ input data sequence. Here, the roll-off point locates at 400 MHz as expected. If we set half the peak current as a threshold, the useful working range would be as large as 3.7 GHz, which is well exceeding the VCO tuning range (≈ 1.1 GHz). Note that it is more or less subject to change for different initial phase conditions. The wide pull-in range ensures correct operation for the loop.

III. BUILDING BLOCKS

A. VCO and Clock Buffer

The LC-tank VCO incorporated in this design is shown in Fig. 11. The bias current, inductors, and device sizes are properly chosen such that it reaches optimal performance. The resistor R_{SS} is used to slightly lift up the output common-mode level of CK_{out} by $100 \sim 200$ mV so as to relax the voltage headroom of the subsequent buffers (realized as differential pairs).

While the 20-GHz clock design is straightforward, clock distribution is relatively challenging. Simulation shows that the
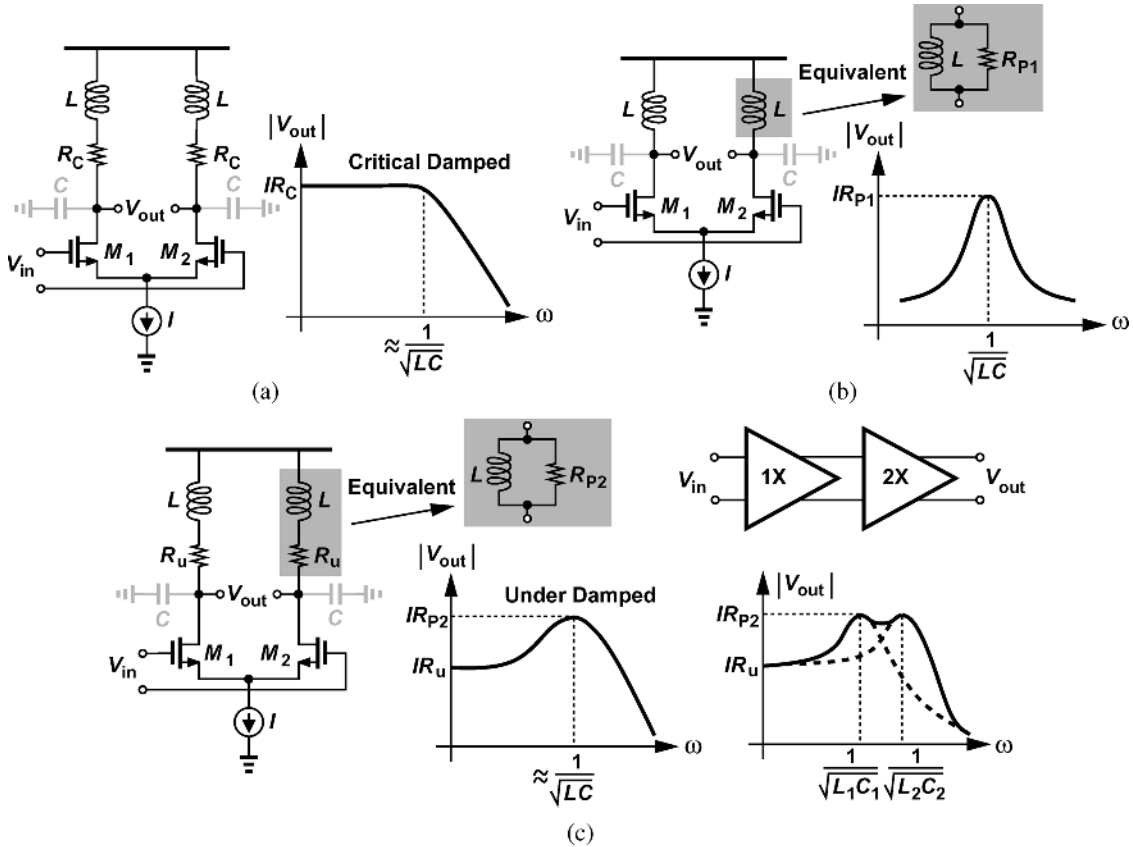
Fig. 12.   Clock buffer with (a) critical inductive peaking, (b) pure inductive loads, and (c) underdamped peaking with cascade realization [7].

clock buffer needs to drive a total capacitance of more than 120 fF including the routing. A conventional CML buffer with resistive loads fails to provide large swings due to the bandwidth limitation. Inductive peaking can improve the bandwidth to some extent, but it also suffers from other tradeoffs. Denoting the peaking inductor, the loading resistor, and the parasitic capacitance as L, R_C, and C, respectively, we obtain the transfer function of a regular differential pair with inductive peaking [Fig. 12(a)] as [12]

$$\frac{V_{\text{out}}}{V_{\text{in}}} = g_{m1,2} R_C \cdot \frac{s + 2\zeta\omega_n}{s^2 + 2\zeta\omega_n s + \omega_n^2} \cdot \frac{\omega_n}{2\zeta} \qquad (10)$$

where $\omega_n^2 = (LC)^{-1}$ and $\zeta = (R_C/2)\sqrt{C/L}$. Generally speaking, we need $\zeta \approx 0.7$ to reach a flat response, making the bandwidth approximately equal to ω_n. Similar to the small-signal analysis, the differential pair steers the tail current I completely under large input and presents a flat $|V_{\text{out}}|$ of IR from dc to approximately the bandwidth ω_n. Note that the large signal behavior resembles the small-signal response and it can be verified by simulation. The key point is, since the flip-flops require a swing of at least 500 mV, R must be $150\,\Omega \sim 200\,\Omega$ or larger. Otherwise I must be increased, which in turn leads to bigger device sizes and larger C. If we were to keep the optimal ζ by increasing L, the bandwidth would be decreased. In other words, it is difficult to realize a large swing output by using inductive peaking only. In fact, it is a waste to keep the voltage gain all

the way down to dc, because the clock buffer only operates at around 20 GHz. We thus resort to other resonance techniques with higher efficiency.

Another possible approach attempting to deliver large swing is to employ pure inductive loads, which resonates out the parasitic capacitance C at the desired frequency [Fig. 12(b)]. Indeed, this method produces a swing of IR_{P1} in the vicinity of $1/\sqrt{LC}$, where R_{P1} represents the loss of L as an equivalent resistance in parallel. With a quality factor Q above 4, the buffer in Fig. 12(b) can create a large output swing easily. However, it is very challenging to precisely line up the resonance frequencies of the VCO and the buffer. For instance, if $Q = 5$, then a 50% magnitude degradation would occur if the two resonance frequencies deviate from each other by 17%. Practical application thus becomes very hard to implement due to PVT variations. The output swing is not predictable since the Q of the on-chip inductors is hard to control.

The above difficulties can be alleviated by introducing an underdamped peaking. As depicted in Fig. 12(c), we keep the loading resistors of Fig. 12(a) but reduce the value to R_u. The output swing starts with a lower value of IR_u from dc and presents a gradual peaking of IR_{P2} at $1/\sqrt{LC}$. Here, we convert the series L-R network into an equivalent parallel combination L-R_{P2}. The difference between Fig. 12(b) and (c) is that the R_{P2} in Fig. 12(c) now becomes predictable, because the physical resistor R_u is fully under our control. In other words,
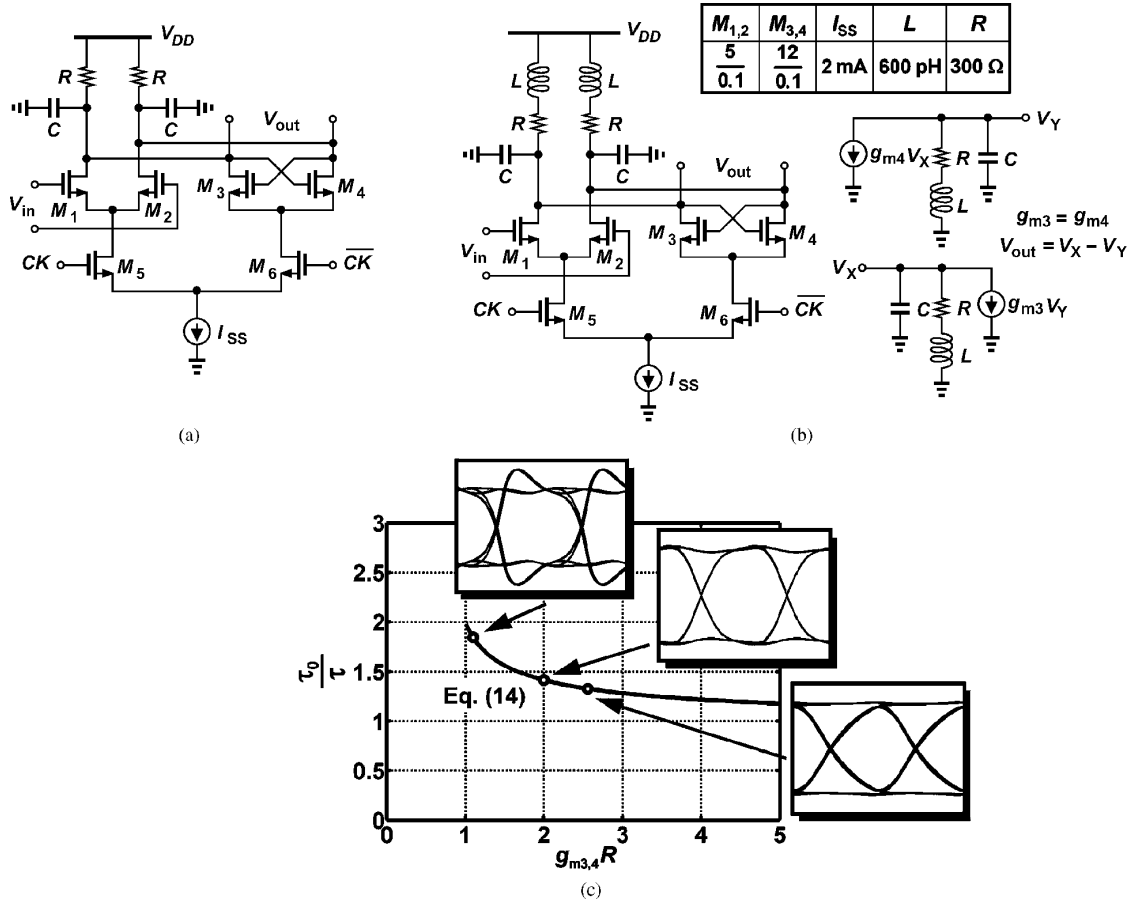
Fig. 13. (a) Regular CML latch. (b) CML latch with inductive peaking. (c) Regeneration speed improvement as a function of $g_{m3,4}R$.

we degenerate the tuned amplifier of Fig. 12(b) in such a way that its peaking and bandwidth become well-behaved to accommodate the desired operation points. As compared with that in Fig. 12(a), this buffer allows more efficient optimization of the gain and bandwidth. For example, if we choose $\zeta = 0.25$, then the equivalent $Q = 2$ and $R_{P2} = 4R_u$. To be more specific, this method plays a compromising role between the resistive and inductive loadings, alleviating bandwidth limitation and providing accurate swing control. Note that this structure is totally different from the purely inductive designs such as [18] and [19].

To further increase the bandwidth, we cascade two stages with different peaking frequencies [Fig. 12(c)]. The split peaks enlarge the operation range significantly. In this design, we realize the peaking moderately to ensure a stable operation, i.e., the maximum peak only exceeds the dc gain by 2.7 dB. As a result, the −3-dB bandwidth of this clock buffer is about 24.6 GHz, which provides adequate margin for PVT variations. Note that the two-stage topology also achieves a good isolation for VCO, protecting it from being disturbed by the sampling flip-flop and the frequency detector. A reverse isolation (S_{12}) of −74 dB is observed in simulation.

B. Retiming Flipflop

The high data rate needs all blocks to operate in current mode. In this work, we also incorporate inductive peaking in each block to further speed up the circuits. Here we plot the conventional and proposed latches in Fig. 13 to make a comparison. While the peaking technique improves the sampling bandwidth of the M_1-M_2 pair, it also helps the regeneration of M_3-M_4. To gain more insight, we first consider the ordinary CML latch without inductors [Fig. 13(a)]. Suppose the latch samples the input data at a position very close to the transition, i.e., the initial output $V_{out,0}$ before the regeneration is very small. It can be shown in [20] that after a half-cycle regeneration, the output V_{out} becomes

$$V_{out} = V_{out,0} \exp\left[\frac{(g_{m3,4}R - 1)T_{CK}}{2RC}\right]. \quad (11)$$

Here, T_{CK} denotes the clock period and $g_{m3,4}$ the transconductance of $M_{3,4}$. That is, the output increases exponentially with a time constant $\tau_0 = RC/(g_{m3,4}R - 1)$. With the help of inductive peaking, this process is expedited. Redrawing the latch with peaking inductor L and the equivalent model in regeneration mode [Fig. 13(b)], we calculate the output V_{out} ($= V_X - V_Y$) again:

$$LC\frac{d^2V_{out}}{dt^2} + (RC - g_{m3,4}L)\frac{dV_{out}}{dt} + (1 - g_{m3,4}R)V_{out} = 0. \quad (12)$$

For the most flat response [the damping factor $\zeta = (R/2)\sqrt{C/L} = 0.7$], we obtain an explicit solution for
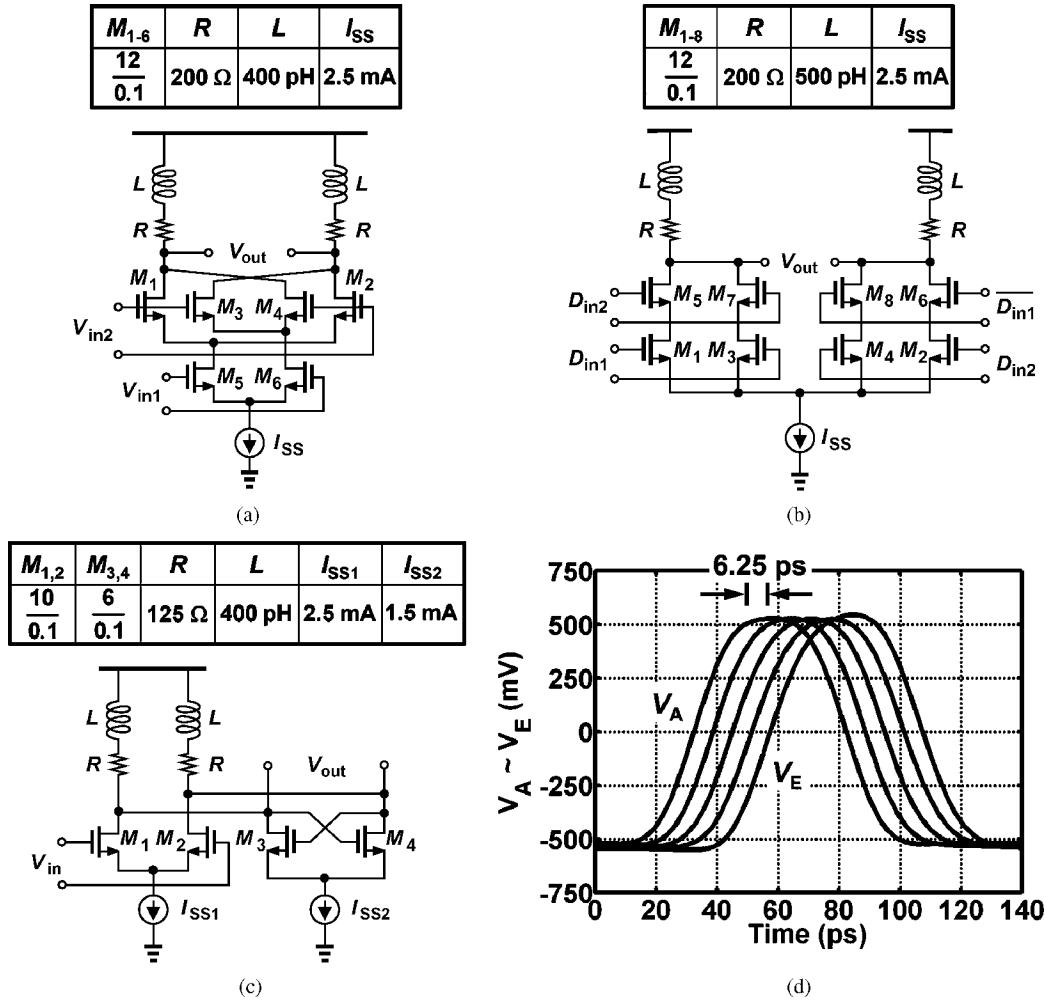
Fig. 14.   Implementation of (a) mixer, (b) XOR gate, (c) delay cell, and (d) simulated waveforms.

$V_{out}(t)$, which grows up exponentially with a new time constant τ:[3]

$$\tau = \frac{2RC}{g_{m3,4}R - 2 + \sqrt{g_{m3,4}^2 R^2 + 4g_{m3,4}R - 4}}. \quad (13)$$

As compared with τ_0, the positive-feedback process is accelerated by a factor of

$$\frac{\tau_0}{\tau} = \frac{g_{m3,4}R - 2 + \sqrt{g_{m3,4}^2 R^2 + 4g_{m3,4}R - 4}}{2(g_{m3,4}R - 1)} \geq 1. \quad (14)$$

Note that $g_{m3,4}R$ must be greater than unity to guarantee positive feedback. Fig. 13(c) plots the speed improvement as a function of $g_{m3,4}R$, demonstrating that the inductive peaking unconditionally improves the regeneration. However, aggressive peaking not only risks the regeneration but leads to significant ringing on the output data. Two cases of time domain waveforms with $g_{m3,4}R = 1.1$ and 2 are shown in the insets of Fig. 13(c) to illustrate such a trade-off. As a result, an improving factor

[3]The second solution of $V_{out}(t)$ decays.

of 1.414 as $g_{m3,4}R = 2$ has been chosen as an optimal point in this design. The device sizes in our design are also listed in Fig. 13(b). Note that in actual design, the speed may be boosted to a lesser extent due to some other considerations such as power consumption and routing convenience.

C. Mixer, XOR Gate, Delay Cell, and V/I Converters

The mixer, XOR gate, and delay cell are implemented as inductively-peaked CML topologies as well. Fig. 14 depicts the designs as well as the device sizes. The delay cell is realized as a hysteresis buffer [17] to provide sharp transition with significant delay. As compared with the delay chain in [21] that utilizes eight purely differential pairs with power dissipation of 36 mW, this work generates the same 25-ps delay while consuming only 24 mW. Note that although not necessary in this prototype, external delay tuning could be achieved by adjusting the two tail currents I_{SS1} and I_{SS2}. Similar to that in [21], the inductors are implemented as stacked spirals to minimize the area and alleviate the routing. The four delay cells are carefully laid out with balanced routing and loading to minimize static error in the retiming flip-flop. Fig. 14(d) illustrates the delayed waveforms simulated in transistor level.
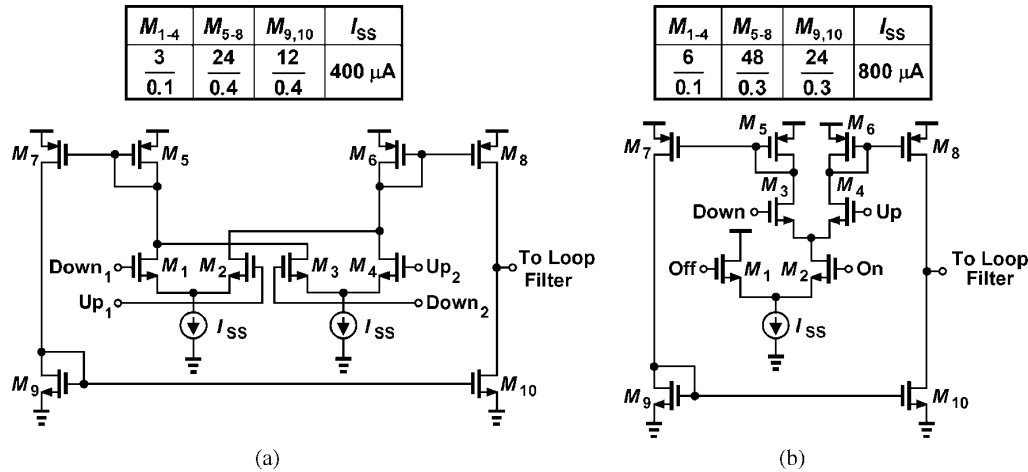
$M_{1\text{-}4}$	$M_{5\text{-}8}$	$M_{9,10}$	I_{SS}
$\dfrac{3}{0.1}$	$\dfrac{24}{0.4}$	$\dfrac{12}{0.4}$	400 μA

$M_{1\text{-}4}$	$M_{5\text{-}8}$	$M_{9,10}$	I_{SS}
$\dfrac{6}{0.1}$	$\dfrac{48}{0.3}$	$\dfrac{24}{0.3}$	800 μA

(a)   (b)

Fig. 15. Realization of (a) $(V/I)_{PD}$ and (b) $(V/I)_{FD}$.

(a)   (b)

(c)
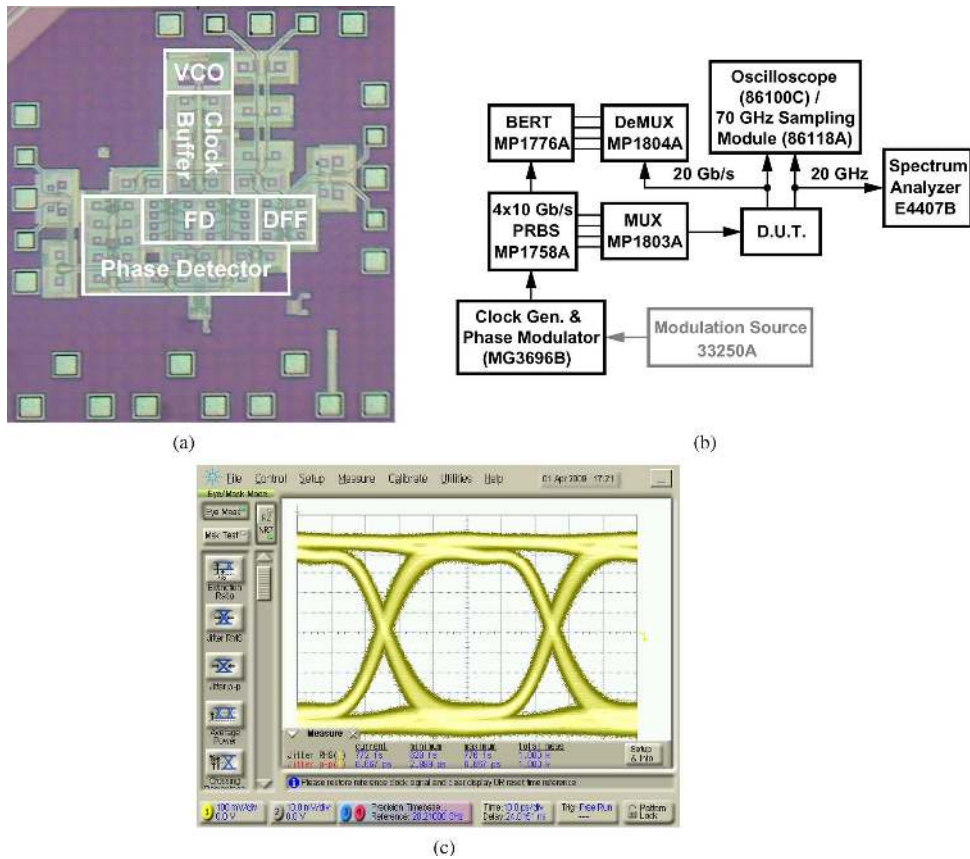
Fig. 16.   (a) Chip micrograph. (b) Testing setup. (c) Input waveform (horizontal scale: 10 ps/div, vertical scale: 100 mV/div).

The $(V/I)_{PD}$ design is depicted in Fig. 15(a), where two differential pairs M_1-M_2 and M_3-M_4 steer two identical current sources. As mentioned in Section II.A, the duty cycle error can be fully tolerated if the two pairs and the associated current sources are matched. Unlike the conventional (XOR-based) linear PDs that potentially suffer from imbalanced clock skews, this approach alleviates the VCO design requirements substantially. Fig. 15(b) depicts the $(V/I)_{FD}$ design. Note that the asymmetry of the up and down paths in Fig. 15(a) and (b) is

not an issue, since they are dealing with near-dc signals here. Resembling that in [15] and [17], $(V/I)_{FD}$ uses a pumping current twice that in $(V/I)_{PD}$ to ensure the FD loop dominates during frequency acquisition.

IV. EXPERIMENTAL RESULTS

The CDR circuit has been designed and fabricated in 90-nm CMOS technology. Fig. 16(a) reveals a photo of the die, which occupies 0.97×0.88 mm^2 including pads. With the help of the
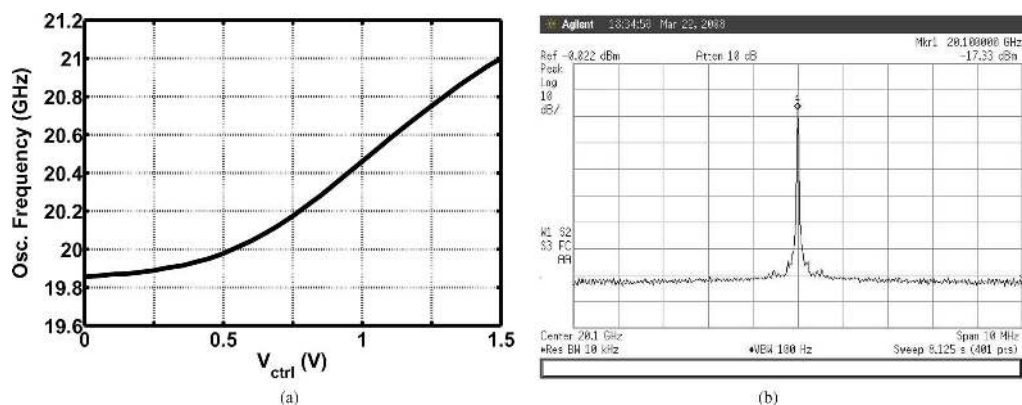
Fig. 17. (a) VCO tuning range. (b) Output clock spectrum under locked condition.

TABLE I
CDR PERFORMANCE SUMMARY

	[22]	[23]	[24]	This Work
Data Rate	12.5 Gb/s	25 Gb/s	12.5 Gb/s	20 Gb/s
Operation Range	300 Mb/s	6.1 Mb/s	2.55 Gb/s	950 Mb/s
PD Type	Linear, Half−Rate	Linear, Half−Rate	Linear, Half−Rate	Linear, Full−Rate
Rec. Clock Jitter	N/A	N/A	N/A	480 fs,rms 4.22 ps,pp (with 2^{31}−1 PRBS)
Rec. Data Jitter	N/A	2.11 ps,rms 5.9 ps,pp (with 2^7−1 PRBS)	6 ps,pp	1.22 ps,rms 7.56 ps,pp (with 2^{31}−1 PRBS)
BER	$< 10^{-12}$ (with 2^{31}−1 PRBS)	$< 10^{-12}$ (with 2^7−1 PRBS)	$< 10^{-12}$ (with 2^{31}−1 PRBS)	$< 10^{-13}$ (with 2^{31}−1 PRBS)
Jitter Tolerance	Exceeds OC−192 mask by 0.24 UI$_{PP}$	Exceeds extrapolated SONET mask by 0.2 UI$_{PP}$	Exceeds OC−192 mask by 0.2 UI$_{PP}$	Exceeds extrapolated OC−192 mask by 0.43 UI$_{PP}$
Freq. Acquisition	Yes (External ref. clock)	No	Yes (External ref. clock)	Yes (Referenceless)
Supply	1.2 V	1.1 V	1.5 V	1.5 V
Power Diss.	351 mW	172 mW	400 mW	154 mW*
Chip Area	N/A	0.32 x 0.22 mm^2	1.5 x 0.75 mm^2	0.97 x 0.88 mm^2
Technology	0.1−μm CMOS	90−nm CMOS	0.13−μm CMOS	90−nm CMOS

* Without I/O buffers.

frequency acquisition loop, the CDR is capable of oper-
ating over a range of 950 Mb/s without any external adjust-
ment, across which no performance degradation is observed.
Error-free operation (BER $< 10^{-12}$) for 2^{31}−1 PRBS input is
achieved for supply voltage varying from 1.3 to 1.7 V (nomi-
nally it is 1.5 V). The circuit is carefully designed to make sure
no more than 1.2 V is applied across any device. The circuit
consumes a total power of 154 mW from a 1.5-V supply, of
which 65 mW is dissipated in the PD, 66 mW in the VCO
and clock buffers, and 23 mW in the FD. The single-ended
input sensitivity is about 1 V$_{PP}$. Fig. 16(c) reveals the input
waveform captured from the output of MP1803A. The jitter
measures 772 fs,rms and 6.67 ps,pp, respectively.

Fig. 17 shows the tuning range of the VCO and the locked
spectrum of the output clock. The phase noise measures
−105 dBc/Hz at 1-MHz offset. The phase noise plots of the
free-running VCO, the phase-locked VCO, and the input data
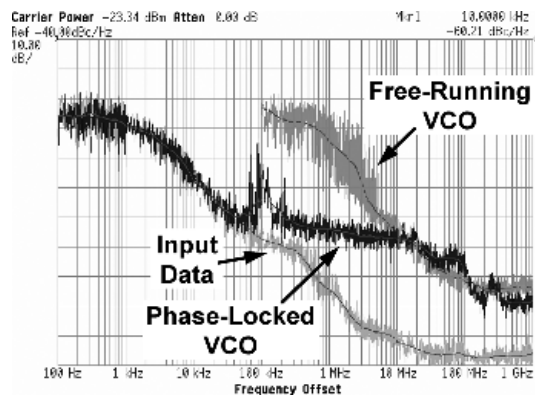are shown in Fig. 18. The loop bandwidth of the CDR can be

Fig. 18. Phase noise plots.

clearly identified as 15 MHz, which corresponds to an optimal
jitter performance in our design. As expected, the noise of the
recovered clock follows that of the input tightly at low offset
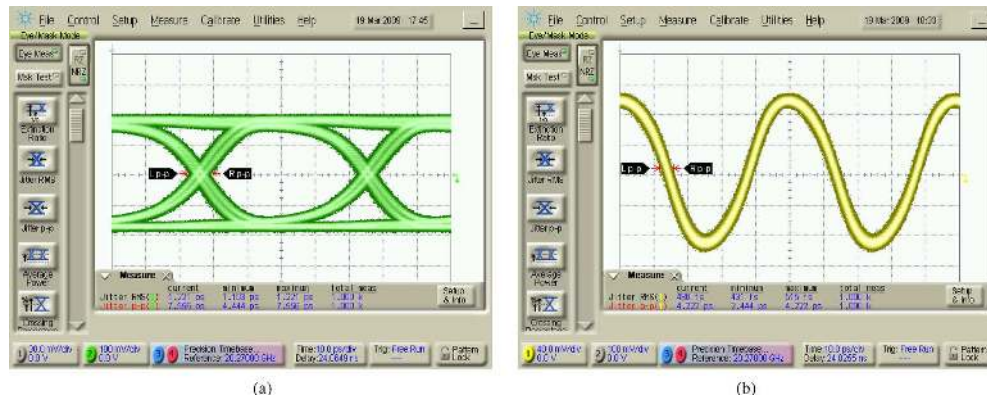frequencies, and approaches the free-running profile after the

Fig. 19. Recovered (a) data and (b) clock in response to a $2^{31}-1$ PRBS [horizontal scale: 10 ps/div, vertical scale: 100 mV/div (left) and 40 mV/div (right), 1 k measurements].
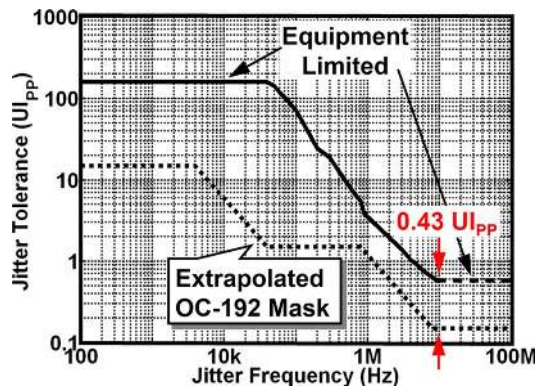

Fig. 20. Jitter tolerance.

loop bandwidth. Jitter generation is obtained as 475 fs,rms by integrating the phase noise from 100-Hz to 1-GHz offset frequencies (the maximum range for our equipments.) If we extrapolate the OC-192 specifications, the jitter generation from 100-kHz to 160-MHz offsets would be 351 fs,rms. Fig. 19(a) depicts the waveforms of the recovered data in response to a PRBS of $2^{31}-1$, suggesting data jitter of 1.22 ps,rms and 7.56 ps,pp, respectively. The recovered clock is also shown in Fig. 19(b), where the rms and peak-to-peak jitter measure 480 fs and 4.22 ps, respectively. The time-domain measurement verifies the jitter generation.

To verify the signal integrity, we have also conducted jitter tolerance testing. Due to the lack of 20-Gb/s jitter tolerance tester, we measure it manually by modulating the system clock generator (MG3696B) and capture the BER accordingly. Note that an external modulation source (33250A) is required for high-frequency modulation, since MG3696B only allows internal modulation up to 1 MHz. With the error threshold set to 10^{-12}, we plot the jitter tolerance profile in response to an input data of 2^7-1 PRBS along with the extrapolated OC-192 specification mask. As plotted in Fig. 20, the measured jitter tolerance exceeds at least 0.43 UI_{PP} for all jitter frequencies. Table I summarizes the performance of this work and some other similar CDR circuits recently published in the literature.

V. CONCLUSION

This work proposes a CDR with mixer-based phase detector that allows linear phase comparison up to 20 Gb/s without causing dead zone. A new frequency monitoring loop makes use of the existing data phases from the phase detector to dynamically ensure frequency locking. Peaking and mm-wave techniques are extensively utilized to extend the bandwidth and maintain the signal integrity. Providing superior performance with reasonably low power consumption, this chip demonstrates a promising future for CMOS high-speed communications.

REFERENCES

[1] P. Yue and M. Rodwell, "mm-wave IC design: The transition from III-V to CMOS circuit techniques, short course, RF and high speed CMOS," in *Proc. IEEE Compound Semiconductor Integrated Circuit Symp. (CSICS)*, Nov. 2006.

[2] Y. M. Greshishchev and P. Schvan, "SiGe clock and data recovery IC with linear-type PLL for 10-Gb/s SONET application," *IEEE J. Solid-State Circuits*, vol. 35, no. 9, pp. 1353–1359, Sep. 2000.

[3] J. D. H. Alexander, "Clock recovery from random binary data," *Electron. Lett.*, vol. 11, pp. 541–542, Oct. 1975.

[4] C. R. Hogge, "A self-correcting clock recovery circuit," *J. Lightw. Techol.*, vol. 3, no. 12, pp. 1312–1314, Dec. 1985.

[5] J. Savoj and B. Razavi, "A 10-Gb/s CMOS clock and data recovery circuit with a half-rate linear phase detector," *IEEE J. Solid-State Circuits*, vol. 36, no. 5, pp. 761–768, May 2001.

[6] H. Noguchi *et al.*, "A 40-Gb/s CDR circuit with adaptive decision-point control based on eye-opening monitor feedback," *IEEE J. Solid-State Circuits*, vol. 43, no. 12, pp. 2929–2938, Dec. 2008.

[7] Y. Amamiya *et al.*, "A 40 Gb/s multi-data-rate CMOS transceiver chipset with SFI-5 interface for optical transmission systems," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2009, pp. 358–359.

[8] *40 Gb/s and 100 Gb/s Ethernet Task Force*, IEEE P802.3ba [Online]. Available: http://www.ieee802.org/3/ba/index.html

[9] A. Pottbacker *et al.*, "A Si bipolar phase and frequency detector for clock extraction up to 8 Gb/s," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, pp. 1747–1751, Dec. 1992.

[10] S. B. Anand and B. Razavi, "A 2.75 Gb/s CMOS clock recovery circuit with broad capture range," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2001, pp. 214–215.

[11] J. Savoj and B. Razavi, "A 10-Gb/s CMOS clock and data recovery circuit with a half-rate binary phase/frequency detector," *IEEE J. Solid-State Circuits*, vol. 38, no. 1, pp. 13–21, Jan. 2003.

[12] B. Razavi, *Design of Integrated Circuits for Optical Communications*. New York: McGraw-Hill, 2002.

[13] J. C. Scheytt *et al.*, "A 0.155, 0.622, and 2.488 Gb/s automatic bit rate selecting clock and data recovery IC for bit rate transparent SDH systems," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 1999, pp. 348–349.

[14] J. Lee and B. Razavi, "A 40-Gb/s clock and data recovery circuit in 0.18-μm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 38, no. 12, pp. 2181–2190, Dec. 2003.

[15] J. Lee, "High-speed circuit designs for transmitters in broadband data links," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1004–1015, May 2006.

[16] L. DeVito *et al.*, "A 52 MhZ and 155 MHz clock-recovery PLL," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 1991, pp. 142–143.

[17] J. Lee *et al.*, "A 75-GHz phase-locked loop in 90-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 43, no. 6, pp. 1414–1426, Jun. 2008.

[18] S. C. Chan *et al.*, "Distributed differential oscillators for global clock networks," *IEEE J. Solid-State Circuits*, vol. 41, no. 9, pp. 2083–2094, Sep. 2006.

[19] A. P. Jose and K. L. Shepard, "Distributed loss-compensation techniques for energy-efficient low-latency on-chip communication," *IEEE J. Solid-State Circuits*, vol. 42, no. 6, pp. 1415–1424, Jun. 2007.

[20] B. Razavi, *Principles of Data Conversion System Design*. Piscataway, NJ: IEEE Press, 1995.

[21] J. Lee and M. Liu, "20-Gb/s burst-mode clock and data recovery circuit using injection-locking technique," *IEEE J. Solid-State Circuits*, vol. 43, no. 3, pp. 619–630, Mar. 2008.

[22] J. Takasoh *et al.*, "A 12.5 Gbps half-rate CMOS CDR circuit for 10 Gbps network applications," in *Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2004, pp. 268–271.

[23] C. Kromer *et al.*, "A 25-Gb/s CDR in 90-nm CMOS for high-density interconnects," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2921–2929, Dec. 2006.

[24] Y. Ohtomo *et al.*, "A 12.5-Gb/s parallel phase detection clock and data recovery circuit in 0.13-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 41, no. 9, pp. 2052–2057, Sep. 2006.

Jri Lee (S'03–M'04) received the B.Sc. degree in electrical engineering from National Taiwan University (NTU), Taipei, Taiwan in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Los Angeles (UCLA), both in 2003. His current research interests include high-speed wireless and wireline transceivers, phase-locked loops, and data converters.

After two years of military service (1995–1997), he was with Academia Sinica, Taipei, Taiwan from 1997 to 1998, and subsequently Intel Corporation from 2000 to 2002. He joined National Taiwan University (NTU) since 2004, where he is currently Associate Professor of electrical engineering. He is now serving in the Technical Program Committees of the International Solid-State Circuits Conference (ISSCC), Symposium on VLSI Circuits, and Asian Solid-State Circuits Conference (A-SSCC).

Prof. Lee received the Beatrice Winner Award for Editorial Excellence at the 2007 ISSCC, the Takuo Sugano Award for Outstanding Far-East Paper at the 2008 ISSCC, the Best Technical Paper Award from Y. Z. Hsu Memorial Foundation in 2008, the T. Y. Wu Memorial Award from National Science Council (NSC), Taiwan, in 2008, the Young Scientist Research Award from Academia Sinica in 2009, and the Outstanding Young Electrical Engineer Award in 2009. He also received the NTU Outstanding Teaching Award in 2007, 2008, and 2009. He served as a guest editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS in 2008 and a tutorial lecturer at the 2009 ISSCC.

Ke-Chung Wu was born in Taipei, Taiwan, in 1983. He received the B.S. degree in electrical engineering from National Taiwan University, Taipei, in 2005. He is currently pursuing the Ph.D. degree in the Graduate Institute of Electrical Engineering, National Taiwan University, Taipei.

His research interests include phase-locked loops and wireline transceivers for broadband data communication.