Low-Power CMOS Equalizer Design for 20-Gb/s Systems

Sameh Ibrahim, *Member, IEEE*, and Behzad Razavi, *Fellow, IEEE*

Abstract—The power consumption of wireline circuits has become increasingly more critical as the pin count and data rate rise. This paper describes a power scaling methodology and a new half-rate speculative architecture for decision-feedback equalizers (DFEs) to relax the speed-power trade-offs. Designed in 90-nm CMOS technology, a 20-Gb/s prototype consisting of a linear equalizer and a one-tap DFE compensates for the loss of an 18-in FR4 trace while drawing 40 mW from a 1-V supply.

Index Terms—Bit error rate, CML latch, decision-feedback equalizers, high-speed equalizers, latch sensitivity, latch offset, unrolled DFE.

I. INTRODUCTION

THE rapid increase in the pin count of chips and the resulting routing complexity on printed-circuit boards (PCBs) and backplanes has made the use of serial links attractive. With the pin count eventually limited by the physical dimensions of packages (and their parasitics), the only option to increase the throughput rate is to design each serial link for a higher speed. The parallel-to-serial transformation can also potentially save significant power because it reduces the number of output drivers (while I/O voltage swings and termination impedances remain constant). It is therefore plausible that data rates as high as 20 Gb/s will become common in the near future.

In applications requiring a large number of high-speed serial links, it is necessary to reduce the power consumed by each building block in the transmitter (TX) and receiver (RX). This paper proposes both a methodology for low-power equalizer design and a new architecture for decision-feedback equalizers (DFEs) that alleviates the power-speed trade-offs [1]. A 20-Gb/s prototype realized in 90-nm CMOS technology equalizes data received from an 18-in FR4 trace while consuming 40 mW.

The next section of the paper provides a brief overview of the design challenges and the prior art. Section III formulates the power scaling limits of DFEs and proposes a methodology for minimizing their power consumption. Section IV introduces the DFE architecture and Section V, the circuit details. Section VI presents the experimental results.

II. BACKGROUND

At speeds of tens of gigabits per second, the loss of FR4 boards poses a great challenge, requiring heavy equalization. From the circuit design point of view, it is simpler to employ linear equalization (in TX and RX) but from the system design perspective, two serious issues plague this approach: the amplification of crosstalk and the lack of ability to equalize for impedance discontinuities (deep notches in the channel frequency response).[1] A receiver would preferably perform the entire equalization by means of a DFE. However, DFEs cannot equalize for pre-cursor inter-symbol interference (ISI) and also require a very large number of taps for high-loss channels. For these reasons, some linear equalization is inserted in the TX path and/or the RX path.

Extensive work on DFEs has produced a multitude of architectures [2]–[5], which can be broadly categorized as "direct" or "unrolled" (speculative) DFEs with "full-rate" or "half-rate" clocking. Fig. 1 conceptually summarizes these developments (only the first tap is shown for simplicity). Targeting a unit interval (UI) of 50 ps, we make the following observations:

1) The direct full-rate DFE of Fig. 1(a) requires that

$$t_{cq} + t_{setup} + t_{FB} < UI \qquad (1)$$

where t_{cq} denotes the clock-to-output delay of the flipflop (FF), t_{setup} its setup time, and t_{FB} the "feedback delay" (arising from the time constant at the summing node). Even with inductive peaking, it is difficult to guarantee this condition in 90-nm CMOS technology for UI = 50 ps, especially with nearly sinusoidal clocks and at low power consumption levels.

2) In the unrolled full-rate DFE of Fig. 1(b) [2], the first tap creates an offset of $\pm IR$ in the two paths, and the MUX selects one of the two based on the previous bit decision. Here,

$$t_{cq,FF} + t_{setup} + t_{sq,MUX} < UI \qquad (2)$$

where the first and last terms on the left refer to the FF clock-to-output and the MUX select-to-output delays, respectively. Even though this architecture replaces the feedback delay with the MUX delay,[2] it still does not reach 20 Gb/s in 90-nm technology. Note that $t_{sq,MUX}$ in (2) is usually smaller than t_{FB} in (1) because adding higher-order taps increases the capacitance at the summing nodes.

[1]Linear equalization (de-emphasis) in the TX also leads to smaller swings, making the received signal more sensitive to noise.

[2]While the architectures considered here employ the MUX in the "analog" domain, they do not require the MUX to be linear. That is, the differential pairs sensing the two inputs can provide some slicing.
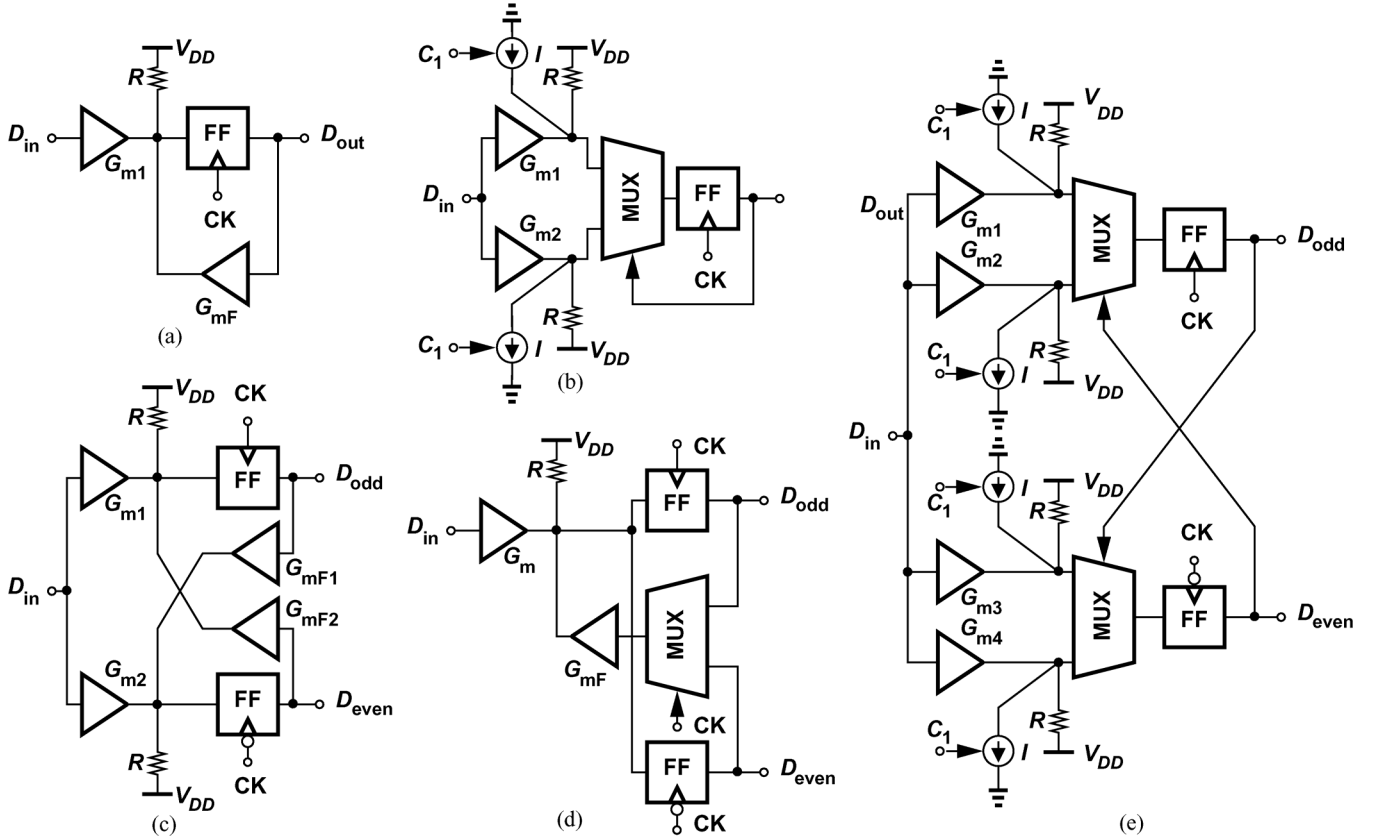
Fig. 1. DFE architecture development: (a) direct full-rate DFE, (b) unrolled full-rate DFE, (c) direct half-rate DFE, (d) multiplexed half-rate DFE, and (e) unrolled half-rate DFE.

3) In the direct half-rate DFE of Fig. 1(c) [3],

$$t_{cq} + t_{setup} + t_{FB} < UI. \tag{3}$$

Note that, despite half-rate operation, the timing here is as stringent as that expressed by (1) for the full-rate counterpart. The principal advantage of this architecture is the simpler design of the CDR circuit and, in particular, the clock buffer.

4) In the multiplexed half-rate DFE of Fig. 1(d) [4],

$$t_{cq,FF} + t_{setup} + t_{p,MUX} + t_{FB} < UI \tag{4}$$

where $t_{p,MUX}$ is the propagation delay in the data path of the MUX.[3] This timing limitation is similar to that expressed by (3).

5) In the unrolled half-rate DFE of Fig. 1(e) [5],

$$t_{cq,FF} + t_{setup} + t_{sq,MUX} < UI. \tag{5}$$

Compared with its full-rate counterpart, this architecture allows a simpler CDR design. However, the implementation is quite complex, demanding a high power dissipation (and numerous inductors if inductive peaking is necessary).

A number of CMOS solutions for the rates around 20 Gb/s have been reported, e.g., [6] employs duobinary signaling to equalize for 14 dB of loss at 10 GHz, [7] exhibits a high bit error

[3]In [4], half UI is allocated to the CDR circuit. Here, we have used 1 UI for consistent comparisons.

rate (BER) (10^{-8}) for a loss of 11 dB at 10 GHz, and [8] exploits 20 dB of linear equalization for a loss of 21 dB at 10.5 GHz. The need therefore exists for an NRZ DFE solution that compensates most of the loss.

III. DFE POWER SCALING LIMITS

A. General Considerations

Consider the generic receiver equalizer shown in Fig. 2(a). The linear equalizer power consumption can be reduced by reverse scaling [9] and/or by scaling its constituent devices (e.g., all bias currents and transistor widths are multiplied by $1/M$ and all load resistors by M, $M > 1$)—to the point where the thermal noise of the chain still remains negligible. Unfortunately, the DFE does not lend itself to reverse scaling because the feedback transconductance, G_{mF}, must bear a certain ratio with respect to the input transconductance, G_{m1}. Thus, only uniform scaling of the devices can reduce the DFE power consumption. Our objective is to determine the scaling limits of G_{m1}, R_1, G_{mF}, and the FF.

The FF considered here employs the current-steering latch depicted in Fig. 2(b), but the methodology can be applied to other topologies as well. It is assumed that a "reference" design has been reached wherein the device dimensions and bias currents are chosen according to the voltage headroom, the required speed and output swing, and the available clock swings. The design is now scaled as follows: $W_j \rightarrow W_j/M$, $I_{SS} \rightarrow I_{SS}/M$,
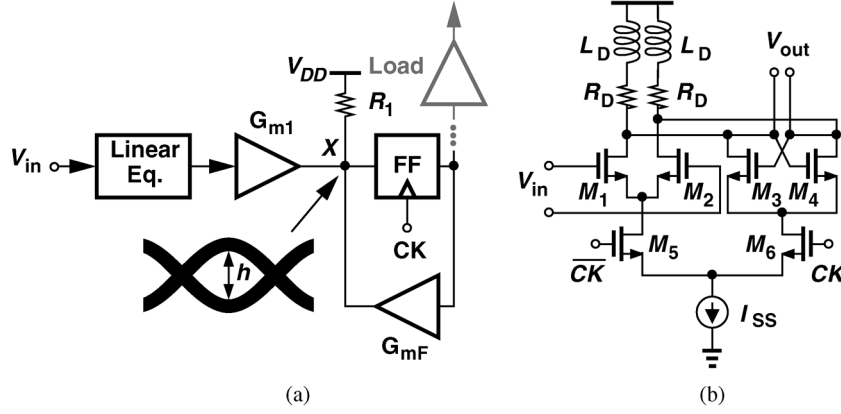
Fig. 2. (a) Generic receiver equalizer. (b) Current-steering latch.

$R_D \to MR_D, L_D \to ML_D$. Consequently, the power dissipation and the data and clock input capacitances fall by a factor of M but so does the drive capability. Thus, for a given, unscalable load (e.g., the CDR and DMUX input), the scaling of the DFE chain is possible to the point where the load can still be driven with reasonable rise and fall times.[4]

We must now determine how large M can be. The factors limiting the scaling include (1) the residual ISI, i.e., the actual vertical eye opening, h, at the summing node in Fig. 2(a) after the equalization is completed; (2) the total offset voltage referred to node X resulting from device mismatches, V_{os}; (3) the total electronic noise referred to node X arising from the linear equalizer, the G_m stages, and the FF, V_n; (4) the sensitivity of the FF, V_{sens}. As seen below, scaling exacerbates all four factors.

B. Effect of Offset and Sensitivity on BER

The minimum eye opening, h, in Fig. 2(a) must be large enough to yield the required BER in the presence of the above nonidealities. With only noise present, the peak-to-peak input swing of the FF, $V_{pp,eq}$,[5] and the total noise referred to this input, $V_{n,eq}$, lead to

$$BER = Q\left(\frac{V_{pp,eq}}{2\sqrt{\overline{V_{n,eq}^2}}}\right) \qquad (6)$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{u^2}{2}\right) du. \qquad (7)$$

For example, a BER of 10^{-14} translates to $V_{pp,eq}/2\sqrt{\overline{V_{n,eq}^2}} = 7.6$.

If the total offset referred to the summing node in Fig. 2(a) is V_{os}, then, on the average, half of the bits see an effective *peak*

[4]Our scaling method can encompass CDR and DMUX as well, but for simplicity, we consider only the DFE.

[5]To avoid confusion, h denotes the minimum and $V_{pp,eq}$ the actual vertical eye opening.

swing of $V_{pp,eq}/2 - V_{os}$ and the other half, $V_{pp,eq}/2 + V_{os}$. The overall BER is thus given by

$$BER = \frac{1}{2}Q\left(\frac{V_{pp,eq}/2 - V_{os}}{\sqrt{\overline{V_{n,eq}^2}}}\right) + \frac{1}{2}Q\left(\frac{V_{pp,eq}/2 + V_{os}}{\sqrt{\overline{V_{n,eq}^2}}}\right). \qquad (8)$$

As will be explained in Section III-G, in a scaled design V_{os} tends to be quite larger than $\sqrt{\overline{V_{n,eq}^2}}$, making the second term on the right-hand side negligible:

$$BER \approx \frac{1}{2}Q\left(\frac{V_{pp,eq}/2 - V_{os}}{\sqrt{\overline{V_{n,eq}^2}}}\right). \qquad (9)$$

In order to ensure high yield, a V_{os} of roughly five standard deviations is chosen. While not critical to the proposed methodology, the choice of five standard deviations is to ensure that one equalizer does not limit the yield of a complex transceiver. For example, a four-lane 100-Gb/s system employing four equalizers and CDR circuits on the receive side and four PLLs and equalizers on the transmit side would suffer a loss of 4% in the yield if each of these building blocks is designed for 3-sigma yield but less than 0.2% for 5-sigma yield.

Let us now consider the sensitivity of the FF, V_{sens}. We define V_{sens} as the minimum input voltage that guarantees regeneration to approximately 80% of the full FF output swing. While somewhat arbitrary, this value can be relaxed to 70% or tightened to 90% with little effect on our derivations.

To see how the FF sensitivity must be taken into account, suppose first that the peak input swing is *equal* to V_{sens}, i.e., $V_{pp,eq}/2 = V_{sens}$. Then, in the presence of noise, half of the FF output levels are degraded. We therefore conclude that $V_{pp,eq}/2$ must be sufficiently greater than V_{sens} so that the noise values in the "tail" of the Gaussian distribution degrade the FF output with negligible probability. Fig. 3 summarizes our findings, suggesting that

$$BER = \frac{1}{2}Q\left(\frac{V_{pp,eq}/2 - V_{os} - V_{sens}}{\sqrt{\overline{V_{n,eq}^2}}}\right). \qquad (10)$$
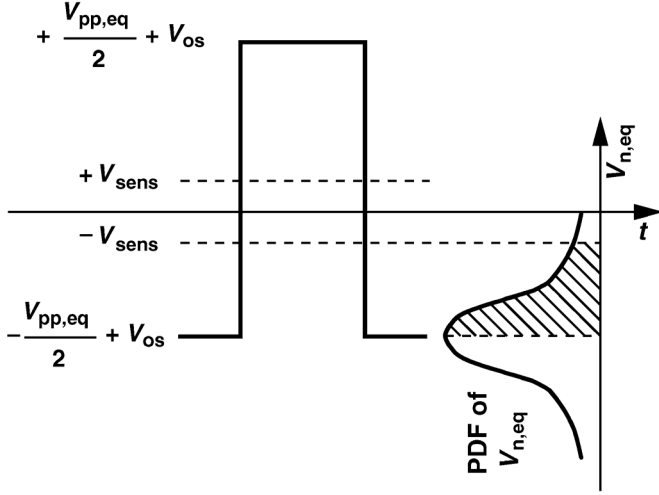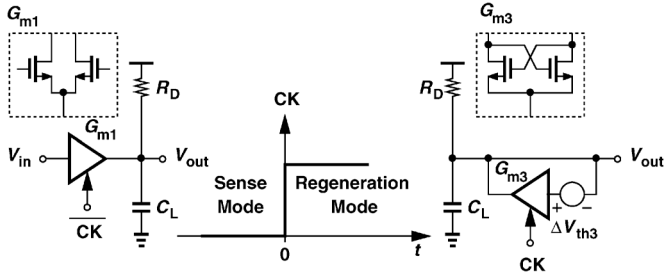
Fig. 3. Offset and sensitivity setting limits for BER.

Fig. 4. Latch small-signal model for sense and regeneration modes.

For BER = 10^{-14}, the argument of Q must reach 7.5, requiring a vertical eye opening of

$$h = V_{pp,eq} \geq 15\sqrt{\overline{V_{n,eq}^2}} + 2(V_{os} + V_{sens}). \tag{11}$$

C. Latch Offset

The FF offset arises primarily from that of the master latch. In the latch topology of Fig. 2(b), the differential pair, the load resistors, and the regenerative pair contribute offset. The first two contributions can be readily formulated in the sense mode, but the third entails a more complex calculation. We proceed as follows: the latch input-referred offset is defined as that input voltage which causes complete metastability (no regeneration). This definition allows the use of a small-signal model for both the input and the regenerative pairs as shown in Fig. 4. Here G_{m1} and G_{m3} denote the transconductances of the two pairs, respectively. We examine the offset contribution of $G_{m3}(= \Delta V_{TH3})$ in two cases: an abrupt clock edge and a gradual clock edge.

With an abrupt clock edge, at the end of the sense mode,

$$V_{out}(0) = G_{m1}R_D V_{in}. \tag{12}$$

Sensing this initial condition, the regenerative pair responds according to the following differential equation:

$$G_{m3}(V_{out} + \Delta V_{TH3}) = \frac{V_{out}}{R_D} + C_L\frac{dV_{out}}{dt}. \tag{13}$$

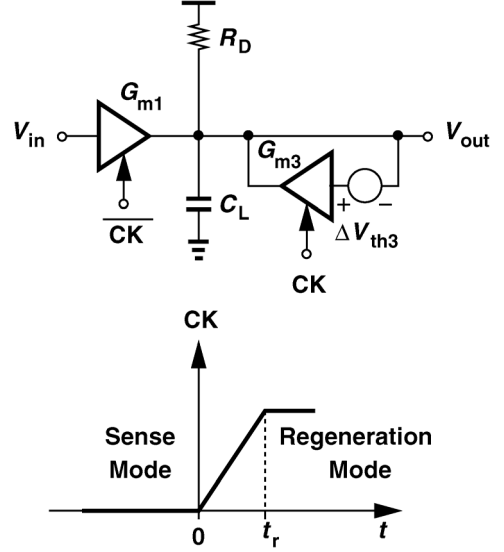
Fig. 5. Latch small-signal model with a gradual clock edge.

Thus,

$$V_{out}(t) = \frac{G_{m3}R_D\Delta V_{TH3}}{G_{m3}R_D - 1}\left(\exp\frac{G_{m3}t}{C_L}\exp\frac{-t}{R_DC_L} - 1\right)$$
$$+ G_{m1}R_D V_{in}\exp\frac{G_{m3}t}{C_L}\exp\frac{-t}{R_DC_L}. \tag{14}$$

To obtain complete metastability, we have

$$\frac{G_{m3}R_D}{G_{m3}R_D - 1}\Delta V_{TH3} = -G_{m1}R_D V_{in} \tag{15}$$

and hence

$$V_{in} = \frac{-\Delta V_{TH3}}{G_{m1}R_D - \frac{G_{m1}}{G_{m3}}}. \tag{16}$$

This result disagrees with our basic understanding that the regenerative pair offset should be simply divided by the gain of the differential pair. This is because the initial voltage necessary at the output node to avoid regeneration is equal to $G_{m3}\Delta V_{TH3}/(G_{m3} - R_D^{-1})$ rather than simply ΔV_{TH3}.

Let us now assume a gradual clock edge. As shown in Fig. 5, both G_{m1} and G_{m3} are partially on during the clock rise time, t_r, and can be approximated as

$$g_{m1}(t) = G_{m1}\sqrt{1 - \frac{\alpha t}{t_r}}$$

$$g_{m3}(t) = G_{m3}\sqrt{\frac{\alpha t}{t_r}} \tag{17}$$

where G_{m1} and G_{m3} are the peak values, $\alpha = V_{CK,pp}/(\sqrt{2}V_{ov,CK})$, $V_{CK,pp}$ is the peak-to-peak voltage swing of the clock, and $V_{ov,CK}$ is the equilibrium overdrive voltage of the clock switches [transistors M_5 and M_6 in Fig. 2(b)]. Equation (17) is valid for $0 < t < t_r/\alpha$, where t_r/α is approximately equal to the time necessary for M_5 and M_6 to steer I_{SS} from left to right.

As explained in Appendix I, we can modify (16) to

$$V_{in} = \frac{-\Delta V_{TH3}}{G_{m1}R_D - \frac{G_{m1}}{G_{m3}}\exp\frac{(G_{m3}R_D-2)t_r}{2\alpha R_DC_L}}. \tag{18}$$
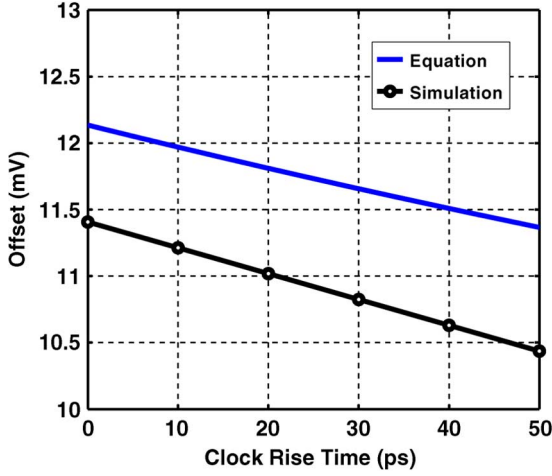
Fig. 6. Simulated and calculated input-referred offset contribution of the regenerative pairs.

Fig. 7. General CML latch small-signal model with all nonidealities included: (a) Sense mode, and (b) regeneration mode.

Fig. 6 plots the input-referred contribution of the regenerative pair as predicted by (18) and by circuit simulations, suggesting a reasonable agreement.

The other sources of offset can be incorporated in a similar manner. As illustrated in Fig. 7(a), a current source I_1 can represent the offset in the sense mode if

$$I_1 = \begin{cases} G_{m1}\Delta V_{TH1} \\ G_{m1}\frac{(V_{GS}-V_{TH})_2}{2}\frac{\Delta\beta_1}{\beta_1} \\ \frac{I_{SS}}{2}\frac{\Delta R_D}{R_D} \end{cases} \quad (19)$$

where the subscript 1 refers to the input differential pair. In the regeneration mode [Fig. 7(b)],

$$I_2 = \begin{cases} G_{m3}\Delta V_{TH3} \\ G_{m3}\frac{(V_{GS}-V_{TH})_3}{2}\frac{\Delta\beta_3}{\beta_3} \\ \frac{I_{SS}}{2}\frac{\Delta R_D}{R_D}, \end{cases} \quad (20)$$

where the subscript 3 refers to the regenerative pair. Note that, in accordance with (18), the effect of I_2 must be divided by $G_{m3}G_{m1}R_D - G_{m1}\exp[(G_{m3}R_D - 2)t_r/(2\alpha R_D C_L)]$. The overall input-referred offset of the latch is therefore equal to

$$\overline{V_{OS,L}^2} = (\Delta V_{TH1})^2 + \frac{(V_{GS}-V_{TH})_1^2}{4}\left(\frac{\Delta\beta_1}{\beta_1}\right)^2$$
$$+ \frac{\Delta V_{TH3}^2 + \frac{(V_{GS}-V_{TH})_3^2}{4}\left(\frac{\Delta\beta_3}{\beta_3}\right)^2}{\left[G_{m1}R_D - \frac{G_{m1}}{G_{m3}}\exp\frac{(G_{m3}R_D-2)t_r}{2\alpha R_D C_L}\right]^2}$$
$$+ \left(\frac{\Delta R_D}{R_D}\right)^2 \frac{(V_{GS}-V_{TH})_3^2}{4\left[G_{m1}R_D - \frac{G_{m1}}{G_{m3}}\exp\frac{(G_{m3}R_D-2)t_r}{2\alpha R_D C_L}\right]^2}$$
$$+ \left(\frac{\Delta R_D}{R_D}\right)^2 \frac{(V_{GS}-V_{TH})_1^2}{4}. \quad (21)$$

D. Latch Noise

The latch noise analysis can draw upon the offset studies carried out above. Again, we define the input-referred noise as that input (random) *waveform* which keeps the latch metastable
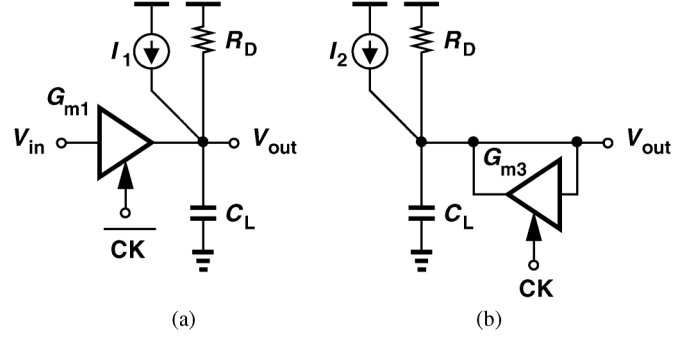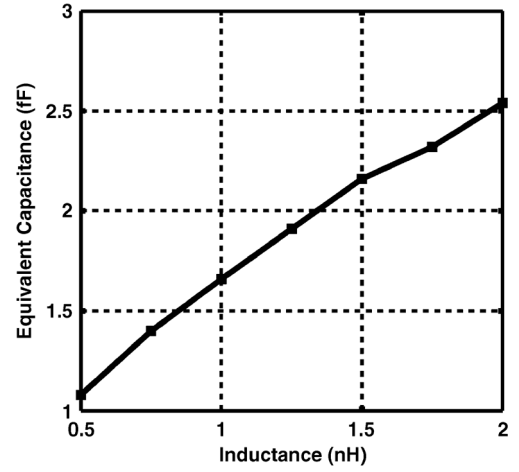

Fig. 8. Equivalent capacitance of stacked inductors.

throughout the regeneration phase. The small-signal model still provides a reasonable approximation because the effect of noises that are added *after* the start of regeneration falls exponentially with time, negligibly affecting the final binary decision [10]. The current sources in Fig. 7 now assume the following forms for the circuit of Fig. 2(b):

$$\overline{I_1^2} = \begin{cases} 2\frac{4kT}{R_D} \\ 4kT\gamma g_{m1} + 4kT\gamma g_{m2} \end{cases} \quad (22)$$

$$\overline{I_2^2} = \begin{cases} 2\frac{4kT}{R_D} \\ 4kT\gamma g_{m3} + 4kT\gamma g_{m4}, \end{cases} \quad (23)$$

where g_{mj} denotes the transconductance of M_j. The input-referred thermal noise voltage of the latch is then given by

$$\overline{V_{n,in}^2} = \frac{8kT\gamma B_n}{g_{m1}} + \frac{\frac{8kT\gamma B_n}{g_{m3}} + 8kTR_D B_n}{\left[g_{m1}R_D - \frac{g_{m1}}{g_{m3}}\exp\frac{(g_{m3}R_D-2)t_r}{2\alpha R_D C_L}\right]^2} \quad (24)$$

where B_n denotes the noise bandwidth at the output node ($\approx \pi/2$ times the −3-dB bandwidth).

E. Latch Sensitivity

For the latch of Fig. 2(b), the output voltage in the regeneration mode is expressed as

$$V_{out}(t) = G_{m1}R_D V_{in0}\exp\frac{(G_{m3}R_D - 1)t}{R_D C_L} \quad (25)$$
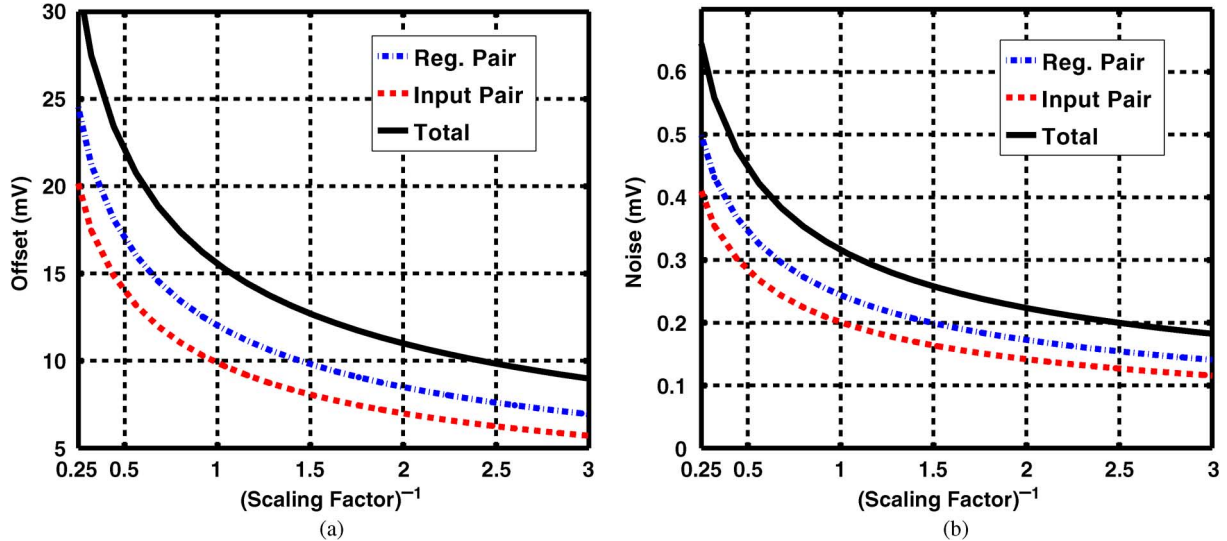
Fig. 9. Latch input-referred (a) 5σ offset and (b) noise voltages as a function of scaling factor.

where V_{in0} is the input voltage at the end of the sense mode. As mentioned in Section III-B, we define the sensitivity as the input voltage that yields an output equal to 80% of the final value. It follows that

$$V_{sens} = \frac{0.8 I_{SS}}{G_{m1}} \exp\left(-\frac{G_{m3}R_D - 1}{2R_D C_L f_{CK}}\right). \qquad (26)$$

In (25), we have assumed that $V_{out}(t)$ completely settles to $G_{m1}R_D V_{in0}$ at the end of the sense mode, ignoring the effect of the previous value of the output. An identical pervious bit generates a larger value at the end of the sense mode and improves the latch sensitivity. On the other hand, an opposite previous bit prolongs the the latch's overdrive recovery, degrading the sensitivity. For the latter case, analysis yields a sensitivity of

$$V_{sens} = \frac{0.8 I_{SS}}{(1 - \kappa)G_{m1}}$$
$$\times \left(1 + \kappa \exp\frac{G_{m3}R_D - 1}{2R_D C_L f_{CK}}\right) \exp\left(-\frac{G_{m3}R_D - 1}{2R_D C_L f_{CK}}\right) \qquad (27)$$

where

$$\kappa = \exp\left(-\frac{1}{2R_D C_L f_{CK}}\right). \qquad (28)$$

The sensitivity thus has some dependence on the bit pattern. However, we neglect this dependence and assume the average sensitivity is given by (26). A conservative approach may include the worst-case sensitivity given by (27). Nonetheless, this choice does not alter the derivation of power scaling limits.

F. Resistor and Inductor Scaling

As explained in Section III-A, linear scaling of the FF means that the load resistors and inductors must be multiplied by M if the transistor widths and bias currents are divided by M. The resistors can be scaled by placing M units in series, thus *improving* their matching.

Inductor scaling, on the other hand, entails two issues: (a) larger dimensions and hence longer interconnects between the latches, and (b) greater parasitic capacitances—an effect
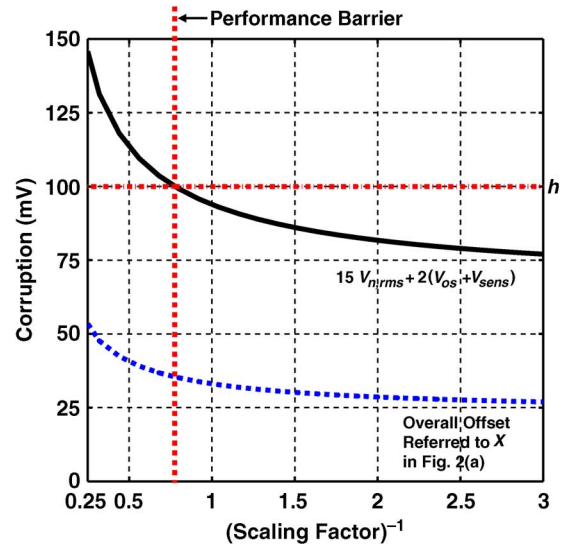

Fig. 10. Overall corruption and the performance barrier.

that conflicts with the linear scaling scenario and degrades the equalized eye. As an example, Fig. 8 plots the equivalent capacitance of a stacked inductor [11] consisting of metal 9, metal 6, and metal 3 spirals as the inductance varies from 0.5 nH to 2 nH.

G. Scaling Limits

As the FFs in a DFE are scaled down linearly,[6] the noise and offset rise and the intrinsic sensitivity degrades due to the higher parasitic capacitance of the inductors. (In this study, we neglect resistor mismatch.) A reference 10-Gb/s one-tap current-steering DFE is designed and subsequently scaled up and down. The design of the reference circuit proceeds as follows: (1) based on the supply voltage and the estimated headroom consumed by the transistors, a reasonable voltage swing is

[6]Since the "reference" design is optimized for headroom, voltage swings and speed, only linear scaling can maintain optimality.
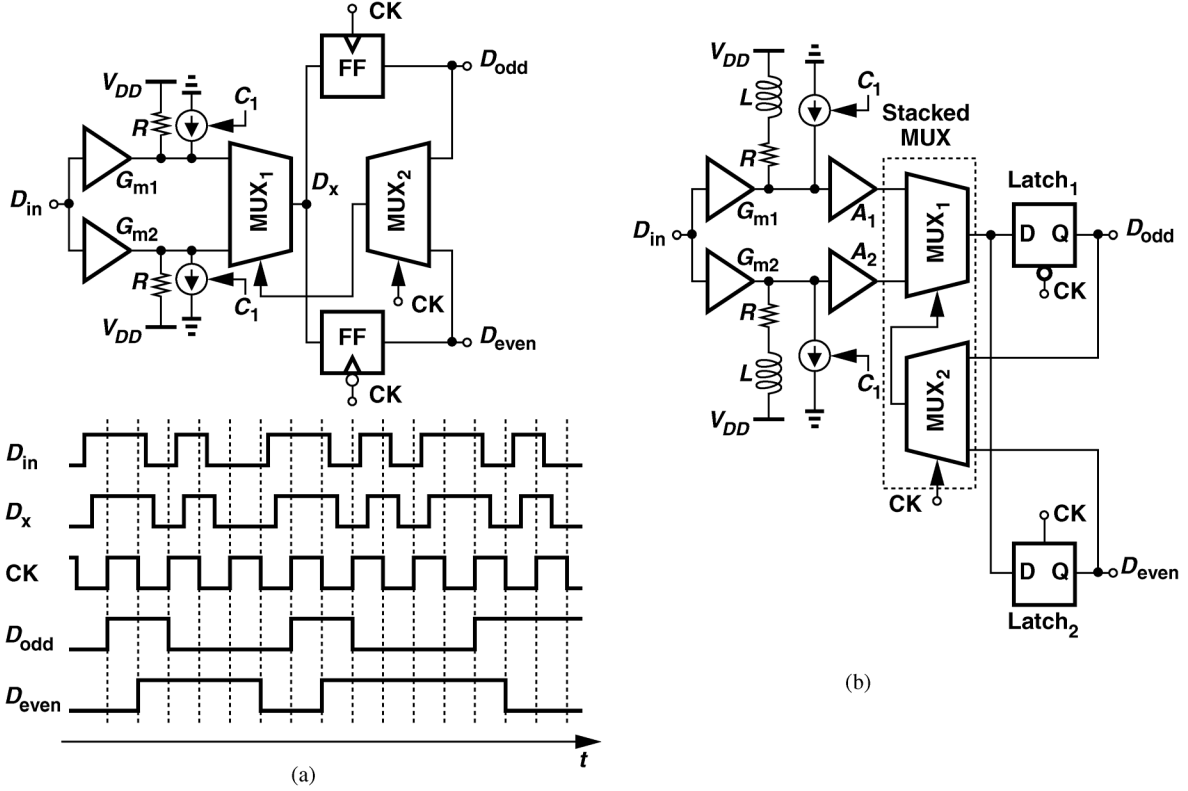
Fig. 11. Multiplexed unrolled half-rate DFE evolution: (a) Idea and waveforms, and (b) implementation.

chosen, (2) to optimize the speed for a given fanout, the widths of the differential pair and the regenerative pair are chosen, (3) for a given clock swing and to allow complete current switching, the width of the clocked transistors is chosen, (4) if the DFE loop settling requirements are not met, inductive peaking is added. The reference design employs the latch in Fig. 2(b) with $W_1/L - W_4/L = 25 \mu m/90$ nm, $R_D = 200 \Omega$, $L_D = 500$ pH, and $I_{SS} = 2$ mA.

Fig. 9(a) plots the simulated input-referred offset resulting from both the input and the regenerative pairs. Note that the standard deviation of the offset is multiplied by approximately a factor of 5 to avoid yield degradation. Fig. 9(b) plots the input-referred noise contributions. The noise simulations follow the technique described in [10]. While the latch rms noise in this example is about 1/50th of the offset, it is important to note from (11) that the noise contribution rises by a factor of 7.5 compared to the offset contribution. Moreover, if offset cancellation is used, the noise may become the design bottleneck.

Fig. 10 shows the outcome of this study. The minimum acceptable eye opening, h in (11), is also shown, revealing a *performance barrier*. This occurs at a point where the total corruption exceeds h, where h is the actual eye opening at the DFE summing node. For the reference design considered here, the barrier emerges for a scaling factor of 1.25, i.e., for $W_1 = W_2 = 20 \mu m$, $R_D = 250 \Omega$, $L_D = 625$ pH, and $I_{SS} = 1.6$ mA.

The plot of Fig. 10 implies that the offset contributes about two-thirds toward the performance barrier, providing great impetus for offset cancellation. However, the complexity and speed penalty resulting from the additional offset-canceling devices in the signal path must be carefully considered. The DFE

timing budget in all of the architectures studied in Section II is prohibitively tight and is likely to worsen with the addition of offset cancellation. For example, the ~20-Gb/s designs in [6]–[8] do not employ offset cancellation.

IV. PROPOSED DFE ARCHITECTURE

The architecture studies in Section II suggest that it is desirable to retain half-rate operation so as to simplify the CDR design. We propose to merge the half-rate architectures of Figs. 1(d) and (e) to obtain a more power-efficient solution. Called the "multiplexed unrolled half-rate" (MUHR) DFE here, the architecture has evolved as illustrated in Fig. 11. We begin with the two speculative paths of Fig. 1(b) but apply the multiplexed input data to two half-rate FFs as shown in Fig. 11(a). The outputs of these FFs must alternately control the selection of the input data, a task performed by MUX_2. The operation of the DFE can be understood with the aid of the waveforms depicted in Fig. 11(a). Upon traveling through two speculative paths, G_{m1} and G_{m2}, the data appears in two versions at the inputs of MUX_1. Based on the previous bit value, MUX_1 selects one of two. In order to correctly reproduce the previous bit by means of a half-rate clock, the two FFs sample D_X on opposite edges of CK, and MUX_2 recreates the full-rate data.

The architecture of Fig. 11(a) merits two remarks. First, the critical path delay constraint is now expressed as

$$t_{cq,FF} + t_{setup} + t_{sq,MUX_1} + t_{p,MUX_2} < UI. \qquad (29)$$

This constraint is tighter than that given by (5) for the unrolled half-rate architecture of Fig. 1(e), the price paid for nearly
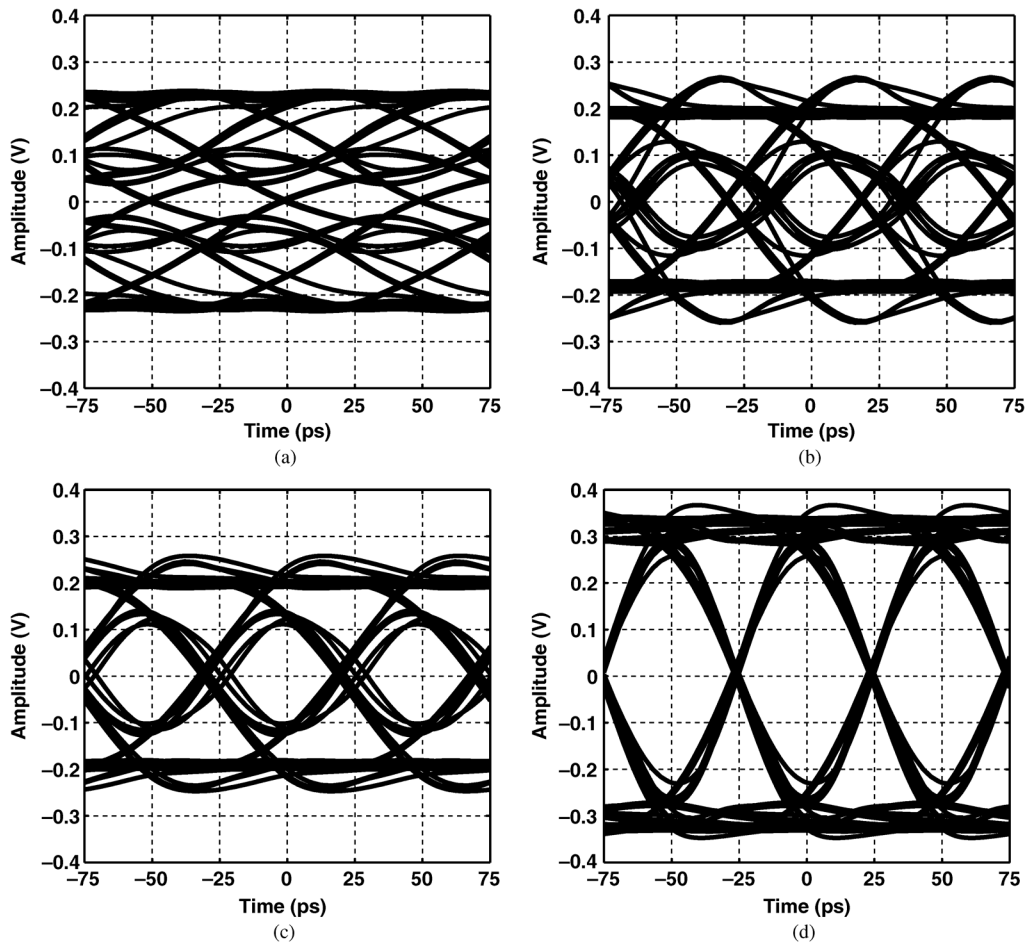
Fig. 12. Equalized eye diagrams (at the input of the FFs): (a) multiplexed half-rate DFE of Fig. 1(d), (b) original MUHR DFE of Fig. 11(a), (c) MUHR DFE with stacked multiplexers, and (d) final MUHR DFE of Fig. 11(b) using inductive peaking.

halving the power and hardware. Second, the other equalizer taps can be included by following the FFs with additional latches, multiplexing their outputs, and using the results to return currents to the summing nodes. Appendix II presents the addition of higher taps to the proposed architecture.

In order to improve the speed of the proposed DFE, we employ three techniques, arriving at the architecture shown in Fig. 11(b). First, the two (current-steering) multiplexers are stacked, thus reducing $t_{sq,MUX_1} + t_{p,MUX_2}$ to approximately t_{sq,MUX_1}. With a 1-V supply, this stacking poses circuit design issues that are discussed in Section V. Second, since the delay of the speculation paths is not critical (for the first tap), amplifiers A_1 and A_2 precede MUX_1, increasing the voltage swings and allowing faster current steering in the critical path. It is important to note that the architecture of Fig. 1(d) can also employ gain stages before the FFs [4] but it must deal with their additional delay. Our architecture, by contrast, faces the delay of the gain stages for only higher taps. Fortunately, circuit simulations indicate that the speed improvement afforded by these stages outweighs their delay contribution, thus aiding the higher taps as well.

The third technique for improving the speed is to replace the two FFs with latches [12]. To understand this point, we note that (a) when $Latch_1$ in Fig. 11(b) is in the regeneration mode,

$Latch_2$ is in the sense mode (and vice versa), (b) MUX_2 operates as a slave for each latch, e.g., when $Latch_1$ is in the regeneration mode, MUX_2 senses D_{odd} and ignores D_{even}. With this modification, the critical path delay constraint reduces to

$$t_{D,latch} + t_{d,SMUX} < UI \qquad (30)$$

where $t_{D,latch}$ is the data-to-output delay of each latch and $t_{d,SMUX}$ is the delay of the stacked MUX.

In order to demonstrate the efficacy of the techniques described above, Fig. 12 shows the transistor-level circuit simulation results for a progression of the DFE design. Here the equalized eye diagram is plotted for the multiplexed half-rate architecture of Fig. 1(d), the original MUHR DFE of Fig. 11(a), the MUHR DFE with stacked multiplexers, and the final MUHR DFE of Fig. 11(b). The higher swing in Fig. 12(d) is due to the addition of amplifiers A_1 and A_2.

Shown in Fig. 13, the overall equalizer architecture consists of a linear equalizer and a 1-tap DFE. For testing simplicity, all of the linear equalization (a boost of 9 dB) is placed on the receive side. In practice, some of this boost can be accommodated on the transmit side. (The design in [8] incorporates 9 dB of boost in the TX.) Also, a demultiplexer is added to the output of MUX_1 to facilitate testing.[7]

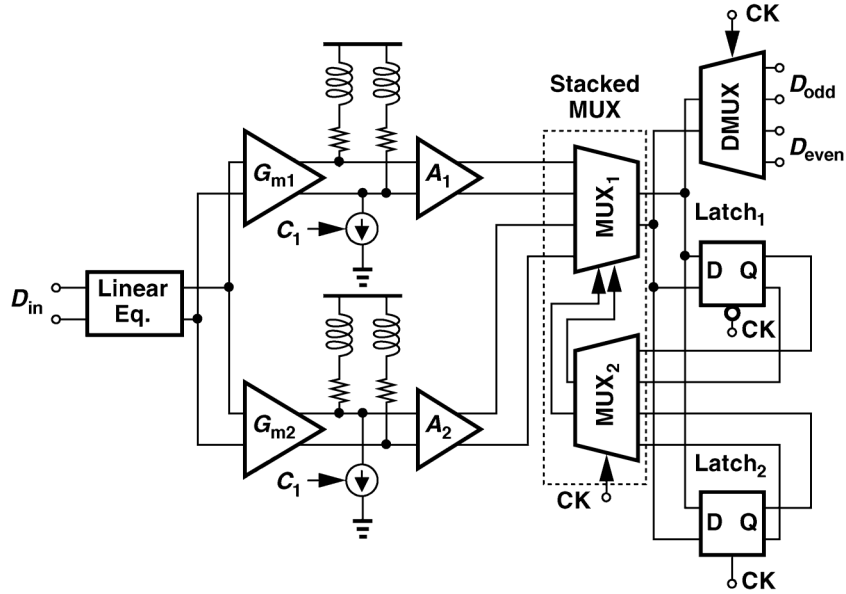[7]The two latch outputs are frozen for only half of each clock cycle.

Fig. 13.   Overall equalizer architecture.

V. BUILDING BLOCKS

The equalizer architecture of Fig. 13 has been implemented in 90-nm CMOS technology. This section describes the design of the critical building blocks.

A. Linear Equalizer

The design of linear equalizers must deal with trade-offs among bandwidth, boost factor, power dissipation, and gain (dc loss) [9]. Fortunately, our system requires a maximum boost factor of about 9 dB, allowing a small number of peaking stages. Shown in Fig. 14(a), the linear equalizer employs a high-pass path and an all-pass path for adaptation to the loss of the channel. The former path consists of a passive peaking stage with a boost (or more accurately, de-emphasis) factor of 6 dB [9], a differential pair, and a capacitively-degenerated output stage. The latter path is simply a resistively-degenerated transconductor.

The design of the passive peaking stage is governed by two issues: 1) it must provide a bandwidth of at least 10 GHz while driving the input capacitance of the differential pair, and 2) it must guarantee an input return loss of -10 dB at 10 GHz along with the capacitances of the input pads and the all-pass path. The differential pair compensates for the loss of this and the output stage, yielding an overall dc gain of a few dB.

The adaptation of the linear equalizer can be performed by an analog loop [9]. In this work, however, the boost factor is adjusted externally in discrete steps so as to allow greater flexibility in testing. To this end, the tail current sources of the two degenerated stages are decomposed into a segmented array of eight units [Fig. 14(b)] and controlled through a serial bus.

With the two-path adaptation scheme, the dc gain tends to vary with the boost factor. This is because the sum of the gains through the high-pass and all-pass paths varies even though the total current in the output stage remains relatively constant. To alleviate this issue, the all-pass transconductor is kept partially on for all settings [by means of the bottom transistors in

Fig. 14(b)]. Additionally, for the minimum-boost setting, M_x turns on, shorting the degeneration resistor and increasing the dc gain.

Fig. 15 plots the equalizer's boost profile for different thermometer-code settings, indicating a maximum boost of 13 dB at 7 GHz. This bandwidth would be inadequate for a stand-alone 20-Gbs linear equalizer, but the DFE used here corrects for the residual errors.

B. Analog Summer

The analog summers at the DFE input are realized as shown in Fig. 16. For a one-tap system, only a logical ONE or ZERO need be speculated, a task performed by drawing a constant current from one of the output nodes. The sign of the tap is controlled by M_3 and M_4, and the magnitude by their tail current. With a 6-bit digital control, the tap coefficient can vary in steps of 30 μA.

C. Feedback Latch

The feedback latches in the architecture of Fig. 13 serve as both storage elements and slicing devices. As such, they must exhibit a short delay with high sensitivity. Depicted in Fig. 17, each latch incorporates class-AB clocking [13] and inductive peaking to maximize the speed. The choice of device dimensions and bias values is not straightforward because the latch both is driven and drives the multiplexer in Fig. 13. That is, the latch/MUX loop must be optimized as one entity. The use of inductive peaking in both facilitates this optimization.

Simulations indicate that the latch of Fig. 17 exhibits a sensitivity of 10 mV and a clock-to-output delay of 17 ps. This sensitivity is obtained for the typical corner of the process at the room temperature. In practice, a larger input level is required to guarantee operation at other corners and temperatures. The circuit provides a differential output swing of 400 mV$_{pp}$ with a common-mode (CM) level of 650 mV. Created by R_1, this CM level is necessary for the stacked MUX input. The level shift

$R_1 = 300 \Omega \quad C_1 = 300 \text{ fF}$
$R_2 = 500 \Omega \quad C_2 = 900 \text{ fF}$
$R_3 = 200 \Omega \quad L_D = 1.8 \text{ nH}$
$R_4 = 160 \Omega \quad W_1 - W_2 = 15 \, \mu\text{m}$
$R_5 = 700 \Omega \quad W_3 - W_6 = 40 \, \mu\text{m}$
$R_6 = 1 \text{ k}\Omega \quad L_1 - L_6 = 90 \text{ nm}$
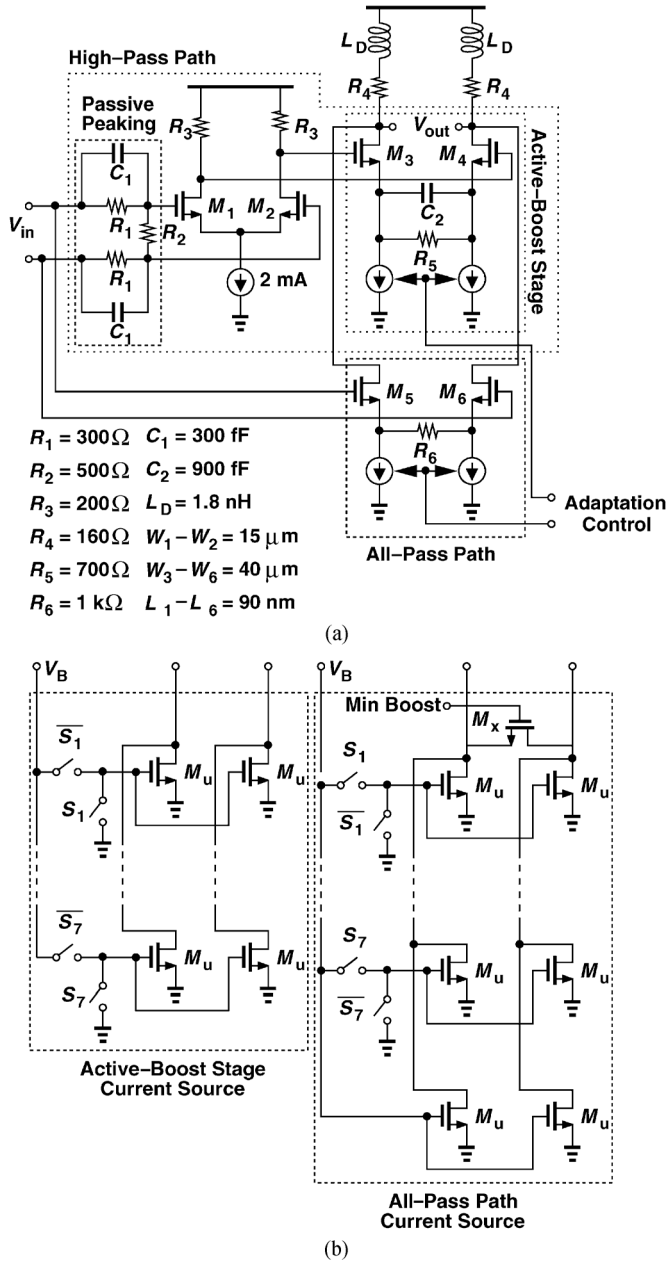
(a)

(b)

Fig. 14.   Linear Equalizer: (a) circuit realization and (b) adaptation control.

provided by R_1 is determined by the average current drawn by the class-AB pair, which itself is set by a current mirror.

D. Stacked Multiplexer

The stacked MUX must operate with moderate data swings and a 1-V supply. Shown in Fig. 18, the circuit ensures that the differential pair transistors, $M_1 - M_8$, do not enter the triode region so that they can steer their tail currents without rail-to-rail data swings. As mentioned in the previous section, the outputs of the two latches are shifted down for this purpose. Class-AB clocking both improves the speed and alleviates the headroom limitation.

The MUX must drive the feedback latches and the DMUX in Fig. 13, demanding inductive peaking. According to simula-
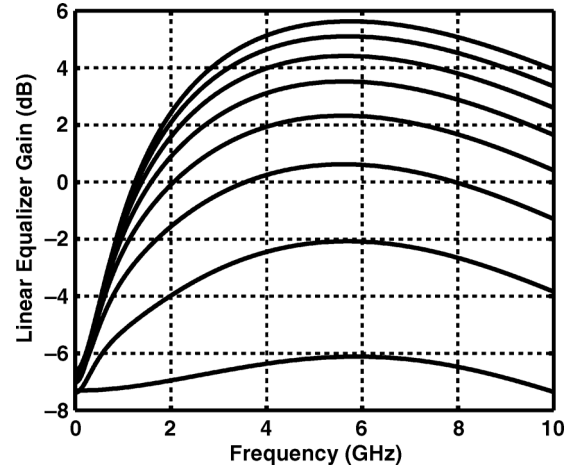

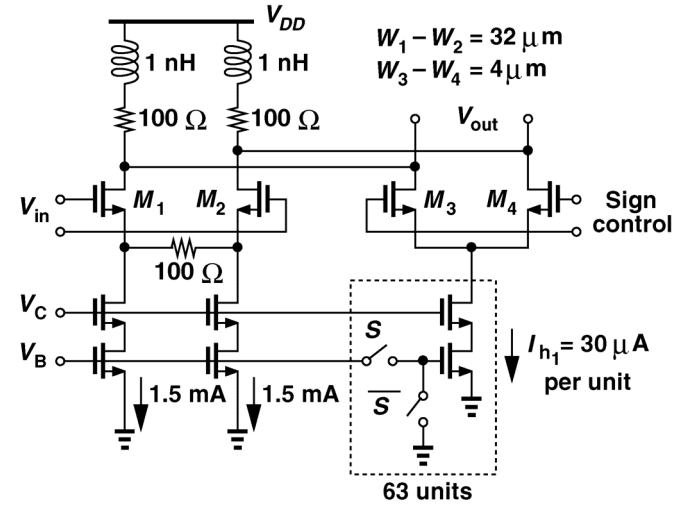Fig. 15.   Simulated gain settings of the linear equalizer.
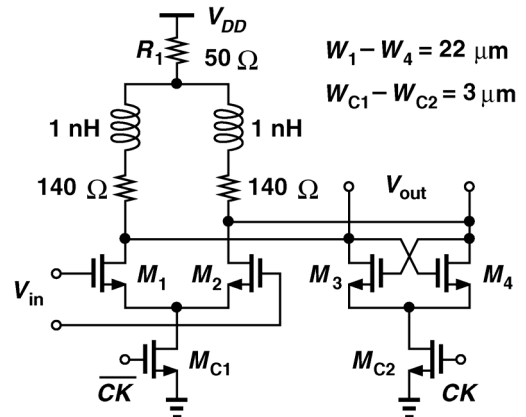

Fig. 16.   DFE analog summer.

Fig. 17.   DFE feedback latch.

tions, the MUX has a delay of 10 ps from the lower data inputs (the gates of $M_5 - M_8$) to the output while consuming 5 mW.

E. Effect of Offset and Noise

The analysis in Section III has been extended to the overall equalizer design so as to quantify its power and BER bounds
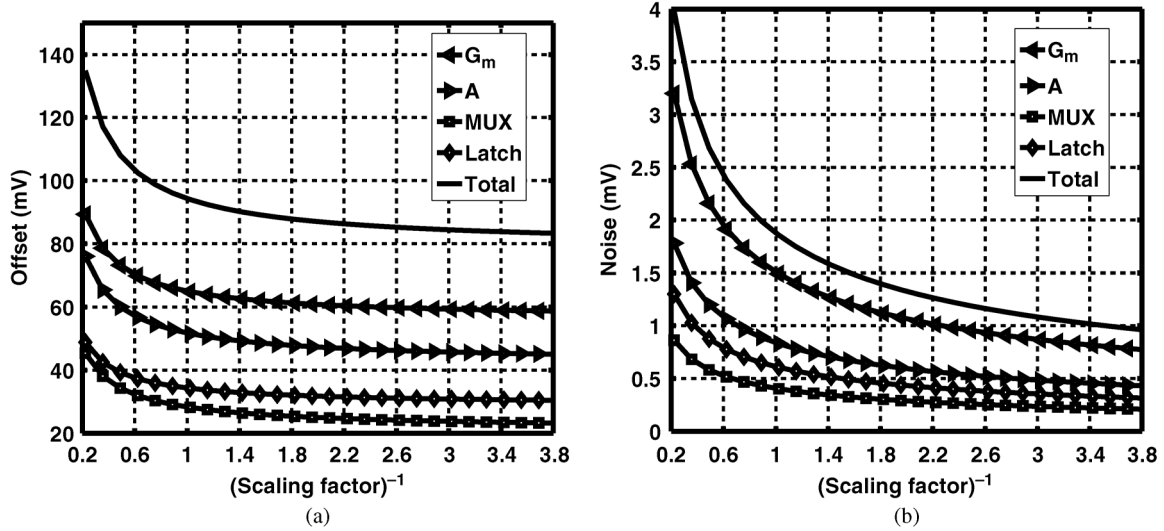
Fig. 19. (a) MUHR DFE offset voltage and (b) MUHR DFE thermal noise voltage at the SMUX output as a function of the scaling factor.
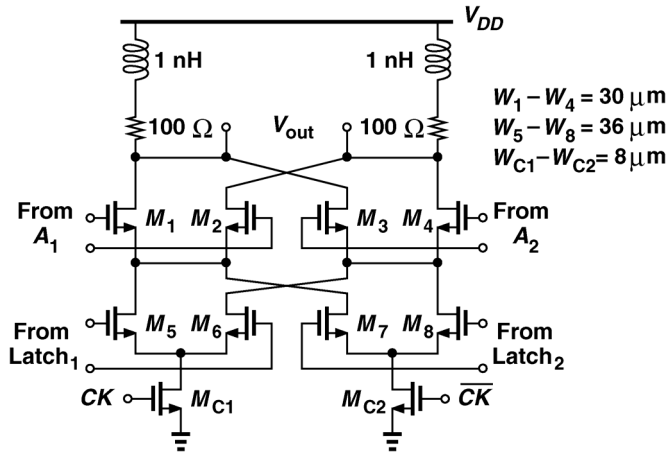

Fig. 18. DFE stacked MUX.

Fig. 20. MUHR DFE required vertical eye opening as a function of the scaling factor. Note from Eq. (11) that the sensitivity and offset are multiplied by two here.

(Appendix III). In this analysis, the offset and noise of the building blocks are referred to the stacked MUX output, where the final equalized eye is to be sensed.

Fig. 19(a) plots the offset contributions at the MUX output as a function of the scaling factor. To obtain these plots, the reference design is simulated at the transistor level and the results are subsequently scaled. It is observed that the linear equalizer contributes the largest offset owing to the gain stages within and following it. Fig. 19(b) repeats the analysis for the noise contributions.

Based on these results, the minimum acceptable eye opening at the MUX output can be computed. Fig. 20 shows the trend versus the scaling factor. In this work, a scaling factor of 1 is employed along with an eye opening of $400 \, \mathrm{mV}_{pp}$ to leave sufficient margin for successful testing.

VI. EXPERIMENTAL RESULTS

The prototype has been fabricated in TSMC's 90-nm CMOS technology and tested at 20 Gb/s in a chip-on-board assembly. Fig. 21 shows the core of the die and identifies the building blocks. The core occupies an area of about $300 \, \mu\mathrm{m} \times 300 \, \mu\mathrm{m}$.
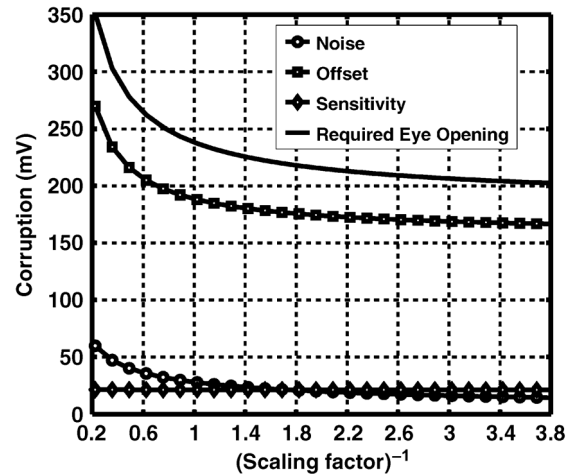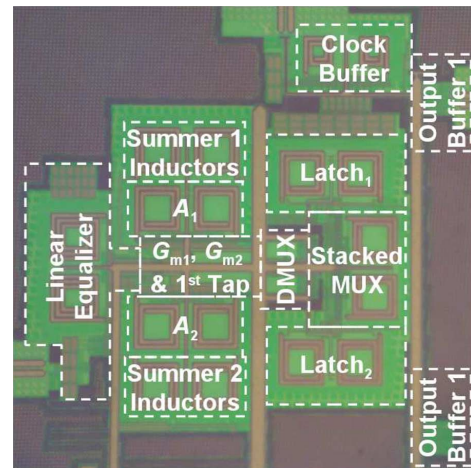

Fig. 21. Equalizer core die photo.

The equalizer has been tested with 6-in and 18-in FR4 traces. Fig. 22 plots the measured frequency response of each trace
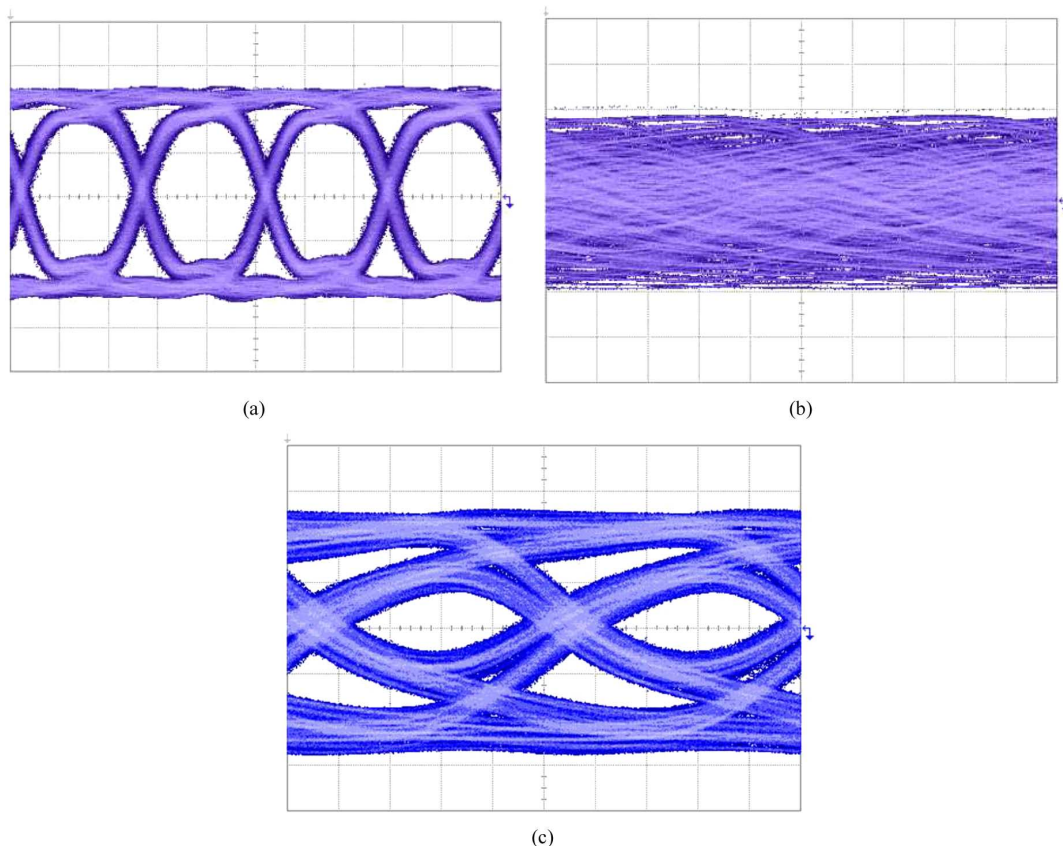
Fig. 23.   Measured eye diagrams: (a) PRBS generator output, (b) 18-in FR4 channel and cables output, and (c) half-rate DMUX output (horizontal scale: 20 ps/div, vertical scale: 100 mV/div for (a) and (b) and 50 mV/div for (c)).
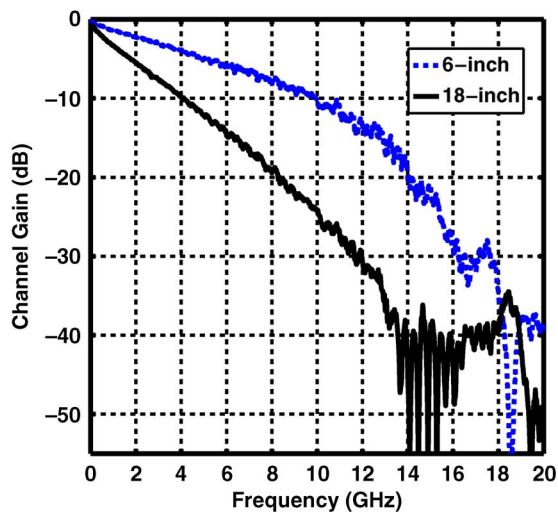

Fig. 22.   Measured frequency response of the channel.

with the cables and dc blocks used in the test setup. The loss at 10 GHz reaches 10 dB and 24 dB for the 6-in and 18-in traces, respectively.

Fig. 23(a) shows the PRBS generator output at 20 Gb/s, revealing a peak-to-peak jitter of 10 ps at the input of the channel. This, together with 7 ps of jitter in the external clock (provided by an RF generator), limits the horizontal eye opening and hence the clock phase margin that the prototype can tolerate.

Fig. 23(b) shows the eye diagram at the end of the 18-in trace, and Fig. 23(c) the half-rate output produced by the on-chip DMUX. The output buffer bandwidth limitations cause some eye closure, but the opening is adequate for the bit error rate tester (BERT).

Fig. 24 shows the bathtub curves with 20-Gb/s PRBS data of length of $2^7 - 1$ for 6-in and 18-in traces. To achieve BER $< 10^{-12}$, the equalizer allows a clock phase margin of 0.44 and 0.36 unit interval for the two traces, respectively. In these tests, the linear equalizer provides no boost for the 6-in trace and 9 dB of boost for the 18-in trace.

The overall equalizer draws 40 mW from a 1-V supply, of which 5 mW is consumed by the linear equalizer, 17 mW by the Gm and amplifying stages, 6 mW by the MUX, and 12 mW by the two latches.

Table I summarizes the performance and compares the results with those of the prior DFEs running at data rates around 20 Gb/s. While our work has a similar power efficiency to those of [7] and [8], a few remarks help create a perspective here. Among the NRZ systems, the design in [7] suffers from a high BER and a narrow clock phase margin (horizontal eye opening). Moreover, the 65-nm design in [8] employs 20 dB of linear equalization for a loss of 21 dB, greatly relaxing the DFE requirements. By comparison, our 90-nm prototype allows only 9 dB of linear equalization for 24 dB of loss.

In addition to speed and power consumption, the specifications of equalizers must reflect the amount of channel loss that

TABLE I
PERFORMANCE SUMMARY AND COMPARISON TO PRIOR ART

| Reference	[6]	[7]	[8]	This work
Data Rate (Gb/s)	18	19	21	20
Architecture	4-tap DFE 1st tap assisted by CDR	1-tap Speculative DFE	Linear equalizer + 1-tap DFE	Linear equalizer 1-tap speculative DFE
Signaling	Duobinary	NRZ	NRZ	NRZ
Clocking	Quarter Rate	Half Rate	Full Rate	Half Rate
Channel Loss (dB)				
Compensated in RX	14	11	11.7	24
BER	10^{-12}	10^{-8}	10^{-12}	10^{-12}
Horizontal eye opening	12% UI	9% UI	N/A	36% UI
Supply (V)	1.2	1	1.2	1
Power (mW)	100.2 (Full RX)	38	42	40
Area (mm^2)	0.2 (Full RX)	0.02	0.04	0.09
Power efficiency (pJ/bit)	5.56 (Full RX)	2	2	2
FOM (pJ/bit)	1.111 (Full RX)	0.564	0.52	0.126
FOM (pJ/bit/dB)	0.398 (Full RX)	0.182	0.171	0.083
Technology	90-nm	90-nm	65-nm	90-nm
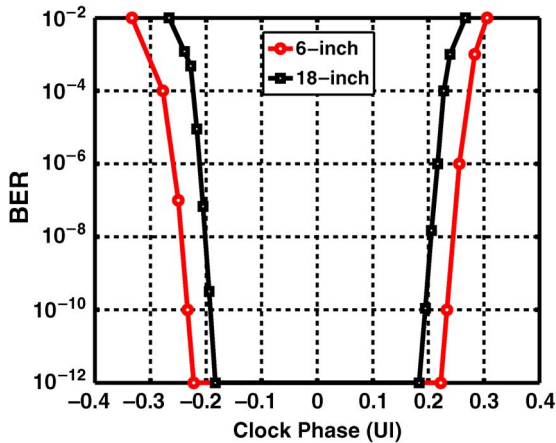
Fig. 24. Bathtub curves for 20-Gb/s PRBS7 data stream.

they compensate. To this end, a figure of merit (FOM) can be defined as

$$FOM = \frac{P}{r_b \times L} \quad (31)$$

where P is the power dissipation, r_b the data rate, and L the channel loss at "Nyquist" frequency. The channel loss can be included as a numerical value (as suggested by [14] for continuous-time equalizers) or a logarithmic value. We employ the latter as it provides a more conservative FOM. This FOM, of course, is only a rough measure of the performance because, for example, a 6-dB increase in the channel loss may not exactly translate to twice the power consumption. The FOM shown in the table is for the equalizer section of each design unless, otherwise stated. Another advantage of the proposed architecture over the others in Table I is that the half-rate operation also saves power in the clock and data recovery circuit.

VII. CONCLUSION

This work has formulated limits of scaling for equalizers and introduced a new DFE architecture that achieves high speed with low power consumption. The architecture merges multiplexed and half-rate speculative topologies to reduce the complexity and power consumption. The power scaling methodology is applied to a 1-tap prototype to allow operation at 20 Gb/s with 40 mW while compensating for a total loss of 24 dB.

APPENDIX I
DIFFERENTIAL EQUATION FOR LATCH WITH GRADUAL CLOCK

With the approximations given by (17) for $g_{m1}(t)$ and $g_{m2}(t)$, the regeneration differential equation appears as

$$G_{m3}\sqrt{\frac{\alpha t}{t_r}}(V_{out} + \Delta V_{TH3}) + G_{m1}\sqrt{1 - \frac{\alpha t}{t_r}}V_{in}$$
$$= \frac{V_{out}}{R_D} + C_L\frac{dV_{out}}{dt}. \quad (32)$$

The solution is as follows

$$v_{out}(t) = \left[\int \left(\frac{G_{m1}V_{in}}{C_L}\sqrt{1 - \frac{\alpha t}{t_r}} + \frac{G_{m3}\Delta V_{TH3}}{C_L}\sqrt{\frac{\alpha t}{t_r}}\right) \times \exp m(t)dt + \kappa\right]\exp\left[-m(t)\right] \quad (33)$$

where

$$m(t) = \int \left(\frac{1}{R_D C_L} - \frac{G_{m3}}{C_L}\sqrt{\frac{\alpha t}{t_r}}\right)dt. \quad (34)$$

No closed-form solution exists here, but the result can be solved numerically for the typical values used in the latch design. Inspection of the numerical solutions suggests the simple modification shown in (18).
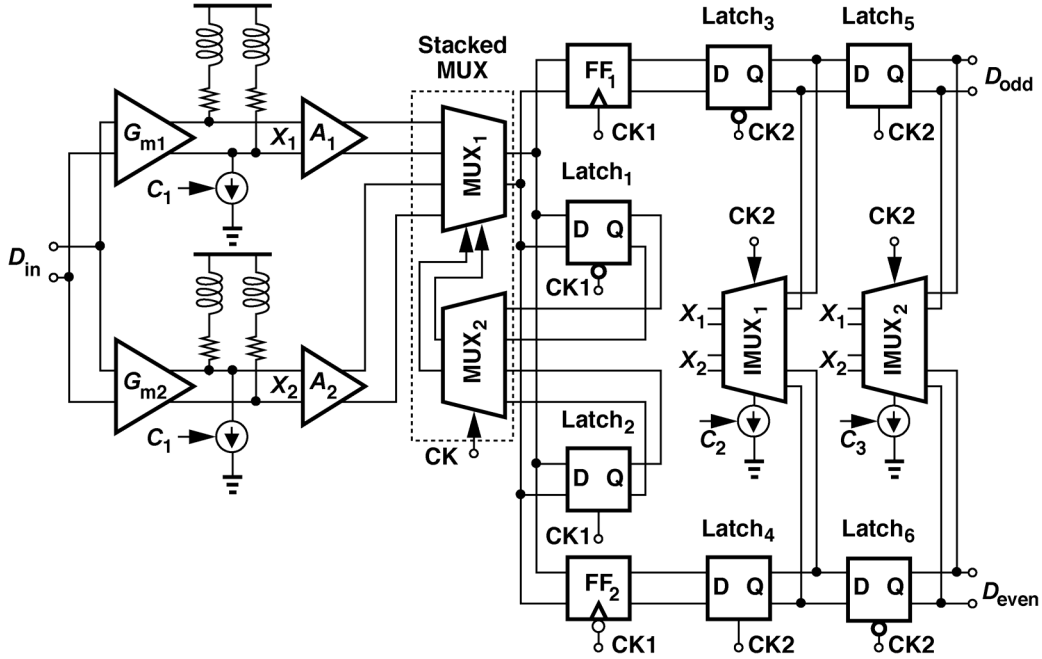
Fig. 25. Proposed MUHR DFE extended to 3 taps.

APPENDIX II
ADDING HIGHER TAPS TO THE PROPOSED ARCHITECTURE

For channels with higher losses, more taps are required to cancel higher-order post-cursor ISI. The speed requirements are not as critical as the first tap, though, because signals with greater swings and at least 2 UIs relax the settling. Fig. 25 shows the proposed MUHR DFE when extended to three taps. Flipflops FF_1 and FF_2 sample the data and $Latch_1 - Latch_4$ and current multiplexers $IMUX_{1,2}$ generate the feedback currents based on the previous bits. The feedback currents are then injected to the summing nodes X_1 and X_2. Note that these stages need not use inductive peaking.

A different clock (CLK2) is used to alternately latch and select the previous bits. This is to ensure that the crossing point of the feedback signal coincides with that of the input signal [4]. A programmable delay is necessary to adjust the timing between the two clocks [4]. In the design, the delay must vary from 10 ps to 25 ps.

A channel with about 32 dB of loss at Nyquist was used to simulate the above DFE. Fig. 26 shows the simulated eye diagram at the output of the SMUX. With the aid of 9 dB of linear equalization, the 3-tap DFE produces an eye opening of 260 mV and 42 ps.

APPENDIX III
SCALING LIMITS OF THE MUHR DFE

As explained in Section III, offset, noise, and latch sensitivity set a limit on scaling down a DFE. The MUHR DFE proposed in Section IV is also affected by the noise and offset voltages of the transconductors G_{m1} and G_{m2}, the amplifiers A_1 and A_2, the stacked multiplexer and the feedback latches shown in Fig. 11(b). When referred to the input of the latches, these nonidealities determine the required vertical eye opening given in (11).
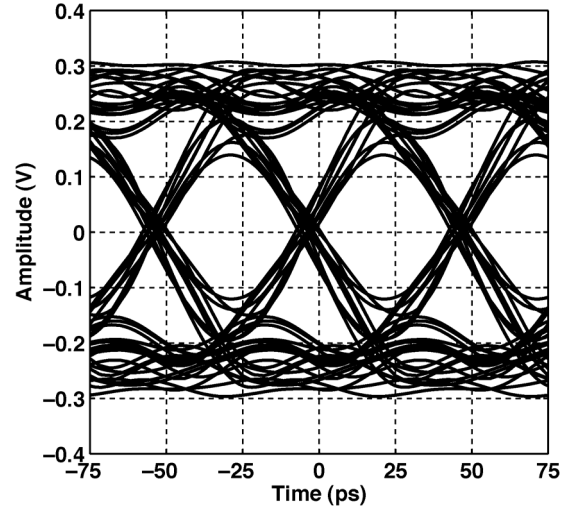

Fig. 26. Simulated eye diagram at the output of the SMUX for the 3-tap MUHR DFE.

At the input of the latches (the output of the stacked MUX), the input-referred offset voltage is given by

$$\overline{V_{OS,Eq}^2} = \overline{V_{OS,G_m}^2}\left(A_{G_m}A_{AMP}A_{SMUX}\right)^2$$
$$+ \overline{V_{OS,A}^2}\left(A_{AMP}A_{SMUX}\right)^2$$
$$+ \overline{V_{OS,SMUX}^2}\left(A_{SMUX}\right)^2 + \overline{V_{OS,L}^2} \qquad (35)$$

where $\overline{V_{OS,G_m}}$, $\overline{V_{OS,A}}$, $\overline{V_{OS,SMUX}}$ and $\overline{V_{OS,L}}$ are the input-referred offset voltages of the transconductor, the amplifier, the stacked MUX and the feedback latch, respectively. A_{G_m}, A_{AMP} and A_{SMUX} are the voltage gains of the transconductor, the amplifier and the stacked MUX, respectively.

The offset of the latch is given by (21). For the transconductor, the amplifier and the stacked MUX, the offsets are given by

$$
\overline{V_{OS,G_m}^2} = (\Delta V_{th,in,G_m})^2
$$
$$
+ \left(\frac{V_{ov,in,G_m}}{2}\right)^2 \left(\frac{\Delta \beta_{in,G_m}}{\beta_{in,G_m}}\right)^2
$$
$$
+ (\Delta V_{th,cs,G_m})^2 \left(g_{m,cs,G_m}\frac{R_{S,G_m}}{2}\right)^2
$$
$$
+ \left(\frac{\Delta \beta_{cs,G_m}}{\beta_{cs,G_m}}\right)^2 \left(\frac{I_{SS,G_m}R_{S,G_m}}{4}\right)^2
$$
$$
+ \left(\frac{V_{ov,in,G_m}}{2}\right)^2 \left(\frac{\Delta R_{D,G_m}}{R_{D,G_m}}\right)^2
$$
$$
\times \left(1 + g_{m,in,G_m}\frac{R_{S,G_m}}{2}\right)^2 \tag{36}
$$

$$
\overline{V_{OS,A}^2} = (\Delta V_{th,in,A})^2 + \left(\frac{V_{ov,in,A}}{2}\right)^2
$$
$$
\times \left[\left(\frac{\Delta \beta_{in,A}}{\beta_{in,A}}\right)^2 + \left(\frac{\Delta R_{D,A}}{R_{D,A}}\right)^2\right] \tag{37}
$$

$$
\overline{V_{OS,SMUX}^2} = (\Delta V_{th,in,SMUX})^2 + \left(\frac{V_{ov,in,SMUX}}{2}\right)^2
$$
$$
\times \left[\left(\frac{\Delta \beta_{in,SMUX}}{\beta_{in,SMUX}}\right)^2\right.
$$
$$
\left. + \left(\frac{\Delta R_{D,SMUX}}{R_{D,SMUX}}\right)^2\right]. \tag{38}
$$

These offsets are shown in Fig. 9(a) as a function of the scaling factor.

Similarly the thermal noise voltage at the output of the stacked MUX is given by

$$
\overline{V_{n,Eq}^2} = \overline{V_{n,G_m}^2}\left(A_{G_m}A_{AMP}A_{SMUX}\right)^2
$$
$$
+ \overline{V_{n,A}^2}(A_{AMP}A_{SMUX})^2 + \overline{V_{n,SMUX}^2}(A_{SMUX})^2 + \overline{V_{n,L}^2} \tag{39}
$$

where $\overline{V_{n,G_m}}$, $\overline{V_{n,A}}$, $\overline{V_{n,SMUX}}$ and $\overline{V_{n,L}}$ are the input-referred thermal noise voltages of transconductor, the amplifier, the stacked MUX and the feedback latch, respectively.

The thermal noise of the latch is given by (24). For the transconductor, the amplifier and the stacked MUX, the noise voltages are given by

$$
\overline{V_{n,G_m}^2} = 2\left[\frac{4kT\gamma B_{n,SMUX}}{g_{m,in,G_m}}\right.
$$
$$
+ \frac{4kT\gamma g_{m,cs,G_m}R_{S,G_m}^2 B_{n,SMUX}}{4}
$$
$$
+ \frac{4kTR_{D,G_m}B_{n,SMUX}}{(g_{m,in,G_m}R_{D,G_m})^2}
$$
$$
\left. + \frac{4kTR_{S,G_m}B_{n,SMUX}}{2}\right] \tag{40}
$$

$$
\overline{V_{n,A,in}^2} = 2\left[\frac{4kT\gamma B_{n,SMUX}}{g_{m,in,A}}\right.
$$
$$
\left. + \frac{4kTR_{D,A}B_{n,SMUX}}{(g_{m,in,A}R_{D,A})^2}\right] \tag{41}
$$

$$
\overline{V_{n,SMUX,in}^2} = 2\left[\frac{4kT\gamma B_{n,SMUX}}{g_{m,in,SMUX}}\right.
$$
$$
\left. + \frac{4kTR_{D,SMUX}B_{n,SMUX}}{(g_{m,in,SMUX}R_{D,SMUX})^2}\right] \tag{42}
$$

where $B_{n,SMUX}$ is the noise bandwidth defined from the noise source to the output of the stacked MUX and is equal to

$$
B_{n,SMUX} = \frac{\pi}{2}BW_{SMUX}. \tag{43}
$$

Here, BW_{SMUX} is the stacked MUX bandwidth, which is quite smaller than that of the amplifier and transconductor because of the high capacitive load. The dominant pole of the stacked MUX frequency response therefore determines the overall noise bandwidth for the different noise sources in the chain. These thermal noise voltages are shown in Fig. 19(b) as a function of the scaling factor.

Fig. 20 shows the required vertical eye opening given by (11) as the design scales. It also shows the contribution of $\overline{V_{OS,Eq}}$, $\overline{V_{n,Eq}}$ and $V_{sens,L}$ to the required eye opening.

REFERENCES

[1] S. Ibrahim and B. Razavi, "A 20-Gb/s 40-mw equalizer in 90-nm CMOS technology," in *IEEE Int. Solid-State Circuits (ISSCC) Dig. Tech. Papers*, 2010.

[2] A. Garg, A. C. Carusone, and S. P. Voinugescu, "A 1-Tap 40-Gb/s look-ahead decision feedback equalizer in 0.18- μm SiGe BiCMOS technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 10, Oct. 2006.

[3] Y.-S. Sohn, S.-J. Bae, H.-J. Park, and S.-I. Cho, "A 1.2 Gbps CMOS DFE receiver with the extended sampling time window for application to the SSTL channel," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2002.

[4] R. Payne, P. Landman, B. Bhakta, S. Ramaswamy, S. Wu, J. D. Powers, M. U. Erdogan, A. Yee, R. G. L. Wu, Y. Xie, B. Parthasarathy, K. Brouse, W. Mohammed, K. Heragu, V. Gupta, L. Dyson, and W. Lee, "A 6.25-Gb/s binary transceiver in 0.13- μm CMOS for serial data transmission across high loss legacy backplane channels," *IEEE J. Solid-State Circuits*, vol. 40, no. 12, pp. 2646–2657, Dec. 2005.

[5] J. F. Bulzacchelli, M. Meghelli, S. V. Rylov, W. Rhee, A. Rylyakov, H. A. Ainspan, B. D. Parker, M. P. Beakes, A. Chung, T. J. Beukema, P. K. Pepeljugoski, L. Shan, Y. H. Kwark, S. Gowda, and D. J. Friedman, "A 10-Gb/s 5-tap DFE/4-tap FFE transceiver in 90-nm CMOS technology," *IEEE J. Solid-State Circuits*, vol. 41, no. 12, pp. 2885–2900, Dec. 2006.

[6] K. Sunaga, H. Sugita, K. Yamaguchi, and K. Suzuki, "An 18 Gb/s duobinary receiver with a CDR-assisted DFE," in *IEEE Int. Solid-State Circuits (ISSCC) Dig. Tech. Papers*, 2009.

[7] D. Z. Turker, A. Rylyakov, D. Friedman, S. Gowda, and E. Sanchéz-Sinencio, "A 19 Gb/s 38 mW 1-Tap speculative DFE in 90 nm CMOS," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2009.

[8] H. Wang, C. Lee, A. Lee, and J. Lee, "A 21-Gb/s 87-mW transceiver with FFE/DFE/linear equalizer in 65-nm CMOS technology," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2009.

[9] S. Gondi and B. Razavi, "Equalization and clock and data recovery techniques for 10-Gb/s CMOS serial-link receivers," *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 1999–2011, Sep. 2007.

[10] J. R. B. S. Leibowitz, J. Kim, and C. J. Madden, "Characterization of random decision errors in clocked comparators," in *Proc. IEEE Custom Integrated Circuits Conf.*, 2008.

[11] A. Zolfaghari, A. Chan, and B. Razavi, "Stacked inductors and transformers in CMOS technology," *IEEE J. Solid-State Circuits*, vol. 36, no. 4, pp. 620–628, Apr. 2001.

[12] A. Emami-Neyestanak, A. Varzaghani, J. F. Bulzacchelli, A. Rylyakov, C. K. Yang, and D. J. Friedman, "A 6.0-mW 10.0-Gb/s receiver with switched-capacitor summation DFE," *IEEE J. Solid-State Circuits*, vol. 42, no. 4, pp. 889–896, Apr. 2007.

[13] 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.

[14] D. H. Shin, J. E. Jang, F. O'Mahony, and C. P. Yue, "A 1-mW 12-Gb/s continuous-time adaptive passive equalizer in 90-nm CMOS," in *Proc. IEEE Custom Integrated Circuits Conf.*, 2009.

Behzad Razavi (F'03) received the BSEE degree from Sharif University of Technology in 1985 and the MSEE and PhDEE degrees from Stanford University in 1988 and 1992, respectively. He was with AT&T Bell Laboratories and Hewlett-Packard Laboratories until 1996. Since 1996, he has been Associate Professor and subsequently Professor of electrical engineering at University of California, Los Angeles. His current research includes wireless transceivers, frequency synthesizers, phase-locking and clock recovery for high-speed data communications, and data converters.

Prof. Razavi was an Adjunct Professor at Princeton University from 1992 to 1994, and at Stanford University in 1995. He served on the Technical Program Committees of the International Solid-State Circuits Conference (ISSCC) from 1993 to 2002 and VLSI Circuits Symposium from 1998 to 2002. He has also served as Guest Editor and Associate Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, and INTERNATIONAL JOURNAL OF HIGH SPEED ELECTRONICS.

Prof. Razavi received the Beatrice Winner Award for Editorial Excellence at the 1994 ISSCC, the best paper award at the 1994 European Solid-State Circuits Conference, the best panel award at the 1995 and 1997 ISSCC, the TRW Innovative Teaching Award in 1997, the best paper award at the IEEE Custom Integrated Circuits Conference in 1998, and the McGraw-Hill First Edition of the Year Award in 2001. He was the co-recipient of both the Jack Kilby Outstanding Student Paper Award and the Beatrice Winner Award for Editorial Excellence at the 2001 ISSCC. He received the Lockheed Martin Excellence in Teaching Award in 2006, the UCLA Faculty Senate Teaching Award in 2007, and the CICC Best Invited Paper Award in 2009. He was also recognized as one of the top 10 authors in the 50-year history of ISSCC.

Professor Razavi is an IEEE Distinguished Lecturer, a Fellow of IEEE, and the author of *Principles of Data Conversion System Design* (IEEE Press, 1995), *RF Microelectronics* (Prentice Hall, 1998) (translated to Chinese, Japanese, and Korean), *Design of Analog CMOS Integrated Circuits* (McGraw-Hill, 2001) (translated to Chinese, Japanese, and Korean), *Design of Integrated Circuits for Optical Communications* (McGraw-Hill, 2003), and *Fundamentals of Microelectronics* (Wiley, 2006) (translated to Korean and Portuguese), and the editor of *Monolithic Phase-Locked Loops and Clock Recovery Circuits* (IEEE Press, 1996), and *Phase-Locking in High-Performance Systems* (IEEE Press, 2003).

Sameh Ibrahim (M'08) received the B.Sc. and M.Sc. degrees in electrical engineering from Ain Shams University, Cairo, Egypt, in 2001 and 2005, respectively. He received the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 2009.

In January 2010, he joined Marvell Semiconductor Inc., Santa Clara, CA, where he is a Senior Analog Design Engineer in the high-speed serial-links group. His research interests include analog/mixed-signal IC design for wireline and wireless applications, high speed serial links using analog DFEs, multi-tone signaling and linear equalization, and system design for wireline and wireless applications.