A 56 Gb/s 6 mW 300 um^2 inverter-based CTLE for short-reach PAM2 applications in 16 nm CMOS

Kevin Zheng[*], Yohan Frans[†], Ken Chang[†], Boris Murmann[*]

[*]Department of Electrical Engineering, Stanford University, CA USA

[†]Xilinx Inc. San Jose, CA USA

Abstract— **A 56 Gb/s inverter-based continuous time linear equalizer (CTLE) in 16 nm FinFET CMOS is presented. The inverters are biased with a regulated ground supply, using a ring oscillator based feedback loop that stabilizes the inverters' unity gain frequency f$_u$ and thus tracks PVT variations. The CTLE core measures only 20 um x 15 um and consumes 6 mW. Using the CTLE as the only means of RX equalization, the transceiver achieves 31% UI margin at BER of 1e-12 for a channel with 8 dB loss at 28 GHz.**

Keywords—CTLE; inverter; CMOS; short reach; ring oscillator; ground regulation; LDO

I. INTRODUCTION

Recent standards, such as CEI-56G-XSR-NRZ, drive the demand for short-reach, high speed wireline interfaces for die-to-die and chip-to-chip links with short PCB traces (see Fig. 1). As a result of the short trace, impedance discontinuities have a relatively small impact on the channel, as seen by the smooth S21 roll-off and pulse response. This type of channel can be equalized effectively using only a continuous-time linear equalizer (CTLE) at the receiver side. Despite these relaxed requirements and low channel loss (<10 dB), it is still challenging to implement the CTLE due to the high bandwidth requirements.

A conventional CML-style CTLE uses an RC source degeneration network to realize low frequency de-emphasis and high frequency peaking (see Fig. 2). Shunt inductors or T-coils are needed to meet the bandwidth specifications and result in a large area overhead, which is a significant issue given today's stringent I/O density requirements. Recent work on inverter-based equalizers [1] has shown potential for area reduction while maintaining low power. However, the work

of [1] does not address the issue of process-voltage-temperature (PVT) variation.

This paper presents a PVT-hardened, inductor-less, inverter-based, single-stage CTLE optimized for short-reach PAM2 applications. Using inverters as transconductors, resistive loads, and active inductors, the CTLE core measures only 20 µm x 15 µm (13X smaller than a typical CML CTLE) and consumes 6 mW. The inverters are operated with an LDO-regulated ground supply, utilizing a ring oscillator based feedback loop to track and absorb PVT variations.
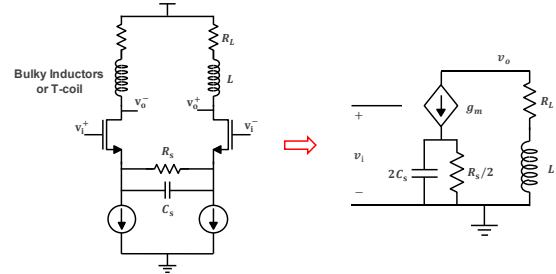

Fig. 2. Conventional CTLE schematic and small signal model.

II. INVERTERS AS ANALOG ELEMENTS

Conventional CML-style CTLEs do not take full advantage of technology scaling and are difficult to operate at low supplies. On the other hand, inverters scale directly with process and become power efficient transconductor (g$_m$) cells when biased appropriately [2]. In advanced processes, NMOS and PMOS devices have nearly the same mobility, and thus using PMOS doesn't incur extra speed penalty, leading to designs that are truly limited by the technology's peak f$_T$.

As shown in Fig. 3, inverters can function as various linear circuit elements in different configurations. An inverter with foot switches works as a switchable g$_m$ cell. A diode-connected inverter behaves as a self-biased 1/g$_m$ resistive load. When the NMOS and PMOS devices' drive strengths are identical, the nominal output voltage is half way between supply and ground. This natural bias point puts all devices in saturation, thus ensuring sufficient output swing and linearity. Active inductors [3] can be realized by adding two resistors between the inverter's output and the devices' respective gates. Same-valued resistors can be used due to symmetry of an inverter in FinFET technology. This technique leverages the gate capacitance of inverters for a high frequency boost that extends bandwidth, thus eliminating the need of bulky passive inductors. For our design, the inverters are directly
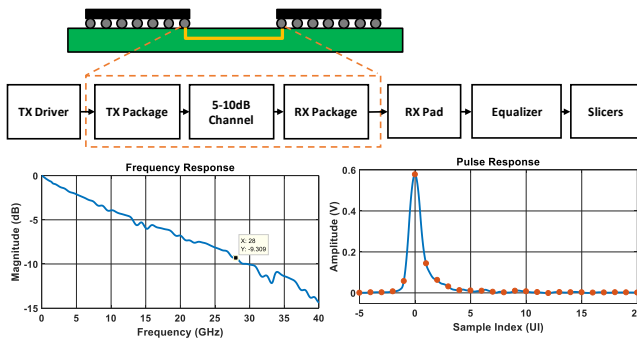

Fig. 1. Short-reach link application block diagram with channel responses.

connected to a 1.2 V supply and an on-chip LDO regulator supplies an elevated ground voltage, VSS_REG. Since all inverters operate around the half supply point, there is no issue with device stress (also during start-up) and thus ultra-low V_T, thin-oxide devices are employed to achieve the highest possible speed.
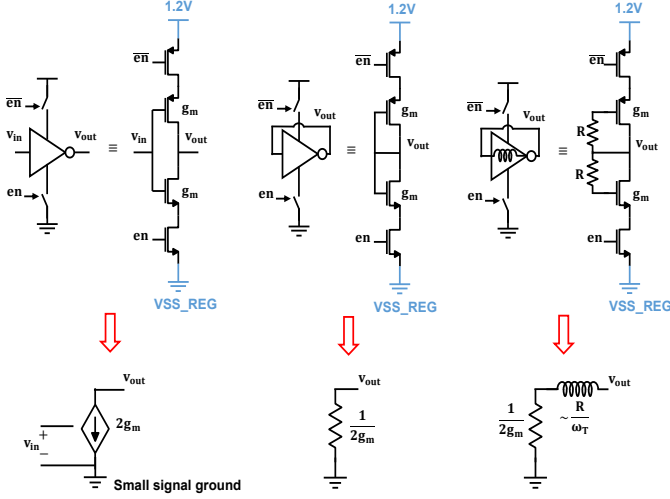

Fig. 3. Inverters in different configurations and their corresponding small-signal equivalent circuits.

III. INVERTER CTLE

Conventional CML-style CTLEs achieve the desired frequency response with source degeneration, but having a source node network in an inverter would affect its bias point. Therefore, additive two-path CTLEs [4] are a better option for inverter-based circuits. Fig. 4 shows the single-ended schematic and simulated frequency responses of our design. The low-frequency gain is determined by the bottom path inverter ratio, g_{m1}/g_{ml}. The high-frequency gain is approximately the ratio of total active and load transconductances, $(g_{m1}+g_{m2})/(2g_{ml})$. The coupling capacitor C is implemented with a fingered MOM device. Active inductors are used in both low- and high-frequency paths for
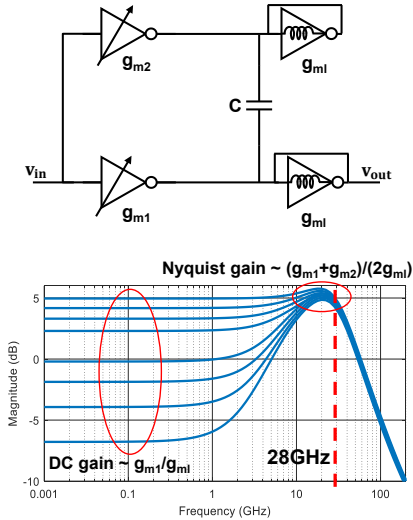

Fig. 4. Single-ended CTLE schematic and simulated frequency responses.

bandwidth extension. Both g_{m1} and g_{m2} are tunable and their sum is kept constant, which results in de-emphasis in the CTLE's transfer function and provides peaking in a power efficient manner. The inverter ratios are also tuned to compensate for gain reduction due to finite output resistance of the inverters. The transfer function is given by

$$\frac{v_o}{v_i} = -\frac{g_{m1}}{g_{ml}} \frac{1 + s\frac{g_{m1}+g_{m2}}{g_{m1}}\frac{C}{g_{ml}}}{1 + s\frac{2C}{g_{ml}}} \cdot P(s) \qquad (1)$$

Here, P(s) contains the bandwidth-extending zero and pole from active inductors, as well as parasitic poles determined by load g_m, drain parasitics (C_{dd}), the subsequent slicers' input gate capacitance (C_{gg}), and any wiring capacitance. P(s) is an important term since it determines the peaking strength and bandwidth. Detailed analysis shows that bandwidth and peaking (approximately) scales with the ratio $g_m/(C_{gg}+C_{dd})$, which corresponds to the inverter small-signal unity gain frequency ω_u. The biasing approach presented in Section IV hence attempts to stabilize this particular ratio.

The CTLE layout (see Fig. 5) resembles a standard cell style and dummies are added to enable source/drain diffusion sharing (to minimize area and reduce systematic mismatches and other layout effects). The layout benefits from the digital-like density scaling, and reduces parasitic effects due to wiring when compared to conventional CML-style CTLEs.
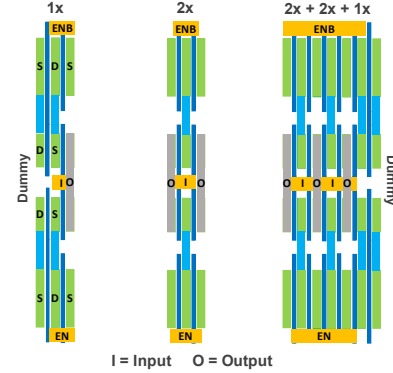

Fig. 5. Example layout diagram.

IV. INVERTER BIASING

Though inverter-based circuits can have relatively stable voltage gain due to their ratiometric nature, their frequency response, including parasitic pole location, is determined by the absolute values of the transconductances and capacitances in the circuit, which are heavily dependent on PVT conditions.

To address this issue, this work employs a replica biasing technique using a ring oscillator, which is widely used for process monitoring. Traditional constant-g_m biasing circuits consume static current and typically, device sizes and power must be scaled up to reduce the impact of random mismatch. On the other hand, a ring oscillator consumes only dynamic power and this power is nearly independent of the number of stages. Thus, it is possible to use a large number of stages to minimize random variations in oscillation frequency.

For a ring oscillator with inverters of equal PMOS and NMOS driving strengths, the oscillation frequency is

$$f_{osc} = \frac{1}{2N \cdot t_p} \qquad (2)$$

where N is the number of stages and t_p is the average inverter propagation delay. Due to symmetry in the inverters, there is no significant difference between rising and falling edge delays. Assuming that the gate delay is dominated by slewing and that the transistors obey the square-law (for simplicity), we can express the ring oscillator inverter delay as

$$t_p = \frac{V_{DD}}{2} \frac{C_{gg}+C_{dd}}{\frac{W}{2L}\mu C_{ox}(V_{DD}-V_{TH})^2} \qquad (3)$$

For the inverters in the analog signal path, the gate bias voltages are at $V_{DD}/2$ and thus their transconductance is

$$g_m = \frac{W}{L}\mu C_{ox}\left(\frac{V_{DD}}{2}-V_T\right) \qquad (4)$$

We note that V_{DD} directly affects the analog inverter's g_m, and thus also the ratio $\omega_u = g_m/(C_{gg}+C_{dd})$, Now, expressing (2) with (3) and (4), we obtain the following expression for oscillation frequency in terms of g_m.

$$f_{osc} = \frac{1}{2N} \frac{(V_{DD}-V_T)^2}{V_{DD}\left(\frac{V_{DD}}{2}-V_T\right)} \frac{g_m}{C_{gg}+C_{dd}} = \frac{\pi}{N}\alpha f_u \qquad (5)$$

In this expression, α is a function of V_{DD}. However, when $V_{DD} \gg V_T$, α approaches 2 and f_{osc} becomes directly proportional to f_u. In our design, this condition is met since V_T is about 150 mV and the nominal V_{DD} is 700 mV. Fig. 6(a) confirms this by plotting the simulated α vs. V_{DD}, demonstrating small variations (within ±5%) in the relevant V_{DD} range. As a result, tuning V_{DD} for constant f_{osc} can be exploited to stabilize f_u across corners. Fig. 6(b) illustrates this by plotting the simulated inverter f_u for different process and temperature corners. For fixed V_{DD}, we see large f_u variations across corners. However, when V_{DD} is computed by an adaptive loop that maintains constant f_{osc}, the f_u variations become small. This translates to a relatively fixed parasitic pole frequency of P(s) in (1), thus maintaining high bandwidth for the CTLE. Note that this tuning mechanism does not perfectly stabilize the (less critical) g_m/C terms in (1), but it still helps in counteracting some g_m variations.
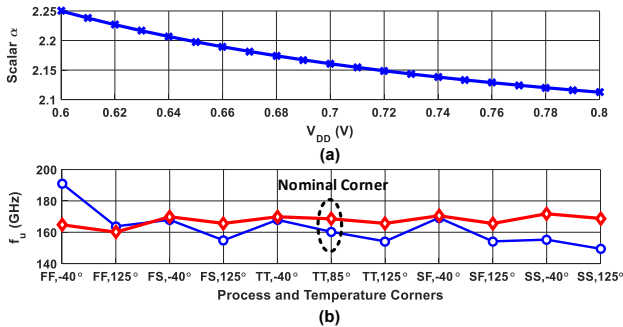

Fig. 6. (a) Simulation of parameter α vs. V_{DD}. (b) Simulated inverter f_u at different process and temperature corners.

Fig. 7 shows the overall implementation of our system along with the adaptive supply loop. The CTLE core has pseudo-differential paths, which provides some supply and common mode noise rejection. Its nominal output common mode is set to 0.85 V to interface with the succeeding NMOS input slicers. The total capacitive load for the CTLE is

approximately 30 fF (including the input capacitance of five slicers and wiring parasitics). A replica diode-connected inverter is used as the common mode reference for input termination. The ring oscillator's output is used as the clock for a finite state machine (FSM) that controls the LDO's output voltage as the ground for the core and reference circuits. The FSM uses an available 100 MHz clock reference to tune the LDO voltage such that the ring oscillator clock achieves a programmable frequency target (nominally 740 MHz, corresponding to 5 ps inverter delay) with hysteresis. The frequency target is set externally for optimal CTLE bandwidth, power and performance. The CTLE's replica-bias block (including reference diode, ring oscillator, LDO and FSM) can be shared by multiple transceiver channels, thus amortizing the area and power cost of the biasing circuits.
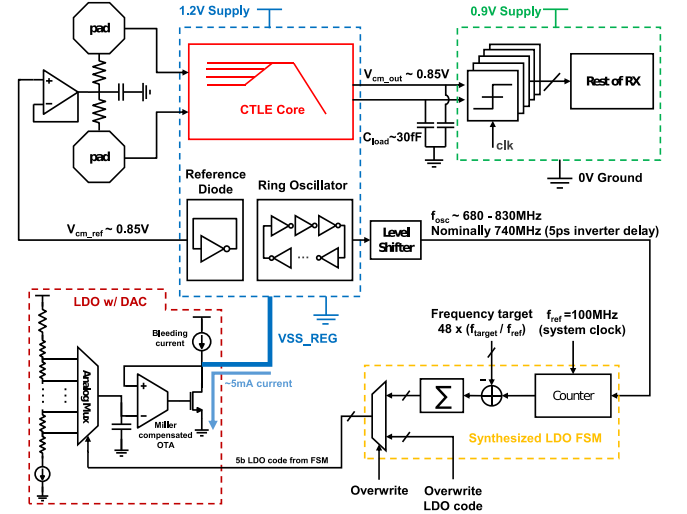

Fig. 7. System block diagram.

V. TEST AND MEASUREMENT

The CTLE is tested as the only means of equalization in a complete transceiver operating at 56 Gb/s with PAM2 modulation. A conventional CML-based CTLE is also fabricated on the same die to compare performance.

As shown in Fig. 8, the inverter-based CTLE has similar performance as CML-based CTLE when tested with the same channel, achieving 31% UI horizontal opening at BER<10^{-12} for a channel with 8 dB loss at 28 GHz. Measured bathtub curves for different LDO modes and various temperatures are shown in Fig. 9, demonstrating the effectiveness of the
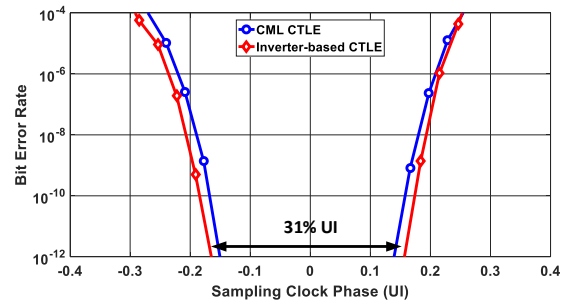

Fig. 8. Bathtub curve comparison between CML-based and inverter-based CTLEs under nominal conditions

employed biasing scheme. When LDO is in adaptive mode, a larger eye width is achieved at higher temperatures as shown in Fig. 10(a). To further validate the function of the replica-bias scheme, the regulated ground voltage is plotted against varying temperature for different LDO modes in Fig. 10(b). In overwrite mode, a fixed LDO code is applied and the ground voltage increases due to its bias resistor's temperature coefficient. In auto mode, the LDO code is adapted and the ground voltage decreases as expected to maintain oscillation frequency at higher temperature.

As indicated in Table I, our CTLE core consumes 6 mW (at room temperature), and measures only 20 µm x 15 µm (see Fig. 11), which is 13X smaller than the CML-based CTLE core on the same test chip. The inverter-based CTLE achieves similar performance with no extra power. Compared to our previous work in [6], the inverter-based CTLE shows significant improvements in both power and area.
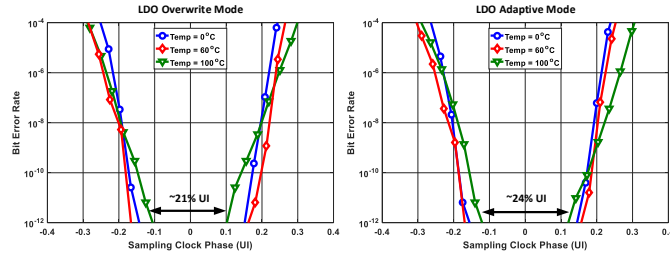

Fig. 9. Bathtub curve comparison between overwrite and auto LDO modes for different temperatures.
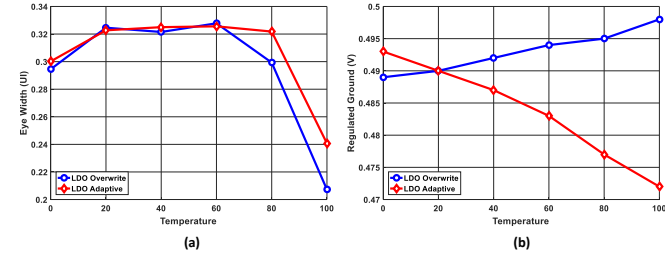

Fig. 10. (a) eye widths vs. temperature and (b) regulated ground voltage vs. temperature in different LDO modes.
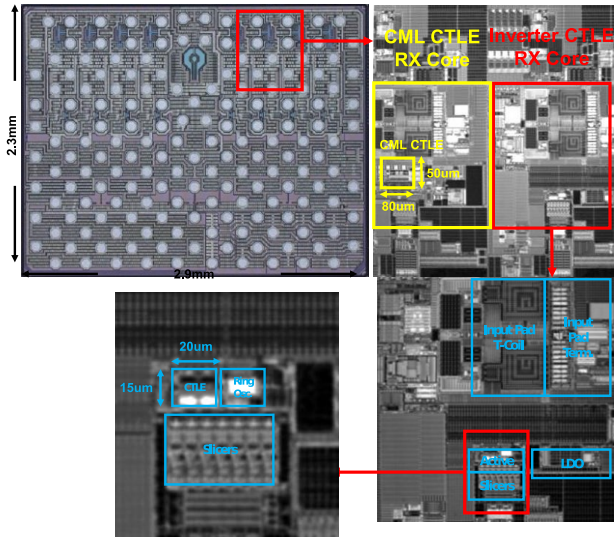

Fig. 11. Chip photos.

TABLE I.   COMPARISON TABLE

Reference	[5]	[6]	This Work	This Work
CTLE Type	2-stage CML	CML	CML	Inverter
Modulation	PAM2	PAM4	PAM2	PAM2
Nyquist Frequency	6.25 GHz	14 GHz	28 GHz	28 GHz
Process	32 nm SOI	16 nm FinFET	16 nm FinFET	16 nm FinFET
Supply	1.1 V	1.2 V	1.2 V	1.2 V + Ground LDO
Max Peaking (DC/Nyq. gain)	-6 dB/12 dB	0 dB/7 dB	-6 dB/6 dB	-6 dB/6 dB
Channel Loss	27 dB[a]	31 dB[a]	8 dB	8 dB
Timing Margin	50% @BER<10^{-12}	n/a	>24% @BER<10^{-12} (100 ºC)	>24% @BER<10^{-12} (100 ºC)
Core Power	5.25 mW[b]	8.4 mW	6 mW	6 mW
Power/Freq.	0.84 mW/GHz	0.6 mW/GHz	0.21 mW/GHz	0.21 mW/GHz
Core Area	Not reported	125 µm x 40 µm	80 µm x 50 µm	20 µm x 15 µm

[a]. Other means of equalizations are also used
[b]. Estimated from power breakdown chart

VI. CONCLUSION

We demonstrated an inverter-based CTLE with constant f_u biasing scheme for short-reach, high-speed links. Both conventional CML and inverter-based CTLE prototypes were fabricated on the same chip within 16 nm FinFET CMOS transceivers. The inverter CTLE is significantly smaller than prior-art circuits, while maintaining competitive power efficiency. The robustness of the biasing approach was proven through measured temperature sweeps.

VII. ACKNOWLEDGEMENT

We would like to thank Xilinx Inc. for their layout and testing support for this project.

REFERENCES

[1] R. Boesch, K. Zheng and B. Murmann, "A 0.003 mm^2 5.2 mW/tap 20 GBd inductor-less 5-tap analog RX-FFE," 2016 IEEE Symposium on VLSI Circuits, Honolulu, HI, 2016, pp. 170-171.

[2] B. Nauta, "A CMOS transconductance-C filter technique for very high frequencies," in IEEE J. Solid-State Circuits, vol. 27, no. 2, pp. 142-153, Feb. 1992.

[3] T. Musah et al., "A 4–32 Gb/s Bidirectional Link With 3-Tap FFE/6-Tap DFE and Collaborative CDR in 22 nm CMOS," in IEEE J. Solid-State Circuits, vol. 49, no. 12, pp. 3079-3090, Dec. 2014.

[4] J. F. Bulzacchelli et al., "A 28-Gb/s 4-Tap FFE/15-Tap DFE Serial Link Transceiver in 32-nm SOI CMOS Technology," in IEEE J. Solid-State Circuits, vol. 47, no. 12, pp. 3232-3248, Dec. 2012.

[5] T. Toifl et al., "A 2.6 mW/Gbps 12.5 Gbps RX With 8-Tap Switched-Capacitor DFE in 32 nm CMOS," in IEEE J. Solid-State Circuits, vol. 47, no. 4, pp. 897-910, Apr. 2012.

[6] Y. Frans et al., "A 56-Gb/s PAM4 Wireline Transceiver Using a 32-Way Time-Interleaved SAR ADC in 16-nm FinFET," in IEEE J. Solid-State Circuits, vol. 52, no. 4, pp. 1101-1110, Apr. 2017.