An Inverter-Based Analog Front-End for a 56-Gb/s PAM-4 Wireline Transceiver in 16-nm CMOS

Kevin Zheng, Yohan Frans, Sai Lalith Ambatipudi, Santiago Asuncion, Hari Teja Reddy, Ken Chang, *Fellow, IEEE*, and Boris Murmann, *Fellow, IEEE*

***Abstract*—This letter describes an inverter-based analog receiver front-end that was evaluated within a complete ADC-based 56-Gb/s four pulse-amplitude modulation wireline transceiver system. The front-end contains hybrid continuous-time linear equalizers with low- and high-frequency peaking as well as inverter-based programmable gain amplifiers that serve as buffers for the 32× time-interleaved SAR ADC. The inductorless design achieves high bandwidth through a reduced number of equalizer stages (with fewer parasitic poles) and active peaking techniques. The transceiver that contains the analog front-end (AFE) achieves $< 10^{-6}$ BER over a backplane channel with 35-dB loss at 14 GHz and with 2-mV_{rms} input crosstalk noise. The AFE core occupies 0.00425 mm^2 in 16-nm fin field effect transistor CMOS and consumes 165 mW.**

***Index Terms*—Continuous-time linear equalizer (CTLE), fin field effect transistors (FinFETs), four-level pulse-amplitude modulation (PAM-4), transceiver, wireline.**

I. Introduction

Increasing data rates in wireline links have led to solutions that employ multilevel modulation (e.g., PAM-4) and ADC-based receivers offering data rates of 56 Gb/s and beyond [1]–[3]. In such systems, a high-performance analog front-end (AFE) that provides wideband gain and pre-equalization is essential for relaxing the ADC requirements and meeting the system requirements at low power dissipation. Current-mode-logic (CML) style circuitry is traditionally used in such AFEs but is beginning to fall out of favor due to decreasing supply voltages and the need for area-hungry inductors or *T*-coils for bandwidth extension. Our proof-of-concept work in [4] and [5] has illustrated the advantages of inverter-based circuits [6] that extract higher speeds for a given technology and can reject process variations through ratiometric design and replica-based tuning. In the present design, we advance these works toward a PAM-4 system with more challenging specifications and validate the performance within a complete 56-Gb/s ADC-based transceiver testbed.

The transceiver architecture is shown in Fig. 1. Utilizing the same transmitter, configurable SAR ADC bank and DSP blocks for equalization (FFE + DFE), adaptation, CDR, and calibration as in [1], our design replaces the CML-style AFE with inverter-based circuits. The AFE block contains inverter-based hybrid continuous-time linear equalizers (CTLEs), programmable gain amplifiers (PGAs), as well the ADC driver stages, which are all implemented with CMOS inverters.

This letter extends our original conference publication [7] and is organized as follows. Section II discusses the pros and cons of the employed subtractive CTLE architecture. Section III describes the complete AFE circuit architecture, while Section IV provides measured results and assesses the design's robustness in presence of supply and temperature variations.

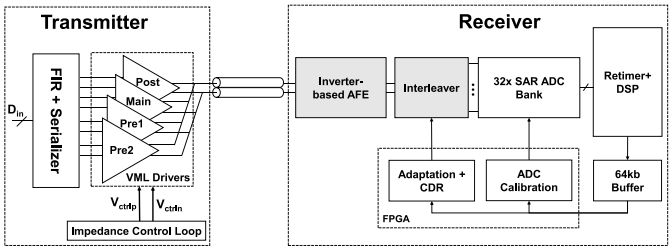

Fig. 1. Transceiver architecture.

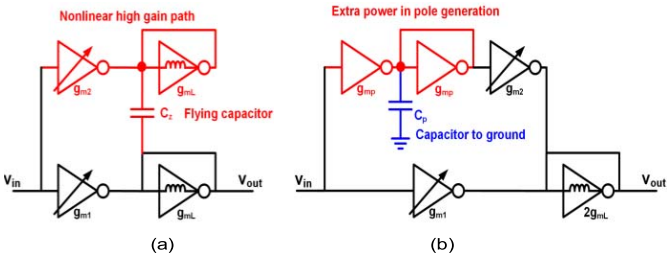

Fig. 2. Comparison of (a) additive and (b) subtractive CTLE circuit.

II. Inverter-Based CTLE Circuit Topology

Our previous work investigated an inverter-based CTLE in a PAM-2 link [4]. However, this topology has limitations when used in a PAM-4 system due to the increased linearity and tunability requirements. In this section, we compare the previously designed additive CTLE with the subtractive topology employed here.

Fig. 2 shows the half circuit schematic of the two CTLE circuits (the actual implementation uses two half circuits for pseudo-differential operation). The inverters are biased near mid-supply to function as analog transconductance cells [6]. To assess the merits of the two approaches, we assume that they are designed for unity voltage gain at low-frequencies and a high-frequency (HF) gain larger than one. This assumption is consistent with the typical requirements in a PAM-4 system, where the high-loss channel necessitates an HF gain boost to restore the attenuated main cursor.

The additive CTLE in Fig. 2(a) achieves peaking through an additive feedforward path with an HF zero. Conversely, in Fig. 2(b), a subtractive path reduces the gain at low frequencies and results in peaking due to an HF pole (g_{mp}/C_p) that blocks the subtraction. For a fair comparison between the two options, the HF output impedance, which sets the HF drive strength, is assumed to be $1/2g_{mL}$ in both circuits. Table I enumerates the required transconductances (normalized to g_{mL}) for flat gains of 0- and 6-dB HF boost. We see that the subtractive CTLE requires extra transconductance and hence additional power to achieve the same gain boosting. This disadvantage is intuitive since the circuit blocks some of the transconductors at high frequencies.

Manuscript received December 1, 2018; accepted January 6, 2019. Date of publication January 23, 2019; date of current version May 17, 2019. This paper was approved by Associate Editor Alvin Leng Sun Loke. *(Corresponding author: Kevin Zheng.)*

K. Zheng was with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. He is now with Xilinx Inc., San Jose, CA 95124 USA (e-mail: kevinz@xilinx.com).

Y. Frans, S. L. Ambatipudi, S. Asuncion, H. T. Reddy, and K. Chang are with Xilinx Inc., San Jose, CA 95124 USA.

B. Murmann is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA.

Digital Object Identifier 10.1109/LSSC.2019.2894933

TABLE I
REQUIRED g_m VALUES (NORMALIZED) FOR BOTH CTLE OPTIONS

CTLE	Additive		Subtractive	
HF gain	0 dB	6 dB	0 dB	6 dB
g_{m1}	1	1	2	4
g_{m2}	1	3	0	2
Σg_{mL}	2	2	2	2
Σg_{mp}	0	0	x	x
Total	4	6	4+ x	8+ x

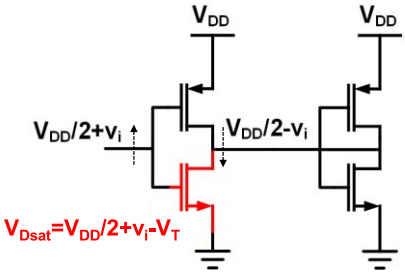

Fig. 3. Inverter-based unity-gain stage for large-signal analysis.

Despite its power disadvantage, we employ the subtractive topology due to signal swing considerations. From Table I, we see that the magnitude of the low-frequency voltage gain in the feedforward path is $g_{m2}/g_{mL} = 3$. On the other hand, the low-frequency feedforward path gain of the subtractive CTLE is $g_{m1}/(2g_{mL}) = 2$. The high gain and resulting large swing at the intermediate node make the additive circuit unattractive for PAM-4 systems, as the signal must be scaled down significantly to prevent clipping and to achieve the required linearity.

Another advantage of the subtractive CTLE circuit pertains to its tuning capacitance. Programming the CTLE's zero location (required by the application) is simpler with the capacitance C_p to ground in Fig. 2(b), as opposed to the flying capacitor C_z in Fig. 2(a). C_p can be a MOS capacitor with high density and its parasitics are unimportant.

Given that signal swing and linearity are the main considerations for choosing the subtractive topology, it is worth reviewing the factors that determine the circuit's nonlinearity. For this discussion, we consider the canonical subcircuit shown in Fig. 3, which is assumed to be biased at mid-supply and sized for unity gain. Nonlinearity in this circuit arises mainly due to three factors: 1) transconductance nonlinearity, 2) MOSFET output conductance nonlinearity, and 3) the onset of signal clipping as the devices' V_{DS} approach V_{Dsat}. Issue (1) is analyzed in detail in [6] and [8]. Provided that the P and N devices have similar I–V characteristics, odd-order nonlinearities are significantly suppressed (and even-order terms are mitigated by the pseudo-differential circuit). Given the high mobility of P devices in FinFET technology, we found that this condition is easily met over PVT so that nonlinearity is not the dominant factor in meeting the PAM-4 requirements (THD $\cong -40$ dB). From a system study that models the nonlinearity errors from a broadband input signal (i.e., PRBS pattern filtered by channel) as an equivalent noise source [8], this design maintains an equivalent system SNR > 17 dB due to nonlinearity, which is sufficient for the application.

Regarding the second issue, we found through simulations that the improved channel control of the FinFETs (smaller drain-induced barrier lowering, DIBL) led to negligible output conductance nonlinearity as long as the devices are saturated. The remaining and most significant issue is therefore to ensure saturation (with some margin) in presence of large signal swings that maximize the front-end's SNR. As illustrated through the arrows in Fig. 3, a key point here is that the device's V_{GS} and V_{DS} are anti-phase, leading to an opposite dependence of V_{Dsat} and V_{DS} on the incremental input signal v_i, i.e., $V_{Dsat} = V_{DD}/2 + v_i - V_T$ and $V_{DS} = V_{DD}/2 - v_i$. Solving for the maximum amplitude for which the device remains saturated, this leads to $v_i < V_T/2$. From this result, we note that the devices' threshold voltage plays the primary role is setting the circuit's linear range. For a standard threshold voltage device in our technology, $V_T \approx 0.4$ V, leading to a maximum peak-to-peak swing of 0.8 V in the pseudo-differential circuit. In reality, one needs to back off from this limit by some margin, as the transition between the saturation and triode regimes is gradual. Our implementation was designed for a differential receiver input swing of ± 300 mV.

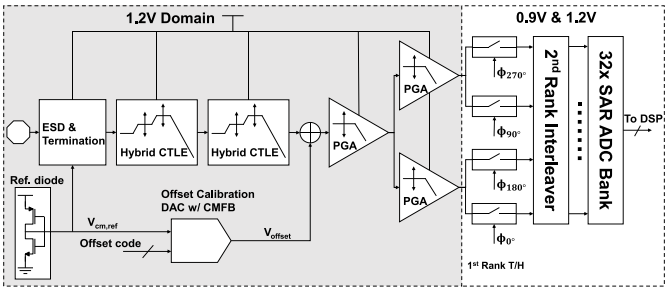

Fig. 4. Receiver AFE block diagram.

Once the device types with the desired V_T are chosen, this also fixes the inverter's supply voltage. For example, assuming threshold voltage magnitudes of 0.4 V and a desired gate overdrive ($V_{GS} - V_T$) of 0.2 V ($g_m/I_D \cong 10$ S/A), the supply voltage should be set to 1.2 V. On the other hand, in a PAM-2 link where linearity and swing requirements are not as stringent, only a 0.7-V supply is needed if ultralow threshold devices with $V_T = 0.15$ V are used.

III. RECEIVER AFE CIRCUIT ARCHITECTURE

Fig. 4 shows the block diagram of the receiver AFE. It contains a cascade of two CTLE stages that are based on the subtractive topology discussed in the previous section. However, in addition to the discussed HF peaking, we employ an additional path for low-frequency peaking as shown in Fig. 5. The resulting transfer function for each CTLE block is therefore given by

$$-\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{g_{m,ap}}{g_{ml}} - \frac{g_{m,hf}}{g_{ml}}\frac{g_{mp1}}{g_{mp}}\frac{1}{1+s\frac{C_{hf}}{g_{mp}}} - \frac{g_{m,lf}}{g_{ml}}\frac{g_{mp2}}{g_{mp}}\frac{1}{1+s\frac{C_{lf}}{g_{mp}}}. \tag{1}$$

We found that the compact two-stage implementation has fewer parasitic poles and can achieve a larger gain-bandwidth product. In the circuit of Fig. 5(a), the low-frequency and HF paths are current-summed with the all-pass path using transconductors ($g_{m,ap}$, $g_{m,hf}$, and $g_{m,lf}$). The CTLE's output impedance stays relatively constant over frequency and is set by the active load (g_{ml}). The resistors added to the gates of the load inverter provide inductive peaking for bandwidth enhancement. The low-frequency and HF poles in the feedforward paths are defined by g_{mp} and their respective programmable MOS capacitor banks. C_{lf} is implemented with five C_{hf} banks connected in parallel. The equalization strengths are tuned by changing the current summing transconductor values. An example implementation of a 4-bit tunable g_m cell is shown in Fig. 5(b). Inverter unit cells are turned on/off with two switches on the P/N sides. Thermometer coding for the 2 MSBs are used to ensure monotonicity when random mismatch is present. The same amount of transconductance (α, β) is added to or subtracted from the all-pass path when LF and HF paths change to minimize the dc gain drop in the frequency responses. However, a small amount of de-emphasis is still desired to scale the peaking. By design, the peaking steps are not linear in dB, but sufficiently small (< 0.3 dB/step) to guarantee adaptation convergence. The simulated frequency response of a single CLTE block is shown in Fig. 6.

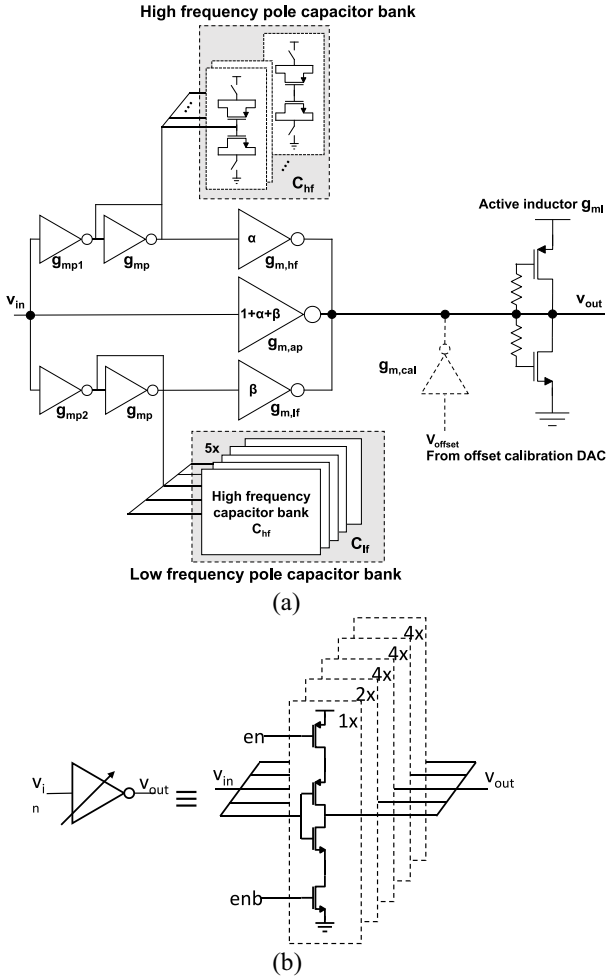

Fig. 5. (a) Single-ended schematic of hybrid CTLE. (b) Example implementation of a 4-bit tunable g_m cell.

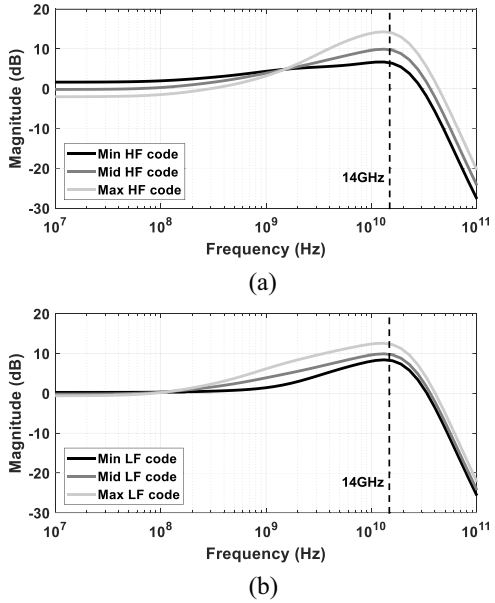

Fig. 6. Simulated CTLE frequency response (post-layout). (a) Mid LF code and different HF codes. (b) Mid HF code and different LF codes.

As indicated in Fig. 4, we perform an offset correction at the output of the second CTLE stage. The subtraction is achieved via a transconductor ($g_{m,cal}$) at the CTLE's current summing node (see Fig. 5).

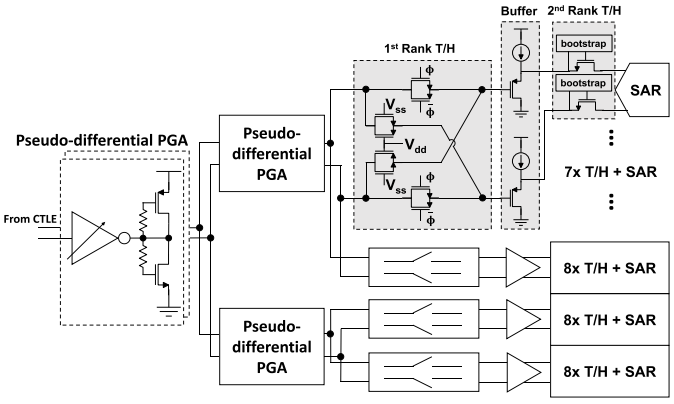

Fig. 7. Schematics of the PGA and ADC interface circuits.

The offset correction voltage (V_{offset}) is generated from a resistor ladder DAC with common-mode (CM) feedback. The received signal is ac-coupled on-board, and a CM reference is generated by a replica diode-connected inverter. Although a pseudo-differential implementation provides first-order cancelation for small CM noise, larger CM variations will exercise mismatches between the two half circuits and convert them to differential-mode nonlinearity. However, we did not find this to be a significant effect in our experiments. The reference diode inverter is sized to have a random CM voltage RMS variation smaller than the RMS offset voltage from the signal chain. The CTLEs are followed by PGAs to optimally clip the ADC. Similar to the CTLE stages, two PGA stages are used to maximize the gain-bandwidth product. The second-stage PGAs also act as the input track-and-hold (T/H) buffers in the ADC. As shown in Fig. 7, the PGAs are realized using programmable inverter transconductors loaded by active inductors, which help boost the bandwidth to 25 GHz. The input T/H switches are CMOS switches with cross-coupled dummies to mitigate capacitive hold-mode feedthrough. Explicit hold capacitors are eliminated to reduce the driver load (the parasitics satisfy the *kT/C* noise requirements). The CMOS switches are intentionally skewed in size so that their charge injection and clock feedthrough provide the necessary drop in output CM voltage to optimally bias the subsequent source followers. The second rank T/H uses bootstrapped switches and the SAR ADC slices are the same design as reported in [1].

The AFE's inverter-based circuits are operated from a 1.2-V supply per the discussion in Section II. With a quiescent point V_{GS} of 0.6 V and peak amplitudes of about 0.2 V, this does not cause any gate oxide reliability concerns for 16-nm FinFet devices. However, an important aspect to ensure is that the supply turn-on dynamics do not lead to large voltages across the oxide. We managed this issue using extensive simulations and by ensuring reasonable capacitive balance for critical nodes between the rails with symmetry in layout.

IV. Measurement Results

The transceiver was tested with very short reach (VSR) and long reach (LR) channels (10 and 35 dB loss, respectively, at 14 GHz). Fig. 8(a) shows the 7b ADC output scans for both channels and eye diagrams after DSP equalization. No transmit-side equalization is used for the VSR channel, which demonstrates sufficient equalization capability and linearity performance of the receiver AFE. The bathtub curves for the LR channel in Fig. 8(b) shows $< 10^{-12}$ BER without crosstalk and $< 10^{-6}$ BER with 2-mV$_{\text{rms}}$ crosstalk.

Different from our previous PAM-2 implementation in [4], this AFE does not have an adaptive supply biasing loop, since the ADC-based architecture can absorb such variations through receiver adaptation. To validate this point, we performed a study on the receiver's sensitivity to supply and temperature changes.

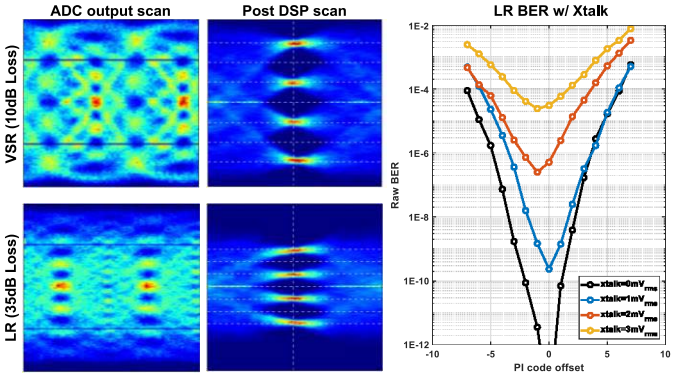

Fig. 8. (a) ADC output scans and post-DSP equalization eye scans for VSR and LR channels. (b) Bathtub curves for LR channel at different crosstalk levels.

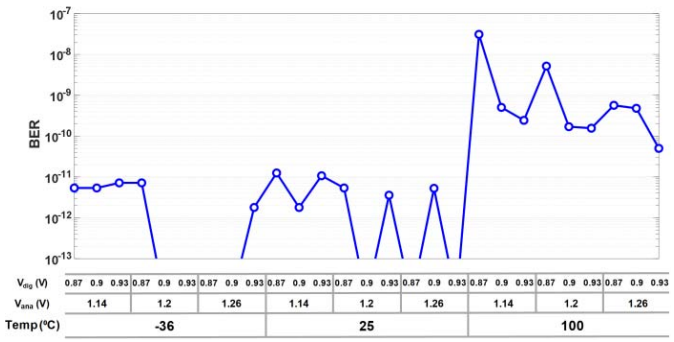

Fig. 9. Transceiver BER performance versus supply and temperature for the LR channel.

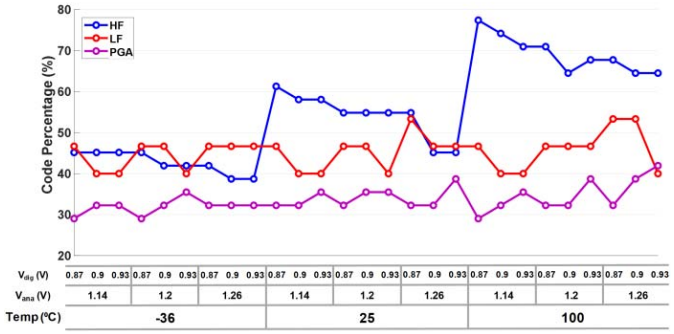

Fig. 10. Equalizer and PGA code percentages versus supply and temperature for the LR channel.

Fig. 9 shows the measured BER for different voltage and temperature corners when the transmit equalizer is fixed to a 3-dB post-cursor boost. It is expected that the BER degrades for very high temperature, since the circuit bandwidth reduces and the thermal noise increases. However, we observe that the system still achieves BER $< 10^{-7}$, which meets the standard requirements with a reasonable margin.

Fig. 10 shows how the AFE coefficients adapt to the different operating conditions. The HF CTLE code increases significantly with rising temperature, showing that the system is combating the extra HF loss at high temperature. Meanwhile, both the LF and PGA codes stay relatively constant across voltage and temperature corners. This is due to the ratiometric nature of inverter-based circuits, specifically the absolute gain is relatively insensitive to the circuit's operating conditions. For the channels under test, the used PGA codes are consistently lower than 50% of the total available codes. This means that the necessary gain range could be reduced for future iterations, which helps to cut down circuit power and complexity.

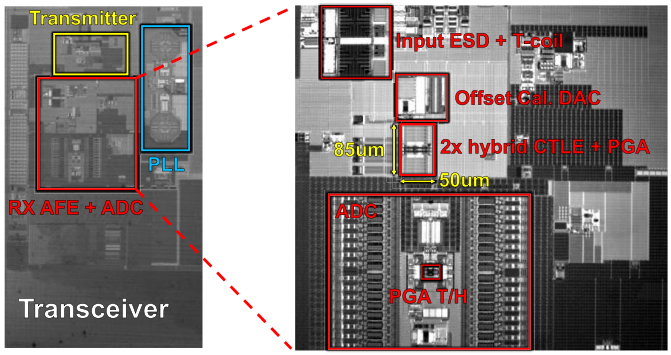

Fig. 11. Die photo.

TABLE II
PERFORMANCE COMPARISON

Reference	[1]	This Work
Technology	16 nm FinFET	16 nm FinFET
Power Supplies	0.9, 1.2, 1.8 V	0.9, 1.2, 1.8 V
CTLE core area	120 µm x 190 µm	50µm x 85µm
CTLE+ADC power	40 mW + 146 mW	34 mW + 131 mW*
Channel loss @14GHz	32 dB	35 dB**
Data Rate	56 Gb/s	56 Gb/s
BER (2mV$_{rms}$ crosstalk)	$<10^{-6}$	$<10^{-6}$

*Power is code dependent. Settings for LR channel are used.
**Different test setup induced additional channel loss.

This letter is compared to [1] and the AFE specifications and transceiver performance metrics are summarized in Table II. The inverter-based CTLEs and PGAs occupy only 50 μm $\times$ 85 μm (see Fig. 11), and the AFE consumes 165 mW total. The CTLE power is reduced because of fewer stages, and the ADC power decreases due to the inverter-based PGA and T/H buffers (whereas [1] is based on CML circuits). Overall, this letter demonstrates that inverter-based AFE circuits are a promising alternative to CML for high-speed links in FinFet CMOS technology.

REFERENCES

[1] P. Upadhyaya *et al.*, "A fully adaptive 19-to-56Gb/s PAM-4 wireline transceiver with a configurable ADC in 16nm FinFET," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2018, pp. 108–110.
[2] Y. Frans *et al.*, "A 56Gb/s PAM4 wireline transceiver using a 32-way time-interleaved SAR ADC in 16nm FinFET," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2016, pp. 48–49.
[3] J. Cao *et al.*, "29.2 A transmitter and receiver for 100Gb/s coherent networks with integrated 4×64GS/s 8b ADCs and DACs in 20nm CMOS," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2017, pp. 484–485.
[4] K. Zheng, Y. Frans, K. Chang, and B. Murmann, "A 56 Gb/s 6 mW 300 um^2 inverter-based CTLE for short-reach PAM2 applications in 16 nm CMOS," in *Proc. IEEE CICC*, Apr. 2018, pp. 1–4.
[5] 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," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2016, pp. 170–171.
[6] B. Nauta, "A CMOS transconductance-C filter technique for very high frequencies," *IEEE J. Solid-State Circuits*, vol. 27, no. 2, pp. 142–153, Feb. 1992.
[7] K. Zheng *et al.*, "An inverter-based analog front end for a 56 Gb/s PAM4 wireline transceiver in 16nm CMOS," in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, Jun. 2018, pp. 269–270.
[8] K. Zheng, "System-driven circuit design for ADC-based wireline data links," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, 2018.