

Fifty Applications of the CMOS Inverter—Part 3

THE ANALOG MIND

In this multipart article, we study applications of the CMOS inverter in today's electronic systems. We quantify the performance of each circuit by simulations in the slow–slow corner of 28-nm technology with $V_{DD} = 0.95$ V at T = 75 °C. The reader is cautioned that inverters suffer from poor supply rejection, requiring onchip voltage regulators.

The Voltage Multiplier

Also called the *"charge pump,"* the voltage multiplier converts a constant voltage to a greater value. For example, energy harvesting systems receive a radio wave and use its energy to charge a capacitor. The voltage thus generated, however, may be excessively small to supply circuit blocks, an issue resolved by a voltage multiplier.

Consider the structure shown in Figure 1(a), where C_X switches between V_{DD} and C_Y under the command of a clock. We observe that V_Y must eventually approach V_{DD} so that C_X and C_Y no longer share charge. We can

Digital Object Identifier 10.1109/MSSC.2024.3498732 Date of current version: 22 January 2025 approximate the settling time of the circuit by viewing C_X , S_1 , and S_2 as equivalent to a resistance of

$$R_{\rm eq} = \frac{1}{f_{CK} C_X} \tag{1}$$

where f_{CK} denotes the clock frequency [Figure 1(b)]. Thus, C_Y charges with a time constant equal to $R_{eq}C_Y = C_Y/(f_{CK}C_X)$.

Let us now turn to the topology of Figure 1(c), where switch S_2 turns on when the bottom plate of C_X jumps from zero to V_{DD} . If C_Y begins with an initial condition of V_0 , we have

$$V_Y = \frac{C_Y}{C_X + C_Y} V_0 + \frac{C_X}{C_X + C_Y} V_{DD}.$$
 (2)

Next, we turn off S_2 , discharge C_X [Figure 1(d)], and repeat the cycle. Equation (2) reveals that each repetition multiplies the previous voltage on C_Y by $C_Y/(C_X + C_Y)$ and adds to it another component equal to $C_X V_{DD}/(C_X + C_Y)$. It can be shown that V_Y approaches V_{DD} . Note that the charge necessary for C_Y is supplied by *CK*.

We now combine the mechanisms presented in Figure 1(a) and (c) so as

to obtain a voltage equal to $2V_{DD}$ on C_X . Depicted in Figure 2(a), the resulting structure operates as follows. First, S_1 turns on, charging C_X to V_{DD} while its bottom plate swings from V_{DD} to zero. Next, S_1 turns off, S_2 turns on, and the bottom plate of C_X travels from zero to V_{DD} . Thus, C_Y receives two charge packets, one due to the initial condition of V_{DD} on C_X , and another due to the rising edge of CK at the bottom plate of C_X .

Given that S_1 and S_2 in Figure 2(a) are activated in opposite phases of the clock, we surmise that we can drive both by *CK* if they are implemented by different types of devices [Figure 2(b)]. In this topology, however, C_X can charge only up to $V_{DD} - V_{TH1}$ when the gate of M_1 is at V_{DD} . Moreover, the circuit resembles a half-wave rectifier in that the charge on C_Y is replenished once per clock cycle. That is, V_Y exhibits a high ripple if the voltage multiplier must deliver a current to a load.

We resolve both issues by adding a complementary path to the structure and driving the gates of



FIGURE 1: (a) A simple two-stage sampler, (b) its equivalent circuit, (c) another mechanism for charging C_Y to V_{DD}, and (d) the circuit of (c) in reset mode.



FIGURE 2: (a) The basic voltage doubler, (b) its actual implementation, and (c) use of two paths to reduce ripple.

the switches by a "boosted" clock. Illustrated in Figure 2(c) [1], the circuit charges C_Y twice per cycle while allowing the dc levels at *X* and *N* to approach V_{DD} .

Assuming $C_x = C_N$ in Figure 2(c), we select the capacitor values as follows. The time constant of the circuit, $\tau = C_Y / (2f_{CK}C_X)$, determines the "startup" time, and C_Y yields a certain ripple voltage for a given load current. For example, $f_{CK} = 100$ MHz, $C_X = 100$ fF, and $C_Y = 2$ pF yield a startup time of roughly $3\tau = 300$ ns. The transistor widths can be near minimum if their on-resistance does not increase the startup time significantly. We select $(W/L)_N = 250$ nm/30 nm and $(W/L)_P = 500$ nm/30 nm.

With the foregoing values and $V_{DD} = 0.25$ V, we simulate the circuit, obtaining the waveforms shown in Figure 3 for V_X and V_Y . We observe that they approach $2V_{DD}$. By cascading *n* such stages, the voltage can be multiplied by a factor of *n* [1].

The Floating-Inverter Amplifier

It is possible to avoid static bias currents in amplifiers through the use of "charge steering." Shown in Figure 4(a) is a differential pair in which the continuous-time tail current and load resistors are replaced with discrete-time networks [2]. When *CK* is low, C_T is tied to ground, and *X* and *Y* are precharged to V_{DD} . After *CK* rises, C_T begins to draw charge from M_1 and M_2 and C_X and C_Y . This continues until V_P reaches the input commonmode (CM) level minus the transistor threshold, V_{TH} . It can be shown that the small-signal voltage gain is approximately equal to $2C_{T/}$ C_X [2]. The steering of charge for a short time interval reduces the power consumption.

To obtain a greater gain, we must select a higher value for C_T , inevitably creating a low CM level at the output [Figure 4(b)], which may be ill-suited to the next stage. We can increase the gain by adding a PMOS pair to the structure, as illustrated in Figure 4(c). Here, the desired output CM level, V_{CM}, is established during reset. In the amplification mode, charge flows from C_{T2} through the transistors to C_{T1} . The key point is that, in response to a differential input, V_X and V_Y now change in opposite directions, thus maintaining a fairly constant CM level.

Let us observe that capacitors C_{T2} and C_{T1} have the same current in the amplification mode because C_X and C_Y carry only differential currents. That is, C_{T2} and C_{T1} appear in series and can be replaced with a single capacitor [Figure 4(d)]. The result is called the "floating-inverter amplifier (FIA)" [3].

It is interesting to note that the output CM level of the FIA cannot change during amplification. Constructing the CM equivalent circuit as in Figure 4(e) [3], we recognize that C_Y (and C_X) draw no CM current.

We simulate the FIA of Figure 4(d) with $C_T = 4$ pF, $C_X = C_Y = 0.2$ pF, $(W/L)_N = 5 \ \mu m/30$ nm and $(W/L)_P =$ 10 $\mu m/30$ nm for the inverters, and half of these values for the switches. In response to a 10-mV differential input the circuit provides the outputs shown in Figure 5, exhibiting a gain of 4 and settling time of about 1 ns. Clocked at 100 MHz, the circuit consumes 160 μ W.

The Stacked Amplifier

In low-power applications, transistors are often biased in the subthreshold region, exhibiting a small gate-source voltage. A self-biased inverter operating in such a regime thus consumes little voltage headroom and makes stacking possible.



FIGURE 3: Voltage multiplier waveforms.



FIGURE 4: (a) A charge-steering circuit, (b) its waveforms for a high gain, (c) the complementary version, (d) floating-inverter amplifier, (e) and its equivalent circuit. CM: common mode.



FIGURE 5: FIA waveforms.

Shown in Figure 6(a) is an example [4], where M_1 - M_4 are stacked. That is, $V_{GS1} + |V_{GS2}| + V_{GS3} + |V_{GS4}| = V_{DD}$. Capacitor C_5 establishes an ac ground at P, but it can be omitted in differential implementations [4]. The inverters amplify V_{in} by a certain factor, producing equal signals at X and Y. The resulting outputs are shorted to each other by C_3 and C_4 . While the overall voltage gain is still that of one inverter, the noise components of the two inverters add in power at the output, reducing the inputreferred noise voltage by a factor of $\sqrt{2}$. This is the principal advantage of this current-reuse method.

We simulate the stacked amplifier with the following values: $(W/L)_N =$ $4 \ \mu m/120 \ \text{nm}$, $(W/L)_P = 8 \ \mu m/120 \ \text{nm}$, and $C_1 = \cdots = C_4 = 10$ pF. To avoid loading the inverters and providing a low high-pass cutoff frequency, the feedback resistors must be tens of megaohms; they are realized by "pseudo resistors," i.e., diode-connected NMOS transistors that sustain little bias voltage.

Figure 6(b) presents the waveforms at *A* and *X*, and Figure 6(c) those at *B* and *Y*. The input is a 1-mV sinusoid at 10 MHz. We note a gatesource voltage of about 240 mV and a voltage gain of 22. The bias current is 20 μ A. Figure 6(c) plots the input-referred noise voltage, revealing a value of about 9 nV/ $\sqrt{\text{Hz}}$. It is possible to stack three inverters so as to further reduce the power [4].

The Wideband Amplifier

The relatively low impedance presented by a self-biased inverter at its input and output leads to useful properties in a cascade of such stages. Considering the topology shown in Figure 7(a), we note that, by virtue of feedback, both Z_{out1} and Z_{in2} assume a low value. Thus, this interface can achieve a wide bandwidth even in the presence of a large capacitance.

In a differential configuration, the bandwidth can be widened by means of negative Miller capacitors [Figure 7(b)] [4], which partially cancel the positive capacitances at *X* and *Y*. We study this circuit with $(W/L)_N = 2 \ \mu m/30 \ nm$ and $(W/L)_P = 4 \ \mu m/30 \ nm$ for Inv₁ and Inv₃, and twice these values for Inv₂ and Inv₄. We have $R_S = 100 \ \Omega$ and $R_{F1} = R_{F2} = 2 \ k\Omega$.

Displayed in Figure 7(c), the differential frequency response suggests a voltage gain of 15 dB. The 3-dB bandwidth rises from 36 GHz to 63 GHz after 3-fF Miller capacitors are introduced, at the cost of 1.3 dB of peaking. The substantial increase in the bandwidth is encouraging, but it also points to high sensitivity of the performance to the value of these capacitors. The power consumption is about 2 mW. The absence of inductors makes this circuit attractive for high-speed applications.

The Source-Series Terminated Driver

In wireline systems, the transmitter (TX) must drive a transmission line having a characteristic impedance, Z_0 , of 50Ω . In addition, the TX output impedance must also be about 50Ω to suppress secondary reflections. The TX driver potentially consumes high power as it must deliver a peak-to-peak voltage swing of about 500 mV to the channel.

Figure 8(a) depicts a current-mode logic (CML) realization of the driver. The output impedance, R_T , is chosen equal to Z_0 , and $I_{SS}(R_T || Z_0)$ yields the desired swing. For a swing of $V_{DD}/2$, this topology draws $P = V_{DD}^2/Z_0$.

Now, consider the structure shown in Figure 8(b), where an inverter serves as the driver, exhibiting an output impedance equal to the on-resistance of M_1 or M_2 . With this resistance set to Z_0 , the output voltage swing is



FIGURE 6: (a) The stacked amplifier, (b) the waveforms at A and X, (c) the waveforms at B and Y, and (d) the input-referred noise.

approximately equal to $V_{DD}/2$. In a differential configuration [Figure 8(c)], the circuit draws a current of $V_{DD}/(2R_T + 2Z_0) = V_{DD}/(4Z_0)$, consuming $P = V_{DD}^2/(4Z_0)$. Introduced in [5], this topology is called the "source-series terminated (SST) driver" and offers a fourfold power advantage over the CML counterpart of Figure 8(a).

The SST driver nonetheless entails one drawback. Unlike the CML circuit, it requires rail-to-rail input swings, an issue at very high speeds. We design the circuit of Figure 8(c) for $Z_0 = 50 \Omega$ with $(W/L)_N = 8 \mu m/30$ nm and $(W/L)_P = 16 \mu m/30$ nm. Figure 9(a) presents the differential output eye diagram at 56 Gb/s, exhibiting little jitter. Next, we precede the drivers by inverters having $(W/L)_N = 2 \mu m/30$ nm and $(W/L)_P = 4 \mu m/30$ nm. The fan-out of 4 does yield some jitter at the input and output of the driver. The resulting data eye delivered to the channel is displayed in Figure 9(b), revealing a peak-to-peak jitter of 400 fs. The overall circuit draws 4 mW.



FIGURE 7: (a) A wideband amplifier, (b) its differential version using negative Miller capacitances, and (c) its frequency response.



FIGURE 8: (a) A CML driver, (b) a source-series terminated (SST) driver, and (c) its differential version.

We should remark that practical designs often place a physical (linear) resistor, R_s , in series with each inverter output. For example, $R_s = 25 \Omega$, requiring that the inverters be twice as wide to provide the other 25 Ω necessary for proper back termination. This method assumes that resistors incur less variability than the on-resistance of MOS transistors. The cost is the higher power consumption of the predrivers.

The PAM4 Driver

The SST driver concept can be extended to four-level pulse-amplitude modulation (PAM4). Shown in single-ended form in Figure 10(a), the circuit is driven by the leastsignificant bit (LSB) and the most significant bit (MSB), operating as a 2-bit digital-to-analog converter (DAC). It can be shown that the $1.5Z_0$ and $3Z_0$ output impedances yield four equally spaced levels at the output [6].

We design the fully differential driver [Figure 10(b)] with $(W/L)_N = 4 \mu m/30 \text{ nm and } (W/L)_P = 8 \mu m/30 \text{ nm}$ for the 2× inverter and half of these values for the 1× counterpart. Each input arrives at 56 Gb/s. Figure 11 plots the differential output eye at 112 Gb/s. The driver consumes 5 mW.

The reader may note that the middle eye in Figure 11 is slightly taller than the bottom and top ones. Arising from the nonlinear output resistance of the inverters' MOS devices, this issue can be alleviated by making the LSB inverters about 10% stronger [8]. Alternatively, resistors can be placed in series with the inverter outputs to reduce the nonlinearity.

The Inductively Peaked Clock Buffer

Clocks driving long interconnects and/or a large number of transistors face heavy capacitive loading. At very high speeds, such clocks suffer from small amplitudes and become prone to noise and jitter. It is possible to improve clock waveforms through the use of inductive peaking. Depicted in Figure 12(a) [7] is an example employing shunt-series



FIGURE 9: Differential output eye diagrams of the SST driver: (a) with an ideal predriver and (b) with an actual predriver.

peaking for driving a large load capacitance, C_L . Inverters Inv_1 and Inv_2 serve as clock buffers, and Inv_3 and Inv_4 provide positive feedback (or a negative resistance) between A and B, speeding up the transitions. If excessively strong, Inv_3 and Inv_4 can cause oscillation, i.e., the topology behaves as an injection-locked oscillator with a limited lock range.

Let us design the circuit to drive $C_L = 100$ fF at a clock frequency of 56 GHz. With the values shown in Figure 12(a), and assuming an inductor Q of 10 at 56 GHz, we obtain the A and B waveforms displayed in Figure 12(b) and the X and *Y* waveforms in Figure 12(c). The single-ended output swing is about 740 mV_{pp} and the power consumption around 6 mW. With no inductive peaking, on the other hand, the theoretical power would be given by $2f_{CK}C_LV_{DD}^2 = 10$ mW. The narrowband nature of the buffer also reduces the clock jitter [7].

The Time-to-Digital Converter

Digital phase-locked loops incorporate a time-to-digital converter (TDC) to digitize the phase difference between their reference and the feedback signal. Similar to an analog-to-digital converter (ADC), the TDC must provide: 1) sufficient resolution for acceptable quantization noise and 2) a full-scale commensurate with the maximum phase error. The latter is dictated by the maximum phase fluctuation of the feedback signal in a fractional-N environment. A common TDC topology employs two chains of inverters with different delays. Called the "vernier TDC", the structure is shown in Figure 13(a),



FIGURE 10: (a) A basic PAM4 driver and (b) its differential version.



FIGURE 11: Output eye diagram of PAM4 driver.



FIGURE 12: (a) An inductively-peaked clock buffer, (b) the waveforms at A and B, and (c) the waveforms at X and Y.

where the inverters are nominally identical and the additional delay, ΔT , is created by means of C_1 - C_n . As periodic inputs A and B travel through the two paths, both positive and negative phase errors.

The lower bound on the phase resolution, ΔT , is dictated by random mismatches between the inverters in

the two paths. The full scale is determined by the number of stages.

We simulate the vernier TDC of Figure 13(a) with $(W/L)_N = 0.5 \ \mu m/30 \ nm$ and $(W/L)_P = 1 \ \mu m/30 \ nm$ for the inverters and $C_1 = \cdots = C_n = 1$ fF. For a zero input phase difference, we observe the waveforms shown in Figure 13(b), obtaining $\Delta T \approx 1.4$ ps per stage.



FIGURE 13: (a) The vernier TDC and (b) its waveforms. FF: flipflop.

Suppose a fractional-N synthesizer is designed for an output frequency of 7 GHz. The input phase error can reach $\pm 3T_{VCO}$, where $T_{VCO} = 1/(7 \text{ GHz}) =$ 143 ps. Thus, the TDC full scale must be as wide as 430 ps, requiring 430 ps/ $\Delta T = 306$ stages in each path. Such long chains introduce considerable phase noise.

The Digital-to-Time Converter

The maximum phase error of $\pm 3T_{VCO}$ mentioned in the previous section can be reduced by means of a digital-to-time converter (DTC), thereby allowing a narrower full scale for the TDC. Illustrated in Figure 14(a), the idea is to insert a programmable phase change in the reference path that matches the phase jumps in V_F .

The DTC acts as a digitally controlled variable-delay line and can be realized as depicted in Figure 14(b), where the thermometer code $D_1 \dots D_m$ controls the capacitors and, hence, the delay. Viewed as an DAC, the structure suffers from differential and integral nonlinearity due to mismatches among the capacitors and their corresponding switches. Noting that D_{cont} in Figure 14(a) contains shaped quantization noise, we recognize that the DTC nonlinearity folds high-frequency noise to inband noise.

We simulate the structure of Figure 14(b) with $(W/L)_N = 250 \text{ nm}/30 \text{ nm}$, $(W/L)_P = 500 \text{ nm}/30 \text{ nm}$, and $C_u =$ 0.5 fF. Displayed in Figure 14(c), the resulting waveforms at X indicate a phase resolution of 350 fs as the thermometer code increases. We now appreciate the difficulty of using the DTC in Figure 14(a): For a full scale of, e.g., 430 ps, the capacitance at X must increase by a factor of $430/0.35 \approx 1,230$, and so must the transition times at this node. Consequently, the waveform incurs an enormous amount of jitter.

To resolve this issue, we can partition the programmable delay among multiple stages [Figure 15(a)]. However, mismatches between the



FIGURE 14: (a) Use of a DTC in a fractional-N synthesizer, (b) its basic implementation, and (c) its waveforms.

inverter strengths lead to nonlinearity. As shown in Figure 15(b), the delay characteristic changes its slope due to this mismatch.

The Analog Feedforward Equalizer

Lossy channels encountered in wireline systems demand equalization in both the TX and the receiver (RX). This is accomplished on the TX side by a feedforward equalizer (FFE). Shown in Figure 16(a) is a generic FFE, which delays D_{in} by multiples of the bit period, T_b , and produces a weighted sum of the results at D_{out} . Coefficients $\alpha_1, ..., \alpha_n$ are so selected as to cancel the postcursors of the channel's impulse response. From another perspective, the FFE provides a high-pass response that partially compensates for the low-pass behavior of the channel.

For example, suppose we wish to transport 56-Gb/s data through a channel having the impulse response shown in Figure 16(b). The first postcursor is equal to 8%, requiring $\alpha_1 = -0.08$.

Realized by latches, the delay elements in Figure 16(a) prove power hungry. Alternatively, they can be simply formed by "analog" delay stages, i.e., inverters [8]. Variations of this delay with PVT does alter the FFE frequency response, but the compensation still remains effective.

Let us return to the SST driver described previously and add a 56-Gb/s analog FFE to it for the channel represented by Figure 16(b). Figure 17(a) shows the result. We choose $(W/L)_N = 1 \ \mu m/30 \ nm$ and $(W/L) = 2 \ \mu m/30 \ nm$ for the FFE inverters that drive the channel.



FIGURE 15: (a) A two-stage DTC and (b) its characteristic.

The main inverters are eight times as strong, i.e., $\alpha = 12.5\%$. This value is greater than the first postcursor in Figure 16(b), but it yields optimum results. With differential signal paths, the negative values of α_j can be readily accommodated. For $T_b = 17$ ps, we construct the delay elements as four inverters.



FIGURE 16: (a) A basic FFE and (b) a channel impulse response example.



FIGURE 17: (a) Differential driver with FFE, (b) its output eye with no FFE, and (c) its output eye with FFE.

(continued on p. 159)



FIGURE 1: (a) and (b) Dr. Vivienne Sze presents her lecture at the University of Texas at Austin.

For the first part, Dr. Sze addressed the issue of increased power and compute demands as technology advances. Similar to overbooking by airlines, overbooking data can be used to achieve greater data speeds and energy efficiency than standard models. Tools such as compute in memory architecture can help achieve this goal. She emphasized the importance of modeling for designspace exploration, support circuits, and emerging devices and the importance of co-design across the hardware stack.

The second part of the presentation covered co-design algorithms and hardware. Dr. Sze spoke about the importance of building and storing occupancy maps as an essential part of the autonomy for localization, navigation, and obstacle detection. Her key takeaways included the importance of co-design of hardware and algorithms as hardware improvements are limited. Hardware can influence algorithm design and help shift bottlenecks.

The final section of the lecture covered co-design across systems. Dr. Sze used low-power 3D time-of-flight imaging as an example to explain depth estimation. She discussed the concept of DecTrain: deciding when to start and stop training on the fly. The idea is to decide when to conduct training based on the margin to improve and the ability to improve. Dr. Sze summarized this as a need to include the choice of energy allocations across compute, sensing, and actuation.

For her conclusion, Dr. Sze emphasized the importance of co-design and collaboration between the different stakeholders in system design. Thank you very much, Vivienne, for this wonderful talk!

> —Mikko Sojonen and Stefano Pietri 555

THE ANALOG MIND (continued from p. 20)

Figures 17(b) and (c) plot the channel output eye diagram without and with FFE, respectively. We note that the eye is opened to some extent. Further performance improvement is delegated to the receiver.

References

- M.-D. Ker, S.-L. Liu, and C.-S. Tsai, "Design of charge pump circuit with consideration of gate-oxide reliability in low-voltage CMOS processes," *IEEE J. Solid-State Circuits*, vol. 41, no. 5, pp. 1100–1107, May 2006, doi: 10.1109/JSSC.2006.872704.
- [2] J. Jung and B. Razavi, "A 25-Gb/s 5-mW CMOS CDR/deserializer," *IEEE J. Solid-State*

Circuits, vol. 48, no. 3, pp. 684–697, Mar. 2013, doi: 10.1109/JSSC.2013.2237692.

- [3] X. Tang, L. Shen, B. Kasap, and N. Sun, "An energy-efficient comparator with dynamic floating inverter amplifier," *IEEE J. Solid-State Circuits*, vol. 55, no. 4, pp. 1011–1022, Apr. 2020, doi: 10.1109/ JSSC.2019.2960485.
- [4] S. Mondal and D. Hall, "A 13.9-nA ECG amplifier achieving 0.86/0.99 NEF/PEF using AC-coupled OTA-stacking," *IEEE J. Solid-State Circuits*, vol. 55, no. 2, pp. 414–425, Feb. 2020, doi: 10.1109/JSSC.2019.2957193.
- [5] C. Menolfi, T. Toifl, P. Buchmann, and J. Weiss, "A 16Gb/s source-series terminated transmitter in 65nm CMOS SOI," in Proc. IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers, San Francisco, CA, USA, Feb. 2007, pp. 446–614, doi: 10.1109/ISS-CC.2007.373486.
- [6] Y. Chang, A. Manian, L. Kong, and B. Razavi, "An 80-Gb/s 40-mW wireline PAM4 transmitter," *IEEE J. Solid-State Circuits*, vol. 53, no. 8, pp. 2214-2226, Aug. 2018, doi: 10.1109/JSSC.2018.2831226.
- [7] J. Kim, S. Kundu, M. Beach, and S. Kim, "A 224-Gb/s DAC-based PAM-4 quarterrate transmitter with 8-tap FFE in 10-nm FinFET," *IEEE J. Solid-State Circuits*, vol. 57, no. 1, pp. 6–20, Jan. 2022, doi: 10.1109/ JSSC.2021.3108969.
- [8] M. Forghani, Y. Zhao, P. K. Khanna, and B. Razavi, "A 112-Gb/s 58-mW PAM4 transmitter in 28-nm CMOS technology," in *Proc. IEEE Symp. VLSI Technol. Circuits (VLSI Technol. Circuits)*, Kyoto, Japan, 2023, pp. 1–2, doi: 10.23919/VLSITechnologyandCir 57934.2023.10185362.