Behzad Razavi

# The Design of a Transimpedance Amplifier

High-speed transimpedance amplifiers (TIAs) serve in the front end of optical communication receivers (RXs). Despite or because of their simple topologies, TIAs pose rigid tradeoffs among their gain, noise, and bandwidth (BW). In this article, we design a TIA in 28-nm CMOS technology while targeting the following specifications:

- non-return-to-zero (NRZ) data rate: 40 Gb/s
- input-referred noise current $< 10\,\text{pA}/\sqrt{\text{Hz}}$
- transimpedance gain: $1\,\text{k}\Omega$
- power consumption $< 5\,\text{mW}$.

The choice of the noise and gain values becomes clear after we delve into the bandwidth and sensitivity requirements. The circuit is simulated in the slow-slow corner of the process at 75° C and with a worst-case supply of $1\ \text{V} - 5\% = 0.95\,\text{V}$. The reader is referred to the literature [1], [2], [3], [4], [5], [6] for more details.

## General Considerations

Figure 1 shows a typical optical communication receiver front end. A photodiode (PD) senses the light arriving through a fiber and generates a proportional current. The TIA then converts this current to voltage and applies the result to a limiting amplifier.

We must recognize that the TIA bandwidth dictates a tradeoff between intersymbol interference (ISI) and the total integrated noise. For a data rate, $r_b$, of 40 Gb/s, the TIA should preferably provide a bandwidth of about $0.7 \times 40\,\text{GHz} = 28\,\text{GHz}$. This point can be understood by noting that the sinc$^2$ spectrum of NRZ data (Figure 2) carries about 88% of the signal's power in the range of $f = 0$ to $f = 0.7 r_b$. That is, the removal of the signal spectrum beyond $0.7\ r_b$ omits only 12% of the power, causing negligible ISI.

On the other hand, such a bandwidth also leads to a large amount of noise. For example, a first-order system with a pole frequency of $f_p$ exhibits an equivalent noise BW equal to $(\pi/2) f_p$.

The matter of noise encourages us to select a narrower BW, perhaps around $0.5 r_b$ (the "Nyquist frequency"), which contains about 75% of the signal power. We must then quantify the ISI and determine whether it is acceptably low. This is accomplished by examining the TIA output eye diagram in response to random data.

The bandwidth challenges begin at the RX input, as revealed by the three capacitances depicted in Figure 3. The PD exhibits one, $C_{PD}$, that trades with its "responsivity," i.e., the amount of current that it generates for a given input optical power. The greater the PD area is, the more current it delivers but at the cost of a higher $C_{PD}$. For example, the PDs reported in [7] and [8], respectively, display responsivities of 0.15 mA/mW and 0.6 mA/mW with capacitances equal to 2.5 fF and 22 fF.

The PD is typically off-chip and must connect to the TIA through a pad, thus imposing a pad capacitance, $C_{pad}$, as well. The third component, $C_{TIA}$, arises from the TIA itself. We assume $C_{PD} + C_{pad} \approx 50\,\text{fF}$.

## Noise and Sensitivity

Let us now deal with the TIA's noise requirement, and specifically, the receiver sensitivity. Suppose that the received light power in Figure 1 is around $25\,\mu\text{W}\,(\equiv -16\,\text{dBm})$ and the PD provides a responsivity of 1 mA/mW.

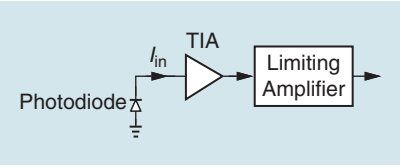


**FIGURE 1:** An optical receiver front end.

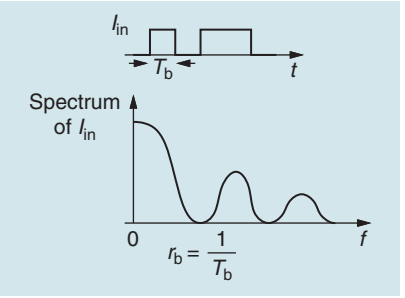


**FIGURE 2:** The sinc$^2$ spectrum of NRZ data.

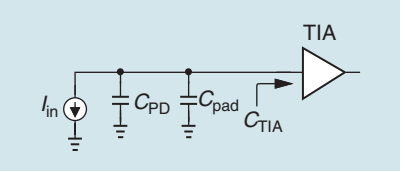


**FIGURE 3:** The capacitances at the input of a TIA.

*Digital Object Identifier 10.1109/MSSC.2022.3219682*
*Date of current version: 18 January 2023*

The PD peak-to-peak output current, $I_{pp}$, is thus equal to 25 $\mu$A, and it must be detected with a sufficiently high signal-to-noise ratio (SNR) to guarantee a certain bit error rate (BER). For NRZ data, we have

$$\text{BER} = Q\left(\frac{I_{pp}}{2I_{in,rms}}\right) \quad (1)$$

where $Q$ denotes the "error function" (the integral of a Gaussian), and $I_{in,rms}$ denotes the TIA input-referred noise. For BER $< 10^{-12}$, we have $I_{pp}/(2I_{in,rms}) \approx 7$, and hence, $I_{in,rms} < 1.8\,\mu A_{rms}$.

Allowing a TIA BW of $r_b/2 = 20$ GHz and approximating its response by a one-pole system, we write the total integrated noise as

$$\frac{\pi}{2} \times 20\,\text{GHz} \times \overline{I_{n,in}^2} \approx (1.8\,\mu A_{rms})^2 \quad (2)$$

where $\overline{I_{n,in}^2}$ is the spectral density of the input-referred noise current and assumed to be constant. It follows that

$$\sqrt{\overline{I_{n,in}^2}} = 10\,\text{pA}/\sqrt{\text{Hz}}. \quad (3)$$

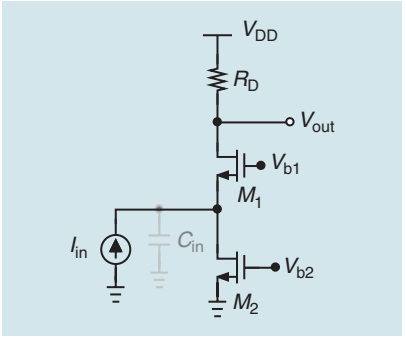


**FIGURE 4:** The common-gate (CG) stage as a TIA.

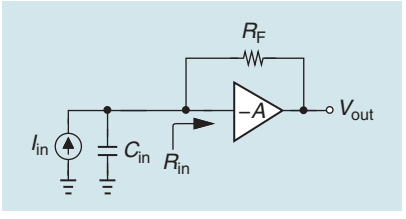


**FIGURE 5:** A feedback TIA.

This result justifies our original TIA noise specification.

## Common-Gate TIAs

To minimize the effect of $C_{PD} + C_{pad}$ in Figure 3, the TIA must provide a low input resistance, $R_{in}$. If we wish the input bandwidth to exceed 20 GHz, then

$$\frac{1}{2\pi R_{in}(C_{PD} + C_{pad} + C_{TIA})} \geq 20\,\text{GHz}. \quad (4)$$

Guessing a value of 25 fF for $C_{TIA}$, we obtain

$$R_{in} \leq 106\,\Omega. \quad (5)$$

This low input resistance may point to a common-gate (CG) TIA, but we can readily see that such a choice fails to meet the noise target. Considering the stage shown in Figure 4 and neglecting channel-length modulation and the body effect, one can show that the input-referred noise current of the CG topology arises from only $R_D$ and $M_2$ if all capacitances are neglected. That is

$$\overline{I_{n,in}^2} = \frac{4kT}{R_D} + 4kT\gamma g_{m2}. \quad (6)$$

This amount must remain lower than $10\,\text{pA}/\sqrt{\text{Hz}}$ but with a bandwidth of about 20 GHz at the input. From (5), we must choose $g_{m1} \approx 1/(100\,\Omega)$, and hence

$$\frac{2I_{D1}}{V_{GS1} - V_{TH1}} = \frac{1}{100\,\Omega}. \quad (7)$$

We select some typical values and estimate the noise. For example, if $I_{D1} = 1$ mA, we require $V_{GS1} - V_{TH1} = 200$ mV. With $V_{DD} \approx 1$ V, we have 800 mV left for $V_{DS2}$ and the drop across $R_D$. If the former is 300 mV and the latter is 500 mV, we have $g_{m2} = 2\,\text{mA}/300\,\text{mV}$ and $R_D = 500\,\Omega$, arriving at

$$\overline{I_{n,in}^2} = \frac{4kT}{500\,\Omega} + \frac{4kT\gamma}{150\,\Omega}. \quad (8)$$

At $T = 350$ K and with $\gamma = 1$, $\sqrt{\overline{I_{n,in}^2}} = 13\,\text{pA}/\sqrt{\text{Hz}}$. Note that the second term on the right-hand side is several times larger than the first. Thus, even with no other imperfections, the CG stage suffers from excessive noise.

The input capacitance, $C_{in}$, in Figure 4 exacerbates the situation. It can be shown that the noise current of $M_1$ now is multiplied by $\pi f_{p,out}$ as it is referred to the input, where $f_{p,out}$ denotes the output pole frequency [6]. In summary, the CG TIA is ill-suited to our purpose.

## Feedback TIAs

The constraints described in the previous section can be avoided through the use of a feedback topology. As exemplified by Figure 5, placing a resistor around an amplifier having a gain of $-A$ yields $R_{in} = R_F/(1 + A)$, which can be chosen around 100 Ω for our target BW. The circuit's input-referred noise current is given by

$$\overline{I_{n,in}^2} = \frac{4kT}{R_F} + \frac{\overline{V_{n,A}^2}}{R_F^2} \quad (9)$$

where $\overline{V_{n,A}^2}$ is the amplifier input-referred noise. The key point here is that $R_F$ carries no bias current and can be greater than $R_D$ in (6), and more importantly, $\overline{V_{n,A}^2}/R_F^2$ can be less than $4kT\gamma g_{m2}$. This property becomes clearer later.

## The Problem of Long Runs

Random data exhibits long "runs," e.g., a long sequence of consecutive ones. In optical receiver design, this attribute makes it difficult to employ ac coupling between the stages.

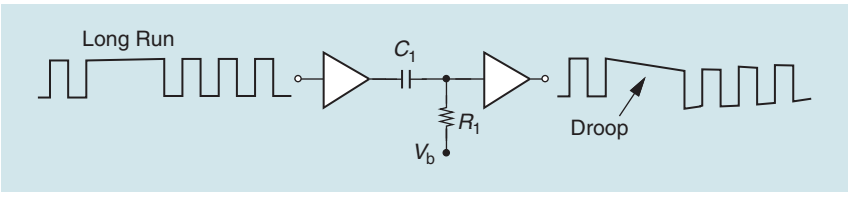


**FIGURE 6:** The effect of long runs in random data.

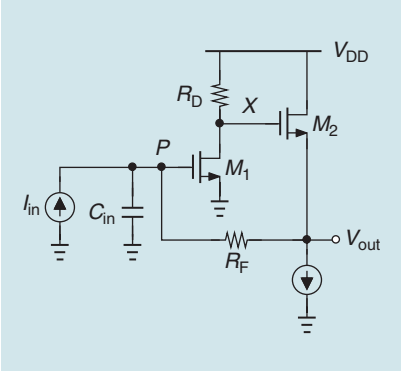


**FIGURE 7:** A feedback TIA implementation.

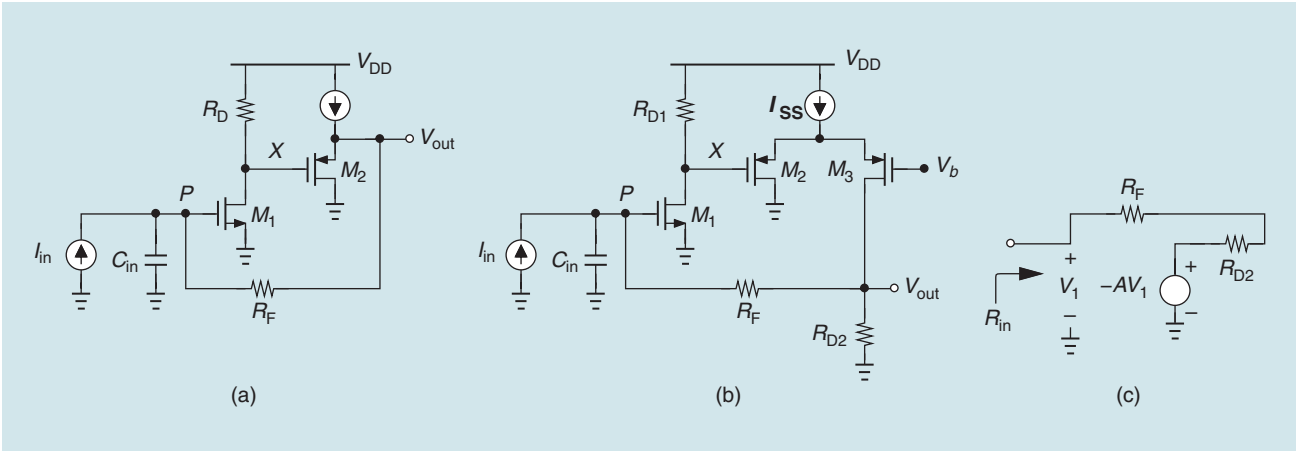


**FIGURE 8:** (a) A feedback TIA using a PMOS follower; (b) the circuit of (a) with a CG stage added; and (c) a circuit model for computing the input resistance.

Illustrated in Figure 6, the issue is that the data experiences a "droop" during long runs, possibly precluding proper detection. For this reason, the interfaces in such receivers must typically rely on dc coupling, a condition that makes low-voltage design challenging.

For example, consider the feedback realization in Figure 7, which incorporates a source follower to drive $R_F$. We have $V_X = V_{GS2} + V_{GS1}$. Resistor $R_D$ can therefore sustain a dc drop of only $V_{DD} - (V_{GS2} + V_{GS1})$, limiting the voltage gain of the common-source (CS) stage. If node $X$ is capacitively coupled to the gate of $M_2$, this issue is resolved, but the droop manifests itself.

## TIA Topology

We wish to replace the source follower in Figure 7 with a circuit that does not create a large positive or negative level shift. For the CS stage to provide a reasonable voltage gain, the dc drop across $R_D$ imposes a relatively low value for $V_X$ (equal to the overdrive voltage of $M_1$). We then surmise that this node must be sensed by the gate of a PMOS device. Depicted in Figure 8(a) is one possibility, where the PMOS source follower $M_2$ introduces a positive level shift. It appears that $V_{out}$ is now compatible with the gate of $M_1$, but a closer look reveals otherwise. We have $V_{DS1} + |V_{GS2}| = V_{GS1}$, and hence, $V_{GS1} - V_{DS1} = |V_{GS2}|$. Since $|V_{GS2}|$ is typically greater than $V_{TH1}$, $M_1$ inevitably resides in the triode region.

This calculation suggests that the voltage at the source of $M_2$ must be shifted down before reaching the gate of $M_1$. This can be accomplished by a CG stage, as illustrated in Figure 8(b). We can also view $M_2$ and $M_3$ as a differential pair, expecting some voltage gain from $X$ to $V_{out}$.

Let us examine the circuit's bias conditions. With a high loop gain, $V_X$ adjusts itself to remain around $V_b$, defining $I_{D1}$ as $(V_{DD} - V_b)/R_{D1}$. According to this current, $M_1$ develops a certain $V_{GS}$, which also appears at the drain of $M_3$. That is, $|I_{D3}| = V_{GS1}/R_{D2}$ and $|I_{D2}| = I_{SS} - |I_{D3}|$. To maximize the voltage gain of the differential pair, we must have $I_{D2} \approx I_{D3} \approx I_{SS}/2$, and therefore, select $R_{D2} = V_{GS1}/(I_{SS}/2)$.

We should remark that the expression $R_{in} = R_F/(1 + A)$ obtained for the topology of Figure 5 tacitly assumes a zero-output impedance for the core amplifier. In Figure 8(b), on the other hand, the open-loop circuit exhibits $R_{out} \approx R_{D2}$. Using the model shown in Figure 8(c), we have

$$R_{in} = \frac{R_F + R_{D2}}{1 + A} \tag{10}$$

where $A$ denotes the unloaded open-loop gain.

## The Preliminary Design

We begin with the design shown in Figure 9, where the sources of $M_2$ and $M_3$ are tied to their n-well so as to avoid the rise in their threshold voltage due to the body effect. Since $V_b = 300\,\text{mV}$, we have $I_{D1} \approx (V_{DD} - V_b)/R_{D1} \approx (0.95\,\text{V} - 0.3\,\text{V})/1\,\text{k}\Omega = 0.65\,\text{mA}$. The CS stage and the differential pair display voltage gains equal to 3.5 and 1.1, respectively, offering a loop gain of 3.9.

How do we select the value of $R_F$? We surmise from (9) that the term $4kT/R_F$ is dominant, dictating $R_F \geq 200\,\Omega$ for $\overline{I_{n,in}^2} = 10\,\text{pA}/\sqrt{\text{Hz}}$. But such a low transimpedance gain

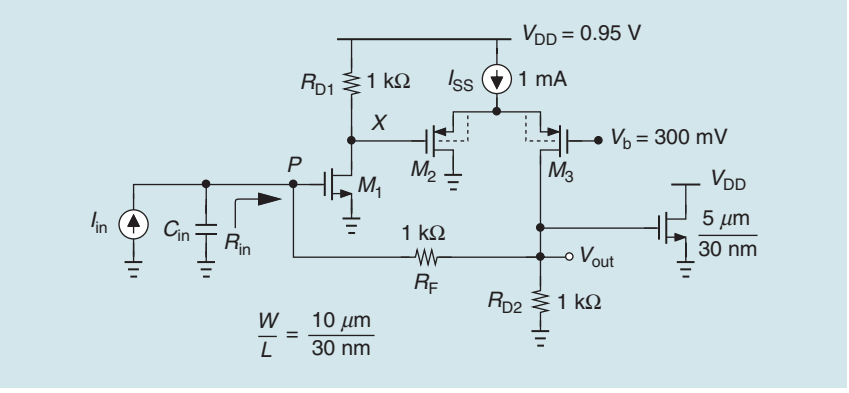


**FIGURE 9:** The preliminary TIA implementation.

produces only a peak-to-peak voltage of $25\,\mu A \times 200\,\Omega = 5\,mV$ at the circuit's output and makes the noise of the next stage significant. We therefore raise $R_F$ to $1\,k\Omega$, expecting $R_{in} = (R_F + R_{D2})/(1 + 3.9) = 200\,\Omega$. This value appears to contradict the bound prescribed by (5), implying inadequate bandwidth. Fortunately, the actual BW is quite larger, a point to which we return later.

To evaluate the performance of the TIA, we perform ac, noise, and transient simulations. Figure 10 plots the frequency response from the input current to the voltage at node $P$ and to the output voltage. The former represents the input impedance and takes on a value of $46\,dB\Omega\,(\equiv 200\,\Omega)$ at low frequencies.

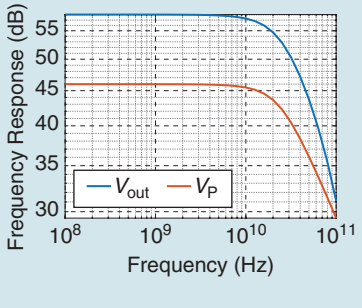


**FIGURE 10:** The TIA ac response.

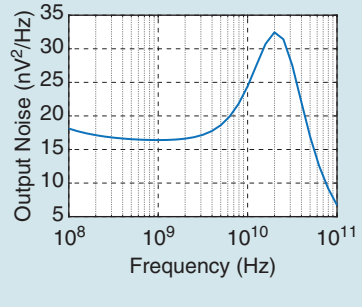


**FIGURE 11:** The TIA output noise spectrum.

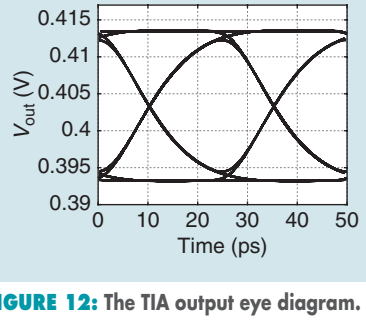


**FIGURE 12:** The TIA output eye diagram.

The latter suggests a transimpedance gain of $58\,dB\Omega\,(\equiv 800\,\Omega)$ and displays a 3-dB bandwidth of 19 GHz.

It is interesting to note that $R_{in} = 200\,\Omega$ and $C_{in} = C_{PD} + C_{pad} + C_{TIA} \approx 75\,fF$ yield only a bandwidth of 11 GHz in Figure 9. Why, then, is the overall TIA bandwidth greater? This phenomenon can be understood from two different perspectives. First, due to the pole at $X$, the circuit is at least of second order, and its closed-loop poles depart from the real axis. It can be proved that, with two such poles, the BW can increase by about 40% compared to a first-order system [5], [6].

Second, we can say that the poles at $X$ and at the output node drop the open-loop gain at high frequencies, thus causing the closed-loop input impedance to rise. This inductive behavior partially cancels $C_{in}$ and increases the BW.

Plotted in Figure 11 is the output noise voltage spectrum. At low frequencies, the noise is dominated by $R_F$ and is calculated by multiplying the input-referred value, $4kT/R_F$ by the square of the transfer function's magnitude, $(800\,\Omega)^2$. At high frequencies, on the other hand, $C_{in}$ nearly shorts the gate of $M_1$ to ac ground, allowing the output noise to reach its open-loop value. For example, the circuit simply amplifies the noise current of $M_1$ $(= 4kT\gamma g_{m1})$ by $R_{D1}$ and the gain of the differential pair.

The input-referred noise current is computed in two steps. First, we integrate the output noise spectrum from 100 MHz to 100 GHz, obtaining a value of $1.8 \times 10^{-6}\ V^2 (\equiv 1.34\ V_{rms})$. Second, we assume a first-order response and ask how much input noise current would generate $1.34\,mV_{rms}$ at the output with a TIA bandwidth of 19 GHz. That is,

$$\overline{I_{n,in}^2}\left(\frac{\pi}{2} \times 19\,GHz\right) \times (800\,\Omega)^2 = (1.34\,mV)^2 \qquad (11)$$

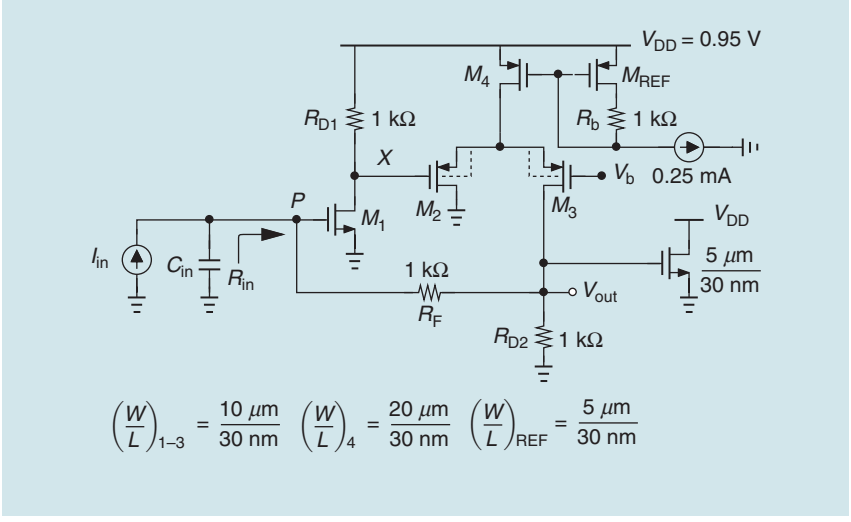


**FIGURE 13:** The complete TIA.

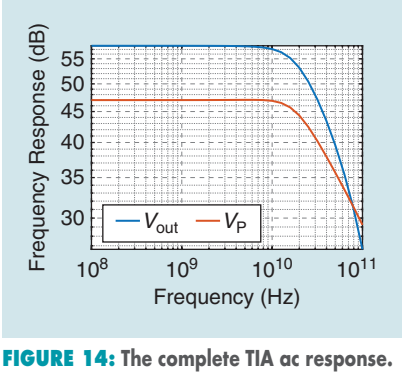


**FIGURE 14:** The complete TIA ac response.

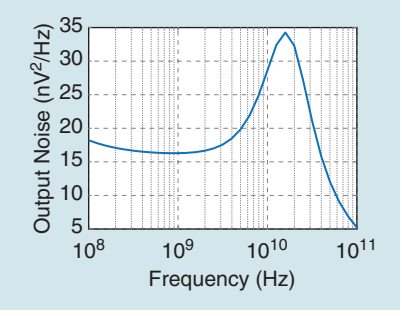


**FIGURE 15:** The complete TIA output noise spectrum.

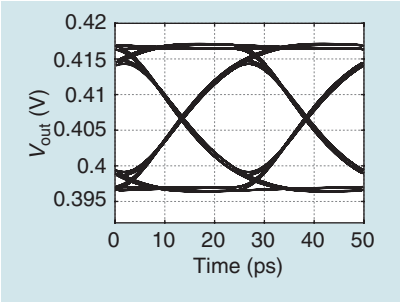


**FIGURE 16:** The complete TIA output eye diagram.

where $(\pi/2) \times 19\,\text{GHz}$ represents the equivalent noise bandwidth for a first-order system. It follows that $\sqrt{\overline{I_{n,\text{in}}^2}} = 9.7\,\text{pA}/\sqrt{\text{Hz}}$.

Our third simulation examines the output in response to an input NRZ current having a peak-to-peak value of 25 $\mu$A. Figure 12 plots the resulting eye diagram, exhibiting a voltage swing of 20 mV, which agrees with $25\,\mu\text{A} \times 800\,\Omega$. The eye is reasonably open.

## The Complete Design

In the final stage of the design process, we implement the tail current source in Figure 9 by a mirror arrangement. Shown in Figure 13 is the result, where a wide tail transistor is necessary for a $|V_{DS}|$ of lower than 150 mV. Resistor $R_b$ ensures that $M_{REF}$ sustains a $V_{DS}$ roughly equal to that of $M_4$. The circuit also includes a 5-$\mu$m/30-nm transistor as the output load to model the input device of the next stage. Note that the bias current of this transistor tracks that of $M_1$ as they share the same $V_{GS}$. The TIA draws a supply current of 2 mA.

The 20-$\mu$m tail transistor in Figure 13 introduces substantial capacitance in the signal path. We thus repeat the previous simulations to assess the performance. Figure 14 plots the magnitude response from $I_{in}$ to $V_P$ and $V_{out}$, revealing a 3-dB bandwidth of 17 GHz. The response at $P$ also indicates that the input resistance has risen to $47\,\text{dB}\Omega = 224\,\Omega$. This occurs because the output resistance of $M_4$ slightly lowers the loop gain.

Figure 15 plots the output noise spectrum, the area under which amounts to $1.3 \times 10^{-6}\,\text{V}^2$. Dividing this value by $(800\,\Omega)^2$ and $(\pi/2) \times 17\,\text{GHz}$, we obtain an input-referred noise current of $8.7\,\text{pA}/\sqrt{\text{Hz}}$. More than 90% of the output noise still originates from $R_F$.

Depicted in Figure 16 is the new eye diagram. As expected, the narrower BW degrades the time response to some extent.

We may ponder the use of series or shunt inductive peaking to improve the TIA' speed. Unfortunately, the necessary inductance values fall in the range of several nanohenries, presenting large capacitances to the signal path around the feedback loop. According to simulations, such remedies only degrade the output eye.

Given that layout parasitics further limit the bandwidth, we predict that a data rate of 40 Gb/s cannot be accommodated by the TIA unless the noise specification is relaxed. Interestingly, the BW–noise tradeoff in this topology cannot be eased by consuming greater power. We also note that the circuit's transimpedance gain is short of the 1-k$\Omega$ target.

## References


[1] B. Moeneclaey et al., "A 40-Gb/s transimpedance amplifier for optical links," *IEEE Photon. Technol. Lett.*, vol. 27, no. 13, pp. 1375–1378, Jul. 2015, doi: 10.1109/LPT.2015.2421521.

[2] A. Awny, R. Nagulapalli, M. Ko, and D. Kissinger, "A linear differential transimpedance amplifier for 100-Gb/s integrated coherent optical fiber receivers," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 2, pp. 973–986, Feb. 2018, doi: 10.1109/TMTT.2017.2752170.

[3] H. Kim, S. Chandrasekhar, and J. Bauman, "A Si BiCMOS transimpedance amplifier for 10-Gb/s SONET receiver," *IEEE J. Solid-State Circuits*, vol. 36, no. 5, pp. 769–776, May 2001, doi: 10.1109/4.918914.

[4] C. Kromer, G. Sialm, D. Erni, and H. Jackel, "A low-power 20-GHz 52-dB$\Omega$ transimpedance amplifier in 80-nm CMOS," *IEEE J. Solid-State Circuits*, vol. 39, no. 6, pp. 885–894, Jun. 2004, doi: 10.1109/JSSC.2004.827807.

[5] E. Sackinger, *Analysis and Design of Transimpedance Amplifiers for Optical Receivers*. Hoboken, NJ, USA: Wiley, 2017.

[6] B. Razavi, *Design of Integrated Circuits for Optical Communications*, 2nd ed. Hoboken, NJ, USA: Wiley, 2012.

[7] Y. Baumgartner, M. Seifried, P. Stark, and L. Czornomaz, "Novel CMOS-compatible ultralow capacitance hybrid III-V/Si photodetectors tested up to 32 Gbps NRZ," in *Proc. Opt. Fiber Commun. Conf. Dig.*, Mar. 2019, p. Th3B-3, doi: 10.1364/OFC.2019.Th3B.3.

[8] M. Huang et al., "56GHz waveguide Ge/Si avalanche photodiode," in *Proc. Opt. Fiber Commun. Conf. Dig.*, Jun. 2018, pp. 1–3, doi: 10.1364/OFC.2018.W4D.6.


SSC