4

Transimpedance Amplifier Design

4.1 Introduction

The transimpedance amplifier (TIA) is without a doubt the most critical building block of the optical receiver. It converts the current generated by the photodiode into an output voltage. The design of this block involves many trade-offs between noise, bandwidth, gain and stability. This chapter tries to reveal all subtleties and challenges encountered during the design of low-noise high-bandwidth TIAs.

A summary of the TIA specifications regarding transimpedance gain, bandwidth, noise and overload currents is given in Section 4.2. Section 4.3 tackles the design of a TIA with shunt-shunt feedback. Design equations are derived for the transimpedance gain, bandwidth, open-loop gain, loop gain and the noise performance. Implementations published in open literature are discussed in Section 4.4. Two main topologies are distinguished: the TIA with common-source input stage and the TIA with regulated cascode input stage. Also some interesting work presented at ISSCC is described. Finally, the designs of three different TIAs, of which two are implemented in a 0.18 µm CMOS and one in a 90 nm CMOS technology, are explained more thoroughly in Section 4.5. These case studies clearly demonstrate the compromises to be made during the design of low-noise high-speed TIAs.

4.2 Performance Requirements

This section presents the main performance requirements for a TIA: high transimpedance gain, high bandwidth, low noise and high overload current.

Transimpedance Gain

The transimpedance gain of the TIA, Z_{TIA}, is defined as the ratio of the small-signal output voltage to the small-signal input current:

$$Z_{TIA} = \frac{v_{out}}{i_{in}} = |Z_{TIA}(f)|e^{j\theta(f)}. \tag{4.1}$$

The higher this value, the more output signal is produced for a given input signal. The transimpedance gain is specified either in units of Ω or $\mathrm{dB}\Omega$. The value $\mathrm{dB}\Omega$ is calculated as $20 \cdot \log_{10}(Z_{TIA}/\Omega)$. The transimpedance gain is a complex quantity, with frequency-dependent magnitude $|Z_{TIA}(f)|$ and frequency-dependent phase shift $\theta(f)$. The transimpedance gain at low frequencies is usually flat, and represented by $Z_{TIA,0}$.

The first reason for having a TIA with high gain is to create a signal with an amplitude large enough to drive the post-amplifier (PA). But there is an additional reason which might be even more important: noise. As the TIA is the first stage in the optical receiver (Fig. 2.1), the noise of the next stages like the PA will be suppressed by the TIA gain. So a lower transimpedance gain (for example to obtain a higher bandwidth (4.11)) cannot simply be exchanged for a larger post-amplification. The total gain remains constant, but the total input-referred noise of the receiver will increase.

Bandwidth

The upper frequency at which $|Z_{TIA}(f)|$ (4.1) has dropped 3 dB below its DC value, is defined as the TIA bandwidth, BW_{TIA}.

As discussed in Section 2.5, a limited bandwidth causes ISI and degrades the opening of the eye diagram. To receive data with a certain bitrate R_b, the bandwidth must be as high as possible to minimize the ISI. But on the other hand, Section 4.3.3 will demonstrate that a large bandwidth increases the noise picked up by the TIA. As a compromise between noise and ISI, a TIA bandwidth equal to $0.7R_b$ is commonly used [Raz03, Säc05].

Noise

The input-referred current noise is one of the most critical TIA parameters. Often the noise of the TIA dominates all other noise sources and therefore determines the sensitivity of the receiver. The equivalent input-referred noise current is the current source that, together with the ideal noiseless TIA, reproduces the output noise of the actual noisy TIA. As stated before in Section 2.4.2, it is a fictitious quantity that cannot be observed in the actual circuit.

To determine the input-referred noise current, the noise power spectral density at the output for each noise source is calculated first. Typical noise sources are transistors, resistors and diodes. Assuming these sources are not correlated, they add up to form the total output noise power spectral density. The power spectral density of the input-referred noise current, $\overline{di^2_{n,TIA}}$, can then be found by taking the frequency-dependent transimpedance gain into account:

$$\overline{di_{n,TIA}^2} = \frac{\overline{dv_{n,TIA}^2}}{|Z_{TIA}(f)|^2}. \tag{4.2}$$

The input-referred rms noise current, $i_{n,rms}$ (2.7), also called the total integrated input-referred noise current of the TIA, $i_{n,TIA}$, is determined by dividing the rms output noise voltage by the DC value of the transimpedance gain. The rms output noise voltage is obtained by integrating the output-referred noise spectrum and taking the square root.

$$i_{n,TIA} = \frac{1}{Z_{TIA,0}} \sqrt{\int_0^\infty \overline{dv_{n,TIA}^2} df}. \tag{4.3}$$

Note that this definition is different from integrating the input-referred noise current to a certain bandwidth. For analytical calculations, the integration has to be carried out to infinity. For simulations and measurements, a few times the bandwidth BW_{TIA}, or simply the bandwidth (for higher order TIAs with a steep roll-off) is taken as upper limit.

For completeness, the definition of noise bandwidth BW_n is repeated here. It is the bandwidth, multiplied by the DC value of the power spectral noise density, which gives the same result as the integration of the power spectral noise density to infinity. For a first order circuit, BW_n equals $\frac{\pi}{2}$ times the 3-dB bandwidth [San06]. For a third order circuit with a steeper slope, the noise bandwidth and the 3-dB bandwidth nearly coincide.

Overload Current

A TIA may receive large signal currents, for example when the distance between transmitter and receiver is very short and the photodetector has a high responsivity. As the input level increases, the TIA will introduce nonlinearities in the signal. Due to the binary nature of transmitted data, some nonlinearity can be tolerated. However, too much nonlinearities may corrupt the signal levels and/or distort the zero crossings, increasing the BER. Naturally, this has to be avoided.

Following the simulations of Chapter 3, the responsivity of integrated CMOS diodes is rather low, because the standard silicon technology is not optimal for optical applications. So even for large optical signals, the currents at the input will remain below the overload current.

4.3 Design of the Shunt-Shunt Feedback TIA

The TIA topology used in this work is the so-called shunt-shunt feedback TIA. A negative feedback network senses the voltage at the output and returns a proportional current to the input. This type of feedback has the advantage that it lowers both the input and output impedance. As a result, the closed-loop bandwidth is increased by the loop gain. In this section, the main design equations for the shunt-shunt feedback TIA are derived.

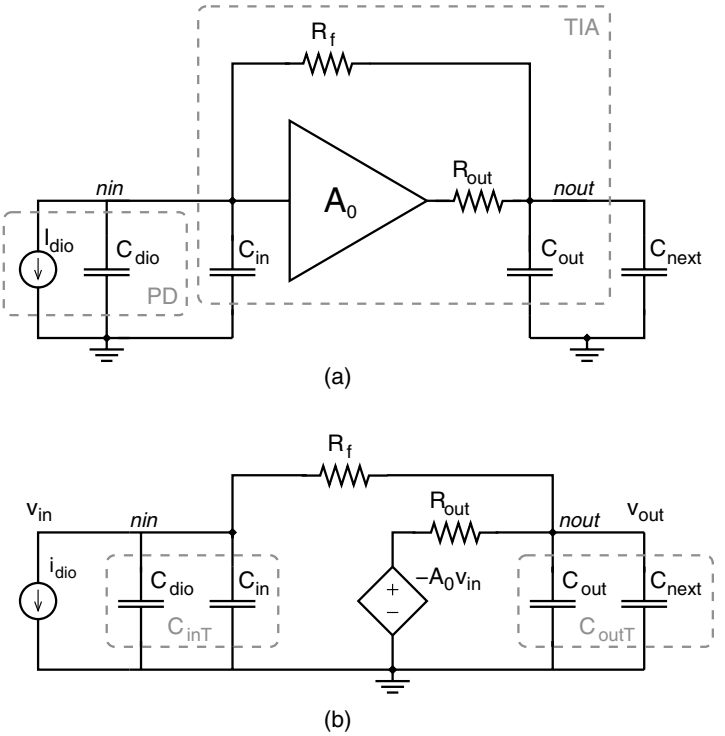


Fig. 4.1. The shunt-shunt feedback TIA: (a) general schematic, (b) small-signal equivalent circuit.

4.3.1 Transimpedance Gain and Bandwidth

Fig. 4.1(a) shows a general schematic of the shunt-shunt feedback TIA. The photodiode is represented by the current source I_{dio} in parallel with the junction capacitance C_{dio}. The TIA consists of a voltage amplifier with DC gain A_0 and feedback resistance R_f. The input capacitance C_{in} and output capacitance C_{out} are determined by the sizes of the transistors that constitute the voltage amplifier. C_{next} is the input capacitance of the next stage. R_{out} is the TIA output resistance and is usually much smaller than R_f.

The analysis of the small-signal equivalent circuit, depicted in Fig. 4.1(b), results in following transimpedance gain:

$$Z_{TIA} = \frac{v_{out}}{i_{dio}} = \frac{Z_{TIA,0}}{s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1}\right) + 1}, \tag{4.4}$$

with:

$$Z_{TIA,0} = \frac{R_f A_0}{A_0+1} - \frac{R_{out}}{A_0+1} \approx R_f, \tag{4.5}$$

$$C_{inT} = C_{dio} + C_{in}, \tag{4.6}$$

$$C_{outT} = C_{out} + C_{next}. \tag{4.7}$$

According to (4.5), the transimpedance gain at low frequencies equals R_f for large values of A_0 and small values of R_{out}.

To find the bandwidth of (4.4), two separate cases are considered. First, suppose the second pole of the TIA has a much higher magnitude than the first pole. Taking into account following approximations:

$$A_0 + 1 \approx A_0, \tag{4.8}$$

$$R_{out} \ll R_f, \tag{4.9}$$

$$R_{out}C_{outT} \ll R_f C_{inT}, \tag{4.10}$$

the well-known expression for the TIA bandwidth BW_{TIA} can be found:

$$BW_{TIA} = \frac{A_0}{2\pi R_f C_{inT}}. \tag{4.11}$$

The dominant pole, which determines the bandwidth in this case, is located at the input node *nin*. Due to the feedback loop, the resistance at this node is divided by the loop gain, which results in a factor A_0 increase of bandwidth. Therefore, A_0 has to be maximized during the design process. Other factors that influence bandwidth are the feedback resistor R_f and the total input capacitance C_{inT}. R_f usually cannot be made too small for gain (4.5) and noise considerations (4.36). C_{inT} consists of two parts: C_{dio}, determined by the photodiode topology, and C_{in}, which increases for larger transistor dimensions. This expression for the bandwidth reveals two important conclusions:

- As the photodiode junction capacitance directly appears in the denominator of (4.11), a photodiode-TIA co-design is mandatory.
- A design aiming for high bandwidth implies an optimization of the voltage gain A_0.

The non-dominant pole is located at the output node *nout*. It coincides with the bandwidth BW_{VA} of the voltage amplifier, which has only one pole in the simplified model of Fig. 4.1.

$$f_{nd,TIA} = BW_{VA} = \frac{1}{2\pi R_{out}C_{out}}. \tag{4.12}$$

However, the two poles of a second-order system can seldom be treated as two real poles. Mostly, they will be part of a complex conjugated pair. In basic control theory, the denominator of a second-order system is written as [Dor98]:

$$\left(\frac{s}{\omega_n}\right)^2 + 2\zeta\left(\frac{s}{\omega_n}\right) + 1, \tag{4.13}$$

where ζ is the dimensionless damping ratio and ω_n is the natural pulsation of the system. When $\zeta = 1$, the system is critically damped. Complex conjugated

poles occur when $\zeta < 1$. The relation between 3-dB bandwidth ω_{3dB}, natural pulsation ω_n and damping ratio ζ can be calculated by setting the magnitude of (4.13) equal to $2/\sqrt{2}$:

$$\frac{\omega_{3dB}}{\omega_n} = (1 - 2\zeta^2) + \sqrt{2}\sqrt{2\zeta^4 - 2\zeta^2 + 1}. \tag{4.14}$$

A smaller ζ will result in a larger 3-dB bandwidth, but also a higher overshoot in the time domain and a larger resonance peaking in the frequency domain. This causes degradation of the high and low levels and jitter in the eye diagram, despite the large bandwidth. In [Dor98] the percent overshoot P.O. is defined by:

$$\text{P.O.} = 100e^{-\zeta\pi/\sqrt{1-\zeta^2}}. \tag{4.15}$$

Setting ζ equal to $\sqrt{2}/2$ is an attractive solution. The percent overshoot is only 4 %, while the 3-dB bandwidth equals ω_n. It also generates a system with a maximally flat frequency response, which corresponds to a Butterworth filter. This way, a compromise is found between bandwidth and overshoot, leading to a high-quality eye diagram.

Equating (4.13) with the denominator of (4.4) gives following results for the natural pulsation and damping ratio of the second-order TIA depicted in Fig. 4.1:

$$\omega_n = \sqrt{\frac{A_0 + 1}{R_f R_{out} C_{inT} C_{outT}}}, \tag{4.16}$$

$$\zeta = \frac{1}{2} \frac{\frac{R_f C_{inT}}{R_{out} C_{outT}} + 1}{\sqrt{(A_0 + 1)\frac{R_f C_{inT}}{R_{out} C_{outT}}}}. \tag{4.17}$$

For sufficiently large gain A_0, the last equation can be rewritten as:

$$\frac{1}{R_{out} C_{outT}} = 4\zeta^2 \frac{A_0}{R_f C_{inT}}. \tag{4.18}$$

This is the condition at the input node *nin* and output node *nout* of Fig. 4.1 which has to be fulfilled to design a TIA with a certain damping ratio ζ. Combining (4.18) and (4.16), the corresponding natural pulsation can be found. More interesting is the 3dB-bandwidth, which can be calculated using (4.14). The results for some typical values of ζ are summarized in Table 4.1.

Comparing the fourth column of Table 4.1 with the earlier derived expression for the TIA bandwidth, (4.11), an increase in bandwidth is noticed. The system under study is unchanged, only the basic assumptions for the hand calculations are different. In the case of (4.11), the TIA has two well-separated real poles while in the other case, the poles are complex conjugates. Due to these complex poles, the resulting 3dB-bandwidth is larger, but some overshoot will occur in the eye diagram. However, this is tolerable as long as the

Table 4.1. Design equations for the second-order TIA of Fig. 4.1.

ζ	**Condition** (4.18)	ω_n (4.16)	ω_{3dB} (4.14)	**P.O.** (4.15)	**P.M.** (4.27)
$\frac{1}{2}$	$\frac{1}{R_{out}C_{outT}} = \frac{A_0}{R_fC_{inT}}$	$\frac{A_0}{R_fC_{inT}}$	$\frac{1.6A_0}{R_fC_{inT}}$	16 %	45°
$\frac{\sqrt{2}}{2}$	$\frac{1}{R_{out}C_{outT}} = \frac{2A_0}{R_fC_{inT}}$	$\frac{\sqrt{2}A_0}{R_fC_{inT}}$	$\frac{\sqrt{2}A_0}{R_fC_{inT}}$	4 %	63°
$\frac{\sqrt{3}}{2}$	$\frac{1}{R_{out}C_{outT}} = \frac{3A_0}{R_fC_{inT}}$	$\frac{\sqrt{3}A_0}{R_fC_{inT}}$	$\frac{1.07A_0}{R_fC_{inT}}$	0.4 %	72°

damping ratio is high enough and no ringing occurs. The smaller ζ, the higher the bandwidth, but the more overshoot and ringing. Also the phase margin decreases with ζ. An expression for the phase margin (P.M.) will be derived in the next subsection. Most important is that the main conclusions of (4.11) are still valid: the TIA bandwidth BW_{TIA} is inversely proportional to the input capacitance C_{inT} (which contains the photodiode capacitance C_{dio}), inversely proportional to the feedback resistor R_f, and directly proportional to the DC gain A_0 of the voltage amplifier. Note also that for $\zeta = \sqrt{3}/2$, the expression for the bandwidth almost equals (4.11). The first two designs discussed in Section 4.5 will use a rather safe phase margin of at least 72°. The last design will allow more overshoot in exchange for a higher bandwidth, so ζ will be more around $\sqrt{2}/2$. This overshoot can be tolerated, as the TIA is followed by a limiting amplifier, which will be discussed in Chapter 5.

To compare the performance of different TIAs, a parameter which takes into account several of its characteristics is needed. In analogy with the gain-bandwidth product GBW of voltage amplifiers [San06], the transimpedance-bandwidth product ZBW is used to evaluate the TIA performance. It is defined by the product of its transimpedance gain $Z_{TIA,0}$ and its 3-dB bandwidth BW_{TIA}:

$$ZBW = Z_{TIA,0} \cdot BW_{TIA} \approx \frac{A_0}{2\pi C_{inT}}. \tag{4.19}$$

The transimpedance-bandwidth is independent of the feedback resistor R_f. To some extent, transimpedance gain can simply be exchanged for bandwidth and vice versa, by changing the value of R_f. The only limitation is that the damping ratio has to be high enough, or equivalently, that there needs to be sufficient phase margin. (Table 4.1). The transimpedance-bandwidth ZBW (4.19) again emphasizes the importance of the voltage gain A_0. As the total input capacitance is largely determined by the photodiode capacitance, whose parameters are dictated by technology rules and fiber size, the only way to optimize ZBW is by maximizing A_0.

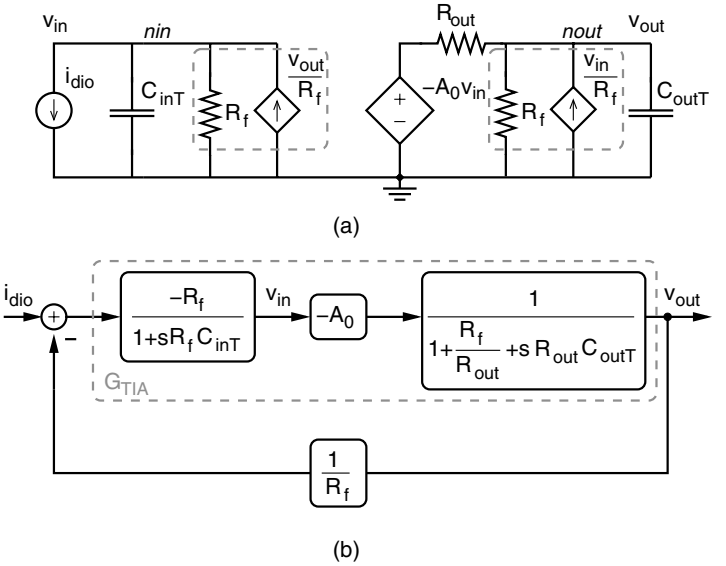


Fig. 4.2. TIA representations to determine open-loop gain and loop gain: (a) Coates transformation of Fig. 4.1(b), (b) block diagram.

4.3.2 Open-Loop Gain and Loop Gain

Another way of looking at the TIA with shunt-shunt feedback, is depicted in Fig. 4.2(a). The feedback resistor R_f between nodes *nin* and *nout* is replaced by its Coates transformation: resistor R_f and voltage-controlled current source v_{out}/R_f between node *nin* and ground; and resistor R_f and voltage-controlled current source v_{in}/R_f between node *nout* and ground. Starting from this equivalent circuit, the block diagram of Fig. 4.2(b) can be derived. It represents the feedback system by its open-loop gain G_{TIA} and loop gain GH_{TIA}.

In the block diagram of Fig. 4.2(b), the feedforward voltage controlled current source v_{in}/R_f has been neglected. This is allowed if $AR_f \gg R_{out}$, which is almost always true. The expression for the open-loop gain G_{TIA} is then given by:

$$G_{TIA} = \frac{A_0 R_f}{(1 + sR_f C_{inT})(1 + \frac{R_{out}}{R_f} + sR_{out}C_{outT})}. \tag{4.20}$$

If it also can be assumed that $R_f \gg R_{out}$, the open-loop gain further reduces to:

$$G_{TIA} = \frac{A_0 R_f}{(1 + sR_f C_{inT})(1 + sR_{out}C_{outT})}. \tag{4.21}$$

The loop gain GH_{TIA} equals the open-loop gain G_{TIA} times the feedback factor $1/R_f$, or:

$$GH_{TIA} = \frac{A_0}{(1 + sR_fC_{inT})(1 + sR_{out}C_{outT})}. \tag{4.22}$$

The phase margin is often used in basic control theory to study the stability of feedback systems [Dor98]. It is the amount of extra phase shift of the loop gain (GH_{TIA}) at unity magnitude (or 0 dB) that will result in a phase angle of 180°. To derive an expression for the phase margin, it is assumed that the poles of (4.22) are real and differ a lot in magnitude. So the dominant pole $f_{d,GH}$ and non-dominant pole $f_{nd,GH}$ of the loop gain GH_{TIA} can readily be found:

$$f_{d,GH} = \frac{1}{2\pi R_fC_{inT}}, \tag{4.23}$$

$$f_{nd,GH} = \frac{1}{2\pi R_{out}C_{outT}}. \tag{4.24}$$

Note that the latter is also the pole which determines the bandwidth of the voltage amplifier, BW_{VA} (4.12). The unity-gain frequency of the loop gain $f_{0dB,GH}$ is given by:

$$f_{0dB,GH} = \frac{A_0}{2\pi R_fC_{inT}}, \tag{4.25}$$

assuming $f_{d,GH} \ll f_{0dB,GH} \ll f_{nd,GH}$. Under the same condition, the phase angle at this frequency equals:

$$\theta(f_{0dB,GH}) = -(\frac{\pi}{2} + arctan\Big(\frac{f_{0dB,GH}}{f_{nd,GH}}\Big). \tag{4.26}$$

This results in the expression for the phase margin P.M.:

$$\text{P.M.} = \frac{\pi}{2} - arctan\Big(\frac{f_{0dB,GH}}{f_{nd,GH}}\Big) = arctan\Big(\frac{f_{nd,GH}}{f_{0dB,GH}}\Big). \tag{4.27}$$

For instance, if $f_{nd,GH} = 3 \cdot f_{0dB,GH}$, the phase margin equals 72°. Other examples for smaller ratios are given in Table 4.1. Note however that these are hand calculations and that the expression for the P.M. approaches reality only if $f_{d,GH} \ll f_{0dB,GH} \ll f_{nd,GH}$. For the other equations in Table 4.1, the requirements regarding pole location are less stringent.

The expressions derived so far are valid for the voltage amplifier model depicted in Fig. 4.1, which has only one dominant pole at the output. If the voltage amplifier consists for example of three identical amplifying stages, three more or less identical poles $f_{nd,GH}$ are present. The loop gain can now be approximated by:

$$GH_{TIA} = \frac{A_0}{(1 + sR_fC_{inT})(1 + \frac{s}{f_{nd,GH}})^3}, \tag{4.28}$$

and the phase margin:

$$\text{P.M.} = \frac{\pi}{2} - 3 \cdot arctan\Big(\frac{f_{0dB,GH}}{f_{nd,GH}}\Big), \tag{4.29}$$

with $f_{0dB,GH}$ given by (4.25). To achieve the same phase-margin, the ratio of $f_{0dB,GH}$ to $f_{nd,GH}$ has to be approximately three times smaller than the ratio in (4.27). So assuming the same $f_{0dB,GH}$, the non-dominant pole of the loop-gain, which coincides with the bandwidth of the voltage amplifier, has to be three times higher in frequency for a three-stage amplifier compared to a single-stage amplifier.

A good reason for choosing a multiple-stage approach would be to increase the voltage amplifier's gain A_0. This way, the bandwidth BW_{TIA} is enlarged, or for the same bandwidth, the feedback resistance R_f can be larger (4.11). As will become clear in Section 4.3.3, a larger R_f leads to a better TIA noise performance and thus a better receiver sensitivity. However, due to stability requirement (4.29), the voltage gain of a three-stage amplifier is limited by the ratio of the bandwidth of the TIA (BW_{TIA}) and the f_T of a certain technology. In [Ing04] it is demonstrated that, when $BW_{TIA} \approx 0.1 f_T$ or larger, the maximal achievable gain of a single stage voltage amplifier is larger than the maximal achievable gain of a three-stage voltage amplifier. Applications demanding a high bandwidth compared to the technology's f_T have to implement a single-stage amplifier rather than a three-stage amplifier, as there is simply no room for placing the extra poles which are introduced in a multiple-stage approach. However, for frequencies 20 to 30 times below f_T, a multiple-stage amplifier is the best design choice.

4.3.3 Noise

The noise sources of the shunt-shunt feedback TIA are added to the general schematic in Fig. 4.3. Two TIA noise sources can be distinguished: the noise from the feedback resistor and the noise from the voltage amplifier. The third noise source stems from the noise generated by the photodiode.

Noise Densities of the Three Noise Sources

The thermal noise of the resistor is modeled by a current noise source $\overline{di_{Rf}^2}$ in parallel:

$$\overline{di_{Rf}^2} = \frac{4kT}{R_f} df. \tag{4.30}$$

The noise current power is proportional to the absolute temperature in Kelvin, and inversely proportional to the value of the feedback resistor R_f. It is white noise, as it does not change with frequency.

The noise of the voltage amplifier is represented by the noise of an equivalent input transistor M_{x} with equivalent transconductance g_{mx}. In most designs, where the input stage features large gain, the equivalent transistor M_{x} simply corresponds to the input transistor of the voltage amplifier. For a MOS

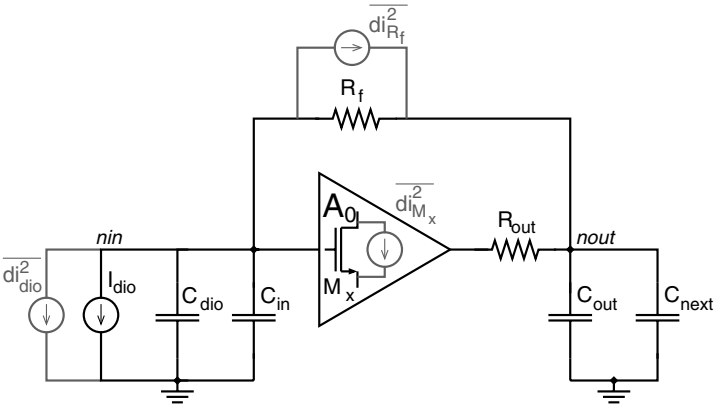


Fig. 4.3. The shunt-shunt feedback TIA with the most important noise sources.

transistor, the classical thermal channel noise $\overline{di^2_{Mx}}$ is still the most important noise source. It is given by:

$$\overline{di^2_{Mx}} = 4kT\gamma g_{mx} df, \tag{4.31}$$

where γ is the excess noise factor. This noise current power is also white and proportional to the absolute temperature and to the transconductance g_{mx}. For long-channel devices, $\gamma = 2/3$. In short-channel devices the effective temperature of the carriers is significantly larger due to the high electric field in the channel and γ can be as high as 2 or 3. The pMOS transistor usually exhibits lower γ values than its nMOS counterpart.

Finally, the last noise source is the noise from the photodiode $\overline{di^2_{dio}}$. Mostly, it consists only of shot noise, which is given by:

$$\overline{di^2_{dio}} = 2qI_{dio} df. \tag{4.32}$$

The shot noise current power is also white, and proportional to the current flowing through the junction. This noise power is data-dependent: when a photodiode receives a zero, no light power is received, and the current I_{dio} is (almost) zero. When a one is received, the generated noise depends on the received light power and the photodiode responsivity. As the noise current varies with time, it is called non-stationary. However, in most cases, the photodiode shot noise can be neglected.

Output Noise Spectral Density and Input-Referred Noise Spectral Density

The power spectral density of the output noise voltage is given by the noise current power of each noise source, multiplied by the square of its transfer function to the output.

$$\overline{dv_{n,TIA}^2} = \frac{\overline{di_{dio}^2} Z_{TIA,0}^2}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2} + \frac{\overline{di_{Rf}^2} \left| \frac{R_f A_0}{A_0+1} \right|^2 \left| 1 + s\frac{R_{out}C_{inT}}{A_0} \right|^2}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2} + \frac{\overline{di_{Mx}^2} \left| \frac{R_{out}}{A_0} \right|^2 \left| 1 + sR_f C_{inT} \right|^2}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2}. \tag{4.33}$$

This expression can be simplified using following assumptions:

- The shot noise of the photodiode is negligible small compared to the other noise sources.
- $A_0 + 1 \approx A_0$.
- The zero $\frac{A_0}{2\pi R_{out} C_{inT}}$ is located at a frequency beyond the frequency range of interest. This corresponds to neglecting the feedforward current of R_f, injected by its noise source in node *nout*.
- Following the one-transistor equivalent of the voltage amplifier, $g_{mx} R_{out} = A_0$.

The resulting expression for the output noise spectral density is:

$$\overline{dv_{n,TIA}^2} = \frac{4kTR_f df}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2} + \frac{\frac{4kT\gamma df}{g_{mx}} \left| 1 + sR_f C_{inT} \right|^2}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2}. \tag{4.34}$$

This expression is plotted versus $\log(f)$ in Fig. 4.4(a). At low frequencies, the noise is dominated by the feedback resistor's thermal noise. At higher frequencies and for a sufficient high gain A_0, the amplifier noise may become dominant. Note that the integration of the power spectral density in this case will lead to an integrated output noise voltage which is dominated by the noise of the amplifier. The representation in Fig. 4.4(a) may suggest something else, but it is misleading due to the logarithmic frequency axis which overemphasizes the integrated low-frequency noise.

In a design aiming primarily for large bandwidth, the dominant noise source might be the feedback resistor and not the amplifier. This is the case when the zero and poles of (4.34) are located close to each other, which is usually true for high BW_{TIA}/f_T ratios. As the bandwidth is comparable to

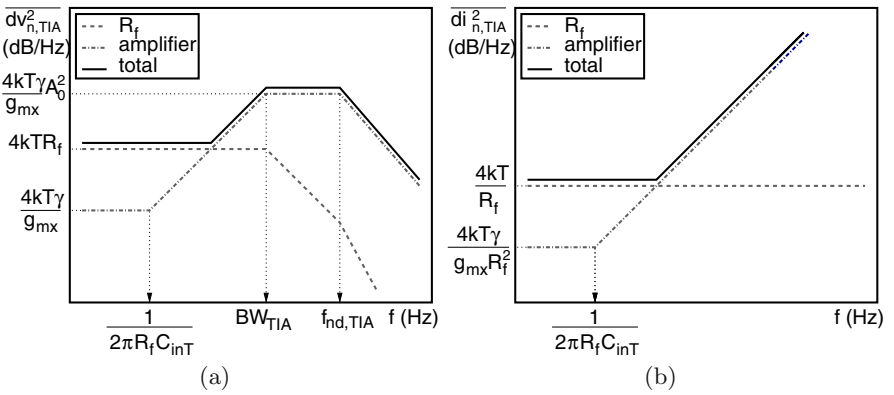


Fig. 4.4. Power spectral densities of: (a) TIA output noise, (b) TIA input-referred noise.

f_T, there is not much headroom left to place the other poles. An additional advantage is that complex conjugated poles increase the bandwidth, as discussed in Section 4.3.1. Moreover, the noise generated by the feedback resistor is quite large, as the TIA bandwidth is inversely proportional to its resistance value. This value is a few orders of magnitude smaller than in traditional designs, where a bandwidth of a few megahertz is sufficient. Finally, also the voltage gain A_0 cannot be very large as too much loop gain can cause stability problems. As a result, the maximum value for the amplifier output noise power spectral density depicted in Fig. 4.4(a) and given by:

$$\frac{4kT\gamma}{g_{mx}}A_0^2, \tag{4.35}$$

will not be reached and the actual maximum will stay below $4kTR_f$. Examples of different output noise spectral densities will be discussed in Section 4.5.

To compare the output voltage noise with the input photodiode current, the output noise is referred to the input. According to (4.2), the power spectral density of the input-referred current noise is given by the division of (4.34) by the square of (4.4).

$$\overline{di_{n,TIA}^2} = \frac{4kT}{R_f}df + \frac{4kT\gamma df}{g_{mx}R_f^2}\left|1 + sR_fC_{inT}\right|^2. \tag{4.36}$$

This expression is plotted versus frequency in Fig. 4.4(b). The first term is caused by the feedback resistor R_f. It is frequency independent and equal to the thermal noise of R_f (4.30). Neglecting the feedforward noise current, only the feedback current at the input remains which directly adds up with the photodiode current. The larger R_f, the smaller this noise, but also the smaller the TIA bandwidth BW_{TIA} (4.11). The second term is the noise contribution

of the voltage amplifier. It features a zero at $1/2\pi R_f C_{inT}$, which is a factor A_0 below the bandwidth BW_{TIA} (4.11). Due to this zero, the amplifier's noise may become dominant at higher frequencies. The noise contribution of both the feedback resistor and the amplifier result in an input-referred noise spectrum which is flat at low frequencies, and rises with f^2 at higher frequencies.

Noise Optimization

As the noise of the TIA directly determines the receiver's sensitivity, the TIA design plan must include a minimization of noise. Since the TIA is a broadband amplifier, not only the power spectral density of the noise is important, but the integrated noise is even more important. Different noise analyses can be found in literature which are based on slightly different assumptions. However they have two approximations in common: first, the noise of the light detector is neglected and second, the noise of the post-amplifier is neglected. In Section 4.5 some design examples will show that the first assumption is not always valid. The latter assumption it true when the transimpedance gain is always sufficiently high, but in reality this is not always the case. However, since an increased transimpedance gain goes hand in hand with a reduced TIA noise, neglecting the noise of the post-amplifier rarely affects the design optimum. An overview of the most important results is given here.

A thorough noise analysis of wide-band amplifiers in bipolar and CMOS technologies is given in [Cha91]. An expression is derived for the equivalent input noise density of a wide-band amplifier with capacitive source, which belongs to the same amplifier category as the transimpedance amplifier. The noise reaches a minimum value when the input capacitance of the voltage amplifier equals all other capacitances at the input node. Assuming that the thermal noise of the input transistor is dominant, the optimal input transistor width is calculated.

A more specific noise analysis of a shunt-shunt feedback transimpedance amplifier is presented in [Ing04]. The output noise power density is calculated, but in contrast to (4.34), two separate poles, a dominant pole (4.11) and non-dominant pole (4.12) are assumed in the denominator. The equivalent input-referred noise current (4.36) is minimized, taken into account following assumptions:

- $g_m R_f > 1$
- C_{in} of the amplifier is mainly determined by the gate-source capacitance C_{gs} of the input transistor. C_{gd} and the Miller effect on this capacitance, are neglected.
- The excess noise factor γ equals 2/3.

The optimal gate-source capacitance C_{gs}, corresponding to the minimal input-referred current noise, is found to be equal to the diode capacitance: $C_{gs} = C_{dio}$. This is the same result as derived by [Cha91].

However, this result is not unconditional, but based on the assumption that the feedback resistor R_f is independent of the input capacitors. In high-speed receivers, the maximal feedback resistor is limited and its noise contribution becomes relatively more important compared to that of the amplifier. As the receiver's bandwidth is limited by both C_{gs} and R_f, there is inevitably some trade-off between them, also for the noise performance. The optimal ratio X_N between C_{gs} and C_{dio} is derived in [Ing04] based on a maximization of the signal to noise ratio at the output. The noise power at the output is obtained by integration of the output noise power spectrum (4.34), represented in Fig. 4.4(a). The signal power at the output depends on the spectrum of the data signal and its data coding scheme. In [Ing04], a flat signal spectrum is assumed. The optimal ratio corresponding with the best signal to noise ratio is than given by:

$$X_{N,opt} = \frac{C_{gs,opt}}{C_{dio}} \approx \frac{1}{\sqrt{1 + \frac{9}{4}\frac{\mu(V_{gs}-V_T)}{A_0 L^2 F_{BW} BW_{TIA}}}}, \tag{4.37}$$

with factor F_{BW} defined as:

$$F_{BW} = \frac{\pi}{2}\frac{BW_{VA}}{BW_{TIA}} - 1. \tag{4.38}$$

For typical values of the various parameters this ratio lies between 0.5 and 1. The optimal C_{gs} is thus always smaller than C_{dio}. In [Ing04], the signal to noise ratio of a TIA with a single stage voltage amplifier in a 0.7 μm CMOS technology is plotted versus X_N. The SNR-curve is relatively flat, which means that a deviation from the optimal X_N only results in a slightly lower signal to noise ratio.

In [Ler04] the bandwidth of the TIA is considered as a fixed design constraint for a given application. As in [Ing04], the noise optimum is found by integrating both signal and noise at the output of the TIA and maximizing the SNR. The aim of the optimization is to find the optimal C_{gs} of the input transistor for a given maximum voltage gain A_0 and unity gain frequency f_T. The power spectral density of the NRZ signal S_{out} is now more accurately described by:

$$S_{out} = i_{dio}^2 T_b \left(\frac{sin(\pi f T_b)}{\pi f T_b}\right)^2 \left|\frac{R_f}{(1 + s\frac{R_f C_{inT}}{A_0})(1 + sR_{out}C_{out})}\right|^2, \tag{4.39}$$

where i_{dio} is the amplitude of the diode input current and T_b is the bit period. So again two separated poles are assumed for the transimpedance gain, and its DC value is approximated by R_f. The dominant pole of (4.39), BW_{TIA} (4.11), is set by the bitrate of the application. It is also (approximately) equal to the unity-gain frequency of the loop gain. It can further also be assumed that the non-dominant pole, given by (4.12), is fixed to a minimum value enforced by the stability requirement. Under these assumptions, the output signal power

is solely depending on the DC transimpedance gain R_f. To maximize the SNR it suffices to minimize the input-referred integrated noise power $i^2_{n,TIA}$:

$$i^2_{n,TIA} = \frac{\int_0^\infty \overline{dv^2_{n,TIA}} df}{R_f^2}. \tag{4.40}$$

So the same integration of the output noise power spectral density is performed as in [Ing04]. The major difference is that now C_{gd} and the Miller effect are taken into account:

$$C_{in} = C_{gs}(1 + M_i \alpha_{gd}), \tag{4.41}$$

with:

$$\alpha_{gd} = \frac{C_{gd}}{C_{gs}}. \tag{4.42}$$

and $M_i = A_0 + 1$. The transconductance g_m is rewritten as:

$$g_m = 2\pi f_T C_{gs} = \frac{2\pi f_T C_{in}}{1 + M_i \alpha_{gd}}. \tag{4.43}$$

The expressions found in [Ler04] demonstrate that the input referred noise increases with the second power of the bandwidth for the R_f contribution and even to the third power for the amplifier contribution. This clearly shows that designing sensitive optical receivers becomes increasingly difficult for larger bitrates. The optimal ratio $X_{N,opt}$ is now given by:

$$X_{N,opt} \approx \frac{1}{\sqrt{(1 + M_i \alpha_{gd})^2 + (1 + M_i \alpha_{gd}) \frac{3\pi}{4} \frac{f_T}{BW_{TIA}} \frac{1}{\gamma A_0 F_{BW}}}}. \tag{4.44}$$

By including the Miller effect, the optimal C_{gs} is even smaller than the optimum (4.37) found in [Ing04]. A numerical example in [Ler04] illustrates that C_{gs} can be as small as $\approx 0.23 \cdot C_{dio}$ (for $\alpha_{gd} = 0.25$, $f_T/BW_{TIA} = 20$, $\gamma = 1$, $M_i = 11$, $F_{BW} = 4$). Note also that according to (4.44), the ratio $X_{N,opt}$ will decrease for newer technologies with increasing f_T and the same bandwidth BW_{TIA}.

4.4 Literature Examples

The main question after studying the high-level design of a TIA is how to realize the voltage amplifier depicted in Fig. 4.1. In literature, two important approaches can be found: the TIA with common source input stage and the TIA with regulated cascode input stage. Both are discussed next, with some paper references as example. Also a survey is given of the latest novelties regarding TIA design presented at the International Solid-State Circuits Conferences (ISSCC).

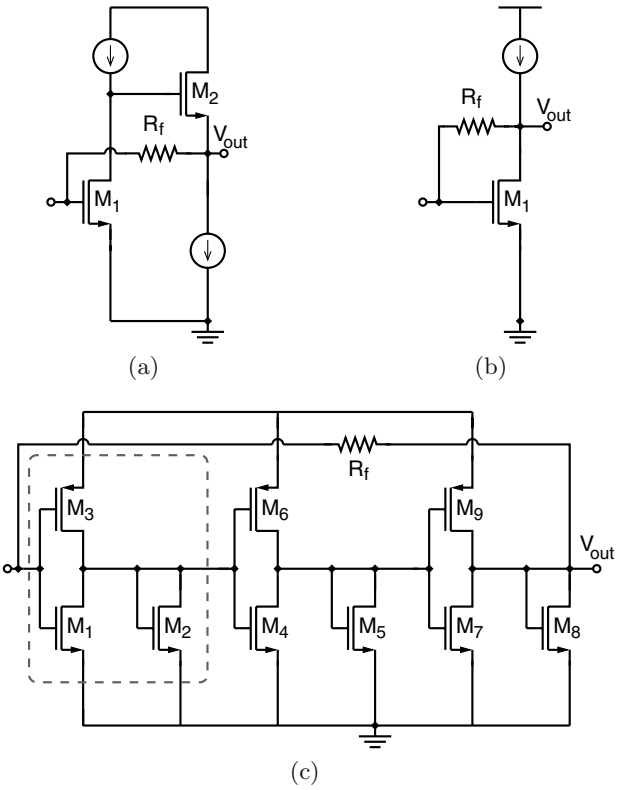


Fig. 4.5. Schematic of (a) a CS TIA with source follower [Kie03, Swo05], (b) a CS TIA [Tsa04, Rad05, Tsa06], (c) a three-stage TIA with g_m/g_m amplifying stage [Ing94].

4.4.1 Common Source TIA

The most widely used topology, implemented in many variants and technologies, is the TIA with common source input stage, depicted in Fig. 4.5(a). Transistor M_1 is the common source transistor which realizes amplification. The source follower (M_2) isolates the drain of M_1 from the loading effect of both R_f and the input capacitance of the subsequent stage. In addition, the output resistance of the source follower is only $1/g_{m,M2}$, which is much less than R_f. The voltage gain A_0 equals $g_{m,M1}/g_{ds,M1}$, assuming the output conductance of the current source can be neglected. This circuit has been used for instance in [Kie03] and [Swo05].

The source follower increases the total loop gain, but consumes a large voltage headroom. One may consider to eliminate this stage, arriving at the circuit depicted in Fig. 4.5(b). In a typical design, the output resistance determined by the parallel combination of the output conductance $g_{ds,M1}$ and the current

source (which might be a simple resistor) is not small anymore compared to the feedback resistance R_f. This topology has been used in [Tsa04, Rad05, Tsa06].

A natural way to implement a common-source amplifier stage in CMOS is adding a pMOS load and connecting the gates, which results in an inverter stage. This way, the transconductance of the input stage is doubled, which leads to a higher voltage gain A_0. As demonstrated in [Ing94], the stability requirement for having sufficient phase margin (4.27) in the traditional common source approach is related to $g_{ds,M1}^2$, which is strongly process-dependent. Therefore, the inverter stage presented in [Ing94] features a diode-connected nMOS load transistor, so that the output resistance is now determined by $1/g_{m,M2}$. The voltage gain A_0 is given by the ratio of transconductances:

$$A_0 = \frac{g_{m,M1} + g_{m,M3}}{g_{m,M2}}. \tag{4.45}$$

The analysis in [Ing94] shows that this ratio is only process dependent through the square root of the ratio of mobilities.

The TIA presented in [Ing94] and shown in Fig. 4.5(c) has three identical stages to create a larger voltage gain A_0. Each of these stages consumes 2 mA from a 5 V supply. A transimpedance gain of 150 kΩ is realized, combined with a bitrate of 240 Mbit/s. This results in a transimpedance-bandwidth of 18 THzΩ. When moving to higher BW_{TIA}/f_T ratios, the intrinsic technology speed limits the placement of the non-dominant poles at higher frequencies, which is needed to maintain stability (Section 4.3.2). Therefore, the TIA presented in [Ing99] is based on a single-stage g_m/g_m voltage amplifier, featuring a high speed combined with an accurate gain. The TIA consumes approximately 5 mA from a single 5 V power supply, and features a transimpedance gain of 1 kΩ. It is realized together with a postamplifier based on a biased inverter chain and achieves bitrates up to 1 Gbit/s. The optical receiver is characterized electrically by replacing the photodiode by its Thevenin equivalent. A large resistor is inserted after the 50 Ω signal source and the photodiode capacitance is modeled by a 500 fF capacitor.

4.4.2 Regulated Cascode TIA

The regulated cascode (RGC) TIA is actually an extension of the common gate (CG) TIA. The latter topology is depicted in Fig. 4.6(a), where the photodiode is connected to the source of the common gate input transistor M_1 [Par97]. This way, the photodiode capacitance sees only a small input resistance determined by $1/g_{m,M1}$. The feedback resistor R_f is inserted between the output (neglecting the source follower M_4) and the drain of M_1. The dominant pole of this feedback amplifier depends on the input capacitance of M_2, the gate and drain capacitance of M_1 and the feedback resistor R_f. In contrast to the common source TIA, the bandwidth is isolated from the photodiode capacitance at the input, which is the major advantage of this topology.

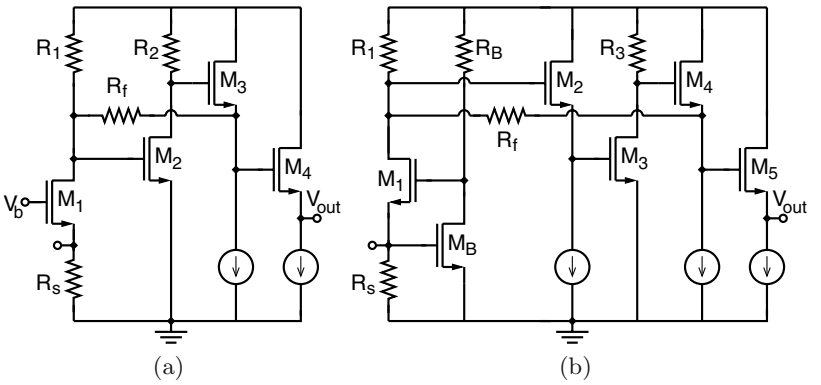


Fig. 4.6. Schematic of (a) a CG TIA [Par97], (b) a RGC TIA [Par00, Par04].

The RGC TIA is shown in Fig. 4.6(b) and presented in [Par00, Par04] . The small-signal input resistance at the source of M_1 is given by:

$$Z_{in,0} = \frac{1}{g_{m,M1}(1 + g_{m,MB}R_B)}, \tag{4.46}$$

where $(1 + g_{m,MB}R_B)$ is the voltage gain of the local feedback stage. One could say that the RGC input stage behaves qualitatively as a CG input stage with a large transconductance G_m of $g_{m,M1}(1 + g_{m,MB}R_B)$. It enables $(1 + g_{m,MB}R_B)$ times better isolation of the input parasitic capacitance from the bandwidth determination. A source-follower buffer stage M_2 is introduced between the RGC input stage and the voltage gain stage $M_3 - R_3$. This follower reduces the total capacitance looking from the high impedance node at the drain of M_1 and thus provides a potential for wider bandwidth. However, it may reduce the total open-loop gain and degrades the circuit linearity.

The measurement results in [Par00] demonstrate a 3-dB bandwidth of 300 MHz and a transimpedance gain of almost 58 dBΩ with a 6 V power supply. The resulting trans-impedance-bandwidth product equals 238 GHzΩ. However, these values are calculated from S-parameter measurements with a network analyzer, where a test power level of −40 dBm emulates a 50 μA current. As a photodiode is a current source with a high output impedance in contrast with the used voltage source with low output impedance (50 Ω), these values should be treated with care.

To facilitate the measurements, the photodiode is modeled as an equivalent electrical circuit in [Par04] and mounted on a test board together with the TIA chip. The frequency response is measured with a network analyzer, achieving a transimpedance gain of 58 dBΩ and 3-dB bandwidth of 950 MHz for a 0.5 pF photodiode capacitance. As a consequence, the transimpedance-bandwidth product equals 755 GHzΩ. Also eye diagrams are presented at 622 Mbit/s, 1.0625 Gbit/s, 1.25 Gbit/s and 1.866 Gbit/s for a $2^{31} - 1$ prbs

data with 125 μA equivalent input current. The chip core dissipates 85 mW from a 5 V supply. No optical measurements are performed, only electrical measurements with a low-impedance voltage source and the electrical equivalent model of the photodiode. For 1.25 Gbit/s operation, the electrical sensitivity is measured to be 5 μA_{pp} for a BER of 10^{-12}.

Although the RGC and CG input stage both have the great advantage that the TIA bandwidth is nearly independent of the photodiode capacitance, there is also a price to pay. The extra transistors and resistors necessary to obtain a low input impedance also present extra noise sources. For instance the resistance current noises of R_S on one hand and the parallel combination of R_1 and R_f on the other hand directly add up with the input signal. In the common source approach, the only thermal resistance noise which appears directly at the input is the noise of R_f. A detailed noise analysis of the RGC TIA can be found in [Par04]. In [San06], a simplified comparison of the noise performance between the TIA with voltage amplifier (common source topology) and the TIA with current amplifier (cascode topology) is made. Both comparisons based on noise densities and based on the integrated noise arrive at the same conclusion: the noise of the TIA with voltage amplifier is always smaller, as long as R_f is large enough. As this is usually the case, the regulated cascode topology is not adopted in this work.

4.4.3 The Latest Trends at ISSCC

To conclude this literature section, an overview of some design trends of the past years at ISSCC is given. In [Sei04a, Sei04b], the problem is addressed when the TIA bandwidth is limited by the parasitic capacitance of a large polysilicon resistor. A conventional 200 kΩ polysilicon feedback resistor has for instance a 3-dB cut-off frequency of 67 MHz and therefore this frequency can be the dominant pole of the circuit. To solve this problem, a capacitive-coupled voltage divider is proposed as feedback network instead of a simple feedback resistor. It consists of a low- and high-frequency path of voltage dividers, connected by a coupling capacitor. A pin photodiode is integrated with the TIA in a BiCMOS technology and has a responsivity of 0.43 A/W for 660 nm light. The bandwidth is measured with a network analyzer modulating a laser and equals 378 MHz. Together with a transimpedance gain of 178 kΩ, this results in a transimpedance-bandwidth product of 67 THzΩ. The power consumption equals 70.5 mW from a 5 V power supply, of which the output driver dissipates 38 mW.

One year later, [Tsa05b] introduces a self-compensated differential SiGe TIA which is tolerant of large capacitances at the input. The effective input capacitance caused by the photodiode is significantly reduced by adding a unity-gain buffer. The voltage at the cathode of the photodiode ideally tracks the RF voltage at the anode of the photodiode. Hence the transient voltage signal across the photodiode is reduced during operation to suppress the effect of its capacitance on the receiver bandwidth. This idea has already been

presented for a single-ended topology in [Tsa04, Tsa05a]. However, a unique feature of the proposed self-compensated differential topology is that it not only significantly suppresses the impact of the photodiode capacitance on the receiver performance, but it also reduces the impact of other capacitances connected from the inputs to ground. When the photodiode is applied externally, these capacitances are due to the bondpads as well as the ESD protection circuits. Optical measurements are performed with the TIA IC integrated with a commercial 1310 nm InGaAs pin photodiode in a chip-on-board assembly. Its parasitic capacitance is around 0.7 pF. A measured 2.5 Gbit/s eye diagram with an optical power of −20 dBm at the input of the TIA IC with ESD protection circuits is shown. The bandwidth of the TIA with ESD equals 1.16 GHz, while the transimpedance gain is adjustable from 1 kΩ to 15 kΩ. No information is provided for which gain setting a 1.16 GHz bandwidth has been measured, so the transimpedance-bandwidth product is situated between 1 THzΩ and 17 THzΩ. The TIA core consumes 4.5 mW from a 3 V supply. Note that this technique is well suited for compensating off-chip capacitances, but difficult to implement with an integrated CMOS photodiode. For a classical or differential n-well photodiode (Chapter 3) the anode corresponds to the p-substrate which inevitably has to be biased at the ground potential. Connecting the anode (p-substrate) to the input of the amplifier and the cathode (n-well) to the output of the unity-gain follower is not possible in this case.

Finally, a compensation technique at the output of the TIA is presented in [Tsa06]. The basic concept is to cancel the loading effects caused by both the output impedance of the amplifier and the feedback resistor, by introducing a compensation element with a negative impedance. Hence the loop gain can be significantly boosted. As a result, the input impedance is reduced and the operating bandwidth is extended. The TIA IC is implemented in a standard 0.35 μm CMOS technology. The differential transimpedance gain is adjustable from 500 Ω to 13 kΩ. The TIA core dissipates 15 mW from a 3 V supply. Optical measurements demonstrate a bandwidth enhancement factor of 3 achieved by the active compensation technique. Eye diagrams are shown at 1.25 Gbit/s and an optical input power of −27 dBm (high-gain mode) and 0 dBm (low-gain mode). No measured values of the bandwidth are reported (except for the enhancement factor of 3), but assuming a bandwidth of at least 625 MHz, the transimpedance-bandwidth product is larger than 8 THzΩ.

To conclude, the transimpedance-bandwidth product of the latest state-of-the-art TIAs equals several THzΩ's. Furthermore, a trend can be observed that it is easier to achieve high transimpedance-bandwidth products by amplifiers with a high gain, but a lower bandwidth, than by high-speed amplifiers.

4.5 Case Studies

This section discusses in depth three different implementations for the voltage amplifier depicted in Fig. 4.1. They all share the common source input stage

topology. As the ultimate goal is to design a high-speed front-end, a single-stage approach is applied in the first two implementations. The last design is fully differential, which enables a two-stage implementation with cross-coupled feedback resistors. The measurement results of these circuits, embedded in a receiver front-end, will be discussed in Chapter 6.

4.5.1 An Inverter-Based TIA for Test Photodiodes in 0.18 μm CMOS

Design Goals and Implementation

As this TIA is part of a test-chip with several photodiode structures, it must be able to generate a stable output voltage, despite the different diode capacitance values at the input. Also, because it is the intention to compare the performance of the diode structures, the TIA should not be the speed limiting factor. Due to the non-optimized responsivity of CMOS photodiodes (see Chapter 3), the equivalent input-referred noise current should be as small as possible.

The test photodiodes are: a classical n-well diode, a quasi-fractal n-well diode, a differential n-well diode, a p^+ n-well diode with guard and an n^+ p-substrate diode. The layout of these structures will be discussed in detail in Section 6.2. Their junction capacitances are respectively 660 fF, 585 fF, 292 fF, 4.6 pF and 6.8 pF.

The schematic of the TIA, with different diode configurations is depicted in Fig. 4.7(a) and Fig. 4.7(b). The n-region of the classical n-well diode, quasi-fractal n-well diode, differential n-well diode and n^+ p-substrate diode respectively is connected to the input, while the p-substrate is connected to ground (Fig. 4.7(a)). The p^+ n-well diode with guard topology consists of 2 diodes: the p^+ n-well diode (PD) to detect the signal, and the n-well p-substrate diode (guard) to remove the diffusing substrate carriers (Fig. 4.7(b)). The p^+region is connected to the input of the TIA, while the n-well region is biased at a higher voltage V_{b3}. In a 0.18 μm CMOS technology, this voltage should be limited to 1.8 V. A higher biasing voltage would generate a larger depletion region and a smaller junction capacitance, but is not applied for reliability reasons. The p-substrate is connected to ground.

The TIA consists of a voltage amplifier with inverter topology (nMOS transistor M_1 and pMOS transistor M_2) and a variable feedback resistance (resistor R_{fb} and pMOS transistor M_{Rfb}). The diode-connected load used for instance in [Ing94, Ing99, Roo00] is omitted. The advantage is that the gain A_0 is much higher, the disadvantage is that R_{out} is now completely determined by the output conductances of transistors M_1 and M_2. These are susceptible to process variations, which can lead to stability problems. To avoid unstabilities during measurements, transistor M_3 is added as load. It is biased in its linear region and can be switched on or off by changing the bias voltage V_{b2}. Dummy stage M_{1d}-M_{2d} is added to create the same biasing conditions for transistor

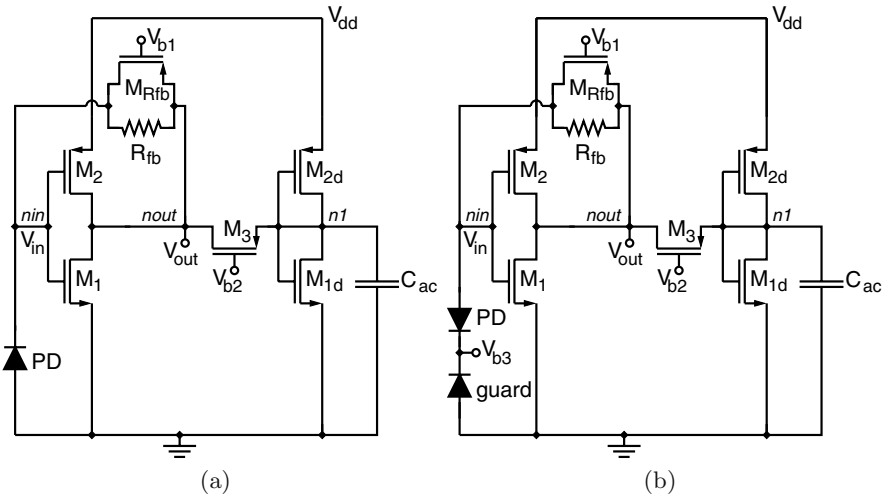


Fig. 4.7. Schematic of the inverter-based 0.18 µm CMOS TIA: (a) configuration for the classical n-well diode, quasi-fractal n-well diode, one half of the differential n-well diode and n^+ p-substrate diode, (b) configuration for the p^+ n-well diode with guard.

Table 4.2. Design parameters of the inverter amplifier and dummy stage.

	M_1	M_2	M_{1d}	M_{2d}
type	nMOS	pMOS	nMOS	pMOS
L	0.18 µm	0.18 µm	0.18 µm	0.18 µm
W	40 µm	60 µm	4 µm	6 µm
V_{DSAT}	0.24 V	−0.46 V	0.24 V	−0.46 V
I_{DS}	3.3 mA	−3.3 mA	0.33 mA	−0.33 mA
g_m	15.7 mS	10.3 mS	1.57 mS	1.03 mS
g_{ds}	0.8 mS	0.7 mS	0.08 mS	0.07 mS
C_{gs}	47 fF	75 fF	4.7 fF	7.5 fF
C_{gd}	15 fF	21 fF	1.5 fF	2.1 fF

M_3 as for transistor M_{rfb}: equal drain and source voltages. It is a replica of inverter M_1-M_2, only the widths of the transistors are 10 times smaller to limit the power dissipation. Capacitor C_{ac} is added to remove ac signals from node *n1*. The design parameters of the inverter amplifier and dummy stage are summarized in Table 4.2. The hand calculations and simulation results for different biasing conditions and different photodiode topologies can be found in Table 4.3 and Table 4.4.

DC Operating Point

As no DC current flows through R_f, M_{Rfb} or M_3, following equation should be fulfilled:

$$I_{DS,M1} = |I_{DS,M2}|. \tag{4.47}$$

Using the current expressions for strong inversion [San06], this equation becomes:

$$\begin{aligned} &K'_n\left(\frac{W}{L}\right)_{M1}(V_{GS} - V_T)^2_{M1}(1 + \lambda_n V_{DS})_{M1} = \\ &K'_p\left(\frac{W}{L}\right)_{M2}(|V_{GS}| - |V_T|)^2_{M2}(1 + \lambda_p|V_{DS}|)_{M2}. \end{aligned} \tag{4.48}$$

W is the transistor width, L is the transistor length, V_T is the threshold voltage, V_{GS} the gate-source voltage, K'_n and K'_p are the transconductance parameters for an nMOS and pMOS transistor respectively, λ_n and λ_p are the channel length modulation parameters for an nMOS and pMOS transistor respectively. As a high bandwidth is required, the length of the transistors is chosen minimal (0.18 µm). Rearranging the terms in (4.48) shows that the DC input voltage V_{IN}, which is the same as the DC output voltage V_{OUT}, is completely determined by the ratio of the width of the transistors:

$$\frac{W_{M1}}{W_{M2}} = \frac{K'_p(V_{DD} - V_{IN} - |V_{Tp}|)^2(1 + \lambda_p(V_{DD} - V_{IN}))}{K'_n(V_{IN} - V_{Tn})^2(1 + \lambda_n V_{IN})}. \tag{4.49}$$

As V_{IN} also determines the reverse bias voltage of the photodiode, it should be large to widen the depletion region and to reduce the photodiode junction capacitance. However, the width only changes with the square root of the voltage (3.22). When no ac coupling is used, the output voltage V_{OUT} equals the input voltage of the next stage. Moreover, this voltage determines the dynamic range for the TIA. A signal current from the photodiode flows through R_f and increases the output voltage. Very large signals would steer transistor M_2 into its linear region, resulting in a corrupted eye diagram. As the currents generated by the integrated diodes are very small, large output voltages almost never occur. In the presented design, the ratio W_{M1}/W_{M2} equals 2/3 (Table 4.2), resulting in a simulated $V_{IN} = V_{OUT} = 0.8$ V.

Transimpedance Gain and Bandwidth

To increase the flexibility of the TIA, a variable transimpedance gain is implemented. This is realized with a fixed resistor R_{fb}, in parallel with a pMOS transistor M_{Rfb} biased in its linear region. The advantage of using a fixed resistor is that high-ohmic polysilicon can be used, with a value of approximately 1000 Ω/square. As a result, only 2 squares are needed to realize a 2 kΩ resistor, resulting in a compact layout with small parasitic capacitance values. The 4 µm x 8 µm resistor has a 3-dB bandwidth of 28 GHz. Transistor M_{Rfb}

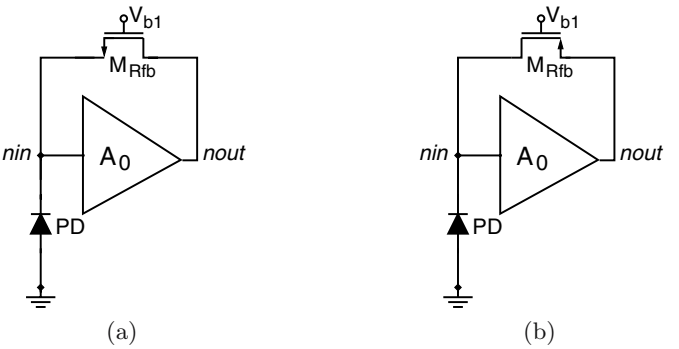


Fig. 4.8. Implementation of the feedback resistor: (a) nMOS-type, (b) pMOS-type.

is turned on when the signal is strong enough to lower the gain, resulting in a higher bandwidth. The overall feedback value R_f is given by:

$$R_f = R_{fb}//R_{MRfb}, \tag{4.50}$$

with:

$$R_{MRfb} = \frac{1}{\mu C_{ox}\left(\frac{W}{L}\right)_{MRfb}(|V_{GS}| - |V_T| - |V_{DS}|)_{MRfb}}. \tag{4.51}$$

A first design choice regarding the feedback transistor is the type of MOS transistor. Fig. 4.8 shows two possible implementations: nMOS-type and pMOS-type. For the nMOS, the source is connected to the input, the drain to the output, the gate to a bias voltage V_{b1} (for instance V_{dd}) and the bulk to ground. When the photodiode generates an input current, the output voltage rises. As a result, V_{gs} remains nearly constant, while V_{ds} increases. For large signals, the transistor may go into saturation, which is undesirable. The pMOS on the other hand has its drain connected to the input and its source connected to the output. The bias voltage at the gate V_{b1} is now for instance ground. When a signal is applied, the source voltage of the transistor increases. So any change in V_{ds} is reflected in an equal change in V_{gs} and the transistor will stay in the linear region, even for large signals. Therefore, a pMOS-type implementation is chosen.

Second, the bulk connection of the pMOS has to be set. As a pMOS transistor is realized in an n-well, the bulk can be connected either to the source or to V_{dd}. In [Ing04], the principles of dynamic signal compression when the bulk is connected to V_{dd} are explained using a diode-coupled ac model. However, due to the bulk-effect, V_T increases, resulting in a larger R_{MRfb} for the same W/L ratio. In this design, the goal is to lower the resistance value of R_{fb} by switching on transistor M_{Rfb} in parallel. To cover a wide range of possible values (by changing V_{b1}), a low starting value when $V_{b1} = 0$ V is

preferred. When the bulk is connected to V_{dd}, a much larger transistor width W is needed to create the same resistance value, leading to higher capacitive parasitics and a lower TIA bandwidth. Moreover, the advantages of using dynamic signal compression are not so large as the input signal is usually very small and thus no compression is needed. Therefore, in this design, the bulk is simply connected to the source. The length is chosen minimal to minimize parasitic capacitances. The width equals 10 μm, which results in a simulated R_f that ranges from 589 Ω ($V_{b1} = 0$ V) to 2000 Ω ($V_{b1} > 0.6$ V).

The output resistance of the voltage amplifier is determined by the output conductances of transistors M_1 and M_2, and the resistance value of M_3. Just like M_{Rfb}, this pMOS transistor is biased in its linear region, and the resistance value R_{M3} is given by a similar expression as (4.51). So the output resistance equals:

$$R_{out} = \frac{1}{g_{ds,M1}} // \frac{1}{g_{ds,M2}} // R_{M3}. \tag{4.52}$$

The voltage gain A_0 is determined by:

$$A_0 = (g_{m,M1} + g_{m,M2}) R_{out}. \tag{4.53}$$

When transistor M_3 is switched off, the output resistance is high, leading to a high gain, which is a nice goal in TIA design (Section 4.3.1). However, generally low output resistances are pursued to avoid output loading. In this design we will have to take the value of R_{out} into account, as it has the same order of magnitude as R_f. Also, because the output conductance of a transistor is susceptible to process variations, the actual value of R_{out} and A_0 might differ. Because the resistance value of a transistor in the linear region shows less deviations after processing, transistor M_3 can be added. However, owing to the lower gain A_0, the TIA performance will be worse. For the transistor dimensions in Table 4.2 and $W_{M3} = 20$ μm, $L_{M3} = 0.18$ μm, the output resistance varies between 644 Ω ($V_{b2} > 0.6$ V) and 253 Ω ($V_{b2} = 0$ V)

Table 4.3 summarizes the results of the main equations derived in Section 4.3 under four extreme biasing conditions. The input photodiode is a classical n-well diode. The transimpedance gain is slightly smaller than R_f due to R_{out}. The maximum value equals 65 dBΩ, while the minimum value equals 54 dBΩ. For the bandwidth, (4.11) gives a too optimistic result. Summing up R_{out} with R_f to calculate the dominant pole, which is justified by (4.4), gives a result which corresponds better with the simulation results. Dependent on the biasing conditions, the bandwidth ranges from 440 MHz up to 1.9 GHz. As expected, a smaller R_f results in a smaller transimpedance gain but a higher bandwidth. Lowering R_{out} leads to a smaller voltage gain and consequently also a smaller transimpedance bandwidth. The transimpedance-bandwidth product is the largest when $V_{b1} = V_{b2} = 1.8$ V and equals 1.1 THzΩ.

Table 4.4 compares the results for different diode topologies under a fixed biasing condition (M_{Rfb} and M_3 both off). The different photodiode junction capacitances cause the bandwidth to vary from 166 MHz up to 910 MHz. The

transimpedance-bandwidth product is the largest (1.6 THzΩ) for the topology with the smallest photodiode capacitance, which is the differential diode.

Open-Loop and Loop Gain

As the assumption $R_{out} \ll R_f$ is not valid for the inverter topology, the equations regarding open-loop and loop gain are not very accurate. However, as can be seen in Table 4.3 and Table 4.4, the simulated unity-gain frequency of the loop gain $f_{0dB,GH}$ equals more or less the TIA bandwidth BW_{TIA}. The DC loop gain is always smaller than A_0 and is more accurately modeled by:

$$GH_{TIA,0} = \frac{A_0}{1 + \frac{R_{out}}{R_f}}. \tag{4.54}$$

Table 4.3 shows that the phase margin is the smallest (99°) when both M_{Rfb} and M_3 are off. Switching on M_{Rfb} lowers R_f and results in a higher $f_{0dB,GH}$, but also a higher dominant pole $f_{d,GH}$ (4.23). Because the ratio R_{out}/R_f becomes larger, the loop gain $GH_{TIA,0}$ drops. The result is a slight increase in phase margin. M_3 is added to compensate for uncertainties in the output conductances of the transistors, which might create an unstable TIA. The simulation results in Table 4.3 show that switching on this transistor increases the phase margin, so stability is ensured. Because R_{out} becomes smaller, $f_{0dB,GH}$ drops while the dominant pole of the loop gain $f_{d,GH}$ remains nearly constant and the non-dominant pole $f_{nd,GH}$ increases in frequency. The result is that the ratio $f_{nd,GH}/f_{0dB,GH}$ becomes larger, thus creating a better phase margin (4.27).

The simulation results in Table 4.4 show that the input photodiodes with the largest parasitic capacitance cause the smallest phase margin. The p^+ n-well diode with guard has a P.M. of 71°. Switching on M_{Rfb} or M_3 will increase this value.

Noise

Fig. 4.9 shows the photodiode model which is used for the noise simulations. It consists of the junction capacitance C_{dio} in parallel with a very large resistor R_p (≈ 1 TΩ) which models the dark current of the photodiode. When the diode is biased forward, $1/R_p$ equals the ac conductance. As stated in the Eldo Device Equations Manual [Men], resistor R_s is added in series 'to include non-idealities'. The value is determined by the parasitic resistance of the anode (for instance the p-substrate), cathode (for instance the n-well), and contacts. It is inversely proportional to the area of the diode: the larger the photodiode area, the smaller the series resistance. While resistor R_p is noiseless, resistor R_s generates thermal noise, given by:

$$\overline{di_{Rs}^2} = \frac{4kT}{R_s} df. \tag{4.55}$$

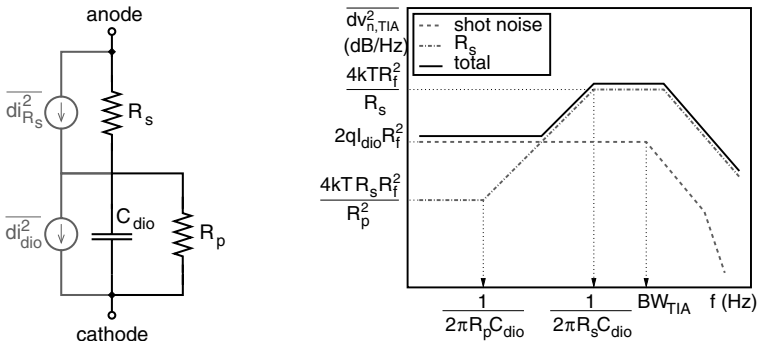


Fig. 4.9. Noise of the photodiode: (a) Eldo diode model [Men], (b) power spectral density of the diode noise at the TIA output.

The noise of the junctions itself is represented by the shot noise current source $\overline{di^2_{dio}}$ (4.32). The spectral density of the noise at the output of the TIA due to the photodiode is given by:

$$\overline{dv^2_{n,TIA}} = \overline{di^2_{dio}} \frac{Z^2_{TIA,0}}{\left| s^2 \frac{R_f R_{out} C_{inT} C_{outT}}{A_0+1} + s\left(\frac{(R_f+R_{out})C_{inT}}{A_0+1} + \frac{R_{out}C_{outT}}{A_0+1} \right) + 1 \right|^2} + \overline{di^2_{Rs}} \frac{\left| \frac{R_s R_f}{R_p} \right|^2 \left| 1 + sR_pC_{dio} \right|^2}{\left| \frac{R_s(R_{out}+R_f)C_{in}C_{dio}}{A_0+1} s^2 + \frac{(RC)_{noise}}{A_0+1} s + 1 \right|^2}. \quad (4.56)$$

with:

$$(RC)_{noise} = (R_f + R_{out} + R_s(A_0+1))C_{dio} + (R_f + R_{out})C_{in}. \quad (4.57)$$

This equation is valid under following conditions:

- The parallel resistance R_p is very large compared to all other resistance values.
- The voltage gain A_0 is large.
- The output capacitance C_{outT} has been omitted for the transfer function of the R_s noise, to keep the equations as simple as possible.
- R_s has been assumed sufficiently low to neglect its effect on the transimpedance gain (4.4) and consequently also on the transfer function of the diode shot noise.

Notice that only the noise at the output of the TIA, originating from the photodiode is modeled here. The influence of the other noise sources is described in Section 4.3.3 and given by (4.34).

The output power spectral density of the noise due to shot noise and R_s noise of the photodiode is depicted in Fig 4.9(b). The DC values of these noise sources are normally much smaller than the DC value of (4.34) and thus not important at low frequencies. At high frequencies (around the TIA bandwidth), the noise of resistor R_s may become very large, and comparable to the noise of the feedback resistor R_f and the equivalent input transistor M_x. When $R_s(A_0 + 1) \gg R_f + R_{out}$ (due to a high A_0 and/or a large R_s), the peak value of this noise power is given by:

$$\frac{4kT}{R_s} R_f^2, \tag{4.58}$$

which is depicted in Fig 4.9(b). The maximum value of the noise power is thus proportional to R_f^2, and inversely proportional to R_s: the smaller R_s, the higher the noise contribution. However, for very small R_s, the assumption $R_s(A_0 + 1) \gg R_f + R_{out}$ does not hold anymore, and the peak value is given by:

$$4kTR_s\Big(\frac{R_f A_0}{R_f + R_{out}}\Big)^2, \tag{4.59}$$

which is directly proportional to R_s.

The simulated output power spectral densities for several cases are depicted in Fig. 4.10. The bias voltages are always $V_{b1} = 1.8$ V and $V_{b2} = 1.8$ V, except in Fig. 4.10(b), where $V_{b1} = 0$ V and $V_{b2} = 0$ V. When both M_{Rfb} and M_3 are turned off, the noise of R_{fb} is dominant at DC (-165 dB/Hz). The DC value due to the photodiode shot noise cannot be seen in the graphs of Fig. 4.10 as the zero of (4.56) is as low as 30 kHz. This DC value is -240 dB/Hz, except in Fig. 4.10(b) where it equals -252 dB/Hz due to the smaller R_f. The noise density of M_1 is always slightly larger than the noise density of M_2. The equivalent $g_{mx,M1}$ for the noise contribution of M_1 in (4.34) is given by:

$$\frac{1}{g_{mx,M1}} = \frac{g_{m,M1}}{(g_{m,M1} + g_{m,M2})^2}, \tag{4.60}$$

and an analog expression holds for $g_{m,M2}$. Because $g_{m,M1} > g_{m,M2}$, the same is true for their noise contributions to the output.

Comparing Fig. 4.10(a) and Fig. 4.10(b) shows the effect on the noise performance when transistors M_{Rfb} and M_3 are turned on. The noise of M_3 can still be neglected, so is not depicted in the graphs. Due to the smaller feedback resistance value of M_{Rfb}, its noise becomes dominant over the noise of R_{fb}. Fig. 4.10(a), Fig. 4.10(c) and Fig. 4.10(d) compare the noise performance for the three n-well photodiodes: the classical, the quasi-fractal and (one half of) the differential photodiode respectively. The larger the diode area, the smaller R_s and the higher the maximum (4.58) of its noise curve. In Fig. 4.10(a), the noise of R_s becomes even larger than the noise of R_{fb}, while in Fig. 4.10(d), the noise curve of R_s stays well below the noise curve of R_f. Finally, Fig. 4.10(e) and Fig. 4.10(f) show the output noise spectral densities

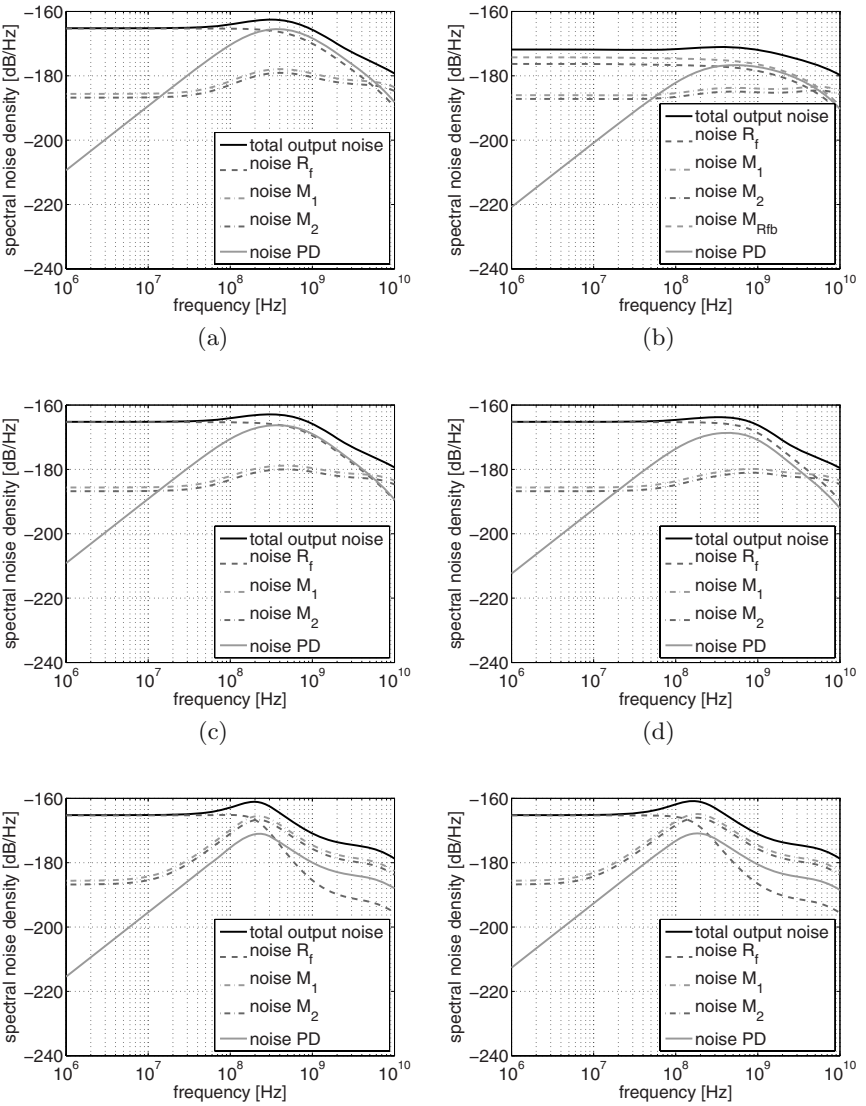


Fig. 4.10. Simulated output noise power spectral densities of:
(a) TIA with classical n-well diode ($R_f = 2000\ \Omega$, $R_s = 1200\ \Omega$),
(b) TIA with classical n-well diode ($R_f = 589\ \Omega$, $R_s = 1200\ \Omega$),
(c) TIA with fractal n-well diode ($R_f = 2000\ \Omega$, $R_s = 1600\ \Omega$),
(d) TIA with differential n-well diode ($R_f = 2000\ \Omega$, $R_s = 3000\ \Omega$),
(e) TIA with p^+ n-well diode with guard ($R_f = 2000\ \Omega$, $R_s = 6\ \Omega$) ,
(f) TIA with n^+ p-substrate diode($R_f = 2000\ \Omega$, $R_s = 5\ \Omega$) .

Table 4.3. Hand calculations (upper part) and simulation results (lower part) of the TIA with classical n-well diode under different biasing conditions.

V_{b1}	**1.8 V**	**0 V**	**1.8 V**	**0 V**
V_{b2}	**1.8 V**	**1.8 V**	**0 V**	**0 V**
R_{out}	644 Ω	644 Ω	253 Ω	253 Ω
A_0	16.8	16.8	6.6	6.6
C_{dio}	660 fF	660 fF	660 fF	660 fF
C_{in}	760 fF	760 fF	395 fF	395 fF
C_{inT}	1.42 pF	1.42 pF	1 pF	1 pF
R_f	2000 Ω	589 Ω	2000 Ω	589 Ω
$\frac{A_0}{2\pi R_f C_{inT}}$	940 MHz	3.2 GHz	500 MHz	1.7 GHz
$\frac{A_0}{2\pi(R_f+R_{out})C_{inT}}$	711 MHz	1.5 GHz	440 MHz	1.2 GHz
$R_f\frac{A_0}{A_0+1}-\frac{R_{out}}{A_0+1}$	65 dBΩ	54 dBΩ	65 dBΩ	54 dBΩ
BW_{TIA}	640 MHz	1.9 GHz	440 MHz	1.4 GHz
$Z_{TIA,0}$	65 dBΩ	54 dBΩ	65 dBΩ	54 dBΩ
ZBW	1.1 THzΩ	952 GHzΩ	782 GHzΩ	702 GHzΩ
$G_{TIA,0}$	88 dBΩ	73 dBΩ	84 dBΩ	70 dBΩ
$GH_{TIA,0}$	22 dB	18 dB	18 dB	15 dB
$f_{0dB,GH}$	692 MHz	2 GHz	510 MHz	1.6 GHz
P.M.	99°	103°	111°	113°
$v_{n,TIA}$	222 μV_{rms}	155 μV_{rms}	169 μV_{rms}	128 μV_{rms}
$i_{n,tia}$	0.12 μA_{rms}	0.3 μA_{rms}	0.1 μA_{rms}	0.27 μA_{rms}
dominant noise source	PD	M_{Rfb}	R_f	M_{Rfb}

for the TIAs with the smallest bandwidth and largest photodiode capacitances at their input: the p^+ n-well diode with guard, and the n^+ p-substrate diode. These diodes also have a much smaller R_s value, so the maximum of its noise curve is given by (4.59) and never becomes dominant. Because the zero of the amplifier noise (4.34) is well separated from the TIA bandwidth, the noise curves of transistors M_1 and M_2 can increase and even rise above the noise curve of resistor R_f at high frequencies.

Table 4.3 and Table 4.4 summarize the results for the integrated output noise and the integrated input-referred noise current. Two times the TIA bandwidth is taken as the upper integration limit. Also the noise source with the largest contribution to the total integrated noise is mentioned. Table 4.3 shows that M_{Rfb} is the dominant noise source when turned on. Otherwise, the dominant noise source is R_{fb} ($V_{b2} = 0$ V) or the photodiode ($V_{b2} = 1.8$ V)

Table 4.4. Hand calculations (upper part) and simulation results (lower part) of the TIA with different photodiodes ($V_{b1} = 1.8$ V, $V_{b2} = 1.8$ V).

	quasi-fractal n-well diode	**differential diode**	**p^+n-well diode with guard**	**n^+p-substrate diode**
R_{out}	644 Ω	644 Ω	644 Ω	644 Ω
A_0	16.8	16.8	16.8	16.8
C_{dio}	585 fF	292 fF	4.6 pF	6.8 pF
C_{in}	760 fF	760 fF	760 fF	760 fF
C_{inT}	1.34 pF	1.05 pF	5.3 pF	7.6 pF
R_f	2000 Ω	2000 Ω	2000 Ω	2000 Ω
$\frac{A_0}{2\pi R_f C_{inT}}$	993 MHz	1.3 GHz	248 MHz	175 MHz
$\frac{A_0}{2\pi(R_f+R_{out})C_{inT}}$	751 MHz	960 MHz	188 MHz	132 MHz
$R_f\frac{A_0}{A_0+1} - \frac{R_{out}}{A_0+1}$	65 dBΩ	65 dBΩ	65 dBΩ	65 dBΩ
BW_{TIA}	741 MHz	910 MHz	230 MHz	166 MHz
$Z_{TIA,0}$	65 dBΩ	65 dBΩ	65 dBΩ	54 dBΩ
ZBW	1.3 THzΩ	1.6 THzΩ	409 GHzΩ	295 GHzΩ
$G_{TIA,0}$	88 dBΩ	88 dBΩ	88 dBΩ	88 dBΩ
$GH_{TIA,0}$	22 dB	22 dB	22 dB	22 dB
$f_{0dB,GH}$	759 MHz	851 MHz	155 MHz	117 MHz
P.M.	97°	91°	71°	75°
$v_{n,TIA}$	224 μV_{rms}	222 μV_{rms}	140 μV_{rms}	140 μV_{rms}
$i_{n,TIA}$	0.12 μA_{rms}	0.12 μA_{rms}	0.08 μA_{rms}	0.08 μA_{rms}
dominant noise source	R_f	R_f	R_f	M_1

owing to the higher integration bandwidth. The table also shows that the input-referred noise is the smallest for the largest value of R_f. Comparing the noise for different diodes at the input, Table 4.4 shows that the input noise of the TIA with p^+ n-well diode with guard and TIA with n^+ p-substrate diode is slightly smaller. This is because the integration bandwidth in these cases is considerably lower than for the TIAs with n-well diodes. The dominant noise source in case the n-well diodes are at the input is R_f. For the TIA with p^+ n-well diode with guard, the noise contribution of M_1 almost equals the noise contribution of R_f, but the latter is still dominant. When the n^+ p-substrate diode is at the input, the noise of of M_1 dominates all other noise sources. Finally note that no attempt has been made to minimize the integrated noise for a certain photodiode input capacitance as described in Section 4.3.3. The large variation in the photodiode capacitance values results in a ratio X_N that

varies between 0.2 for the differential diode and 0.02 for the n^+ p-substrate diode. However, the prime goal was to compare the speed performance of the diode topologies, by amplifying the signals with identical TIA configurations that have a larger bandwidth than the intrinsic bandwidth of the respective photodiodes.

4.5.2 An Inverter-Based TIA for Test Photodiodes in 90 nm CMOS

Design Goals and Implementation

The design of this TIA has a lot in common with the one described in the previous section (Section 4.5.1). Again the aim is to amplify the signal from different types of photodiodes, comparing their speed performance. The major difference is that the design is done in a 90 nm CMOS process, with minimum effective gate-length of 80 nm. This is reflected in a higher level of Eldo transistor model, namely level 54 or BSIM4. This model is an extension of the BSIM3 model (or Eldo level 53) and addresses the MOSFET physical effects into the sub-100 nm regime. BSIM4 has as much as twenty improved and/or new models compared to BSIM3 [UC 04], of which only a few are:

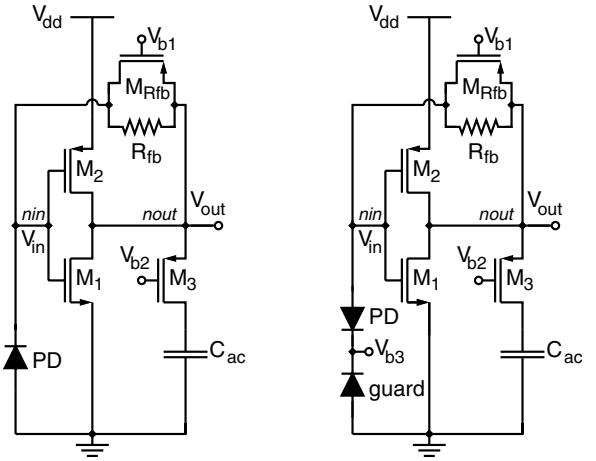
- an accurate new model of the intrinsic input resistance for both RF, high-frequency analog and high-speed digital applications;
- a new accurate channel thermal noise model and a noise partition model for the induced gate noise;
- an accurate gate direct tunneling model for multiple layer gate dielectrics;
- a comprehensive and versatile geometry-dependent parasitics model for various source/drain connections and multi-finger devices;
- improved model for steep vertical retrograde doping profiles.

These highly accurate models need of course several new parameters to describe the physical phenomena due to downscaling. It is clear that the 'intuitive' square law for the saturation current given by:

$$I_{ds} = K' \frac{W}{L} (V_{gs} - V_T)^2 (1 + \lambda V_{ds}), \tag{4.61}$$

becomes more and more out of date [San06]. However, it is still useful to perform raw hand-calculations.

The schematic of the TIA, again with different diode configurations, is depicted in Fig. 4.11(a) and Fig. 4.11(b). The test photodiode structures now are a classical n-well diode and a quasi-fractal n-well diode (Fig. 4.11(a)), and a p^+ n-well diode with guard (Fig. 4.11(b)). In the latter case, besides the detecting diode PD, also a guard diode is present. The bias voltage of the n-well (V_{b3}) should stay below 1.1 V. Because no detailed junction capacitance models are available for this technology, approximate calculations are performed to determine the total junction capacitance, including bottom-plate



(a) (b)

Fig. 4.11. Schematic of the inverter-based 90 nm CMOS TIA: (a) configuration for the classical n-well diode and quasi-fractal n-well diode (b) configuration for the p^+ n-well diode with guard.

as well as sidewall capacitance. For the classical n-well diode, $C_{dio} \approx 1.4$ pF, for the quasi-fractal n-well diode, $C_{dio} \approx 1.6$ pF, and for the p^+ n-well diode with guard, $C_{dio} \approx 6$ pF.

The TIA consists again of a single-stage inverter amplifier (nMOS transistor M_1 and pMOS transistor M_2) and a variable feedback resistance (resistor R_{fb} and pMOS transistor M_{Rfb}). The main difference with the previous design is that the dummy stage is left out. Transistor M_3 operates in the linear region, and can be turned on or off by changing the bias voltage V_{b2}. To have drain and source of transistor M_3 biased at the same DC voltage, capacitance C_{ac} is added, which prevents DC current from flowing through M_3. The function of transistor M_3 remains the same: taking care of stability problems due to process variations on output conductances $g_{ds,M1}$ and $g_{ds,M2}$. The same TIA structure with input inductor is presented in [Ler04]. Due to the series inductor between the diode and the TIA, the bandwidth BW_{TIA} is decoupled from the unity-gain frequency $f_{0dB,GH}$ of the loop gain. This way a TIA can be realized with the same gain and bandwidth compared to the regular TIA in Fig. 4.11(a), but with a smaller input-referred noise current. The detailed analysis falls beyond the scope of this text, but the interested reader is referred to [Ler04].

The design parameters of the inverter amplifier are summarized in Table 4.5. The main simulation results for an 'average' 1.5 pF photodiode are given in Table 4.6. Because of the high resemblance with the previous TIA,

Table 4.5. Design parameters of the inverter amplifier.

	M_1	M_2
type	nMOS	pMOS
L	80 nm	80 nm
W	100 µm	300 µm
V_{DSAT}	0.09 V	−0.12 V
I_{DS}	3.8 mA	−3.8 mA
g_m	44 mS	46 mS
g_{ds}	3.5 mS	3.9 mS
C_{gs}	34 fF	103 fF
C_{gd}	10 fF	36 fF

no exhaustive overview is given anymore, but only the main design issues are discussed briefly.

DC Operating Point

As M_1 and M_2 are the only transistors delivering DC current, both currents should be equal:

$$I_{DS,M1} = |I_{DS,M2}|. \tag{4.62}$$

Assuming the square law is still valid for the 80 nm-channel transistors in strong inversion, (4.49) gives the relationship between the widths of the transistors and the DC input voltage. Note that V_{DD} is now only 1.1 V. In the presented design, the ratio W_{M1}/W_{M2} equals 1/3 (Table 4.5), leading to a simulated $V_{IN} = V_{OUT} = 0.49$ V.

Transimpedance Gain and Bandwidth

The transimpedance gain is made variable in the same way and for the same reason as in Section 4.5.1: to increase flexibility when different types of photodiodes detect light at the input. For the implementation of the fixed 2700 Ω resistance, salicided p^+ polysilicon is used. Due to a compact layout of 1 µm x 9 µm, the RC-constant of the resistor results in a bandwidth as high as 80 GHz. The variable resistance is realized as a pMOS transistor, with bulk connected to source. Its resistance value is given by (4.51). M_{Rfb} has minimal length and a width of 70 µm, so the resulting R_f, which is the parallel combination of R_{fb} and M_{Rfb}, ranges from 120 Ω to 2700 Ω. The nominal design value, resulting in a bandwidth of 500 MHz, is 680 Ω (Table 4.6). This corresponds to a transimpedance-bandwidth product of 340 GHzΩ. This value is considerably smaller than the transimpedance-bandwidth product of the inverter-based design in 0.18 µm CMOS (Section 4.5.1). As ZBW is inversely

proportional to the input capacitance (4.19), this decrease is mainly due to the larger junction photodiode capacitance associated with deep-submicron CMOS technologies.

Due to the ac-coupling capacitor C_{ac}, the output resistor at DC is fixed and determined by the output conductances of the transistors:

$$R_{out} = \frac{1}{g_{ds,M1} + g_{ds,M2}}. \tag{4.63}$$

Only at higher frequencies, determined by R_{M3} and C_{ac}, the resistance value of R_{M3} (4.51) is seen at the output, in parallel with $g_{ds,M1}$ and $g_{ds,M2}$:

$$R_{out,ac} = \frac{1}{g_{ds,M1}} // \frac{1}{g_{ds,M2}} // R_{M3}. \tag{4.64}$$

The advantage of this approach is that the DC voltage gain is now also independent of R_{M3}. It is given by:

$$A_0 = (g_{m,M1} + g_{m,M2}) R_{out}, \tag{4.65}$$

while:

$$A_{0,ac} = (g_{m,M1} + g_{m,M2}) R_{out,ac} \tag{4.66}$$

is the gain when C_{ac} acts as a short circuit. The possible pitfall is that C_{ac} must be high enough to enable the reduction of the output resistance by R_{M3} for the frequencies of interest. The implementation of this large capacitance value in a CMOS technology requires a capacitor with a high density per unit area to limit the (expensive) area. As no MIM-capacitors are available, a metal wall capacitance structure is used [Yao04]. It uses the lateral capacitance instead of the vertical capacitance normally used. In deep-submicron technologies, the lateral spaces between metal lines in the same layer are smaller than the vertical spaces between layers, leading to a higher capacitance density. Also the lateral spacing is better controlled. The calculated capacitance per unit area is around 1.7 fF/μm^2, enabling a $C_{ac} \approx 13$ pF within the restricted area. Table 4.6 summarizes the maximum, minimum, and nominal values of $R_{out,ac}$ and the corresponding $A_{0,ac}$.

Open-Loop and Loop Gain

The behavior of open-loop gain and loop gain is the same as the behavior of A_0. At DC, they do not change when M_3 is turned on or off. At higher frequencies (dependent on the value of C_{ac}), the influence of M_3 will become more important. When instabilities occur it allows to decrease R_{out} and reduce the loop gain at these frequencies, in order to adjust the phase margin and maintain stability. This might be necessary because the output conductances of the transistors, which determine R_{out}, are subject to process variations.

Table 4.6 shows that the simulated unity-gain frequency $f_{0dB,GH}$ equals the TIA bandwidth BW_{TIA}. The phase margin for a photodiode capacitance of 1.5 pF equals 75°.

Table 4.6. Simulation results of the TIA with n-well photodiode for which $C_{dio} = 1.5$ pF.

R_{fb}		2700 Ω
R_f	min-max	120 Ω-2700 Ω
	nominally	680 Ω
R_{out}		130 Ω
$R_{out,ac}$	min-max	60 Ω-130 Ω
	nominally	90 Ω
A_0		11.22
$A_{0,ac}$	min-max	15 dB-21 dB
	nominally	18 dB
BW_{TIA}		500 MHz
$Z_{TIA,0}$		57 dBΩ
ZBW		340 GHzΩ
$f_{0dB,GH}$		500 MHz
P.M.		75°
$i_{n,TIA}$		0.21 μA_{rms}

Noise

As no detailed model is available for the junction capacitances, an estimated value of 10 Ω is taken for R_s (Fig. 4.9(a)). The noise of this resistor will never become larger than the noise of the feedback resistor R_f or input transistors M_1 and M_2. Table 4.6 reveals that the input-referred integrated noise current $i_{n,TIA}$ equals 0.21 μA_{rms}. Assuming the noise models of Section 4.3.3 are valid, R_f is the dominant noise source.

4.5.3 A Differential Bandwidth-Optimized TIA in 0.18 μm CMOS

Design Goals and Implementation

This TIA differs in two major ways from the previous designs discussed in Section 4.5.1 and Section 4.5.2. First, it is optimized for one particular photodiode: the differential photodiode for which the dark and illuminated junctions both have a junction capacitance of 159 fF respectively. Second, the topology is totally different: a differential common-source structure is implemented, and the voltage amplifier consists of two stages. A differential amplifier is actually a very natural choice in combination with the differential photodiode topology: the illuminated junctions are connected to one input while the dark junctions are connected to the other input. Note however that the input

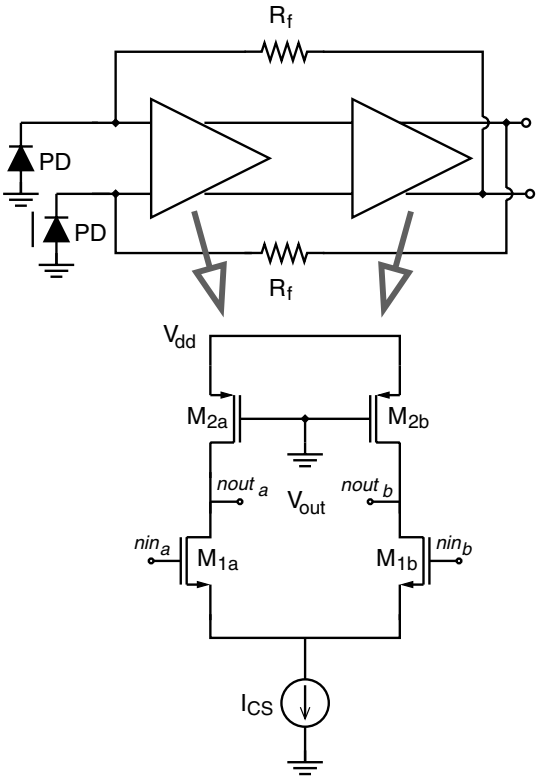


Fig. 4.12. Schematic of the differential 0.18 µm TIA.

current is not truly differential: the current from the illuminated junctions consists of a drift and diffusion component, while the current from the dark junctions only consists of the diffusion component. After this differential TIA stage, the difference signal consisting only of the drift component still has to be constructed. The presented TIA is part of an optical receiver which has the ultimate goal to receive and amplify signals with bitrates as high as a few Gbit/s. The other building blocks of this receiver, together with the measurements, will be discussed in Section 6.4.

Fig. 4.12 shows the circuit schematic of the differential TIA. A benefit of a differential topology is that the number of stages in an amplifier is not limited to an odd number. As a compromise between high bandwidth (one single-ended stage for stability reasons) and high voltage gain (three single-ended stages to increase amplification), this TIA consists of two differential amplifying stages with cross-coupled feedback. Each stage consists of the input nMOS transistors M_{1a}-M_{M1b} and the pMOS load transistors M_{2a}-M_{2b}. These transistors are biased in the linear region by connecting their gates to ground. The feedback resistor has as fixed value. As a compromise between

Table 4.7. Design parameters of the two differential stages.

	first stage		**second stage**	
	M_1	M_2	M_1	M_2
type	nMOS	pMOS	nMOS	pMOS
L	0.18 µm	0.18 µm	0.18 µm	0.18 µm
W	34 µm	12 µm	13 µm	8 µm
V_{DSAT}	0.22 V	−0.94 V	0.25 V	−0.95 V
I_{DS}	1.95 mA	−1.95 mA	1.1 mA	−1.1 mA
g_m	11.9 mS	1.25 mS	5.2 mS	0.68 mS
g_{ds}	0.64 mS	2.2 mS	0.26 mS	1.7 mS
C_{gs}	41 fF	14 fF	16 fF	9 fF
C_{gd}	12 fF	8 fF	5 fF	6 fF

gain and noise on one hand and bandwidth and phase margin on the other hand, the resistance equals 5000 Ω. The resistor is implemented in high-ohmic polysilicon, and with dimensions of 2 µm x 10 µm, its 3-dB bandwidth equals almost 20 GHz.

Table 4.7 shows the transistor design parameters. It is apparent that the two stages are not identical. A more optimal solution is to have a large first stage and a smaller second stage. The hand calculations and simulations results are summarized in Table 4.8.

DC Operating Point

The DC current through the transistors is set by the common-mode current source I_{CS}. Each branch carries one half of its DC current. Owing to the feedback resistor R_f, which does not carry any DC current, the input DC voltage V_{IN} of the TIA is the same as the output DC voltage V_{OUT}. It is determined by:

$$V_{OUT} = V_{DD} - \frac{I_{CS2}}{2} R_{M2}, \tag{4.67}$$

where I_{CS2} is the DC current in the second stage of the amplifier and R_{M2} is the resistance value of M_2 in the second stage, and given by (4.51). The higher W_{M2}, the lower R_{M2} and the higher V_{OUT}. The input voltage is given by:

$$V_{IN} = V_{GS,M1} + V_{DS,CS1}, \tag{4.68}$$

where $V_{DS,CS1}$ is the drain-source voltage of the transistor constituting the current source in the first stage. Any change in W_{M1} will be reflected in a change of $V_{GS,M1}$ when the current is kept constant. However, the current source will adjust its drain-source voltage $V_{DS,CS1}$ such that V_{IN} still equals V_{OUT} given by (4.67).

In the presented design, the simulated $V_{IN} = V_{OUT} = 1.3$ V. This is somewhat higher than the DC voltage in the designs of Section 4.5.1 and Section 4.5.2, where one half of the power supply is taken as guideline. As the signal current from the photodiode is so small, large-signal problems with the output swing will never occur. A higher DC voltage has the advantage that the input photodiode capacitance is lowered. The intermediate voltage in between the two stages equals 1.2 V.

Transimpedance Gain and Bandwidth

The transimpedance gain of this TIA is fixed and determined by the value of R_f, which is 5000 Ω. Due to the loading of R_{out}, the gain is somewhat lower and equals 4566 Ω or 73 dBΩ (Table 4.8).

The voltage gain is maximized by implementing two stages. The total gain $A_{0,tot}$ is than given by:

$$A_{0,tot} = A_{0,1} \cdot A_{0,2}, \tag{4.69}$$

with $A_{0,1}$ and $A_{0,2}$ the voltage gain of respectively the first and the second stage. Their values are determined by:

$$A_0 = g_{m,M1} R_{out}, \tag{4.70}$$

while:

$$R_{out} = \frac{1}{g_{ds,M1}} // R_{M2}. \tag{4.71}$$

As explained several times in this chapter, a large voltage gain is important to achieve a maximization of the TIA bandwidth, in first order given by (4.11). The optimization process shows that the two stages shouldn't be identical. The first stage can be made larger: as long as the input capacitance is small enough compared to the photodiode capacitance, the voltage gain increases with $g_{m,M1}$ and so does the bandwidth. The optimum is reached when increasing W_{M1} results in a lower bandwidth due to the larger input capacitance and decreasing W_{M1} results in a lower bandwidth due to a lower voltage gain A_0. The second stage still gives some additional voltage gain, but cannot be made too large as it determines the dominant pole of the voltage amplifier, and thus also the stability of the TIA.

Table 4.8 shows a large discrepancy between expression (4.11) and the simulated 3-dB bandwidth. This is because the poles of the TIA are complex conjugated and cause some gain peaking. So the actual bandwidth equals 4.3 GHz, which is larger than the value predicted by (4.11). Together with a gain of 73 dBΩ, this results in a transimpedance-bandwidth product of 19 THzΩ. This is more than one order of magnitude larger than the transimpedance-bandwidth product of the designs discussed in Section 4.5.1 and Section 4.5.2.

Open-Loop and Loop Gain

The frequency at which the gain margin in measured, $f_{0dB,GH}$, equals 2.3 GHz, significantly smaller than the simulated BW_{TIA}. The phase margin is 69°. Any change in values which increases the TIA bandwidth will result in a smaller phase margin. For example, a larger W_{M1} in the second stage will increase A_0 and consequently also BW_{TIA}. But also the loop gain $GH_{TIA,0}$ increases, which results in a lower phase margin. Decreasing R_f to extend the bandwidth not only results in a worse noise performance, but also increases $f_{0dB,GH}$ and lowers the phase margin. So allowing sufficient phase margin for stable operation, this TIA is truly optimized with respect to bandwidth.

Noise

The simulated output noise power spectral density for the several noise sources is depicted in Fig. 4.13. At low frequencies, the feedback resistor R_f is clearly the dominant noise source. At higher frequencies, the noise of the series resistance R_s of the photodiode comes into play. Both contributions to the integrated output noise are equivalent: 0.57 $\mathrm{mV_{rms}}$ comes from the photodiode and 0.56 $\mathrm{mV_{rms}}$ comes from R_f. The transistor noise of the second stage can be neglected, as it is suppressed by the gain of the first stage. The noise of the transistors of the first stage never becomes dominant for the frequencies of interest. The DC value of the transistor noise contribution is given by:

$$\left(4kT\gamma g_{m,M1} + \frac{4kT}{R_{M2}}\right)\left|\frac{A_{0,2}R_{out,1}}{1+A_{0,tot}}\right|^2 df \approx \frac{4kT\gamma df}{g_{m,M1}} + \frac{4kTdf}{R_{M2}g_{m,M1}^2}. \quad (4.72)$$

The noise of the linearly biased transistor M_2 is determined by the thermal noise of its equivalent resistance value. As this term is considerably smaller than the first term in (4.72) and provided that the gain $A_{0,tot}$ is large enough, the equivalent g_{mx} as defined in (4.34) of this two-stage differential TIA is simply given by the transconductance of the input transistor, $g_{m,M1}$.

Note also that the optimum (4.44) for the input transistor is reached more or less: for $\alpha_{gd} = 0.3$, $M_i = 5$ (only the gain of the first stage is important for the Miller effect), $f_T/BW_{TIA} = 10.7$, $\gamma = 1$, $A_0 = 11.75$ and $F_{BW} = 2$, the optimum ratio between C_{gs} of the input transistor and he photodiode capacitance C_{pd} equals 0.33. In the presented design, this ratio equals 0.26, which is somewhat lower. However, this deviation is justified for following reasons:

- In [Ing04] it is shown that the optimum is a very flat optimum, so any deviation only leads to a small increase in noise.
- A smaller input transistor lowers the current consumption for the same $V_{gs} - V_T$.
- Formula (4.44) is only an approximation, as it is derived for a single stage, single-ended TIA with common-source topology.

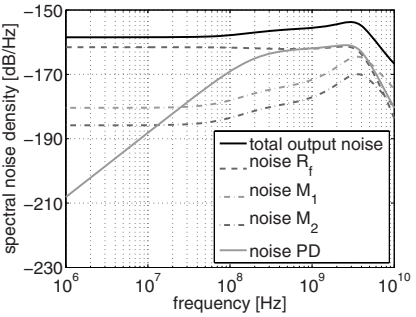


Fig. 4.13. Simulated output noise power spectral density of the differential TIA with photodiode.

Table 4.8. Hand calculations (upper part) and simulation results (lower part) of the differential TIA with photodiode.

R_{out}	first stage	352 Ω
	second stage	531 Ω
A_0	first stage	4.2
	second stage	2.8
	total	11.75
C_{dio}		159 fF
C_{in}		91 fF
C_{inT}		250 fF
R_f		5000 Ω
$\frac{A_0}{2\pi R_f C_{inT}}$		1.5 GHz
$R_f \frac{A_0}{A_0+1} - \frac{R_{out}}{A_0+1}$		73 dBΩ
BW_{TIA}		4.3 GHz
$Z_{TIA,0}$		73 dBΩ
ZBW		19 THzΩ
$G_{TIA,0}$		94 dBΩ
$GH_{TIA,0}$		20 dBΩ
$f_{0dB,GH}$		2.3 GHz
P.M.		69°
$v_{n,TIA}$		1.36 $\mathrm{mV_{rms}}$
$i_{n,TIA}$		0.29 $\mu\mathrm{A_{rms}}$

Finally, the power spectral density of the total output noise is also plotted in Fig. 4.13. Note that for instance at low frequencies, the noise is 3 dB higher than the noise of the dominant noise source, R_f. This is due to the differential nature of the circuit, where every noise source appears twice and consequently also has to be counted twice.

4.6 Conclusions

This chapter has covered an in-depth design analysis of the first electrical circuit of the optical receiver: the transimpedance amplifier. As a starting point, its main performance requirements have been defined. The photodiode current must be converted into an output voltage with high transimpedance gain. The bandwidth of the TIA should be 0.7 times the required bitrate. The equivalent input-referred noise current must be as small as possible, while the overload current must be as large as possible to design a TIA with a large dynamic range. In this work, the focus lies on low-noise TIAs as the current produced by CMOS photodiodes is very small.

The TIA with shunt-shunt feedback has been proposed as basic structure and its performance has been studied in detail. Small-signal analysis of the bandwidth reveals that allowing complex conjugated poles with a minimum amount of overshoot leads to an increase in bandwidth. This bandwidth is inversely proportional to the input capacitance (including the photodiode junction capacitance), inversely proportional to the feedback resistance, and directly proportional to the DC gain of the voltage amplifier. The transimpedance gain is mainly determined by the feedback resistance. Consequently, to maximize the transimpedance-bandwidth product, the input capacitance must be small, while the voltage gain should be maximized. Design equations for the loop gain reveal that the ratio between the dominant pole of the voltage amplifier and the unity-gain frequency of the loop gain must be high enough to have sufficient phase margin (for instance a ratio of 3 to have 72° phase margin). For a multiple-stage voltage amplifier, this ratio even has to be larger. As a result, applications requiring a high bandwidth compared to the technology's f_T will have a single-stage topology rather than a three-stage topology. The noise analysis shows that for designs aiming for a high bandwidth, the dominant noise contributor might be the feedback resistor R_f rather than the input transistor of the voltage amplifier. The input-referred current noise spectrum is flat at low frequencies and mainly determined by the noise current of R_f. At high frequencies, the noise spectrum rises with 20 dB/decade due to the amplifier noise. Different noise optimization techniques have been discussed which derive an optimal C_{gs} for the input stage of the voltage amplifier.

After the high-level analysis, several TIAs have been discussed at the transistor level. Two main topologies have been illustrated with some examples found in literature: the TIA with common source input stage and the TIA

with regulated cascode input stage. Also three interesting TIA designs of the latest years at ISSCC have been considered.

Finally, three TIAs implemented in standard CMOS technologies have been presented in this chapter. The first TIA is based on a single-stage inverter amplifier and used to compare the performance of different types of 0.18 μm CMOS photodiodes. Precautions have been taken to guarantee stability under all circumstances. Depending on the photodiode capacitance, the bandwidth ranges from 166 MHz to 910 MHz with a gain of 65 dBΩ. The TIA with classical n-well diode has a transimpedance-bandwidth product of 1.1 THzΩ, while the TIA with differential diode has a transimpedance-bandwidth product of 1.6 THzΩ. Using the ISSCC designs of Section 4.4.3 as a bench-mark, these transimpedance-bandwidth product values are coming close to present state-of-the art. This design also reveals that at higher frequencies the noise of the photodiode series resistance might become important.

The second design is also based on a single-stage inverter amplifier, but now photodiodes implemented in a 90 nm technology are compared. With an n-well photodiode capacitance of 1.5 pF at the input, the TIA bandwidth equals 500 MHz. Having a gain of 57 dBΩ, this corresponds to a transimpedance-bandwidth product of 340 GHzΩ. The smaller transimpedance-bandwidth product is mainly due to the larger photodiode junction capacitances in nm-scale technologies. Also the voltage gain A_0 does not change favorably with downscaling: in both designs, the TIA is based on an inverter amplifier, where A_0 is determined by the ratio of g_m and g_{ds}. This comes down to a voltage gain that is directly proportional to the product of Early voltage and channel length, and inversely proportional to $V_{GS} - V_T$. As demonstrated by this example, deep submicron CMOS technologies only provide very limited gain. Just this high voltage gain is needed in a TIA to achieve a large bandwidth and a high transimpedance-bandwidth product. Moving to newer technologies with higher f_T's, a different approach should be adopted. In a design for a certain bitrate and consequently a constant bandwidth, the f_T/BW_{TIA} ratio increases with downscaling. As explained in Section 4.3.2, a multi-stage approach can now be used to increase the voltage gain, because there is enough room now to place the extra poles. The optimal ratio for minimum noise $X_{N,opt}$ $(C_{gs,opt}/C_{dio})$, derived in Section 4.3.3, will decrease for increasing f_T/BW_{TIA}. However, this will not result in a lower current consumption, as the rise of C_{dio} will be much higher, requesting higher transistor widths for minimal noise and maximal bandwidth performance.

The last TIA is part of a complete optical front-end receiver implemented in 0.18 μm CMOS. A differential photodiode is used to detect the light signals, which are further amplified by a differential TIA. It comprises a two-stage voltage amplifier with cross-coupled feedback. A bandwidth of 4.3 GHz and a transimpedance gain of 73 dBΩ result in a transimpedance-bandwidth product of 19 THzΩ. Implementing a two-stage voltage amplifier and allowing some gain peaking with complex conjugated poles, results in a TIA with large bandwidth, but with sufficient phase margin and low noise performance. The power

dissipation of the core circuit, without biasing, equals 11.3 mW. These results are really competitive with present state-of-the-art (Section 4.4.3), combining a transimpedance-bandwidth product of several tens of THzΩ with a bandwidth of a few GHz. Furthermore, the TIA is designed in a fully standard CMOS technology with a minimal gate-length of only 0.18 μm and a fully integrated photodiode at its input. The measurement results of all described TIAs will be revealed in Chapter 6.