Getting the best out of photodiode detectors

Advanced Technology Development Group
Gray Cancer Institute

A great number of biomedical and analytical applications require the detection of light; this includes chemiluminescence, bioluminescence, and fluorescence as well as many others. These applications require the use of a detector to convert light into an electrical signal, and to accomplish this there are three basic technologies: photomultiplier tubes (PMTs), avalanche photodiodes (APDs), and photodiodes. Photomultipliers are discussed in a separate document.

When to use a particular detector is not always that straightforward to determine. A photodiode is suitable in applications where there is plenty of light available. In applications with very weak signals, a photomultiplier is the best choice. But there are applications where the choice is not clear. Focusing on solid-state options, this note reviews solid-state detector characteristics and the criteria for detector and amplifier selection.

The silicon photodiode

A silicon photodiode is essentially a P-N junction consisting of a positively doped P region and a negatively doped N region. Between these two regions exists an area of neutral charge known as the depletion region. When light enters the device, electrons in the structure become excited. If the energy of the light is greater than the bandgap energy of the material, electrons will move into the conduction band. The result is creation of holes throughout the device in the valence band where the electrons were originally located. Electron-hole pairs generated in the depletion region drift to their respective electrodes: N for electrons and P for holes, resulting in a positive charge build-up in the P layer and a negative charge build-up in the N layer. The charge is directly proportional to the amount of light falling on the detector. If an external circuit is connected to the electrodes, current will flow in the circuit.

The above describes the photovoltaic method of operation. It is also possible to apply a reverse bias to the photodetector, creating the photoconductive mode. This has the effect of increasing the electric field strength between the electrodes and the depth of the depletion region. The advantages of this kind of operation are lower capacitance, and hence higher speed as well as improved linearity. However, dark current is directly dependent on reverse bias voltage, and thus becomes larger with increasing bias voltage. Generally, PIN photodiodes and APDs are operated in this fashion.

The noise in a photodiode can take two forms. The first is the shot noise of the dark current, which results from the statistical uncertainty in photon arrival rate. It is present in all signals and has the following form:

$$I_{dark} = (2e\ i_{dark}\ B)^{1/2}$$

where I_{dark} = RMS noise current, e = electronic charge, i_{dark} = photogenerated signal current, and B = electrical bandwidth over which noise is measured. The second noise source is the thermal noise of the shunt resistance, also known as Johnson noise, and it takes the form:

$$I_{Rsh} = (4kTB/R_{sh})^{1/2}$$

where I_{Rsh} = RMS noise current resulting from Johnson noise, k = Boltzman's constant, T = absolute temperature of the photodiode, and R_{sh} = shunt resistance of the photodiode. Shot noise will dominate in photoconductive operation, while Johnson noise will do so in the photovoltaic mode.

Avalanche Photodiode

The APD is a specialized silicon PIN photodiode designed to operate with high reverse voltages. These reverse voltages generate high electric fields at the P-N junction. Some of the electron-hole pairs passing through or generated in this field gain sufficient energy (greater than the bandgap energy) to create *additional* electron-hole pairs, a process known as impact ionization. If the newly created electron-hole pairs acquire enough energy, they also create new pairs. This effect, known as avalanche multiplication, is the means by which APDs produce internal gain, typically in the region of 20 to 200 times

Since the APD is always operated in the photoconductive mode, its noise takes the same form as the photodiode dark-current shot noise, with the addition of a few terms:

$$I_{APDdark} = (2e \, . \, i_{dark} M^2 \, . \, F \, . \, B)^{1/2}$$

where M = detector internal gain and F = detector excess noise factor. Of the two additions, the gain simply amplifies the noise as it does the signal and has no net effect on the signal-to-noise ratio. The excess noise factor is noise added to the output signal by the multiplication process of the APD and is strongly dependent on wavelength as well as gain.

So far the only factor to play a role in detector selection has been noise equivalent power (NEP), the amount of light equivalent to the noise level of a device, or, put another way, the light level required to obtain a signal-to-noise ratio of unity. Since the noise level is proportional to the square root of the frequency bandwidth, the NEP is measured at a bandwidth of 1 Hz and expressed in units of $W/Hz^{1/2}$. But as the light level increases, the NEP no longer plays a role in the signal-to-noise ratio. The shot noise of the signal tends to dominate the signal-to-noise ratio.

$$I_{signal} = (2e \, . \, (i_{signal} + i_{dark}) \, . \, M^2 F \, . \, B)^{1/2}$$

where i_{signal} = photogenerated signal before gain. If the application uses relatively high light levels, the shot-noise performance of the detector is the only thing to consider, because the detector's dark noise and amplifier will be insignificant compared to the shot noise of the signal.

Understanding Amplifier Noise

Probably the most overlooked aspect of detector selection is that of an amplifier. The amplifier almost always determines the noise floor of the detector-amplifier combination. Whenever possible, it is generally preferable to use the transimpedance amplifier configuration, with the amplifier placed next to the detector. When the detector must be separated from the amplifier, the detector output must be terminated (generally by 50 Ω) and for high-speed applications, a voltage amplifier configuration is necessary. The two configurations are shown in figure 1.

Amplifier noise can be broken down into three major components. The first two take the familiar form of the photodiode shot noise and Johnson noise. The third term arises from the input voltage noise of the amplifier. It is the shot noise of the input bias current and the Johnson noise of the feedback resistor. It is very strongly related to the frequency bandwidth. The total detector-amplifier noise becomes:

$$I_{tot} = ((i_{dark})^2 + (i_{Rshunt})^2 + (i_b)^2 + (i_f)^2 + (i_v)^2)^{1/2}$$

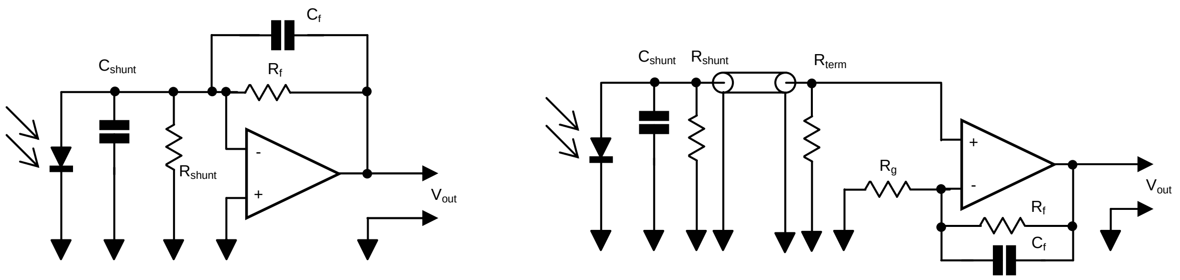

Figure 1: A transimpedance amplifier (left) and a voltage amplifier (right). The voltage amplifier configuration is used more commonly at very high frequencies or when the detector must be placed remotely from the detection electronics – best avoided if possible.

It is now possible to calculate the NEP for a given detector and, based on these calculations, select the best detector for a particular application. Referring to the table below, the ‘Si diode #1’ photodiode is designed for low capacitance, while low dark current is the principle feature for the ‘Si diode #2’; the APD is capable of high gain. The table evaluate the noise performance of each device to illustrate how to choose detectors. The values are typical, and the detectors are of comparable area.

Parameter	**Si diode #1**	**Si diode #2**	**Si APD**
Sensitivity (S)	0.6 A/W @ 960 nm	0.6 A/W @ 960 nm	30 A/W @ 960 nm
DarkCurrent (i_{dark})	10 pA	100 fA	1 nA
Shunt Release (R_{sh})	1 GΩ	50 GΩ	N/A
Shunt capacitance (C_{shunt})	65 pF	4300 pF	40 pF
Excess Noise Factor (F)	1	1	3.7
Gain (M)	1	1	60
I_{dark}	1.8×10^{-14} A	1.7×10^{-15} A	-
I_{apdark}	-	-	2.6×10^{-12} A
I_{Rsh}	4×10^{-14} A	5.7×10^{-15} A	-
I_B	5.6×10^{-15} A	5.6×10^{-15} A	5.6×10^{-15} A
I_f	4×10^{-14} A	4×10^{-14} A	4×10^{-14} A
I_v	3.5×10^{-15} A	2.3×10^{-13} A	3.6×10^{-17} A
I_{tot}	6×10^{-14} A	2.3×10^{-13} A	2.6×10^{-12} A
NEP Detector + Amp	1×10^{-13} W_{rms}	3.8×10^{-13} W_{rms}	8.6×10^{-14} W_{rms}
NEP Detector	7×10^{-14} W_{rms}	9.9×10^{-15} W_{rms}	8.6×10^{-14} W_{rms}

The above are calculated for the following ‘typical’ conditions when dealing with low detector bandwidth schemes: Bandwidth = 100 Hz; Feedback resistor = 1 GΩ; Amplifier bias current = 1 pA; Input noise voltage = 15 nV/Hz $^{1/2}$

If we make the choice on the basis of NEP alone, the total system performance has not been optimized. From the above information, the choice might seem to be the APD. But the APD requires a high-voltage power supply to bias it, is very temperature-sensitive, and generally costs more than a photodiode. So, in the above example, Si diode #1 would seem to be the best choice. Furthermore, when considering the detector's signal-to-noise performance at various light levels, it can be seen from Figure 2 that the photodiode's signal-to-noise ratio will be better than the APD when the amplifier noise is no longer a factor. This is because of the excess noise factor of the APD. Unless the application demands the lowest NEP possible, the photodiode would be the best choice under these conditions.

When should an APD be used? Based on the above equations, it is evident that the amplifier noise is strongly dependent on the bandwidth. Figure 2 was plotted for bandwidths of 100 Hz (upper panel) and 1 MHz (lower panel). When the bandwidth is increased, the results are rather different, as shown in the lower half of Figure 2. Of course, we could choose a photodiode with lower capacitance; the choice then, with respect to an APD, would not have been so clear-cut. However, the crucial point is that the gain of the APD is necessary when the amplifier noise is large, as is the case with wideband applications, or when the light source is weak.

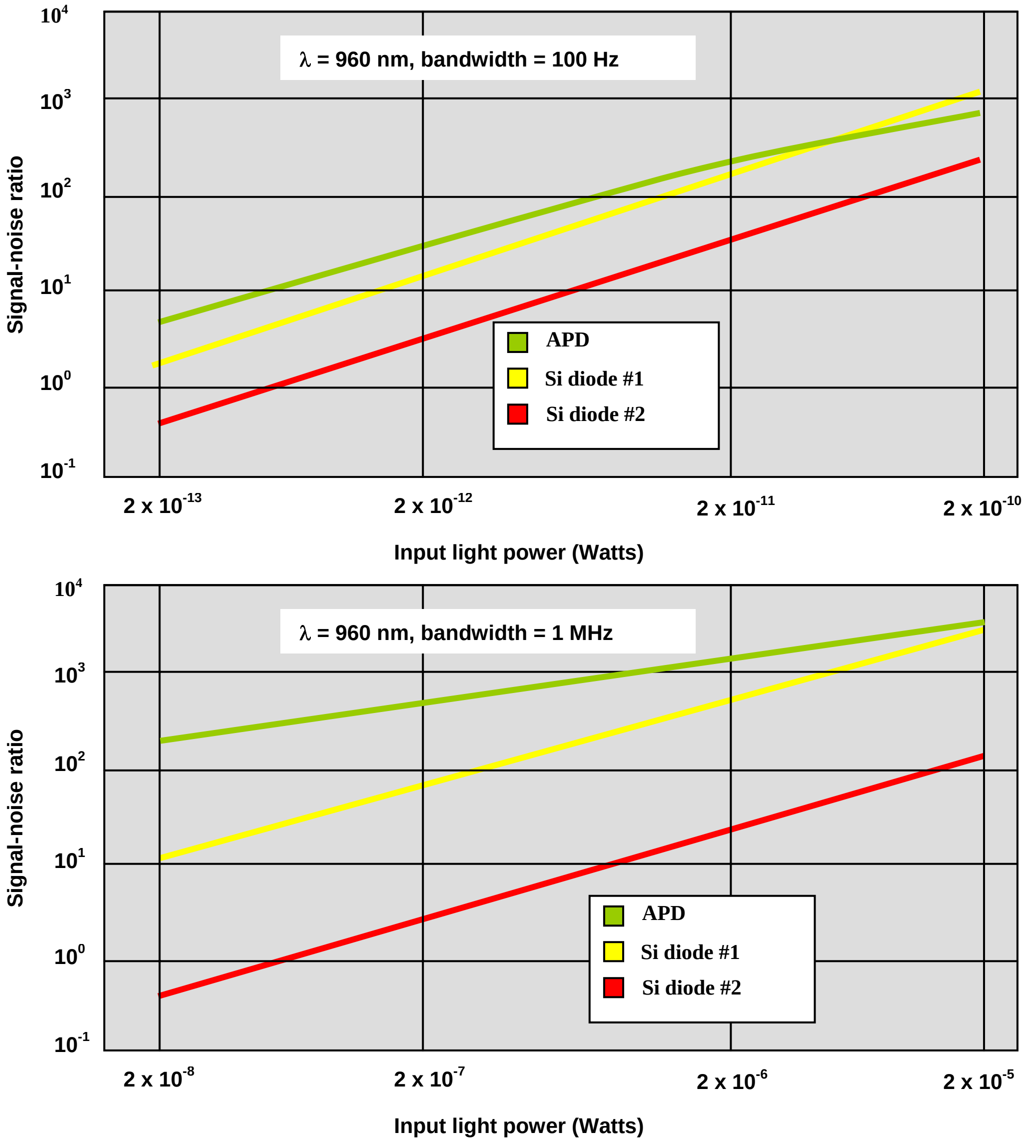

Figure 2:

Signal-to-noise ratios attainable with three different solid-state detectors,

at low electrical bandwidth (upper panel)

and

at high electrical bandwidth (lower panel).

Clearly, the choice must take into account the available light power, the desired signal-to-noise ratio and the bandwidth, as well as the active area, i.e. by the ability for the diode to collect all potentially available photons. It is worthwhile considering the use of a lens to collect available light; although the additional cost, particularly associated with mounting and alignment issues, may seem high, it is often worthwhile as it allows the area, and hence capacitance, to be reduced. The latter can of course also be reduced by reverse-biasing the detector, at the expense of increased dark current.

Choosing the correct detector is very application-specific, but here are some general guidelines.

- Try to use the smallest active area possible. If the light source for the application is diffuse, this might not be practical; however, from the standpoint of noise, small diodes have lower capacitance and dark current. They are also less expensive.
- In most applications, low shunt capacitance will be more important than low dark current. Furthermore, the NEP in the specifications does not take capacitance into account; therefore care should be exercised when comparing detectors by NEP.
- To reduce the contribution of Johnson noise, always use as large a feedback resistor as possible in the transimpedance stage, provided the required bandwidth can be met.
- In small-frequency-bandwidth applications, photodiodes operated in a photovoltaic mode will generally outperform photoconductive devices. The detector's shunt resistance should be much greater than the feedback resistance if noise is to be minimised.
- In wide-frequency-bandwidth applications, PIN photodiodes operating in the photoconductive mode are preferred because of lower terminal capacitance. APDs with their internal gain perform very well in wideband applications as well. They should be considered when the light source is weak and the amplifier noise is large.

Choice of amplifier

For low frequency work, generally below 1 MHz, very low bias current FET amplifiers are almost always preferable. Their DC performance must of course be carefully evaluated. The feedback resistor R_f, within the bandwidth constraints ultimately determined by C_f. If bandwidth considerations dictate the use of a low value of R_f (maximum tolerated response time constant = τ_{max} = R_f x C_f), take care to include the diode's shunt capacitance, C_{shunt}, in the overall component selection: the noise gain of the virtual earth amplifier will increase at high frequencies, potentially causing a peak in the frequency response and affecting amplifier stability. It may thus be necessary to add C_f to ensure a good step response from the system. Beware of testing the amplifier by purely electrical means, i.e. by injecting current into the inverting input from a signal generator voltage pulse applied to a resistor: there is no substitute for testing with an optical pulse. Optical pulses can be readily generated with a light-emitting diodes (LEDs), previously calibrated with a known-response photodetector system: 50 ns risetimes are easily produced, and with a little care in driving the LED, 5 ns risetimes are possible to achieve.

Above around 100 KHz to 1 MHz, a bipolar input stage is preferred: at high frequencies, values of R_f will be in the range of 1 kΩ to 100 kΩ. Fortunately operational amplifiers exhibiting very low voltage noise and capable of high frequency operation, with a wide gain-bandwith product are readily available, as typified by the CLC 425 operational amplifier (Comlinear, now National Semiconductors) and its descendants such as the National LMH6624, which feature input noise densities of <1 $nVHz^{-1/2}$ and <2.5 $nAHz^{-1/2}$.

The use of current feedback operational amplifier topologies seems attractive at very high frequencies. In general, however, it is best to steer clear of these if very broadband systems are needed: although very fast indeed, they tend to exhibit unacceptable tails in their settling times in the 1-100 μs time range. If a 'flat' step response to 0.1% or better is required, take care! Many an hour has been wasted trying to 'fix' a standard step output light source when the detection system was the culprit. Furthermore, current feedback operational amplifiers are very sensitive to stray capacitance at the inverting input – just where the photodiode capacitance appears. Placing a low value (10-100Ω) resistor in series with the input does help in 'isolating' the detector capacitance from the input stage. In the end, however, voltage feedback amplifiers are much easier to use now that they are available with GBW products exceeding 1 GHz.

Of course as the frequency increases, a good DC performance is harder to achieve and it is not unusual for HF amplifiers to 'drift'. In many cases an 'autozero' loop can be used to good advantage. A continuously active loop is perfectly OK if response down to true DC is not required (e.g. low frequency time constant <0.01 Hz can be easily achieved without the use of large value differentiating capacitors). The basic arrangement is shown in Figure 3, where a second, DC-stable operational amplifier is used to continuously monitor the output DC level and apply the appropriate offset to the input stage. If a response down to DC is required, it can often be arranged to close the loop when the detector is in darkness (i.e. when the system output should be zero) and open the loop when the measurement is in progress.

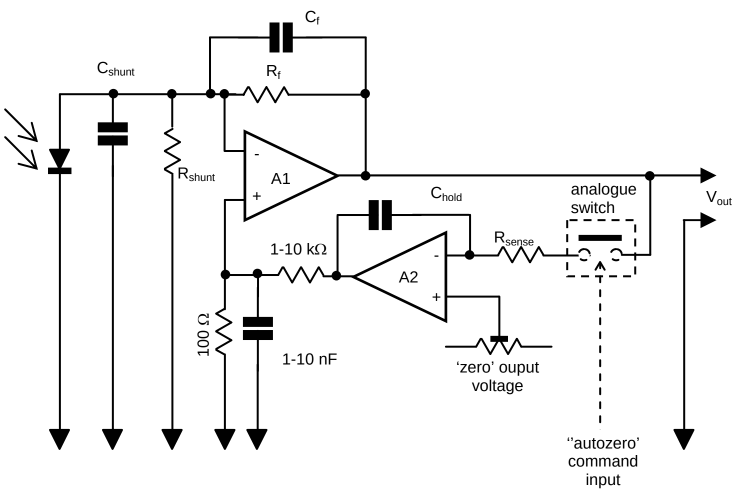

Figure 3: Photodiode transconductance amplifier with 'autozero' feedback loop to maintain output at 0 V DC. Any DC errors present at A1's output cause A2 to supply charge to C_{hold} to charge and thereby apply a correction voltage to A1's non-inverting input to bring A1's output back to zero. Amplifier A2 is a low-frequency, DC-stable device and its already low input offset voltage can be further trimmed out with the preset. The analogue switch is optional and can be eliminated if a low-frequency break point is acceptable, i.e. that the loop is continuously active through R_{sense}, typically several MΩ's to minimise the value of C_{hold}. If the switch is fitted, R_{sense} is made low to ensure fast settling to zero.

Of course it is essential that amplifier A2 has not only a low offset voltage, but also a low bias current, particularly when the analogue switch is used. In practice, it is the leakage current of the analogue switch which determines the drift. Clearly there is a compromise between the settling time achieved when the switch is closed and the drift associated with a practical value of C_{hold} (usually in the range 10-1000 nF). If a triggered autozero period of around 10 ms can be tolerated, reed relays can be used in place of the analogue switch, and negligible drifts over 10's of seconds are possible, particularly if the control range is restricted by attenuating further A2's output before it is presented to A1's non-inverting input. For the ultimate performance, the autozero 'integrator' can be replaced by a triggered A-D/D-A converter combination, where an infinite hold time can be achieved. This rather more complex arrangement will be the subject of a separate 'Baseline Restoration Methods' note.

Finally, care should be taken about how a bias voltage is applied to the photodiode. In the circuits presented here, the cathode would be biased to a positive voltage, but it should be well decoupled to ground with low-inductance ceramic capacitors and applied through a low-value current-limiting resistor, typically around 1 kΩ, since the photodiode output current is likely to be rather low. At high frequencies, all supplies should of course be well decoupled with ceramic capacitors of 10-100 nF and connections should be kept short. Surface-mount devices are in general preferred due to their small size, but unfortunately do not help with failing eyesight…..