Detectors

1. Basics & Nomenclature

Modern astronomical detectors come in several varieties depending on wavelength and use case. The traditional detector introduced in the 1800s, the photographic plate, is too inefficient ($\sim 5\%$) and too inconvenient to convert into digital form to be of use today, though it continued to be used until about 2000. Here, we will not review the physics behind modern detectors except insofar as necessary to explain how to interpret data taken with them.

In the optical, out to about a micron, the most common detector in use is the *charge-coupled device* (CCD). Out to about 28 μm, the most common detectors are infrared detector arrays. Although CCD and infrared arrays are different technologies, both are semiconductor-based detectors that detect photons of energy higher than their band gap between their bound valence electron energies and their conduction band. Silicon is the best-developed CCD technology and has a band gap of 1.1 μm. Germanium CCDs could in principle extend this sensitivity to 1.8 μm, but their technology is not cost-effective today. HgCdTe ("MER-CAH-TEL") infrared detectors have a band gap of that can be designed anywhere between 0.4–12 μm. Si:As detectors can extend to 28 μm (and are for example used in the two redder WISE bands).

Each pixel of a CCD detects individual photons that hit it, each of which usually contributes one electron to the overall signal. For most hardware, the charge is *read out* at amplifiers along the side of the CCD, by transferring the charge from pixel to pixel to the edge of the CCD. At the last pixel the charge is converted to a digital signal by the amplifiers. CCDs may be readout by various numbers of amplifiers. Because of the finite temperature of the devices, some number of electrons are released even when there are no photons entering the device; this phenomenon is known as the *dark current.* To reduce the dark current, CCD detectors need to be cooled; variations in temperature can cause variations in the dark current.

Infrared detector arrays operate similarly, but the infrared-sensitive materials cannot today be used to support the circuits necessary for CCD operation. Instead, the infrared sensitive material is bonded to a silicon detector on which the read out occurs. These detectors are not CCDs but multiplexers, which are read directly out on each pixel. The charge can be read out nondestructively, allowing many reads on the same pixel. However, the read out is intrinsically noisier.

From this basic understanding, there results an interpretation of the digital numbers reported by a semiconductor device:

$$\mathrm{DN} = \frac{n_e}{\mathrm{Gain}} = \frac{n_p + \mathrm{Dark}}{\mathrm{Gain}} \tag{1}$$

where n_e is the number of electrons recorded by the device, n_p is the number of photons actually detected, i.e. which are converted to electrons (which will include all background photon sources). Gain represents the *gain*, the number of electrons per DN reported by the electronics. Dark represents the dark current in units of electrons.

The noise in the DN is due primarily to two sources: Poisson noise around the mean n_e and *read noise* associated with the electronics. The read noise does not usually depend on the signal. Therefore:

$$\sigma^2_{\rm DN} = \frac{n_e}{{\rm Gain}^2} + {\rm Read\ Noise}^2 \tag{2}$$

The Poisson noise includes the "object" signal, the background signal, and the dark current contributions to the expected number of electrons n_e.

Over their usuable dynamic range in number of electrons, CCD devices are remarkably but not perfectly linear. CCD devices are limited in their dynamic range by the number of electrons each pixel can store. If the number of electron approaches or exceed this limit, those electrons typically bleed to neighboring pixels; since the electronics is not isotropic, they typically bleed along the same direction the device is read out. They also may be limited by the dynamic range of their analog-to-digital converter, which is often 16-bits.

CCDs are also sensitive to cosmic rays and these will release electrons that contribute to the signal. A characteristic feature of a cosmic ray is that it will be a sharper feature than the atmosphere and optics allow. The exact nature of the cosmic ray distribution depends on altitude (with more reaching higher altitude and space detectors) and orientation of the detector relative to vertical (since underneath the atmosphere the cosmic rays will be directed preferentially downwards).

In the ultraviolet, other detectors are still in use. For example, GALEX uses a position sensitive proportional counter. These devices are essentially a crossed grid of wires under voltage inside a chamber filled with an inert gas that can be ionized by UV photons. When a UV photon causes a charge track, the voltage change is detected and recorded. Unlike CCDs, these detectors can detect individual photons.

In the X-rays, CCD detectors are now typically used. CCDs lose sensitivity in the UV due to absorption on their surfaces, but at energies > 120 eV (< 100 Å) they become sensitive again. Unlike in the UV and optical, X-ray photons can release many electrons. This fact, and that the photons much rarer, allows X-ray CCDs to be energy sensitive. At high count rates, they become complicated to analyze due to latency effects; for some period of time after a photon is detected subsequent photons cannot be.

In the low frequency radio, antennae of various sorts are generally used as detectors, and the signal is read out as a changing voltage. Radio detectors therefore preserve the phase information in the electromagnetic signal, unlike higher frequency instruments. This allows interoferometry on very wide baselines to achieve very high angular resolution (low λ/D). In the higher frequency radio (sub-mm), superconductor-insulator-superconductor (SIS) devices are used, which are photon-sensing and so do not preserve the phase. In both regimes, feedhorns in the focal plane of the telescope are often used to match the focal plane to the antennae. In some configurations, antennae are used in arrays without any focusing element, and images are reconstructed through cross-correlating in the resulting signals. Noise in radio instruments is often dominated by thermal noise due to the finite temperature of the receiver electronics.

No detectors capture all of the photons hitting each pixel. The fraction of photons detected is known as the *quantum efficiency* of the device. It is a function of wavelength (most notably near the band gap). CCD and infrared devices have quantum efficiencies that are typically quite high, e.g. up to 95%.

2. Commentary

CCDs only became widely used in astronomy until the mid-1990s. Prior to that time most imaging and spectroscopy was performed with photographic plates. Photographic plates have only $\sim 5\%$ quantum efficiency typically, and obviously are harder to translate into digital results. However, they remained competitive for wide field imaging surprisingly long, because the high cost of CCDs kept them from being used to completely cover the focal planes of wide field telescopes.

3. Key References

- *Design and Construction of Large Telescopes*, Bely (2003)
- *Astrophysical Techniques*, Kitchin (2009)

4. Order-of-magnitude Exercises

1. Modern CCDs have read noise can be as low as 2 electrons, whereas infrared detectors often have read noise of ~ 20 electrons. How many independent reads would have to be performed of an infrared detector to beat down the read noise to compete with CCD read noise?

2. The cosmic X-ray background at 1 keV is about $\nu I_\nu \sim 3 \times 10^{-11}$ W m^{-2} sr^{-1} (Fabian & Barcons 1992). The Chandra X-ray telescope often takes exposures of $\sim 10^4$ s, and has a resolution of order an arcsec. What is the expected number of background photons within the FWHM of a source that are contributed by this background for such an exposure? The number of photons should be to an order of magnitude:

$$N \sim \frac{(\nu I_\nu) A \Delta T \mathrm{FWHM}^2}{h\nu} \tag{3}$$

   νI_ν is the energy per unit time per area per solid angle if we assume the width of bandpass in Chandra we are talking about has $\Delta\nu \sim \nu$. $h\nu$ is the energy per photon (i.e. 1 keV). A is the effective area of Chandra—note that because it is grazing incidence optics this effective area is smaller than the diameter of the telescope. Consulting Chandra's web site we find that at 1 keV, $A \sim 400$ cm^2. FWHM^2 represents to order of magnitude the solid angle we are integrating over. We will take FWHM ~ 1 arcsec. ΔT is 10^4 s as the problem asks. Plugging in numbers and changing units a bit:

$$\begin{aligned} N &\sim \frac{(3 \times 10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}\text{sr}^{-1})(400 \text{ cm}^2)(10^4 \text{ s})(1 \text{ arcsec}^2)}{1 \text{ keV}} \\ &\sim \frac{(3 \times 10^{-8} \text{ erg cm}^{-2} \text{ s}^{-1}\text{sr}^{-1})(400 \text{ cm}^2)(10^4 \text{ s})(2.35 \times 10^{-11} \text{ sr})}{1.602 \times 10^{-9} \text{ erg}} \\ &\sim \frac{(3)(4)(2.35)}{1.602} \times 10^{-4} \sim 1.8 \times 10^{-3} \end{aligned} \tag{4}$$

So you expect within any FWHM to have almost no background photons from the cosmic X-ray background, even for a 3 hour exposure!

3. What is the highest angular resolution that can be achieved by a ground-based interferometric radio experiment at ~ 1.6 GHz?

   In interferometry, one can use aperture synthesis to create a larger baseline aperture than traditional telescopes by placing an array of multiple telescopes. The distance between any pair of two telescopes, the baseline, sets the diffraction limit for images produced by correlating the phase signal between the telescopes. The diffraction limit is $\theta \sim \frac{\lambda}{D}$ and the largest possible distance we can separate two ground-based telescopes is of order the diameter of the Earth $D \sim 10^7$ meters. For 1.6 GHz light, we have an angular resolution of

$$\theta \sim \frac{3 \times 10^8}{(1.6 \times 10^9)(10^7)} \sim 2 \times 10^{-8} \text{ rad} \sim 4 \times 10^{-3} \text{ arcseconds.}$$

Author: Trey Jensen

5. Analytic Exercises

1. The detected electrons contributing to the noise include the object, background, and dark current signal. Write the fractional error in the object counts. Separate the terms involving the other sources of noise. For low object signal, how does the fractional error depend on signal? For high object signal, how does the fractional error depend on signal?

   The noise N includes the Poisson noise in the object, sky, dark counts ($n_{\rm obj}$, $n_{\rm sky}$, and $n_{\rm dark}$), added in quadrature with each other and the read noise $N_{\rm read}$:

$$N = \sqrt{N_{\rm obj}^2 + N_{\rm sky}^2 + N_{\rm dark}^2 + N_{\rm read}^2} = \sqrt{n_{\rm obj} + n_{\rm sky} + n_{\rm dark} + N_{\rm read}^2} \tag{5}$$

   The fractional error in the object counts is:

$$\frac{N}{n_{\rm obj}} = \frac{\sqrt{n_{\rm obj} + n_{\rm sky} + n_{\rm dark} + N_{\rm read}^2}}{n_{\rm obj}} = \sqrt{\frac{1}{n_{\rm obj}}\left(1 + \frac{n_{\rm sky} + n_{\rm dark} + N_{\rm read}^2}{n_{\rm obj}}\right)} \tag{6}$$

Defining:

$$\sigma_{\rm other}^2 = n_{\rm sky} + n_{\rm dark} + N_{\rm read}^2, \tag{7}$$

we find that if $n_{\rm obj} \gg \sigma_{\rm other}^2$ (the object-dominated regime), then

$$\frac{N}{n_{\rm obj}} \approx \frac{1}{\sqrt{n_{\rm obj}}}. \tag{8}$$

If $n_{\rm obj} \ll \sigma_{\rm other}^2$ (the background-dominated regime), then

$$\frac{N}{n_{\rm obj}} \approx \frac{\sigma_{\rm other}}{n_{\rm obj}}. \tag{9}$$

Therefore, for the faintest fluxes (below the background) the fractional errors are a stronger function of the flux than they are at brighter fluxes (above the background).

Author: Dou Liu

2. It is a commonly used practice to combine different observations of a quantity (say flux) by using a weighted mean, using weights equal to the inverse variance of each observation. The advantage of doing so is that if the inverse variance is known, this weighted mean is the minimum variance estimator of the quantity itself. For imaging conducted with photon-counting detectors, when the background is minimal, the noise in a data set is the Poisson noise in the expectation value of the signal. However, observationally we often only have access to one realization of the signal, and often the quoted errors are based on the Poisson noise estimated from the signal itself. If we take multiple observations and combine them together weighting by the inverse variance estimated in this way, it leads to a bias. Ignoring any background contribution to the noise, estimate this bias as a function of the true expected number of photons $\bar{n}$.

   Let us define $\bar{n}$ as the expectation value of the number of photons, N as the number of observations, n as the actual number of photons in any given observation, and $\{n_i\}$ as the set of actual observations of numbers of photons.

   The inverse variance weighted mean of the observations is:

$$\hat{n} = \frac{\sum_{i=1}^N n_i/\sigma_i^2}{\sum_{i=1}^N 1/\sigma_i^2}, \tag{10}$$

   where σ_i^2 is the variance we use. If we use $\sigma_i^2 = n_i$ to estimate the variance, we find:

$$\hat{n} = \frac{N}{\sum_{i=1}^N 1/n_i}, \tag{11}$$

   We can ask what the expectation value of this estimate is by averaging over an ensemble. This will tell us what the bias in this estimator is. We find:

$$\langle\hat{n}\rangle = \frac{N}{\left\langle \sum_{i=1}^N 1/n_i \right\rangle} = \frac{1}{\langle 1/n\rangle}, \tag{12}$$

where the last step uses the fact that all of the observations are equivalent, so the expectation value of $1/n_i$ is the same for all of them.

Now for a true Poisson distribution, the value of $\langle 1/n \rangle$ is problematic because of cases when $n = 0$. It is probably a bad idea to weight your data by the inverse variance if your estimate for the variance is zero. In the Jupyter notebook we treat this case, assuming our data analyst sets $\sigma_i^2 = 1$ when $n_i = 0$, which is, well, pretty common practice.

For large $\bar{n}$, we can approximate the Poisson distribution as a Gaussian and find an approximate solution which is pretty good for $\bar{n} > 10$ or so. We write:

$$\left\langle \frac{1}{n} \right\rangle = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \mathrm{d}n \frac{1}{n} \exp\left(-(n-\bar{n})^2/2\sigma^2\right) . \tag{13}$$

Then in the integral we Taylor expand $1/n$ around $\bar{n}$:

$$\frac{1}{n} \approx \frac{1}{\bar{n}} - \frac{1}{\bar{n}^2}(n-\bar{n}) + \frac{1}{\bar{n}^3}(n-\bar{n})^2 + \ldots \tag{14}$$

The first term just integrates to $1/\bar{n}$. The second term is odd around $\bar{n}$ so integrates to zero. The third term integrates to:

$$\frac{1}{\sqrt{2\pi}\sigma} \frac{1}{\bar{n}^3} \int_{-\infty}^{\infty} \mathrm{d}n \, (n-\bar{n})^2 \exp\left(-(n-\bar{n})^2/2\sigma^2\right) . \tag{15}$$

This is a known definite integral and we then find, substituting $\sigma^2 = \bar{n}$, that this term is just $1/\bar{n}^2$. So then combining the first and third terms:

$$\left\langle \frac{1}{n} \right\rangle \approx \frac{1}{\bar{n}} \left(1 + \frac{1}{\bar{n}}\right) \tag{16}$$

and therefore

$$\langle \hat{n} \rangle = \frac{1}{\langle 1/n \rangle} \approx \bar{n} \left(1 - \frac{1}{\bar{n}}\right) . \tag{17}$$

The Jupyter notebook shows this result compared to the full Poisson calculation, showing that the approximation becomes reasonable (i.e. close to the right answer relative to the bias itself) at $\bar{n} > 10$.

What is the consequence of this calculation? It is that you should be careful when you are weighting by an inverse variance that you have estimated from the signal itself. Say we took several images of the same object under the same conditions with the same exposure time—i.e. without any reason to expect that the expectation value of n would change from observation to observation. That case is exactly what we have just considered, and if we take the mean of the observations, and if we expect the variance to be dominated by the object itself, then inverse variance weighting leads to a bias (and in this case, no actual gain in precision). On the other hand, if we take several exposures whose variances we expect to vary for reasons we can predict—like different exposure times, or because we can determine that the background signal is changing—we *can* take those contributions to the difference in variance into account in our weighting, and we *should* to achieve mininum variance. There is no very simple hard-and-fast rule—usually you have to think about your specific data set to figure out the right thing to do.

6. Numerics and Data Exercises

1. CCD images from SDSS are available as corrected frames. Those images are calibrated in physical units, and have the sky background subtracted. Read the documentation there, including the data model. Download a corrected frame in the r-band. Use the formulae in the documentation to translate the image into uncalibrated DN units with the sky background reinstated (the files have all the information necessary to do that). Is it all integers? Use the result to estimate the noise per pixel in a region without stars or galaxies. Can you independently estimate that noise from the distribution of flux values themselves, and do the estimates agree?

REFERENCES

Bely, P. Y. 2003, The Design and Construction of Large Optical Telescopes

Fabian, A. C., & Barcons, X. 1992, Annual Review of Astronomy and Astrophysics, 30, 429

Kitchin, C. R. 2009, Astrophysical Techniques, Fifth Edition