



THE NORWEGIAN ASTRONOMY AND ASTROPHYSICS
OLYMPIAD COMMITTEE

IOAA NOTES

Part 9: Charge Coupled Devices (CCDs)

Author:
Đorđe Veljković

15th February 2026

1 Basic Statistics

Although this handout will be about astronomical instrumentation, specifically pertaining to use of Charge Coupled Devices (CCDs), to best understand some principles behind what images produced by such devices are affected by, we need to know some basic statistics, which is the subject of this section.

1.1 Mean, Variance, Standard Deviation

These three terms are crucial to know, and you probably already know one of them, namely the mean. The mean is simply another name for average in the context we're using it in (in actuality, there are a few different types of mean, so they're not strictly equivalent, but for all our intents and purposes, mean=average). Mean is often denoted with the letter μ , and we can write it out formally as:

$$\mu = \frac{\sum_{i=1}^N x_i}{N}$$

Variance is a measure of how much the data points are spread around the mean value. The formula for variance is:

$$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \mu)^2}{N} \quad (1)$$

where x_i is the value of any datapoint, and μ is the mean. Notice that we *square* this difference before doing all of the summation. The reason for this is to take into account deviations both above and below the mean. If we didn't square the difference, then the positive and negative deviations would cancel each other out.

Now, you might have noticed that we labelled variance as σ^2 , and didn't really use its own symbol, so what might this σ be? In this context, σ is used to denote the so-called standard deviation, which is simply the square root of the variance.

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \mu)^2}{N}} \quad (2)$$

The standard deviation is also a measurement of data spread, just like variance is, but the key difference is in units. Where variance has the units of our measurements squared, the standard deviation has the exact same units our measurements have. This is great when we want to interpret data, and you'll see standard deviation (std) used more often than variance (var).

So, why use variance at all, then? Variance is incredibly useful if we for some reason want to add or subtract multiple independent random variables. In that case, the variance of the sum is equal to the sum of variances, or written out:

$$\text{Var}[A + B] = \text{Var}[A] + \text{Var}[B] \quad (3)$$

1.2 Statistical Distributions

Some processes in nature are quite random, but there are different types of randomness exhibited by different physical processes. These randomnesses often follow so-called distributions, so even though a variable is random, there are some values that are more likely than others, and how likely the different values are is called the distribution of a random variable. The two main distribution types that are of interest to us are Gaussian distributions and Poisson distributions.

You are probably already familiar with Gaussian distributions, also called normal distributions. These are colloquially often referred to as "Bell curves", and look like Figure 1. Now, you might be tempted to simply read off the probability from the y-axis here, but the y-axis doesn't represent probability. What it actually is is the value of the so-called "probability density function", but that is beyond the scope of this work.

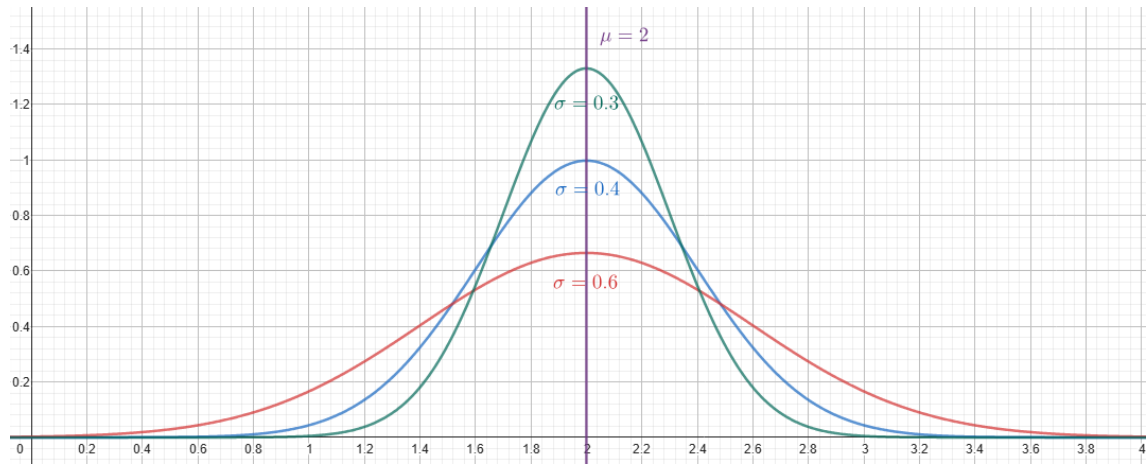


Figure 1: Three normal distributions with the same mean value $\mu = 2$, but different standard deviations specified on the plot.

If you do need to find probability from a bell curve, you will either be given lookup tables, or you will need to estimate the area under the curve. Look up tables provide probabilities for the variable being equal to some specific value based on the mean and std of the distribution, while the area method, as shown in Figure 2 shows the probability that the variable is bigger than or smaller than a certain value, or between two values.

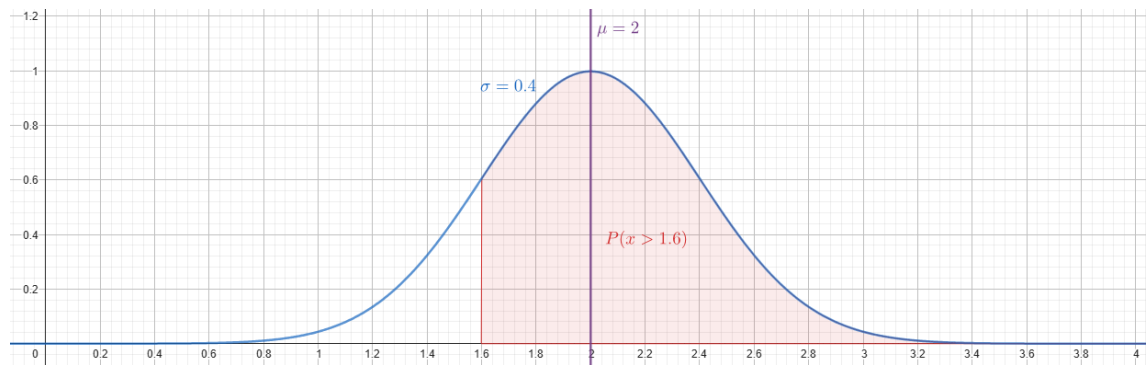


Figure 2: Finding the probability of the random variable following a normal distribution with $\mu = 2$, $\sigma = 0.4$ is greater than 1.6.

Now, the Poisson distribution is a bit different. While the normal distribution looks at values of a random variable, the Poisson distribution looks at the probability of an event happening multiple times in a specified interval. This distribution is described by the parameter λ , denoting how many events are *expected* to happen in some specified interval. The interval doesn't necessarily have to be time - it can be (as in, how many calls am I going to get per hour), but it can also be something like area or volume (how many times will a ball fall on a specific patch of grass). Both the mean and the variance of the Poisson distribution are equal to λ , and some examples of Poisson distributions can be seen in Figure 3.

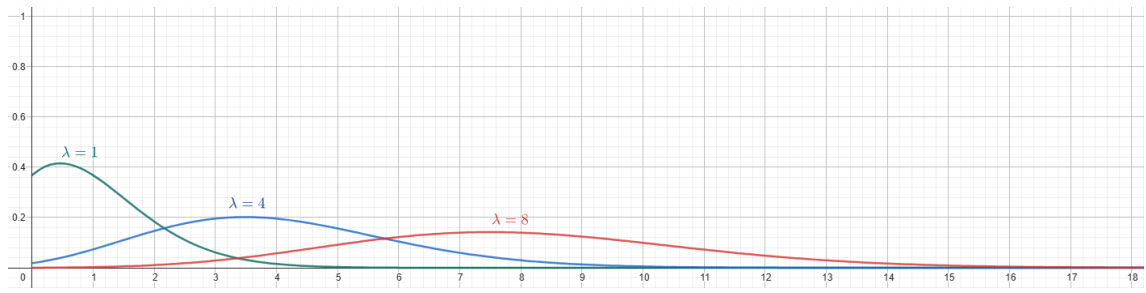


Figure 3: Three Poisson distributions with different numbers of events in an interval, specified next to the appropriate lines.

Unlike the normal distribution, here we can read the probability straight off of the y-axis. If we take the distribution with $\lambda = 1$ as an example, the probability of $n = 2$ events happening in the given interval is 0.1839, or 18.39%.

In a more practical sense, as an example, a normal distribution would describe the measurement error of an instrument, while a Poisson distribution would describe the likelihood of a certain number of photons hitting a specific pixel on a CCD (and we'll get to that later) in a certain amount of time. Using our example above, $\lambda = 1$ means that we expect one photon to hit a CCD pixel in a certain amount of time (say, 1 second, but the length of time doesn't matter). The probability of 2 photons hitting that same pixel in the same amount of time, i.e. 1 second, would then be 18.39%.

2 CCD Theory

2.1 Working Principle

The whole point of using any kind of instrumentation is to capture photons coming from distant light sources and turn them into a digital image that we can view and process. Digital images consist of pixels in a rectangular grid, and analogously, we have pixels on our CCD detectors. These pixels are made out of semiconductor material and are therefore able to collect light and turn it into electrons via the photoelectric effect. These electrons can then be stored inside of the pixel until the exposure is completed and we want to read out the constructed image. CCDs are most often made out of silicon which is efficient at photon absorption at wavelengths between $3000 \text{ \AA}$ and $11\,000 \text{ \AA}$, corresponding to the visual spectrum.

The important components of a CCD to know about are the pixels, the parallel shift register, the serial shift register and the output amplifier. To explain these, the so-called "bucket analogy" is often used. Look at Figure 4 to be able to follow the explanation. Each bucket collects rain, just as pixels collect photons. When the rainfall stops, the buckets are slowly transported along the vertical conveyor belts, until an entire row of buckets gets to the horizontal belt. Then, the vertical conveyor belts pause while the single horizontal conveyor belt transports all of the buckets to the measuring cylinder. Once all the buckets from the horizontal conveyor belt get into the measuring cylinder, the vertical conveyor belts activate again, and the process is repeated until all the buckets are emptied into the cylinder.

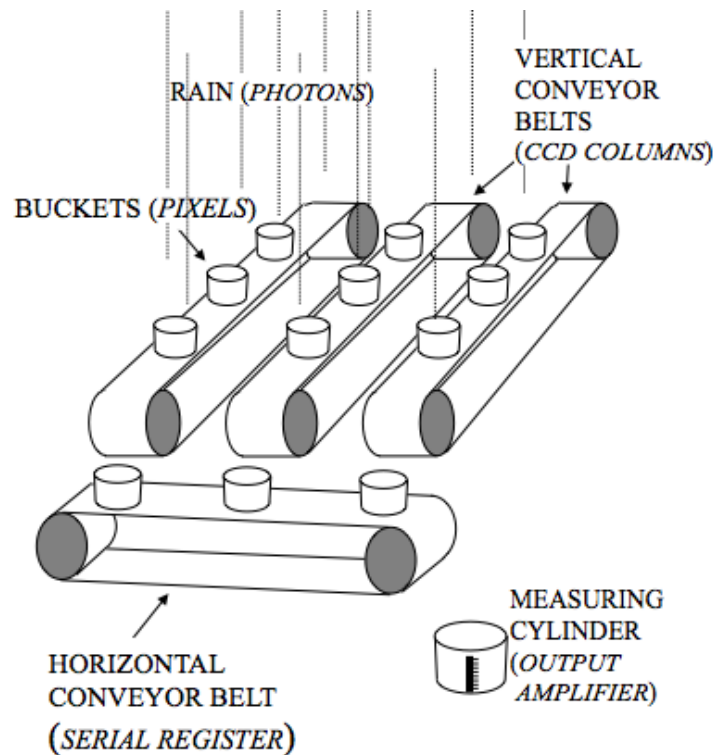


Figure 4: A sketch of the bucket analogy for the working principle of a CCD.

In the analogy above, the rainfall represented photons, and the buckets collecting the rain are analogous to pixels in a CCD. The conversion from photons to electrons happens within the pixel, made out of a semiconductor material, via the photoelectric effect. The vertical conveyor belts are like the parallel CCD shift registers, and there is one parallel CCD shift register for each CCD column. These shift registers are responsible for moving the accumulated charge to the single serial

register, which in turn transports everything to an output node, where the charge is amplified, and then converted to a digital signal. Finally, before a new exposure, the CCD reruns the entire readout process to get rid of potential residual charge. You can see the structure of an actual CCD in Figure 5.

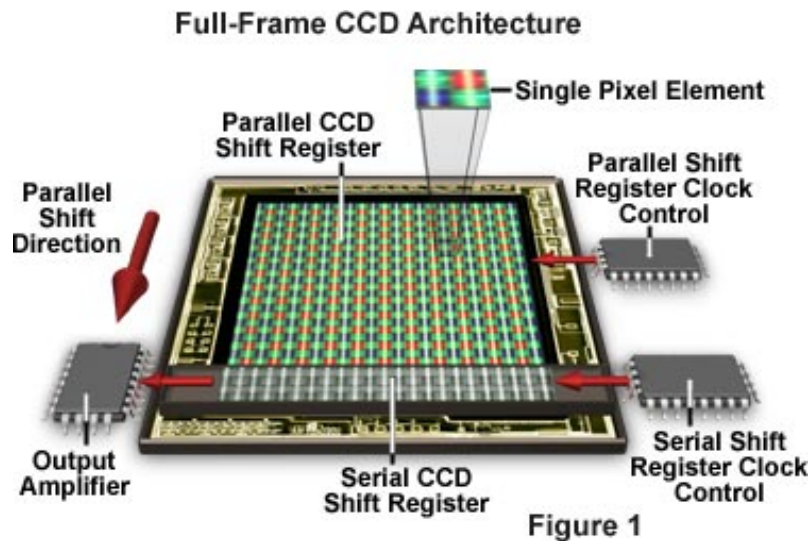


Figure 5: A sketch of an actual CCD with the electronic components in question.

This might seem like a very long process, given that CCDs can have thousands, or even millions of pixels for the really big ones, but the whole readout usually takes less than one second. The efficiency is also higher due to the fact that CCDs often operate in extremely cold environments, and we'll see why when we cover noise sources later.

2.2 Why CCDs? What else is there?

Knowing the working principle, a natural question arises - why exactly are CCDs so popular in astronomical research? This technology stems from the 80s... An incredible property of the pixels on a CCD is that they're incredibly efficient at capturing photons, so much so that close to 100% (in actuality 99.99999%) of photons that hit the detector get captured. In stark contrast, other imaging methods only capture up to 20% of incident photons and manage to turn them into electrons that are then read out at the end of imaging.

A modern alternative to CCDs that are becoming more and more popular are CMOS devices. They operate on the exact same principle, except the fact that instead of having shift registers transport charge to an output amplifier, each pixel in a CMOS has its own readout circuitry, so the region for photon collection on each pixel is smaller than for CCDs, which means a lower quantum efficiency. The tradeoff for the lower quantum efficiency is a much higher sensitivity. However, modern CMOS devices have developed technologies that allow them to rival CCDs in photon collection, having quite high quantum efficiency, although CCDs are still dominating the field.

2.3 Properties

We just covered a very important property of a CCD, namely its quantum efficiency, defined as the percentage of photons that hit the sensor and are captured and turned to charge. Quantum efficiency for silicon CCDs is highest for wavelengths in the visual spectrum. This is because photons with shorter wavelengths either get reflected off the surface or absorbed in thin surface layers, since they can't penetrate far enough down. Photons with longer wavelengths, on the other hand, simply pass right through the sensor and don't get picked up.

Quantum efficiency also depends on the operating temperature of a CCD. The efficiency for the red end of the spectrum doesn't substantially change with temperature, but the bluer we go in the spectrum, the more temperature affects quantum efficiency, with higher temperatures leading to greater quantum efficiencies. Despite this, we operate CCDs at extremely low temperatures, and not room temperature, and you'll understand why once we cover the noise sources.

Plate scale is another important property, and it denotes how many arcseconds of the sky are covered by each pixel. CCDs are used in conjunction with telescopes, and the plate scale therefore depends on the focal length of the telescope! The formula is:

$$P = \frac{206\,265 \frac{\text{arcsec}}{\text{rad}} \cdot \mu \left[\frac{\mu\text{m}}{\text{px}} \right]}{f [\mu\text{m}]}, \quad (4)$$

where μ is the side length of a single pixel, and f is the focal length of the telescope being used. This formula is useful when the optics are simple - straight from the telescope to the CCD. However, in practice, there might be more optics between the two, in which case it is easiest to estimate the plate scale empirically by looking at known separations or sizes of imaged objects.

2.3.1 ADC Dynamic Range and Gain

CCDs are monochrome, meaning that the images captured using them are black-and-white¹, which makes sense, since we're simply counting the number of photons that get collected by a pixel, and don't really look at any of the other qualities. When we convert the number of electrons to a digital signal, we are limited by how many bits we have available. As a reminder, a digital signal is binary and represented by zeros and ones. If we have 16 bits available, then we can represent $2^8 = 256$ different levels of brightness, as an example. The range from 0 to 256, or all the possible values an ADC can output is called the *dynamic range*.

The goal when designing a CCD is to use as much of the dynamic range as possible. A simplified way of looking at what an ADC does is that it counts the number of electrons, and then multiplies that number by some predefined constant to get a value inside of the dynamic range. This multiplier is often referred to as the *gain*. So, if a pixel can hold 10 000 electrons, we would want the gain to be $\frac{256 \text{ ADU}}{10\,000 \text{ e}^-} = 0.0256 \frac{\text{ADU}}{\text{e}^-}$.

Now, remember that the ADU outputs a digital signal, which must be a whole number. This means that if we have 100 electrons, the output "should" be $100 \cdot 0.0256 = 25.6$, and for 101, it should be 25.856, but since this is impossible, both of these will be truncated and the output will be 25. In practice, the uncertainty of the readout is equal to the number of electrons in a single ADU, which is the reciprocal of our gain value, i.e. $\pm \frac{1}{0.0256} \approx \pm 39 \text{ e}^-$. This error is acceptable, but if we had a case of 8 bits with 100 000 photon capacity per pixel, we'd have an error of $\pm 390 \text{ e}^-$, which is quite substantial, especially since we look at small variations in light in research.

In reality, ADCs are a bit more complicated, and the output of an ADC is not perfect. It is usually very linear for most of the range, but at some point, at very high electron counts, the ADC might start to exhibit non-linear behaviour. In addition, due to read noise that we'll cover next, there is a small offset where the ADC outputs values even though no photons fall on a certain pixel. It is therefore vital that we know what the linear area of the ADC is, so that we might take this into account during analyses.

2.4 Noise Sources

Although the noise levels on CCD images are extremely low, so much so that they're not discernible by the naked eye, they still have a significant effect on astronomical analyses. We quantify the amount of noise using a parameter called signal-to-noise ratio, which you will often see written as simply SNR. Most of the noise sources presented here will follow either a normal or a Poisson

¹Despite this, we want full-colour photographs of deep sky objects using CCDs. The modern way to do that is to take photographs through three different filters (red, green and blue) and then combine them in software.

distribution, so make sure you understand what those mean before diving in here! From now on, instead of saying "number of photons/electrons", we will say "count", which is the more commonly used term.

Read noise is due to internal electronics that aren't perfect. As such, they can introduce their own electrons into the readout, contaminating the value that we want to see. The number of these electrons is normally distributed and its standard deviation is expressed in counts per pixel. In CCDs, this noise comes specifically from the circuitry responsible for amplification of the signal and the AD conversion. The standard deviation of read noise is given as $\sigma_{\text{read}} = \sqrt{N_{\text{pixels}} \cdot \text{noise_value}^2}$ (notice that the noise value term is squared).

Dark noise is reflected in electrons that are released from the pixel material due to a rise in temperature. The higher the temperature, the higher the chance for an electron to be freed from the silicon and be captured in the pixel along with the electrons generated from the photons which we're actually interested in. The noise level is usually expressed in counts per pixel per minute (or per second), and is also normally distributed. The standard deviation of dark noise is given as $\sigma_{\text{dark}} = \sqrt{N_{\text{pixels}} \cdot \text{noise_value} \cdot t}$.

The existence of dark noise makes it so that we must have a low enough temperature to minimise it, but high enough temperature to have good quantum efficiency for the bandwidth we're interested in imaging. In this context, the operating temperature is in the order of magnitude of -100°C to -20°C , depending on the CCD.

Hot pixels are closely related to dark noise, and they are pixels that release valence electrons in much higher rates than the other pixels, up to many thousands of times more. As such, they'll appear to be much brighter in longer exposures, but this noise is luckily quite stable over time and straightforward to remove. They exist most often due to contamination of material during manufacturing, but there aren't that many of them.

Shot noise is the noise associated with the incident photons and the fact that we might not always measure the same amount of photons in the same amount of time. As such, this noise is Poisson distributed, reported in units of count, with standard deviation simply being equal to $\sigma_{\text{shot}} = \sqrt{N_{\text{source_photons}}}$. We specify source photons here, as photons can come from non-source objects as well, such as the sky.

Sky noise is very much alike shot noise, and comes from the fact that the sky is not completely black, but has its own intrinsic brightness. It depends on quite a few different factors, with the most significant one being the presence of the Moon. Luckily, this type of noise can be estimated empirically and compensated for in image analysis. The noise associated with the sky is just like the photon noise, with $\sigma_{\text{sky}} = \sqrt{N_{\text{sky_photons}}}$.

Now, we have all the necessary noise sources and can calculate SNR, which is by definition the mean of the signal divided by the standard deviation of the noise.

$$\text{SNR} = \frac{\mu_{\text{signal}}}{\sigma_{\text{noise}}}$$

The noise in the above formula is the sum of all the noises we've mentioned so far. Remember that when we add noises (or any other independent random variables), we can't just add the standard deviations. We must first calculate the variances by squaring the standard deviations, and then we can add them together to find the variance of all the noise together. Finally, we can convert that to standard deviation. In other words:

$$\begin{aligned} \sigma_{\text{noise}} &= \sqrt{\text{Var}[\text{noise}]} = \sqrt{\text{Var}[\text{shot noise}] + \text{Var}[\text{sky noise}] + \text{Var}[\text{read noise}] + \text{Var}[\text{dark noise}]} = \\ &= \sqrt{\sigma_{\text{shot}}^2 + \sigma_{\text{sky}}^2 + \sigma_{\text{read}}^2 + \sigma_{\text{dark}}^2} = \\ &= \sqrt{N_{\text{source_photons}} + N_{\text{sky_photons}} + N_{\text{pixels}} \cdot \text{read_noise}^2 + N_{\text{pixels}} \cdot \text{dark_noise} \cdot t} \end{aligned}$$

And with the mean value of the signal simply being the number of source photons, we have the

full SNR formula as:

$$\text{SNR} = \frac{N_{\text{source_photons}}}{\sqrt{N_{\text{source_photons}} + N_{\text{sky_photons}} + N_{\text{pixels}} \cdot \text{read_noise}^2 + N_{\text{pixels}} \cdot \text{dark_noise} \cdot t}}$$

We can see here that if an object being imaged is extremely bright, the $N_{\text{source_photons}}$ term dominates the others, and the SNR for very bright objects simply reduces to:

$$\text{SNR} = \frac{N_{\text{source_photons}}}{\sqrt{N_{\text{source_photons}}}} = \sqrt{N_{\text{source_photons}}}$$

Other than noise, artefacts exist - they're not taken into account during SNR calculations, but can still have quite a significant effect on an image produced by a CCD.

Empty pixels will still be translated by the ADC into a value normally distributed with a mean of zero. This means that sometimes, these pixels might be read out by the ADC as negative. To combat this, electronics are set up to add a bias voltage to all pixels - in theory, this bias voltage should be different for every pixel, but it actually differs ever so slightly. It's also not uncommon to find "bad" columns that all have a consistently higher bias value than the rest.

Sensitivity variations exist, as no two pixels will be perfectly alike. This can give a sort of a vignetting effect when imaging objects or simply regions of different brightnesses.

Cosmic rays immensely full of photons can also hit the CCD pixels and they're often reported in number of rays per unit of area per unit of time. On a CCD image, a cosmic ray is virtually indistinguishable from a star.

2.5 Reducing Noise

When measuring dark noise, we need to take a sufficiently long exposure that we can be sure that we're not just measuring read noise. Remember, all these noise sources are quite low in intensity, especially compared to the intrinsic noises in other technologies. The example frames given in this section have all been generated by Python code that I wrote myself, and I deliberately exaggerated some properties just to show the essence of the noise reduction process.

One very important thing to notice about SNR is that the number of incident photons obviously depends on time - the longer the exposure, the more photons hit the pixels. As such, we can see that SNR actually grows with $\sqrt{t}$, meaning that the more time we use for our exposure, the greater the overall SNR will be.

Just taking longer exposures doesn't suffice in reducing noise, especially when you remember that even though the SNR increases as a whole, the dark noise, for example, will increase more and more with time. It's therefore important that we calibrate our CCD before taking photographs, and that when we take photographs of the sky, we take multiple and average them. When we average the images we get from the CCD, the SNR increases by a factor of $\sqrt{N}$, where N is the number of images we average. We must take both several calibration images as well as several images of interest.

The reason for taking multiple calibration images is the fact that read noise is completely random and unpredictable, but present in every single image we see from a CCD. As such, it's different for every image taken, so if you simply take one calibration frame for the read noise, and subtract it from any other future image, you will actually add more noise in the process, since the read noise is vastly different between the two. The reason for taking multiple images of the same sky is to average out possible cosmic rays, because unless we have multiple photos to compare, we can't tell what's a star and what's a cosmic ray.

There are several types of calibration frames, of which the relevant ones for us to recognise are bias frames, dark frames, and flats.

Bias frames are frames taken with zero exposure time and with the camera shutter closed, essentially just doing a readout and not taking an image. One usually takes quite a few of these bias

frames (usually 10 or more) and averages them out, as they're pretty quick to take, and you can get an accurate estimate of the bias without the readout noise. An example of a single bias frame can be seen in Figure 6 where the bias level was exaggerated for effect.

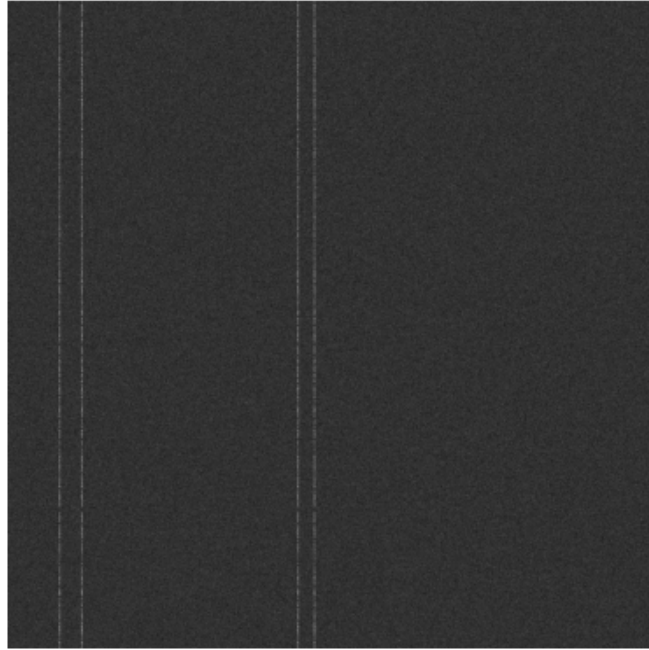


Figure 6: A bias frame with clearly defective columns. In reality, these columns wouldn't be as significant as shown, but they are still common, albeit weak. The fuzzy background is due to read noise.

Dark frames are also taken with the shutter closed, but they're taken for the same amount of time that you plan on imaging the object of interest, meaning that if you're planning to photograph something with an exposure time of 60 seconds, you should ideally have taken a 60-second dark frame. If cooled enough, such as for professional astronomy, dark currents are essentially zero, but if not, they're pretty important. You can also get away with taking shorter dark frames by assuming that the dark current grows linearly with time, although this doesn't always hold. Again, if you can, you should take multiple dark frames and average them out.

You can notice in Figure 7 that the dark frame also contains the bias columns and is therefore also affected by the bias. As such, when calibrating, we can opt to use multiple dark frames, and averaging out enough of them will rid us of the bias, just like with the bias frames.

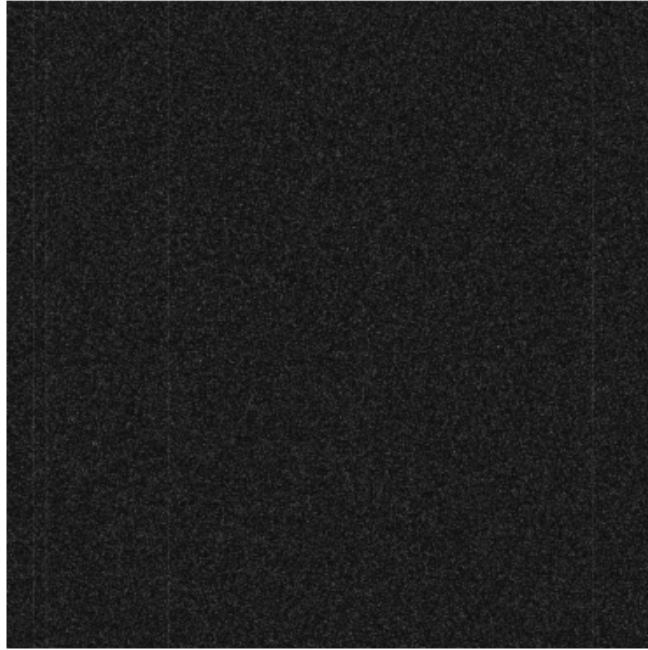


Figure 7: A simulated dark frame containing both read noise, bias and dark noise.

Finally, flat frames are exposures taken of evenly illuminated light sources, such as the twilight sky or some sort of lamp, and they're used to ascertain pixel sensitivity as well as to take into account any possible dust particles that may be on the optics. An example flat frame is shown in Figure 8. Flat frames will vary for any colour filter we might use, since the different pixels will have different sensitivities to different wavelengths.

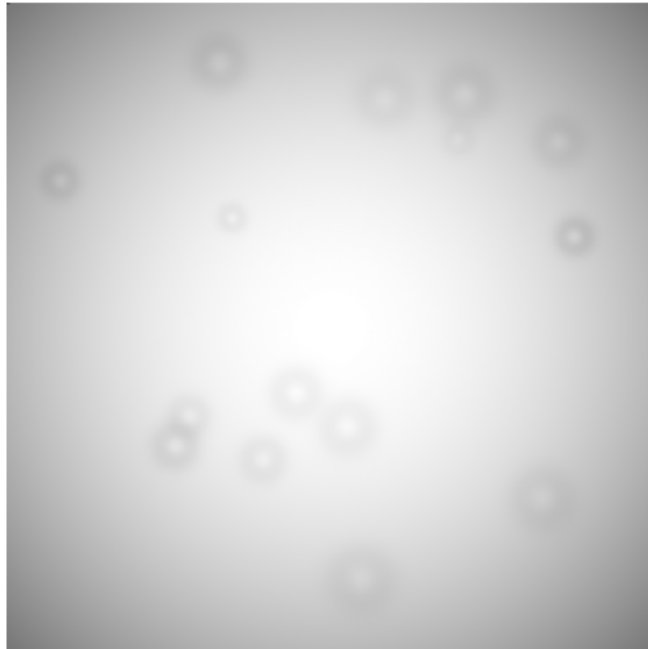


Figure 8: A simulated flat frame with clearly visible vignetting (darker corners, brighter centre) and quite a few dust particles shown as donut-like shapes on the frame.

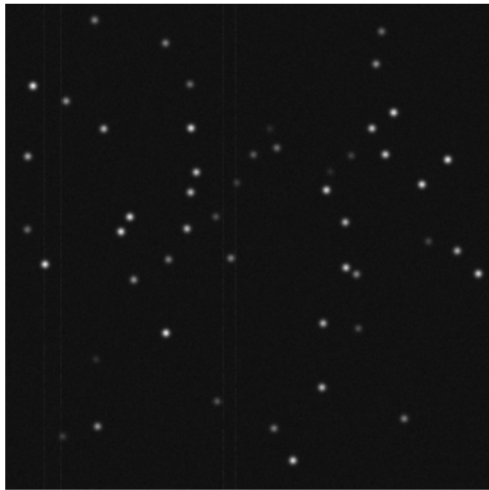
An important thing to note is that the flat frame is not subtracted from the final image, rather

divided by it. The sensitivity variations reduce the amount of light *reaching* the pixels, meaning that it only affects photons coming from external sources. In other words, it doesn't affect anything produced by dark noise, but since we are taking a longer exposure of a light source, dark noise is contained within a flat frame. In essence, when we use calibration to get rid of our noise, the final image will be:

$$\text{image} = \frac{(\text{raw_frame} - \text{bias_frame}) - (\text{dark_frame} - \text{bias_frame})}{(\text{flat_frame} - \text{bias_frame}) - (\text{dark_frame} - \text{bias_frame})}$$

If the dark current is negligible, as it is in professional settings with extremely cold operating environments, we don't need dark frames at all and can just use bias and flat. If the dark frame exposure is the same length as the raw exposure, no bias frames are needed. However, if we've taken shorter dark frames, the bias must be subtracted from those first, after which they must be scaled to the correct exposure time, assuming a linear growth in dark current.

Finally, Figure 9 shows examples of simulated images of stars in an open cluster. Again, the noise levels in these are quite exaggerated, but are just there to prove a point. Even exaggerated, they aren't really visible all that well to the naked eye, but they would have a significant effect on any sort of photometric analysis.



(a) A shorter exposure photograph of the region of interest, short enough that the bias columns are still visible.



(b) A longer exposure photograph of the region of interest. Can you notice the cosmic rays that have snuck in?

Figure 9: Two images taken of the same region of the sky one after the other, one being with longer exposure time, causing a couple of cosmic rays to appear looking like stars not present in the shorter exposure.