PHSX815_Project2: Modeling instrumental noise and astronomical signals

Ryan Low

March 2021

1 Introduction

Modern astronomy relies on Charged Coupled Devices (CCDs) and other such imaging sensors for recording astronomical data. All of these technologies rely on photons exciting the electrons in some semiconducting material. Counting those electrons becomes a proxy for the number of photons detected. Because of this, recording astronomical data is a counting problem, and thus we can expect the number of photons recorded on a CCD to be distributed as a Poisson distribution. As with all electronic measurements, we must also be aware of sources of noise. Since the noise appears in our counts, we can also expect it to be distributed as a Poisson distribution. However, the rate at which noise occurs, λ_{noise}, and the rate at which photons fall onto the detector, λ_{star}, depend nontrivially on other confounding factors. We model detector noise and atmospheric seeing to determine when we can distinguish between an astronomical signal and a noise source.

2 Problem Statement

While we are able to characterize the average noise on a detector, that doesn't mean we will know λ_{noise} exactly throughout the entire observation. For instance, changing dome conditions or poorly maintained equipment can cause nontrivial changes to λ_{noise}. For our present purposes, we will model how the noise characteristics of the detector vary with temperature. Because we are dealing with a semiconducting system, how the electrons are distributed in energy depends on the Fermi-Dirac distribution (Equation 1).

$$P(E) = \frac{1}{1 + \exp\left(\left(E - E_F\right)/k_B T\right)} \tag{1}$$

For silicon, the band gap is about $1.12\,eV$ and the Fermi energy is approximately half of the band gap energy. Using this distribution, we can model the number of noise electrons that we count, which in turn gives us a distribution for λ_{noise}.

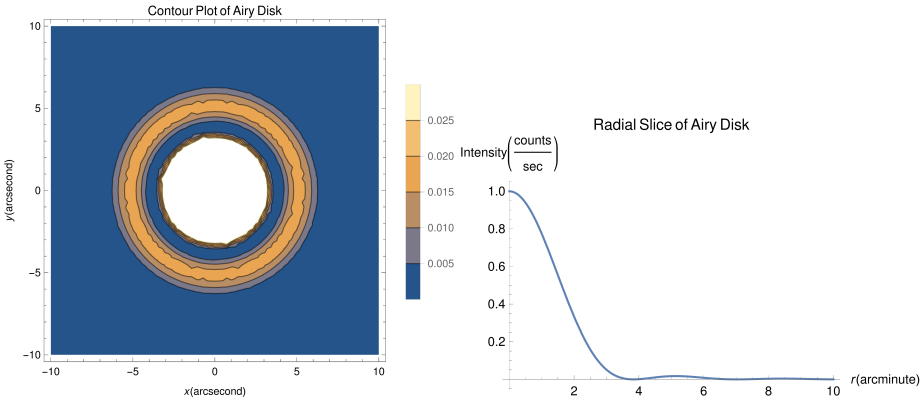

Figure 1: (Left): An analytical contour plot of an Airy Disk. (Right): The Airy disk as a function of radius.

Suppose our detector is a single pixel. This is a small detector area. Because of this, we are collecting light from an area of the sky of fractions of an arcsecond. The starlight from space must pass through the atmosphere, where thermal variations cause the light to stochastically refract. This random refraction, known as atmospheric seeing, causes the apparent position of the star to rapidly change on the sky, which is why the stars twinkle. While we can measure the average seeing for a night of observation, it is a stochastic process and therefore causes the apparent position of an object to vary over an exposure. With our small detector area, seeing will become a significant factor the photon counts we measure. A point source of light passing through a circular aperture produces an intensity pattern given by the Airy disk (Equation 2), where I_0 is the maximum intensity of the source, $J_1(x)$ is the Bessel function of the first kind, q is the radial distance from the observation point to the optical axis, R is the distance from q to the aperture, a is the radius of the aperture, and λ is the wavelength of light.

$$I(\theta) = I_0 \left(\frac{2J_1\left(2\pi a q \lambda^{-1} R^{-1}\right)}{2\pi a q \lambda^{-1} R^{-1}} \right)^2 \tag{2}$$

For simplicity, we will consider an observational configuration such that $2\pi a \lambda^{-1} R^{-1} = 1$, that is the prefactor of q is unity. Analytical plots of Equation 2 are presented in Figure 1. By sampling this intensity pattern at random positions, we can model λ_{star} changing as the image of the star moves across the detector.

3 Algorithm Analysis

For both models, we must eventually perform hypothesis testing. In principle, that would mean integrating over all nuisance parameters. In practice, we perform Gibbs sampling. In Gibbs sampling, this integration is approximated by

taking intermediate samples of these nuisance parameters according to their own distributions. Then, using those samples, we generate samples of our rate parameters and feed those rate parameters into a Poisson distribution. This final set of Poisson-distributed samples is a simulated set of data with these nuisance parameters integrated out. We use this when constructing the simulated data for the Log-Likelihood Ratio (LLR) of each model.

So that our computation time remains reasonable, we use `numpy` methods wherever possible. This includes both array operations and random number generation. `numpy` methods are faster than their pure-Python counterparts because they pass execution to an underlying `C` implementation. Since compiled `C` code is much faster than interpreted Python code, using `numpy` affords greater computational speed, which allows us to get away with some less efficient methods.

In our noise model, a noise electron is counted if it is excited into the conduction band. For our purposes, this occurs when the electron's energy is above the Fermi level. The probability that an electron has this energy is

$$P_{detected} = \int_{E_f}^{\infty} P\left(E\right) dE$$

We can easily perform this integral numerically, and do so using Monte Carlo integration. Since the tail probability in Equation 1 is extremely small, it is sufficient to just integrate up to a reasonable upper bound, in our case $\left(E - E_f\right)/k_B T = 1$. Once this probability is calculated, we can produce a uniformly-distributed number from 0 to 1 for each free electron in the pixel and decide whether the electron is excited or not. The total number of excited electrons is our noise rate parameter, λ_{noise}. For our purposes, we will assume that the number of free electrons is fixed. How the resulting distribution of λ_{noise} looks is presented in Figure 2.

To take samples from Equation 2, we perform Metropolis-Hastings sampling. To do so, we randomly walk around the parameter space. The x and y positions are randomly sampled from a Gaussian distribution centered on the previous x and y position and with a standard deviation equal to the mean seeing. We then evaluate Equation 2 at the proposed position and the old position and take their ratio, forming the acceptance probability. We do not need to obtain the Gaussian probability since they are symmetric with respect to the old position and the new position. By generating a uniformly distributed random number, we accept or reject the new position. This resulting Markov Chain gives us many samples of position on the Airy disk. An example of this sampling compared to its analytical counterpart is presented in Figure 3. Evaluating 2 at a position gives an intensity value, which gives us a sample of λ_{star}. How the resulting distribution of λ_{star} looks is presented in Figure 4.

Finally, we must calculate the LLR. Since we do not have the analytic form of the probability of seeing an event, we must calculate that numerically. To do so, we use our simulation results to form histograms. Normalizing the histograms by their integrals turns the bin heights into probabilities. Using these numerically-

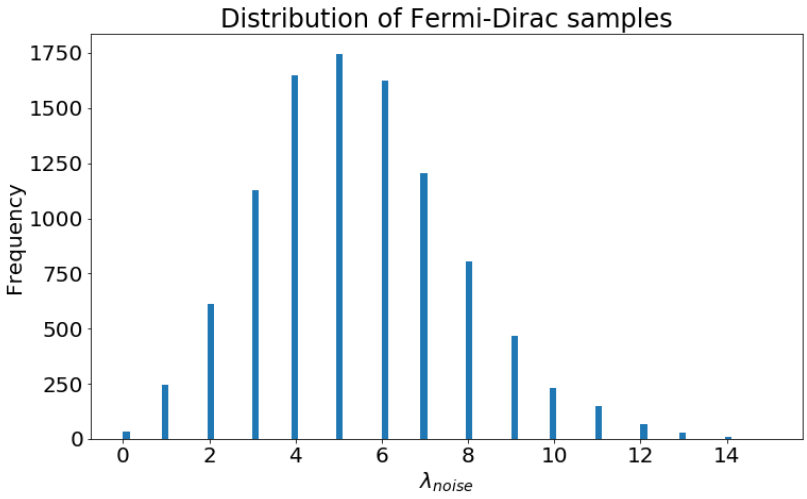

Figure 2: Distribution of λ_{noise} following our algorithm.

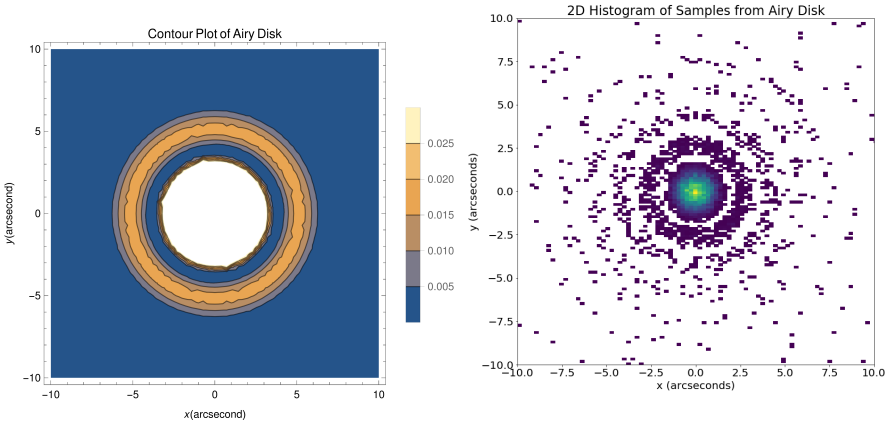

Figure 3: (Left): An analytical contour plot of an Airy Disk. (Right): Position samples from an Airy Disk. Notice how the band structure is retained.

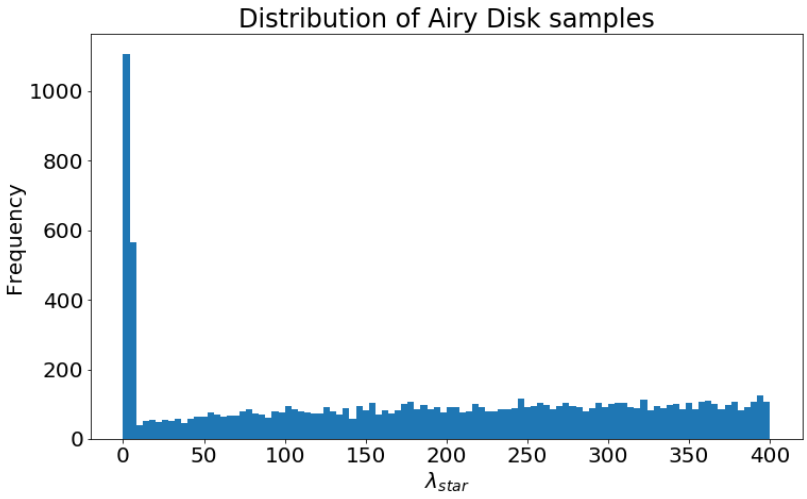

Figure 4: Distribution of λ_{star} using the Airy Disk.

calculated probabilities, we can then calculate the LLR as usual (Equation 3).

$$\text{LLR} = \sum_{i}^{N} \log\left(\frac{P\left(x_i|H_0\right)}{P\left(x_i|H_1\right)}\right) \tag{3}$$

We consider the null hypothesis, H_0, to be the case where we only detect noise. The alternative hypothesis, H_1, is when we detect some signal. We want to reject the null hypothesis when we are observing an astronomical object.

4 Results

We simulate a star with mean intensity of 100 counts per second, mean seeing of 1 arcseconds, and a detector temperature of $300\,K$. For both models, we simulate 100 measurements over 100 experiments. From our simulations, we were perfectly able to distinguish between the signal and the noise. The LLR plot is presented in Figure 5. Typically, the detector is cooled to liquid nitrogen temperatures, so around $77\,K$. Using that temperature, we find that the distributions are still distinguishable (Figure 5). Surprisingly, the two distributions are closer together than they were when the temperature was higher. This is due to the relatively large number of λ_{star} samples near zero, due to how the Airy disk rapidly decays away from the origin. Despite this, it appears that atmospheric seeing does not significantly impede our ability to distinguish a signal source from noise.

We also repeated the analysis for a star with mean intensity of 20 counts per second with the same parameters as above (Figure 6). The two distributions

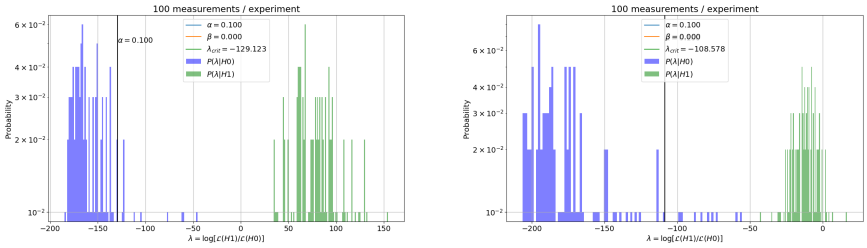

Figure 5: Log-Likelihood Ratio for a star with $I_0 = 100$ counts/second and detector at (left) $T = 300\,K$, (right) $T = 77\,K$.

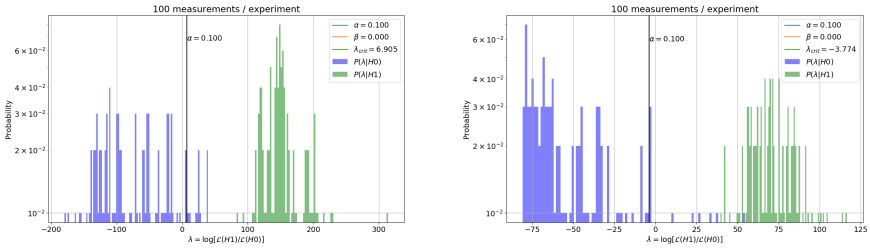

Figure 6: Log-Likelihood Ratio for a star with $I_0 = 20$ counts/second and detector at (left) $T = 300\,K$, (right) $T = 77\,K$.

are just as distinguishable.

However, we must recognize the limitations of this model. In this model, we assumed a single pixel with a very small detecting area. Real detector arrays can have a field of view of several arcminutes. The image moving off of one pixel, with its small detecting area, doesn't affect the measurement as much since the image will move onto a different pixel. Thus, seeing affects our ability to resolve an image rather than detect it. Another major limitation is with our noise model. We assumed that the number of free electrons in the pixel was fixed. This is not realistic, since the number of free electrons in a pixel depends on the doping process, which cannot be controlled exactly.

5 Conclusions

Realistically, this kind of signal and noise modeling does not need to be performed for day-to-day astronomy. We have seen that it is very easy to distinguish between instrumental noise and even a faint signal. It is reasonable enough to characterize the mean seeing and the mean noise to account for their effects. However, this kind of modeling is necessary when designing new instrumentation for revolutionary new observational design or when we are planning observations near the extremes of our sensitivity limits. The methods used in this analysis may be extended to consider two sources in the detection area or for radically different intensity patterns. Those extensions would allow us to model observing close binary systems or observations using a more complicated instrument such as an interferometer respectively.