CCD Laboratory Evaluation
Observational Techniques for Astronomy (AST 6725)

Alyssa Bulatek, Michael Estrada, Rachel Losacco, Sheila Sagear

October 15, 2020 (revised December 15, 2020)

1 General Objectives

Though it is generally not true that a carpenter is only as good as the tools they have at their disposal, it is important for craftspeople to know the characteristics of their tools well before they begin to use them in earnest. The same adage rings true for the astronomer. In order to understand our data, we must understand the tools we use to take our data. In particular, we would like to know how our CCD translates light we see from an object we are observing into electronic signals that we can read. There are four particular characteristics that we focus on in this lab: read noise, dark current, gain, and linearity. All of these characteristics affect the way that the CCD records data, and so the particulars of each effect are important to understand. Each characteristic is discussed in more detail in the sections that follow.

2 Experimental Setup

Sarik ran the lab for us from his office. The CCD was housed within a cardboard box for most of the lab. The box had oval-shaped holes as handles, but they were obscured by a computer case and the wall for most of the lab. When the CCD needed to be illuminated, the box was removed and one (or two) computer monitors were placed the CCD to allow light into the shutter. Sheila ssh'ed into the computer to do the actual data acquisition while the rest of us joined in on the Zoom call.

3 Read Noise

Objective: Our first objective was to estimate the read noise of the CCD, or the statistical noise generated by the output circuitry in the detector while it is not being illuminated.

Methods: In order to estimate the read noise of the detector, we took nine consecutive dark exposures, each 0.1 seconds long. In Python, we stacked these exposures in an array and calculated the RMS of each pixel value across the nine images. **In actuality, we used the standard deviation, but I will use the terms interchangeably in this writeup.** The mean (and median) values of this RMS array should represent the average read noise of the detector in DN, and the RMS value of the RMS array should represent the statistical spread in that average read noise value. When we divide the RMS by the square root of the number of pixels on the detector, we get the uncertainty on the average read noise value. In the data processing process, I called this the "array averaging" method for measuring the read noise.

A more statistically robust way of estimating the read noise and the spread in the read noise is to assume that the read noise values as measured per pixel lie in a Gaussian distribution about some mean value and with some statistical spread around that value. So, we can fit a Gaussian curve to a histogram of the values in our RMS array to get another estimate of the read noise and a value which represents the spread of values around our average estimate. The uncertainty in the average read noise for the Gaussian fit method can also be computed in a similar way to the "array averaging" method, by dividing the RMS value by the square root of the number of pixels.

We also took six more dark exposures with longer exposure times (three at 1 second and three at 10 seconds, which are 10 times and 100 times our original exposure time respectively) to determine if there was a measurable difference in read noise with exposure time. We estimated the read noise via both the "array averaging" method and the Gaussian fit method, comparing our average read noise value, RMS, and uncertainty to the 0.1 second exposure values.

We could attempt to determine is a light leak in the CCD housing by taking several dark exposures of equal exposure time, with half of them taken under normal conditions and the other half taken when a bright light was being shined onto the detector (the room lights or a desk lamp could work). If the median pixel value of the median-combined "bright" frames is significantly higher than the median pixel value of the median-combined "dark" frames, then there may be a light leak. We did not test this due to time constraints.

Results: Our initial estimates for the average read noise of the detector as well as the statistical spread and uncertainty on that value are presented in Table 1, denoted with "array averaging" in the method column. The mean value of our RMS array was 25.376 DN and the median value was 25.078 DN; averaging these two values together, we get an average read noise of 25.227 DN. The RMS of the RMS array, which represents the statistical spread in the average read noise value, is 7 DN. The uncertainty in the average read noise, or the RMS value divided by the square root of the number of pixels (8,487,264), is 0.002 DN. **Our average read noise as measured by the array averaging method from 0.1 second exposures is 25.227 $\pm$ 0.002 DN.**

Our estimates for the average read noise using the Gaussian fitting method are also presented in Table 1, denoted with "Gaussian fit" in the method column. Our fit to the histogram of the values in our RMS array is shown in Figure 1. After performing the fit, we got a mean value of 24.426 DN for the average read noise and 6 DN for the statistical spread in the histogram. We computed the uncertainty in the average read noise by dividing the RMS value by $\sqrt{8487264}$, which gives us 0.002 DN. The read noise value we estimated using the Gaussian fitting method is slightly lower than the read noise from the array averaging method. The two average read noise estimates do not fall within their respective measurement uncertainties. **Our average read noise as measured by the Gaussian fitting method from 0.1 second exposures is 24.426 $\pm$ 0.002 DN.**

The estimated average read noise, statistical spread, and uncertainty for our longer exposure times (1 second and 10 seconds) are also noted in Table 1, with their exposure times noted in the final column. We measured these values using both the array averaging method and the Gaussian fit method, and we note that the Gaussian fit method revealed noticeably poor results. The average read noise estimates are lower for these longer exposure times (19.449 DN and 17.693 DN for the array averaging and Gaussian fit measurements for the 1 second exposures, and 19.556 DN and 17.795 DN for the array averaging and Gaussian fit measurements for the 10 second exposures), and the statistical spreads on those estimates are also higher (11 DN for the array averaging results, and 10 DN for the Gaussian fits). Consequently, the uncertainties are higher on the longer exposure time read noise estimates. The mean of the spread of read noise values for each of the longer exposure times is also lower, and the distributions seem less Gaussian, as we can see in Figures 2 and 3. This is unusual (we would expect the read noise to increase with exposure time), and we do not understand the cause. We speculate on the causes for these results in the next subsection. **Our average read noise as measured by the array averaging method from 1 second exposures is 19.449 ± 0.004 DN, while it is 17.693 ± 0.004 DN as measured by the Gaussian fit method for the same exposure time frames. For 10 second exposures, we found these values to be 19.556 ± 0.004 DN and 17.795 ± 0.004 DN respectively.**

Characteristic	Value	Method (exposure time)
Average read noise	25.227 DN	array averaging (0.1 s)
Statistical spread in average read noise (σ)	7 DN	array averaging (0.1 s)
Uncertainty on average read noise	0.002 DN	array averaging (0.1 s)
Average read noise	24.426 DN	Gaussian fit (0.1 s)
Statistical spread in average read noise (σ)	6 DN	Gaussian fit (0.1 s)
Uncertainty on average read noise	0.002 DN	Gaussian fit (0.1 s)
Average read noise	19.449 DN	array averaging (1 s)
Statistical spread in average read noise (σ)	11 DN	array averaging (1 s)
Uncertainty on average read noise	0.004 DN	array averaging (1 s)
Average read noise	17.693 DN	Gaussian fit (1 s)
Statistical spread in average read noise (σ)	10 DN	Gaussian fit (1 s)
Uncertainty on average read noise	0.004 DN	Gaussian fit (1 s)
Average read noise	19.556 DN	array averaging (10 s)
Statistical spread in average read noise (σ)	11 DN	array averaging (10 s)
Uncertainty on average read noise	0.004 DN	array averaging (10 s)
Average read noise	17.795 DN	Gaussian fit (10 s)
Statistical spread in average read noise (σ)	10 DN	Gaussian fit (10 s)
Uncertainty on average read noise	0.004 DN	Gaussian fit (10 s)

Table 1: A summary of our estimated read noise for the detector along with measures of statistical spread and uncertainty. Our measurements for the read noise and related uncertainties for different exposure times (0.1, 1, and 10 seconds) are shown along with the results we got from using Gaussian fits to measure the average read noise value more robustly.

Summary: We estimated the read noise of the detector as well as the statistical spread in the read noise and its uncertainty using two different methods, one slightly more statistically robust than the other. Using the "array averaging" method, where we calculate statistics from an RMS array of nine 0.1 second dark exposures, we estimate an average read noise of 25.227 ± 0.002 DN. Using the "Gaussian fit" method, where we fit a Gaussian curve to the distribution of pixel values in the RMS array, we estimate an average read noise of 24.426 ± 0.002 DN. These values are not within their respective measurement uncertainties. We also estimated the average read noise using 1 second and 10 second exposure times, but got anomalous results (we measured a lower-on-average read noise for these longer exposure times, contrary to what we would expect). Though we do not know the cause of these anomalous results, we can speculate that they may have occurred because of a faulty CCD or because we average together fewer exposures for the longer exposure times (three) than for the shorter exposure time (nine). The latter explanation could also explain why our distributions of read noise estimates for the longer exposure times were non-Gaussian.

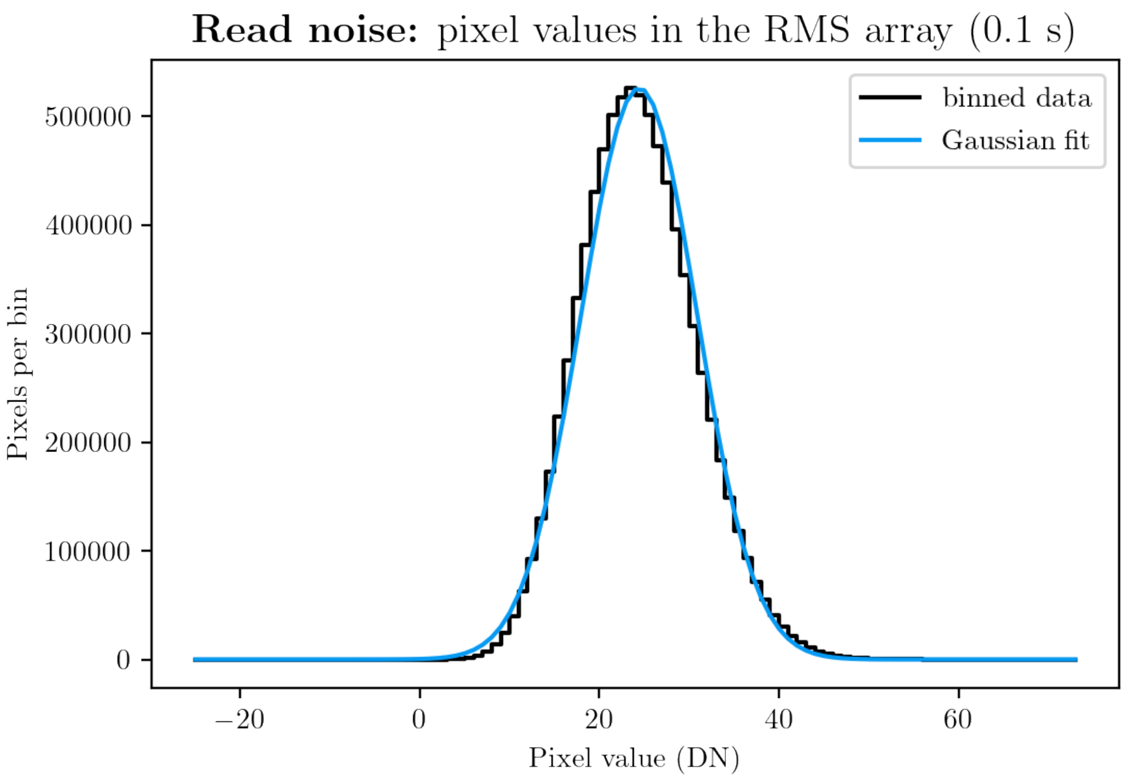

Figure 1: A histogram of the read noise measurements for the 0.1 second exposures with a smooth Gaussian fit to that histogram overlaid. We used the best-fit μ (mean) and σ (statistical spread) values from the Gaussian fit to measure the read noise and uncertainty via this method.

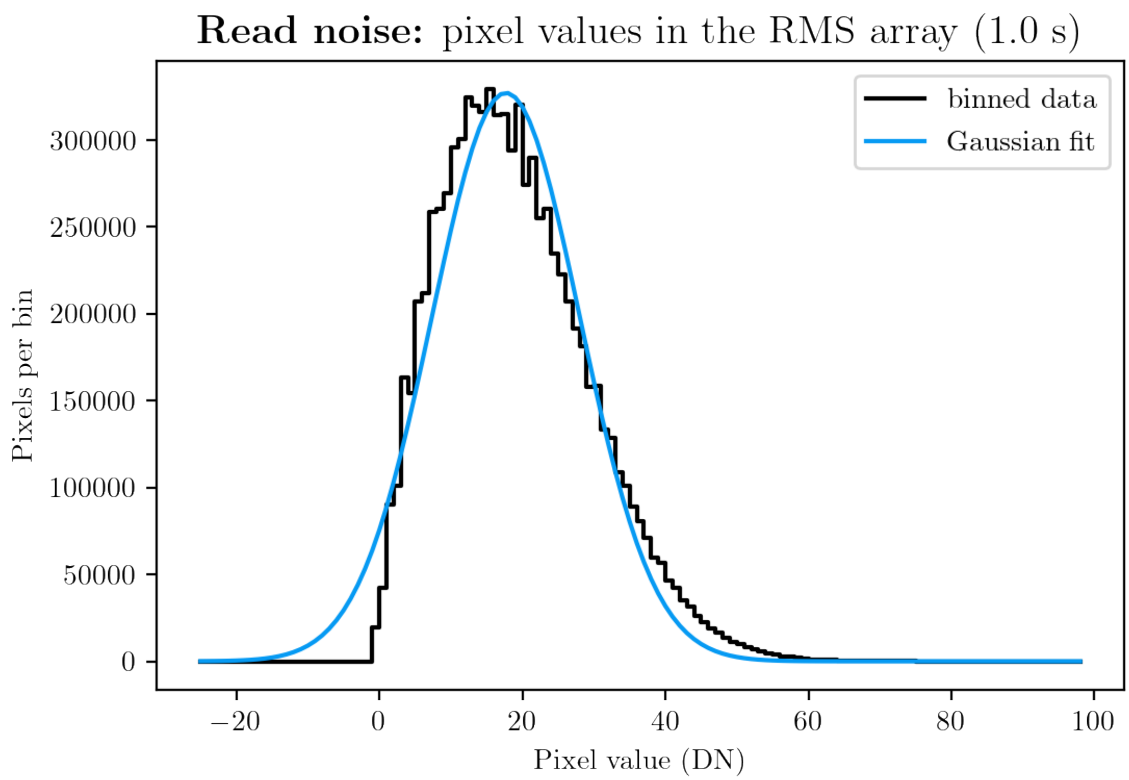

Figure 2: A histogram of the read noise measurements for the 1.0 second exposures with a smooth Gaussian fit to that histogram overlaid. We used the best-fit μ (mean) and σ (statistical spread) values from the Gaussian fit to measure the read noise and uncertainty via this method. The underlying data distribution is not Gaussian. Instead, the data resemble a Poisson distribution more closely.

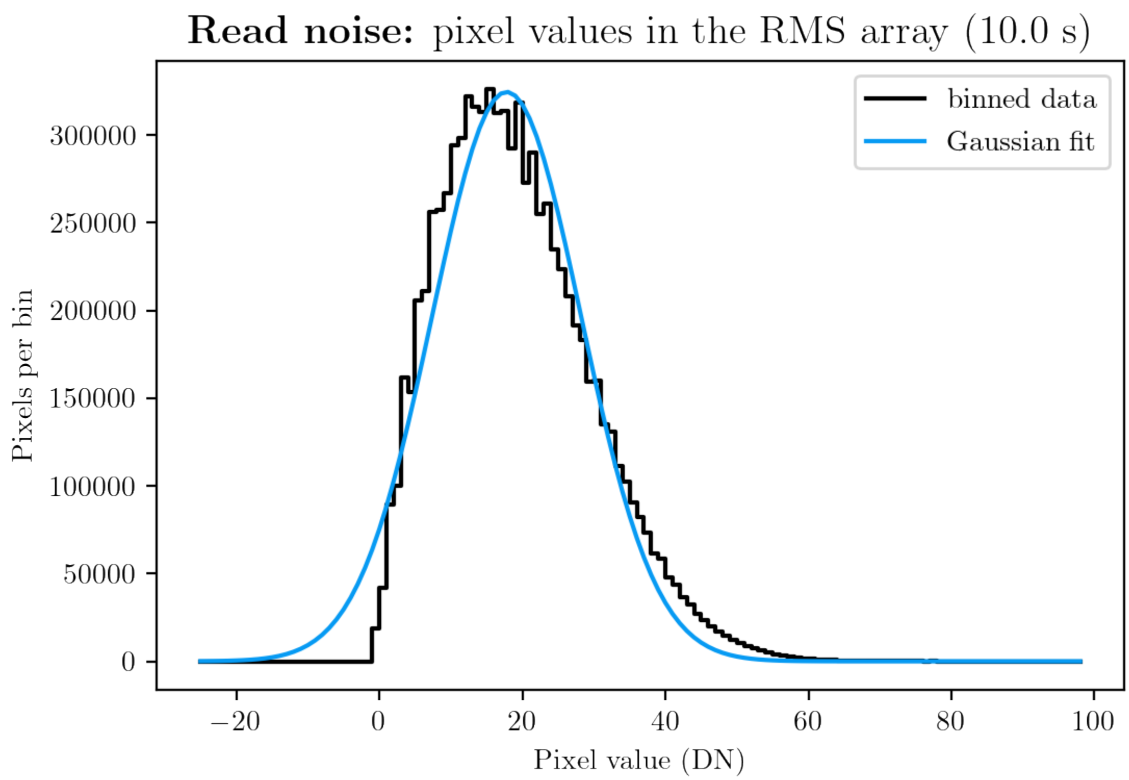

Figure 3: A histogram of the read noise measurements for the 10.0 second exposures with a smooth Gaussian fit to that histogram overlaid. We used the best-fit μ (mean) and σ (statistical spread) values from the Gaussian fit to measure the read noise and uncertainty via this method. The underlying data distribution is not Gaussian. Instead, the data resemble a Poisson distribution more closely.

4 Dark Current

Objective: Our next objective was to measure the dark current in the CCD, or the rate at which the detector builds up signal while unilluminated due to the thermal properties of the detector.

Methods: To measure the dark current in the CCD, we need several exposures at progressively longer and longer exposure times. We used five exposure times: 1, 5, 25, 125, and 500 seconds. Between each of these exposures, we took short (0.2 second) bias frames, but we did not end up incorporating those images into our analysis.

By taking images at increasingly larger exposure times, we can measure dDN/dt, or the slope that governs the (hopefully) close to linear relationship between the number of counts we get at each exposure time and the exposure time itself. This slope will be our value for the dark current.

For this calculation, we also used both an "array averaging" and a "Gaussian fit" method to measure the mean pixel value in each image. In the "array averaging" method, we calculated the mean and standard deviation of each frame at each exposure time. We can then perform a linear fit on the data and use the value of the slope as our measured dark current.

We can be more statistically robust by using a "Gaussian fit" method to calculate the mean and statistical spread of the pixel distributions at each exposure time. We repeated this analysis with the Gaussian fit method, fitting a line to the data and using its slope as the measured dark current value.

Results: Our measured values for the dark current using both measurement methods are reported in Table 2. These measurements were taken when the CCD was at an approximate temperature of 0.02 °C averaged across the five frames. The temperature for these frames took one of two values, either 0.189 °C or −0.232 °C, so it is possible that these two temperature steps were the closest the temperature measurement device on the CCD could get to zero. We can compare our dark current value to the manufacturer-reported value of 0.4 ADU/s at 0 °C (as found by Nazar).

For the array averaging method, our data points represent the average signal value at a given exposure time. The uncertainty on this value is calculated by dividing the standard deviation of the signal values in each frame by the square root of the total number of pixels. This results in the error bars being smaller than the data points in our resulting plot. Their values are as follows, sorted from smallest exposure time to largest: ± 0.03 DN, ± 0.01 DN, ± 0.01 DN, ± 0.07 DN, ± 0.2 DN. In order to measure the slope of these data points, I used a Python package called `lmfit`, which allows for non-linear least-squares model fitting to data. I applied a linear least-squares fit to the data and got a fit with a slope of 0.266 ± 0.004 DN/s. The uncertainty on this slope takes into account the individual uncertainties on each data point. The chi-squared value for this fit is 0.037, and with three degrees of freedom, the reduced chi-squared statistic is 0.012. This reduced chi-squared value being less than one indicates that the model is over-fitting the data. This is not a good fit. I suspect that the CCD being faulty could play a part in this nonlinear behavior. The average pixel values as a function of exposure time, along with their uncertainties, are plotted on Figure 4. Our line of best fit and the boundaries of a 3σ spread on both the slope and intercept of that fit are also shown. **Using array averaging, we derive a value of 0.266 ± 0.004 DN/s for the dark current in the detector. The linear fit we used for this yields a chi-squared statistic of 0.037 (reduced chi-squared 0.012), which indicates that the model is overfitting the data.**

For the Gaussian fit method, we fit a Gaussian function to the pixel values in the images at each exposure time. We recorded the mean values from each of those fits as well as their uncertainties (calculated by dividing the best-fit sigma parameters from our Gaussian fits by the square root of the total number of pixels). Again, the uncertainties on our average values are smaller than the data points in our plot. Their values are (rounded to one significant figure) 0.01 DN for each exposure time. We followed the same line fitting procedure for the Gaussian fit data as for the array averaging data. The slope we measured was 0.25 +/- 0.01 DN/s. The uncertainty on this slope takes into account the uncertainties on each data point. The chi-squared value for this fit was 0.0067, with a reduced chi-squared statistic of 0.0022, which indicates a poor fit. Again, I suspect this poor fit is due to the CCD being faulty. We show the average pixel values and their uncertainties as measured by the Gaussian fit method in Figure 5. We also show the line of best fit for these data as well as a 3σ uncertainty region encompassing 3σ spreads on both the slope and intercept parameters of the fit. **Using the Gaussian fit method, we derive a value of 0.25 $\pm$ 0.01 DN/s for the dark current in the detector. The linear fit we used to derive this parameter has a chi-squared value of 0.0067 (reduced chi-squared 0.0022), which indicates that our linear model is overfitting the data.**

Characteristic	**Value**	**Method**
Average dark current	0.266 DN/s	array averaging
Uncertainty on average dark current (σ)	0.004 DN/s	array averaging
Average dark current	0.25 DN/s	Gaussian fit
Uncertainty on average dark current (σ)	0.01 DN/s	Gaussian fit

Table 2: A summary of our estimated dark current for the detector along with its uncertainty. We show the measurements we made using both the "array averaging" method as well as the more robust "Gaussian fit" method.

Summary: We estimated the dark current in the detector as well as the uncertainty on estimated value using the same two methods we used to estimate the dark noise. In the first method, we calculated the average signal level in dark exposures of progressively longer exposure time, and calculated a linear fit on those signal values versus exposure time. The slope we measured (0.266 DN/s) represents our dark current measurement (with $\pm$ 0.004 DN/s as our uncertainty). For the second method, we created a histogram of the pixel values for each exposure time, fit a Gaussian to the histograms, and plotted the mean and sigma values of each Gaussian versus their exposure time. We then fit a line to the mean values versus exposure time, and that slope (0.25 $\pm$ 0.01 DN/s) served as another measure of the dark current. The reduced chi-squared values for both of our fits were less than 1 (0.012 and 0.0022 for array averaging and Gaussian fitting respectively), which indicates that the linear model is overfitting the data.

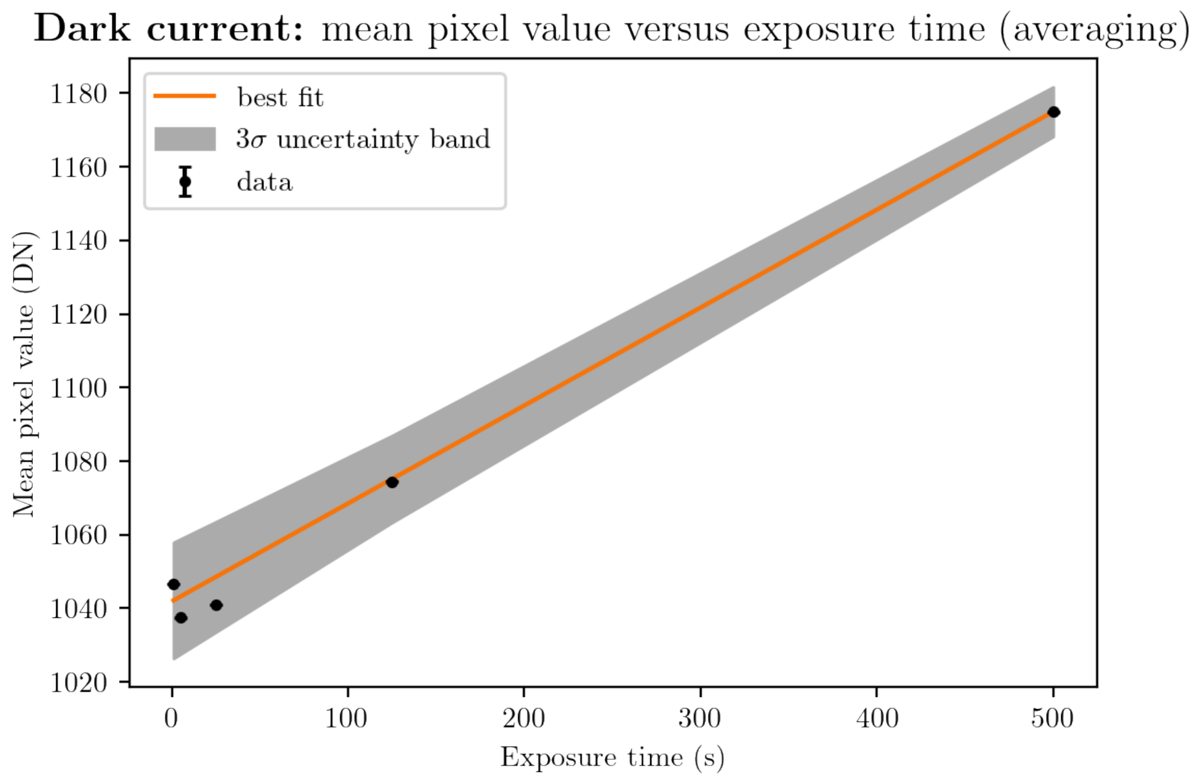

Figure 4: Mean pixel values (measured using the “array averaging” method) for the dark current measurements with progressively longer and longer integration times. The mean values are shown in black dots with error bars representing the uncertainties on the mean values. The error bars are smaller than the data points; their values are given in the text. The orange line is the best linear fit to our data, taking into account the uncertainties. The slope of the line is our dark current value. A 3σ uncertainty on the linear fit, taking into account the uncertainties on both the slope and intercept parameters, is shown as a grey area.

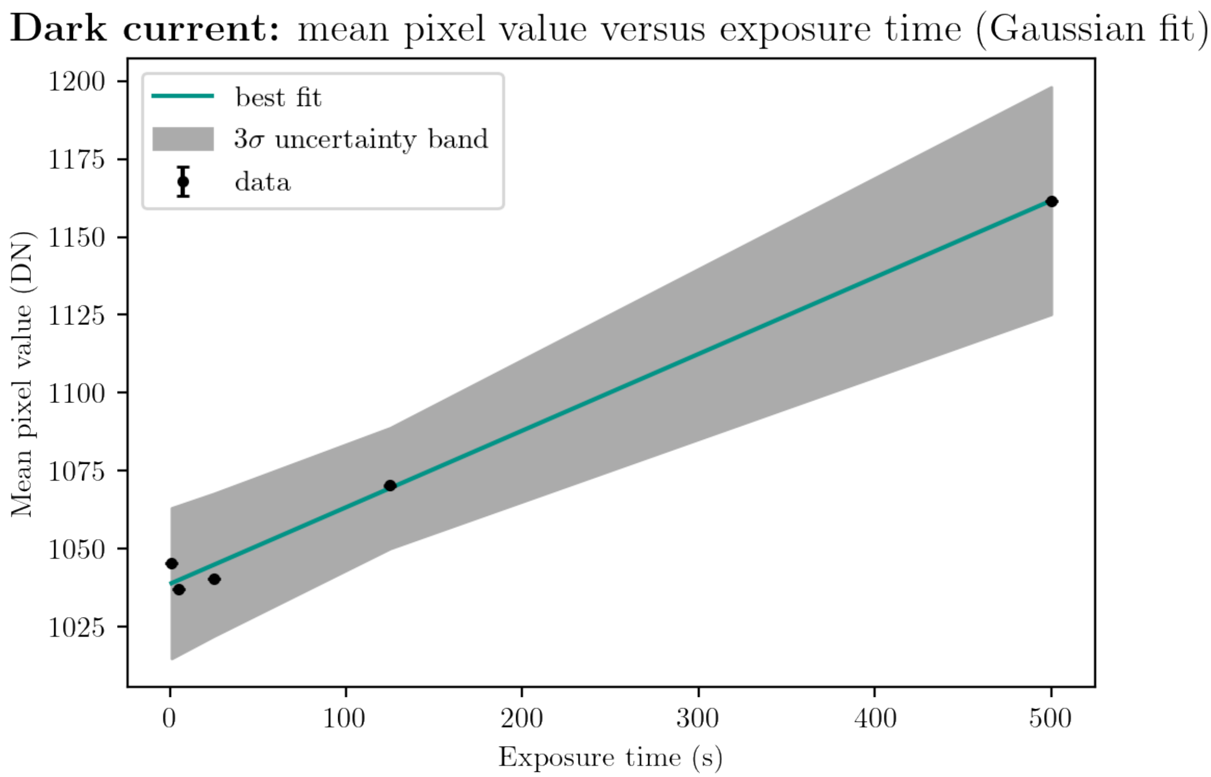

Figure 5: Mean pixel values (measured using the "Gaussian fit" method) for the dark current measurements with progressively longer and longer integration times. The mean values are shown in black dots with error bars representing the uncertainties on the mean values. The error bars are smaller than the data points; their values are given in the text. The teal line is the linear fit to our data, taking into account the uncertainties. The slope of the line is our dark current value. A 3σ uncertainty on the linear fit, taking into account the uncertainties on both the slope and intercept parameters, is shown as a grey area.

5 Gain

Objective: Next, we wished to determine the relationship between the number of counts (DN) measured by the detector in a given pixel and the number of electrons liberated in that pixel, also known as the gain (in units of DN per electron). To do this, we needed a measure of the variance in each pixel as well as the signal, giving us over eight million distinct estimates of the gain.

Methods: As suggested by Sarik, we used Group 4's data for this section since our data were not satisfactory. For this section, Group 4 attempted to uniformly illuminate the detector such that the average pixel value at an exposure time of 100 seconds was about 50,000 DN. They took a series of exposures at increasing exposure times (0.1, 1, 5, 15, 25, 50, 75, and 100 seconds) so they could probe a variety of signal strengths. They took three exposures at each exposure time (though for the 1.0 second and 75.0 second exposure times, they actually took four exposures—I initially only selected the latter three in case something had happened to the first exposure, but I incorporated these two extra images into my analysis for the revised version of this lab).

We imported each of the images at each exposure time and calculated the pixel-wise mean and standard deviation of each pixel across the three images, leaving us with a mean array and a standard deviation array for each exposure time. We then squared the standard deviation arrays to get the variance arrays. Next, we calculated the median of each of the mean arrays to get a measure of the "signal" at each exposure time, and we calculated the median of the variance arrays to get the variance at each exposure time. We plotted these variance values versus signal, and performed a linear fit on those data points. The slope represents our gain in DN per electron.

Results: The results for our gain measurement as well as the uncertainty on the measurement are given in Table 3. I used the same least-squares fitting algorithm discussed earlier to fit a line to the data. The average gain value from the linear fit was 2.2 DN/electron with an uncertainty of 0.2 DN/electron. The chi-squared value on this fit is extremely large: 2.0×10^8 (our reduced chi-squared was 4.8×10^7). This means the fit is extremely poor. Our fit is shown in Figure 6. A systematic effect that could have affected our measurement is the fact that it is unlikely that the CCD was fully illuminated given the laboratory environment it was in. This would cause there to be many dark areas of the CCD, which would draw the average pixel value down artificially. This could result in an underestimate of the gain. Also, we should have taken nine frames at each exposure time instead of three. To try to get a more reasonable gain value, I attempted to use each of the pixels in my fit. However, the code I wrote did not successfully run after thirty minutes, and my computer's fans were running at full blast, so I interrupted the script.

Nazar found a reported value from the manufacturer of 2.78 DN/electron, which is not within the measurement uncertainty of our estimated value. **We estimate a value for the gain of the detector to be 2.2 ± 0.2 DN/electron. However, our linear fit to the signal-variance pairs has an extremely high chi-squared value ($> 10^8$, reduced chi-squared of $> 10^7$), which indicates the data are poorly modeled by a linear fit.**

Characteristic	**Value**	**Method**
Average gain	2.2 DN/electron	array averaging
Uncertainty on average gain (σ)	0.2 DN/electron	array averaging

Table 3: A summary of our estimated gain for the detector along with its uncertainty.

Summary: We (or, really, Group 4) estimated the gain value of the detector by taking a variety of exposures of different signal strengths with the detector illuminated. We then calculated the mean and variance in each pixel at each exposure time, and took the median of each of those values for a given exposure time. Plotting variance versus signal, we fit a line to our eight data points and measured the slope, which gives us an estimate of the gain. We calculated a gain value of 2.2 ± 0.2 DN/electron, which is not within the measurement uncertainty of the gain value reported by the manufacturer of 2.78 DN/electron. Our fit is a poor one, as the chi-squared value for it is greater than 10^8 (reduced chi-squared greater than 10^7).

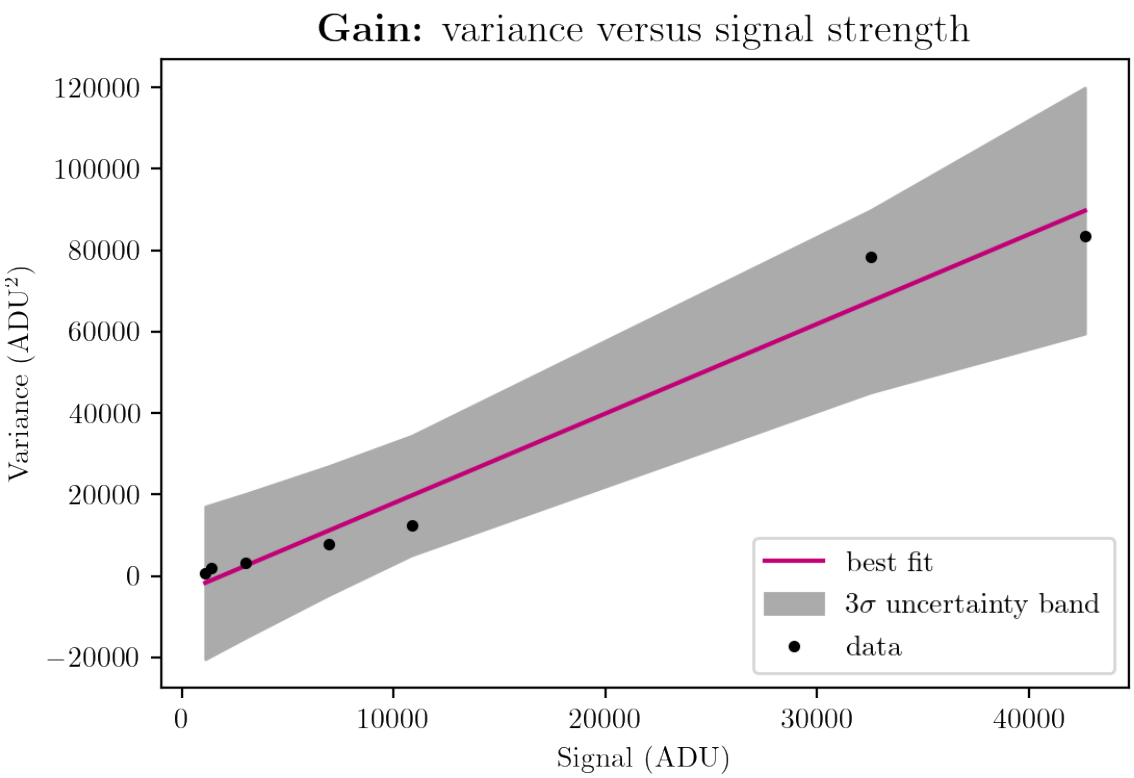

Figure 6: Variance versus signal for a variety of exposure times. The pink line is a linear fit to the data. The slope of the line represents our estimate of the gain for the detector. A band where the fit parameters can vary by 3σ is shown in grey.

6 Linearity

Objective: Finally, we wanted to measure the linearity of the response of the detector as a function of light level. Ideally, the number of counts reported by the detector would be proportional to the number of photons striking the detector. We can see if this is true if we take images of different exposure times that give us a wide range of pixel values while keeping the illumination constant, allowing us to measure the linearity of the detector response as a function of exposure time.

Methods: We also used Group 4's data for this section due to time constraints on the day we took our data. Group 4 used two different illumination setups: the "ambient" setup, in which two computer monitors were used to illuminate the detector, and the "dim" setup, where only one monitor was used. This was used to measure the difference in linear behavior in different light levels. Group 4 took a series of exposures in the ambient light (at 1, 3, 10, 30, 60, 120, and 180 seconds) and in the dim light (at 3, 10, 30, 60, 120, 180, 240, and 300 seconds). We took the median pixel value at each exposure time for each light level. Then, we simply plotted each pixel value as a function of exposure time, fitting a line to the data and qualitatively assessing how linear the data looked.

Results: We plotted the median pixel value for each exposure time for each light level in Figure 7. We can see that for the dim lighting conditions, the relationship between median pixel value and exposure time is described well by a line. For ambient lighting conditions, however, a linear fit does not describe the data well—at longer exposure times, we see a drop in the median value. I used `lmfit` to determine these linear fits. 3σ uncertainty bands are shown for each brightness level. The larger uncertainty band for the ambient brightness level indicates that it is described less well by a linear fit. On the other hand, the dim lighting uncertainty band is not visible in this plot, so a linear model fits those data well. For clarity, the dim lighting data are shown on their own in Figure 8.

Summary: We measured the linearity of the detector response by taking progressively longer exposure times under two different lighting conditions. We note that in dimmer conditions, we see generally linear behavior across all exposure times we took images at. In the brighter conditions, we see deviation from linearity at longer exposure times.

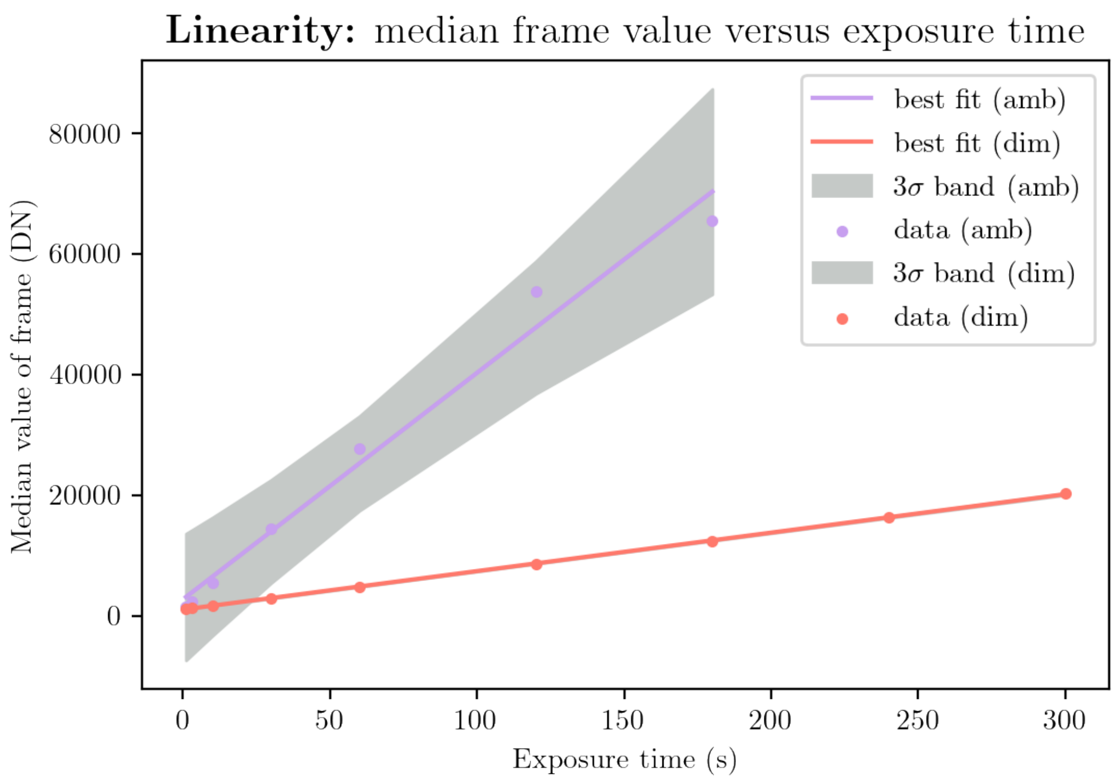

Figure 7: Median pixel value as a function of exposure time for two different light levels. "Ambient" lighting was under the illumination of two computer monitors, while "dim" lighting was under the illumination of a single computer monitor. The ambient lighting conditions show nonlinearity at higher exposure times, as can be seen by the non-ideal linear fit. Conditions are more linear for the dim lighting scenario. The 3σ band is comparable to the width of the line in this plot, so the dim lighting data are shown on their own in Figure 8.

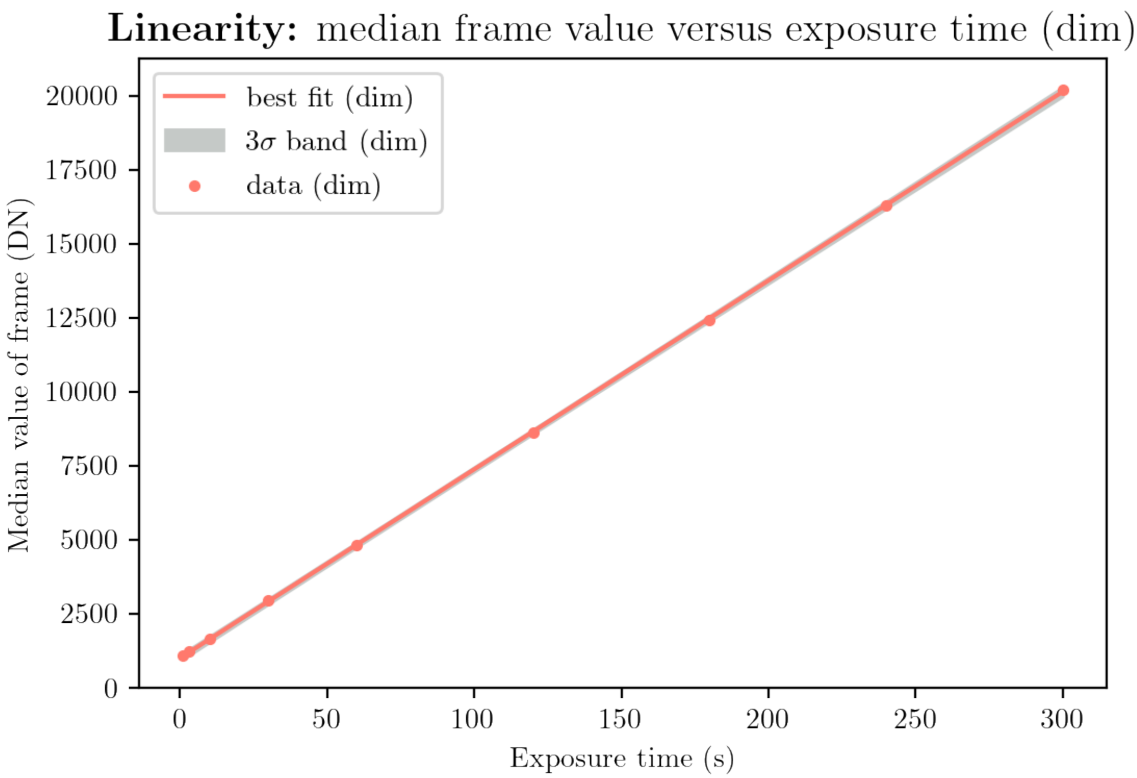

Figure 8: Median pixel value as a function of exposure time for the dim lighting conditions on their own. These data are described well by a linear fit, so the 3σ uncertainty band is very thin.

7 Additional Characteristics

We were unable to measure blooming with the CCD because we did not set up a point source in the lab. Also, we were not able to measure the charge transfer efficiency due to the high quality of the CCD.

8 Conclusion

Table 4 shows a summary of each of the characteristics we measured in this lab, as well as uncertainties on each value. Regarding strange results, we were getting strange results on the first day of lab, likely due to a faulty USB connector on the CCD. It seemed to fix itself on the second day of lab, although we ran out of time and had to use another group's data. An aspect of the lab that could improve our results would be to perform the lab in a windowless room so that changing ambient light conditions from outdoors do not affect our data significantly. Since we were doing the lab around sunset, this may have affected the lighting conditions in the lab over time.

Characteristic	**Method**	**Value**	**Spread (σ)**	**Uncertainty**
Read noise	array averaging (0.1 s)	25.227 DN	7 DN	0.002 DN
Read noise	Gaussian fit (0.1 s)	24.426 DN	7 DN	0.002 DN
Read noise	array averaging (1 s)	19.449 DN	11 DN	0.004 DN
Read noise	Gaussian fit (1 s)	17.693 DN	10 DN	0.004 DN
Read noise	array averaging (10s)	19.556 DN	11 DN	0.004 DN
Read noise	Gaussian fit (10s)	17.795 DN	10 DN	0.004 DN
Dark current	array averaging	0.266 DN/s	—	0.004 DN/s
Dark current	Gaussian fit	0.25 DN/s	—	0.01 DN/s
Gain	array averaging	2.2 DN/electron	—	0.2 DN/electron

Table 4: A summary of our estimates of the read noise, dark current, gain, and linearity of the CCD we will use for our imaging and spectroscopy projects. A dash (—) indicates that the value was not directly measured. This applies to cases in which the uncertainty on a fit was measured directly through a least-squares algorithm.