

# Differential Nonlinearity Analog to Digital Converter Calibration: draft 3

Renee Nichols

April 1, 2024

This work calibrates the analog to digital converters (ADC) on the LSSTCam using a differential nonlinearity informed method. We see that the method developed here allows for informed calibration that falls within the specifications from the manufacturer, including no unbounded integral nonlinearity and differential nonlinearity that does not deviate from expectation. This method provides robust results using two different types of runs from two different portions of electric-optical testing of the LSSTCam. Together it demonstrates the stability of the electronics through time.

## 1 Motivations

Understanding the analog to digital converters allows us to have an understanding of how the electronics are working within the system and thus an understanding of the signal received by the telescope from the sky, which is crucial to all four science themes as put forth by the LSST Science Book [CITE]. Rubin Observatory's LSSTCam uses 18-bit analog to digital converters to record the signal from the bias level all the way to saturation of the silicon charge coupled device (CCD) sensors. We seek to understand how these electronics work across the entire focal plane - over 3024 amplifiers, and verify the specifications set by the manufacturer actually represent the electronics we are using. This is largely motivated by a previous attempt to characterize the ADCs by calculating the differential nonlinearity (DNL) along bins in the ADC resulting in an unbounded integral nonlinearity (INL) that was found over the range of the ADC. There was no other indication these features should be present, suggesting a need for a more accurate calibration to be constructed. In the following sections we will define DNL and INL, examine the original method to calibrate the ADCs, and look at the results of those calibrations in context of the expectation for these electronics.

### 1.1 Definition of Differential and Integral Nonlinearity

The analog to digital converter's purpose is to convert an analog signal to a digital one that can be readout and stored for further use. The output signal is measured in analog to digital units (ADU). The input signal has limits, called edges, which maps input current into the stated ADU value. These edges have a lower bound, called the left edge, and an upper bound, called the right edge. The difference between the right edge and the left edge determines the width of the bin itself, and is determined through the process called calibration of the ADC. As seen in Figure 1, in a perfect ADC (blue), the bins have a width of exactly 1, and are spaced integers apart from one another. Yet in a true ADC (orange), bins may differ from the ideal bin width of 1, and have non integer left and right edges mapping to specific ADU values.

The presence of deviations from an ideal ADC in terms of bin widths, leads to two types of nonlinearities in the ADC called differential and integral nonlinearity. DNL is thought of as the deviation of a bin's width from that of an ideal binwidth. Differential nonlinearity is defined exactly as:

$$DNL = 1 - w \tag{1}$$

where  $w$  is the width of a bin. This value can be computed for any bin along the ADC. Values of the DNL are typically on the order of  $\pm 0.5$ , but any value greater than  $\pm 1$  would indicate a skipped bin. Integral nonlinearity is the deviation of the middle of a true bin from that of an ideal bin. More specifically,

$$INL = m_{ideal} - m_{adc} \tag{2}$$

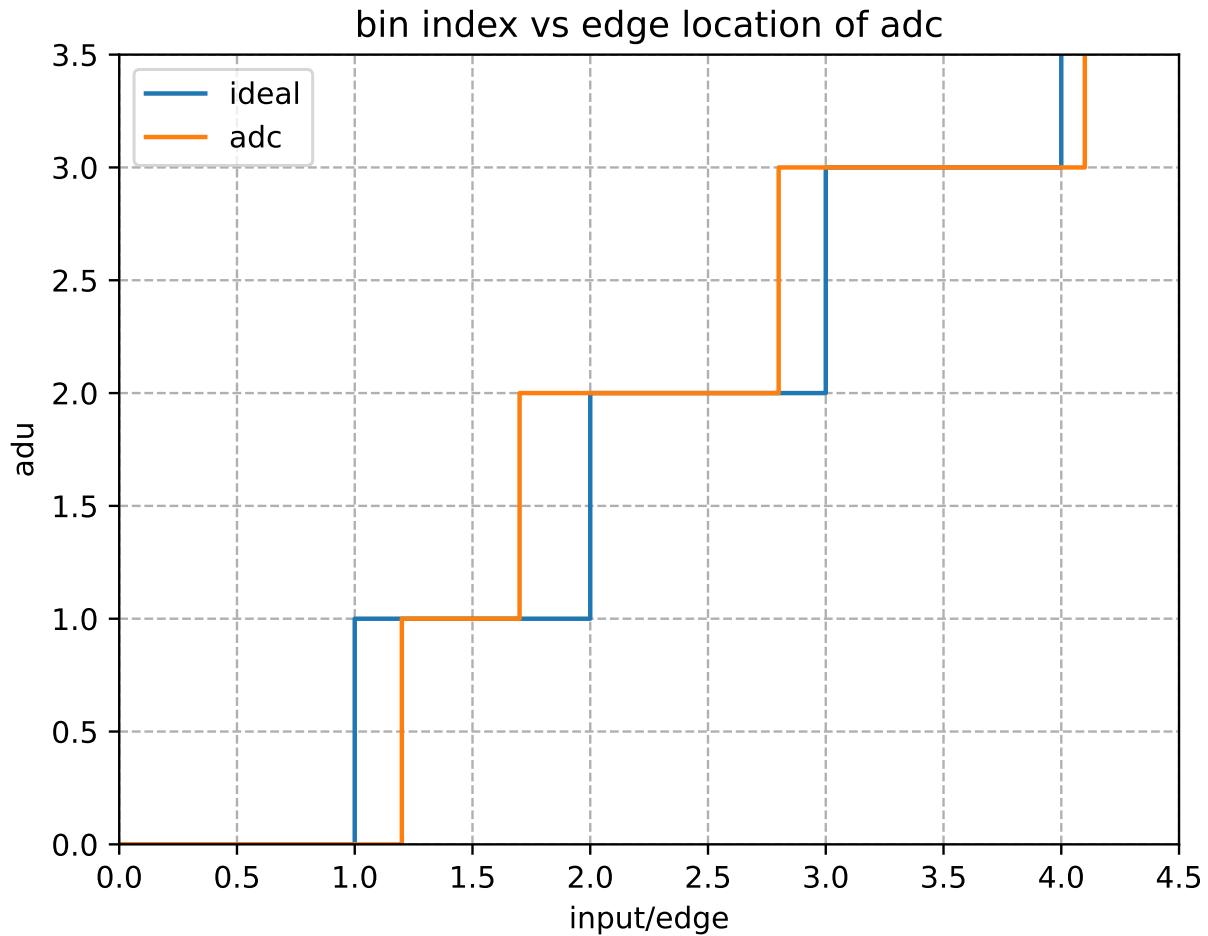


Figure 1: Above is a schematic demonstrating how the analog to digital converter ideally works (blue), and how it truly works (orange). Along the y axis is the ADU value recorded by the ADC, which has integer values. Along the x axis is the edge of the ADC, or the left and right edges of the bin which maps to a given ADU. As tracked along the blue curve, an ideal ADC has bins with integer edges which map to corresponding ADU. The orange curve shows a fabricated true ADC, which has left and right edges that deviate from the ideal, but still maps to a single ADU. The difference in the left and right edges leads to the presence of a differential nonlinearity, a difference between an ideal bin width and a true bin width, and also an integral nonlinearity, a difference between an ideal middle of a bin, and a true middle of a bin.

where  $m_{ideal}$  is the midpoint of an ideal bin, and  $m_{adc}$  is the midpoint of an ADC bin. INL typically yields contributions from each bin on the order of  $\pm 1$ , with deviations above 2 being out of specification [CITE].

In order to understand the signal coming from the ADC, it is important to examine whether any structure of DNL and/or INL is present in the calibrated ADC edges. For instance, a local portion of the ADC with DNL would in turn also have effects on the INL. More cause for concern could arise if INL grows some structure over the range of the ADC without corresponding DNL to match, as this suggests a nonlocal effect is occurring and should be examined. We strive to examine any structure that is present through our calibration, especially if the nonlinearity observed falls outside the scope of the expectation for the ADCs used in the electronics for LSSTCam. This will be further examined in sections 1.3 and 3 of this work.

## 1.2 Runs, Distributions, and Filters

Calibration of the ADC involves using signal from the CCD over its range, from bias through saturation, which is mapped to roughly 100k bins and edges. Since the ADC implies information about the amount of signal input into the electronics, this correlates directly to the amount of light incident on the focal plane. To cover a wide range of ADU, this implies a wide range of exposures is needed in order to represent the whole range of the CCD. To achieve this, runs of many exposures were used for the purpose of calibration. One of these runs is Electro-optical testing run 13144, which is over 600 exposures of flat pairs. We compile all these exposures into a single combined distribution, in counts per ADU, as we are interested in the behavior of the entire ADC over the range of the CCDs. This can be seen in Figure 2.

As seen in Figure 2, there are some fluctuations over the range of the ADC in terms of how much signal is present bin-to-bin. In order to understand how we would expect an ideal ADC to work, we want to preserve the long range behaviors of the full distribution, while minimizing any bin-to-bin deviations. This is accomplished by using a filter, in this case a Savitzky-Golay filter. As seen in figure 3, over the range of the ADC, the filter smooths between the bin to bin behaviors, while preserving the behavior over the long range. The resulting distribution, in counts per ADU, from the filter is assumed to be the expected count distribution if the ADC behaved as an ideal ADC. We note that the individual bins demonstrate deviations in the true distribution from that of the filter, which is direct evidence that DNL is present in our ADC. As motivated in section 1.1, characterizing this DNL is very important to understand how the ADC works.

## 1.3 Original Method of ADC DNL calibration

In the method developed by A. Shestakov, which this work was motivated by, the edges were computed using a DNL informed method. Using run 13144, as shown in Figures 2 and 3, a Savitzky-Golay filter with a window of 65 and order 3 was used to model an ideal ADC's behavior. The DNL of each bin was then computed, using the reformulation of the DNL as

$$DNL = \frac{c_e}{c_o} \quad (3)$$

where the  $c_e$  is the expected number of counts from the filter smoothing the true distribution, and  $c_o$  is the observed number of counts from the true distribution. Reformulating the other equation for DNL, the width,  $w$ , was then found by

$$w = 1 - DNL \quad (4)$$

for each bin along the ADC. Assuming the first left edge correspond to the first ADU with signal, the edges of the bins were placed accordingly as

$$r = l + w \quad (5)$$

where  $r$  is the right edge and  $l$  is the left edge. This allowed the calibration of the ADC over a range of approximately 26-100k ADU, which was deemed sufficient for our science use.

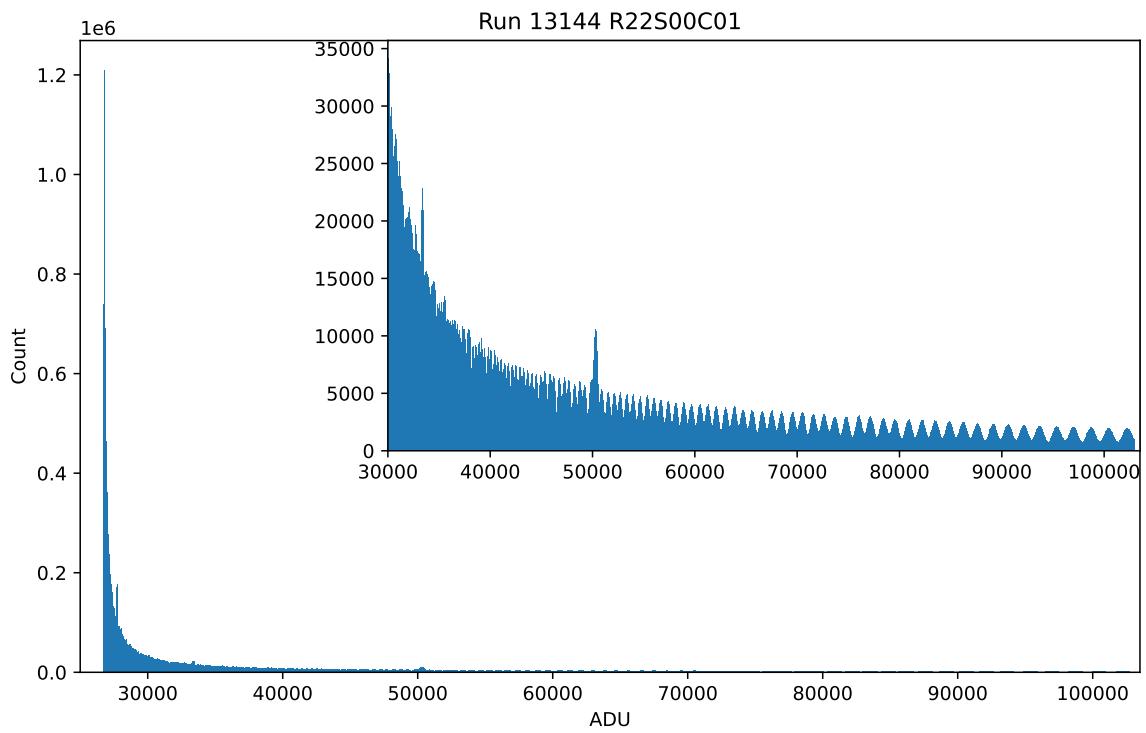


Figure 2: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144. The distribution demonstrates that sub 30,000 ADU is more greatly represented than other regions of the ADC, yielding the distribution that is skewed right with counts on the order of one million. The upper righthand corner zooms in on the later region of the ADC, demonstrating the gentle tapering out of the signal over the full range of the CCDs.

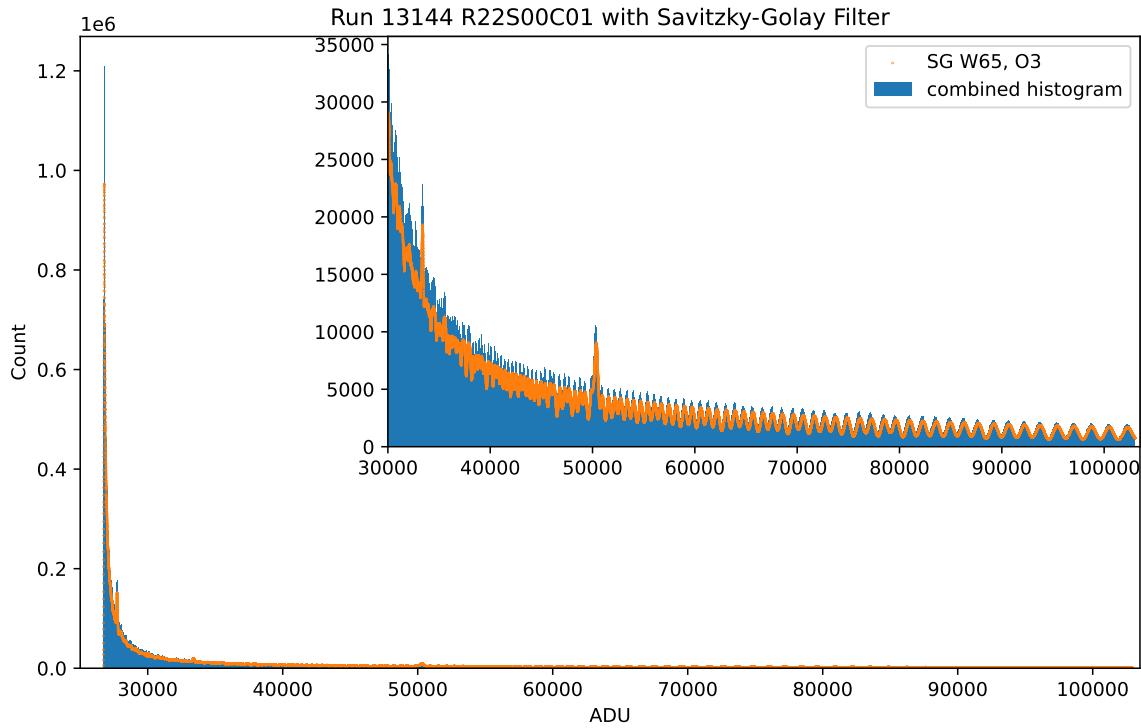


Figure 3: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144 as seen in figure 2. Overlayed in orange is the output from a Savitzky-Golay filter with window size 65, and order 3. The deviations between the filter and the histogram directly demonstrates presence of differential nonlinearity in the ADC. The upper righthand corner zooms in on the later region of the ADC, demonstrating the gentle tapering out of the signal over the full range of the CCDs and demonstrates the filter follows the long range behavior of the distribution.

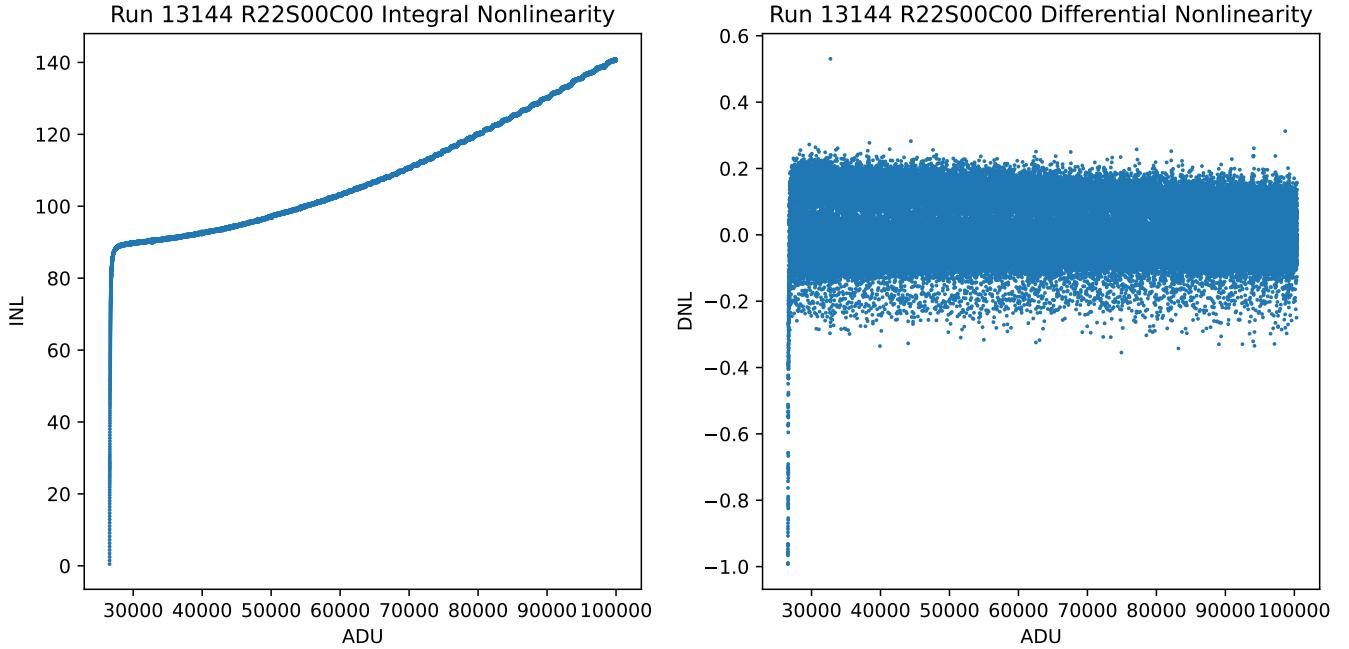


Figure 4: Above on the left demonstrates the contributions to the integral nonlinearity using the method described in Section 1.1. Clearly this demonstrates an unbounded integral nonlinearity that grows progressively throughout the range of the CCD. On the right, the differential nonlinearity is shown over the region of the ADC. For sub 30000 ADU, the DNL sharply peaks, suggesting reason for concern.

#### 1.4 Differential and Integral Nonlinearity blow up

The original method, as discussed in the previous section, Section 1.3, yields both, INL and DNL with deviations from specification. As seen earlier in Figure 1, the method utilized a run of flat pairs where there was significant signal in the early range of the ADC, and an exponential decay of signal outside that region. This signal is on the order of 1 million near bias and approximately 5 thousand near saturation, which is multiple orders of magnitude different over the range of the ADC. Therefore, when the filter was applied and represented the true distribution of an ideal ADC, there were large deviations between the filtered and true distribution. In turn, when computing the DNL, this lead to very large DNL in this early region. These large DNL led to bins on the order of 0.1 ADU, suggesting that the bins near the bias were very narrow, while later bins were much wider. This is seen in Figure 4, where the DNL for lower ADU is on the order of 1, while it stabilizes for higher ADU. This was a direct result of how well the filter does with a skewed distribution and not a real characteristic of the ADC, as seen in Figure 3, where there is significant deviation between the orange and blue curves. When starting the binning procedure from Section 1.3 on later ADU, the issue with large DNL disappears, suggesting DNL dependence on the starting point of the method. This suggests the method should be examined in order to get a consistent calibration of the ADC, independent of starting location.

Another issue arose when examining the INL from this method. As seen in Figure 4, we found an unbounded INL over the range of the ADC, which resulted in a deviation of over 140 ADU from an ideal ADC. Suggesting that this ADC was extremely out of specification. Clever cuts to the distribution lowered the components of the INL to 10 ADU, but no amount of selective cutting of the distribution forced both the DNL and INL to fall within the specifications. This further motivated the need to find a new method for calibration, which was informed by the DNL and INL, as this large of a deviation from the specifications smelled of an artifact of the method.

## 2 Process

After examining the previous method of calibrating the ADC, we see that a need existed for a calibration of the ADC that represents the ADC without inputting fabricated INL or DNL in order to represent the true behavior of the electronics. In this, we will still use entire runs for the calibration, but are more selective on preparing our distribution of counts over the range of the ADC, including regions with more representation than others in terms of density of exposures. We also look at the advantages of using “ramp” type runs to uniformly probe the entire range of the ADC with their choices of exposures, and compare this run type to previously used flat pair runs. We also more carefully consider the use of the filter and develop a robust mathematical treatment of the signal and the filter by examining and constructing a probability distribution function (pdf) or the distribution of counts across pixels. We will use this to construct the bin edges along the full range of the ADC. Embedded is also discussion of removal of particular bins, protection of the distribution from the filter, and other aspects crucial to calibration, as well as discussion of where to find a version of code used in this work.

### 2.1 Preparing Data from Runs

As discussed briefly in section 1.2, the ADC covers a range of 100,000 bins, with some expected distribution of signal to fall into each of these bins. In a more specific treatment than previously discussed, this distribution, called the probability distribution function, informs how much is present in each of the bins. Since electronics behave individually, we have no a-priori knowledge of this distribution and instead need to examine the signal from the ADC that we do have access to. As seen in section 1.4, the shape of the distribution taken over the range of the ADC plays a large role in the parameters such as the DNL and INL. More specifically, utilizing a distribution where there is a large amount of exposures in which fall in a single region, and few in another, the region with more signal creates small on average bins, with large amounts of DNL and an unstable INL. In order to calibrate the ADC successfully then, we need a distribution where all portions of the ADC are equally represented, otherwise we are introducing features and substructure of the ADC that simply do not exist.

This implies that we have two criteria for calibration, relative flatness and the full range of the range. Yet our data come in the form of exposures from EO Runs, and signal in the ADC is proportional to the amount of light incident on the sensor. This implies many exposures, with many different amounts of light is necessary to cover the full range of the ADC. Although, as seen from Figure 2, using all exposures equally in a run may yield a distribution where some regions are over or under represented. Therefore we need to determine the groupings of exposures that demonstrate the distribution of signal over the entire ADC. Of course, a run of images with increasing exposure time with continuous readout would satisfy this requirement, as we could see how the ADC responds over the full range of the CCDs. This was completed during Run 6b of EO Testing as run 13549, and is seen as Figure 5. While this data is very useful for this calibration, due to it being the full range of the ADC and relatively consistent signal in all regions, it is not always practical to request a special run to further verify this calibration, as this run took 5.5 hours to be completed. Instead it is beneficial to be able to also use flat pair runs, such as run 13144, which still record exposures from bias to saturation of the ADC, but have an uneven distribution of the ADC which is represented. To utilize these runs we use a dynamic pre-scale to select portions of all exposures to include in the combined distribution.

The pre-scale method combines all the exposures from a flat pair run into a combined distribution. We examine the region of the ADC where the combined signal is relatively flat, typically higher ADU codes, and select the saturation level for the entire distribution. For run 13144, from EO Testing Run 5, we used 5000 counts per ADU. We then compute a pre-scale factor,  $\chi_i$ , for each subregion of the ADC, with each subregion to include 25 codes, where:

$$\chi_i = \frac{5000}{c_o} \quad (6)$$

where  $c_o$  is the observed counts from the combined histogram (Figure 2). For any regions of the ADC where the observed was less than 5000, then the pre-scale factor was taken to be 1, or all data in that region

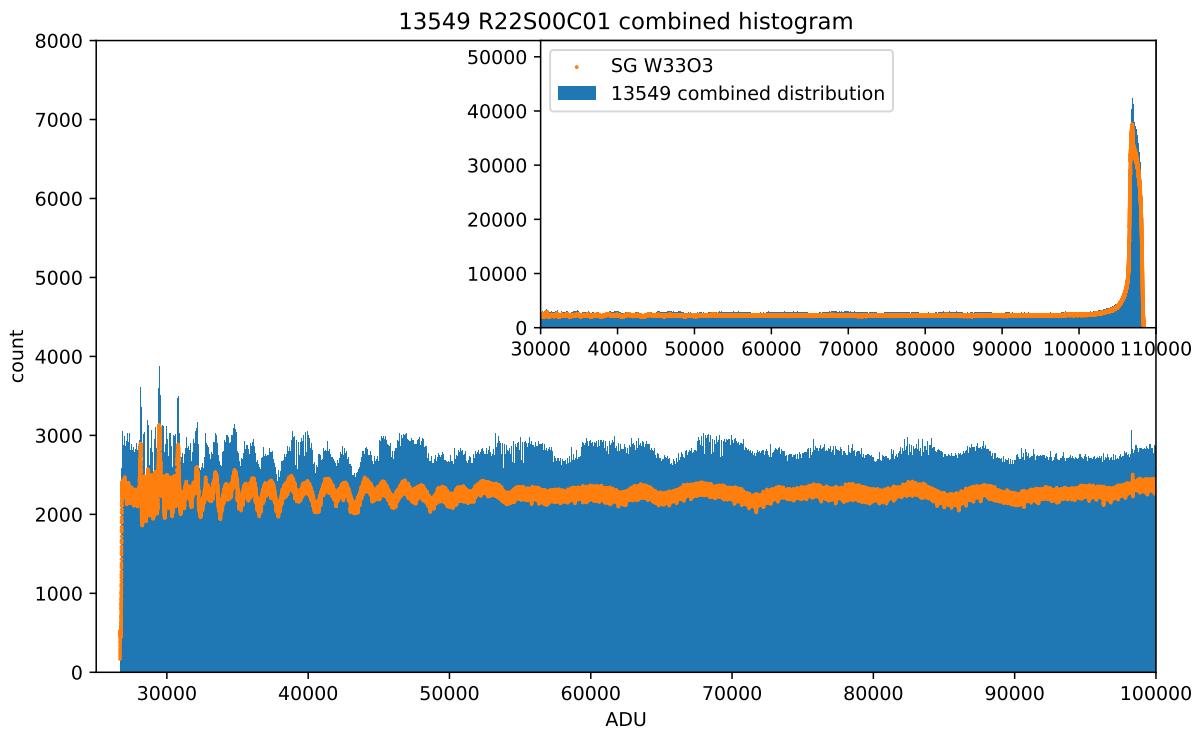


Figure 5: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 234 ramp exposures taken in run electro-optical testing run 13549. Overlayed in orange is the output from a Savitzky-Golay filter with window size 65, and order 3. The deviations between the filter and the histogram directly demonstrates presence of differential nonlinearity in the ADC. In the upper righthand corner demonstrates a peak near the end of the distribution. These peaks are discussed in detail in the Section 2.3.

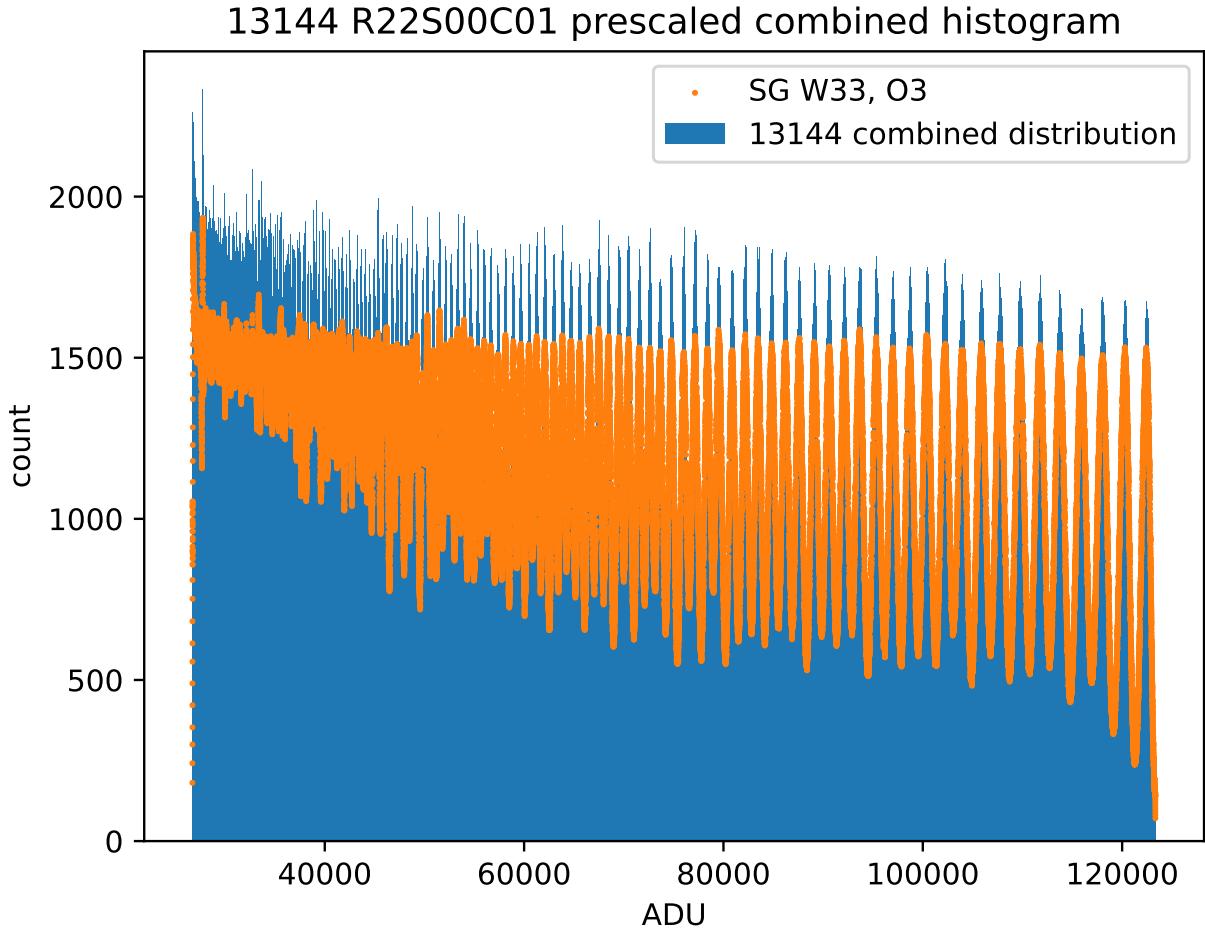


Figure 6: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144. It has then been rescaled, utilizing the pre-scaling process detailed in Section 2.1, in order to reveal this pseudo-flat distribution along the ADC. Overlayed in orange is the output from a Savitzky-Golay filter with window size 33, and order 3. The deviations between the filter and the histogram directly demonstrates presence of differential nonlinearity in the ADC still present despite the rescaling done via the pre-scale method.

was used for the combined distribution. Next, each exposure is pulled, its peak signal in ADU is identified and the corresponding pre-scale factor is selected. Then the number of pixels to select from that exposure, denoted as  $\eta$ , is determined via:

$$\eta = \chi_i * n \quad (7)$$

Where the number of pixels in the amplifier,  $n$ , is dependent on the type of CCD that is connected to that amplifier's ADC, either an E2V or ITL from the two manufacturers of the CCDs. A randomization process selects  $n$  pixel digitizations from the exposure, and these are added to the combined distribution. This is repeated for each exposure in the run, and builds a pseudo-flat distribution over the full range of the ADC that represents the behavior in each exposure in the run, but also cuts out the uneven representation of some ADU regions over others. The resulting distribution, Figure 6, yields a distribution and therefore filtered output which can then be used for the calibration process.

## 2.2 Determination of ADC edges

As motivated in Sections 1.4 and 2.1, a new calibration process is needed using the understanding of the probability distribution function (PDF) of signal in ADU. Each bin of the ADC will have counts which are the integral of the pdf, from the left edge of the bin to the right edge of the bin described as

$$c_o = \int_l^r \varrho dx \quad (8)$$

where l is the left edge of the ADC, and r is the right edge and dx is the direction along the ADC. We have, from Section 2.1, the distribution of counts,  $c_o$ , and neither the PDF nor the location of the edges, although for calibration sake we want to find the edges. There is a relationship between the filtered distribution values and the PDF as well,

$$f_i = \int_i^{i+1} \varrho dx \quad (9)$$

where  $f_i$  is a filtered output per ADU i, that is the filtered value is equivalent to the integral of the pdf from the integer left edge of the pdf over a region of width 1.

In this calibration we change the window of the Savitzky-Golay filter to being 33 instead of 65, as motivated in the appendix. Using our earlier description of the filter, we can utilize the fundamental theorem of calculus via cumulative sums, to examine how an integral of the pdf behaves over larger regions of the ADC.

$$\sum_{i=0}^j f_i = \int_0^{j+1} \varrho dx \quad (10)$$

Since our ultimate goal is to determine the edges of the ADC, we need to interpolate non-integer values of these integrals. This is done by fitting a cubic spline,  $\zeta$ , to the cumulative sum, which allows us to evaluate the integral of the PDF with non integer end points, representing the true edges of the ADCs. Mathematically, we decompose the integral of the observed distribution to be:

$$c_o = \zeta(r) - \zeta(l) \quad (11)$$

Rearranging eq. 11:

$$c_o + \zeta(l) = \zeta(r) \quad (12)$$

Given we choose the first left edge to be our earliest populated bin and  $c_o$  is from the distribution in Section 2.1, we then move to determine the right edge. To determine the right edge, we first assume a width of 1 then via a convergence method we check if the right edge is too small/ large and increase/decrease the bin size by a standard amount and recheck the value. This process is repeated until a convergence interval is reached, typically  $\pm 1$  count. See the appendix for further information about the convergence criteria. Altogether, this yields a calibration of a single ADC in under 10 seconds, without making any calculation or assumptions about the DNL and INL in its computation, and instead utilizes the combined distribution of the amplifier and a filter in order to recover the edges of the ADC.

## 2.3 Details of the Calibration

In the calibration process, some issues arose with minute fixes implemented in order to not throw off the entire calibration process. The first of which involves the implementation of the spline of the cumulative sum as described in Section 2.2. Since the spline is a cubic function, it has the potential to interpolate and yield a negative values. This is an unphysical solution, as a region of the ADC can only yield 0 or a positive number of counts. In order to protect the spline from this, we use padding, or adding more bins with 0 counts before where our ADC first has recorded signal in bins. This inherently protects the spline from yielding negative values, as the function is forced through 0.

A second safeguard is to remove bins that may be noise of the ADC. We remove any bin with less than 100 counts, which are bins with signals that are 1/50 of a typical bin. Cutting these bins differentiates between noise and true behavior of the ADC, as we wish to model the ADC and not the noise.

Third, since we are concerned with the behavior of the ADC and looking at its substructure, we examine whether the combined distribution yields any locations along the ADC with a skipped bin. Since bins are removed from the distribution which are classified as noise, this leads to the possibility of missing codes along the entire ADC. In order to prevent this, the method takes into consideration the location of a skipped bin, and recommends that the calibration begins after a skipped bin. This allows the resulting calibration to not be influenced by noise of the electronics. The method takes this a step further to return the location of the first bin used, in order to compare from amplifier to amplifier where we start the calibration process. This not only preserves the behavior of the ADC, but prevents the accidental termination of the method or interpolation over a bin with insufficient data for calibration.

The last concerns in this process occur in run 13549 which was used as one of the demonstration calibrations. The full distribution on some amplifiers demonstrate "peak" behaviors either near the bias or saturation, or both. The upper portion of Figure 5 demonstrates one of such peaks, a peak near the saturation of the CCD. These peaks have finite widths, suggesting more is going on beyond a "pile up" binning error in the histograms of the exposures. This was troubling to observe, as the distributions in most amplifiers did not have this, and there seemed to be no correlation between amplifiers with this feature, and those without. It became essential to terminate the distributions with these peaks to exclude them from the edge calibration, due to known INL/DNL injection from peaked sources as discussed in Section 1.4. Termination excluded all bins within a range of 50 of the peak, effectively about 100 bins were removed. This became on the order of less than 1 percent remove to prevent INL and DNL injection, and still allowed calibration for the majority of the ADC. In the case of Figure 5, with a peak near the saturation, this meant the edge method stops before reaching where the saturation is achieved, and instead makes edges over the region in the main part of the distribution.

The final problem we encountered was a raft issue in the R4X region of the focal plane in run 13549, which had more digitizations than expected in every single exposure in the run. This implied there were measurements for over 7k "pixels" above that which are expected for ITL sensors. This was only observed in the 3 rafts in this column, although the same method for pulling the images was also utilized in for all amplifiers on all rafts, suggesting an error is present in how the run was saved or accessed. Since a reason for these extra digitizations could not be found, we note this deviation but continued with the calibration process anyways. We also note that there seems to be no discernible differences seen in the distributions or their calibrations as seen in Section 3.

## 2.4 GitHub Access

To access a version of the code which produced these results, use the GitHub repository [rejniche/Focal-Plane-DNL-adc-Calibration](https://github.com/rejniche/Focal-Plane-DNL-adc-Calibration), and follow the three step process to access the data, calibrate the ADCs, and create comparison plots.

## 3 Conclusion

Utilizing the calibration process of the ADC as described previously with the PDF, we are able to calibrate all 3024 non-corner ADCs on the LSSTCam focal plane using both run 13144 and 13549. For each individual amplifier's ADC we can see the substructure of the ADC, Figure 7, by looking at the individual bins and their widths as a function of their location along the ADC. We can see the presence of any structure in the distribution itself, including the downticks present every 128 counts along the ADC. The downticks present in this amplifier are not observed in all amplifiers, although the substructure in the form of macroscopic groupings of the DNL can be seen in all amplifiers. This suggests that the real behavior of the ADC is recovered through the method developed in section 2, and is present regardless of the type of data used in the calibration. Even though as Figure 7 demonstrates, the distribution plays a role in the exact locations of the bin widths over the ADC, there is no statistical evidence (via a Paired T Test) that the widths created from the two runs difference deviates from zero. Also seen in Figure 7, there are some noticeable differences in the locations of the edges of the ADCs when depending on whether run 13144 or run 13549 was used for

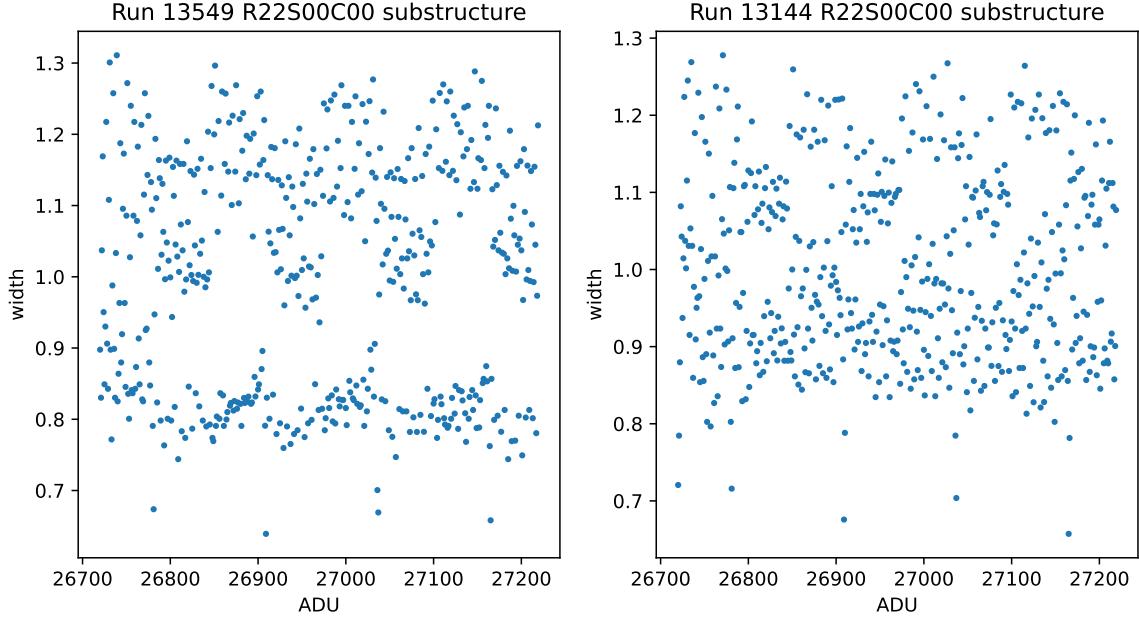


Figure 7: The above figure demonstrates the calibrated bin widths of the the amplifier R22S00C01, near the lowest ADU region. At this zoom level, the substructure of the ADC can be seen clearly. On the left is the result from the ramp run 13549, and on the left is the result from the pre-scaled run 13144. Some slight deviations can be seen between the two runs, due to their individual distributions, but both demonstrate regions of macroscopic grouping of DNL and yield similar bin for bin values of the DNL, as explained in section 3.

the calibration. We examine the locations of the left edge in more detail, and determined that R22S00C00 has left edges that differ on average by 0.2 ADU per bin. This suggests there is a systematic shift of the left edges depending upon the distribution used in the calibration. Given there is no difference in the widths between the two runs nor any systemic differences in INL and DNL, this suggests that there is a slight offset of the binning depending on the calibrating run. This means largely the widths created by both runs yields consistent calibrations of the ADC. It is also noted that over the full range of the ADC, there are not regions where the DNL grows rapidly, as seen in Figure 4, suggesting that any large DNL bin is isolated rather than indicative of an issue in a region.

Beyond a single amplifier, we can further look at the entire focal plane, and the edges made for each of the ADCs. For the sake of space, we show just the calibration of 13549 over the entire focal plane. Figure 8 demonstrates the absolute value maximum DNL for all bins along each ADC on the focal plane. The DNL falls within the allowable 0.85-1.5 range as set by the manufacturer. Figure 9 demonstrates the absolute value maximum INL for all bins along each ADC on the focal plane. The INL falls within  $\pm 2$  for every ADC. This suggests we have all the ADCs within the specifications set by the manufacturer, and the same situation occurs for 13144 although plots are not included.

We note, although it is not explicitly shown, that the issues of instability of the DNL and INL, as discussed in Section 1.4 and the Appendix, are no longer present. This implies we are able to calibrate the ADC regardless of the commencing bin. This suggests that using a pre-scale on a flat pairs run works just as well for this DNL calibration of the ADC as using a special ramp run. An additional thing to note is the stability of our electronics through time, meaning the adcs are delivering consistent binning results from December of 2021 to November of 2023. This is highly beneficial as the survey of LSST is 10 years, and it would be problematic if time changed the DNL calibration from a science perspective.

13549 max dnl

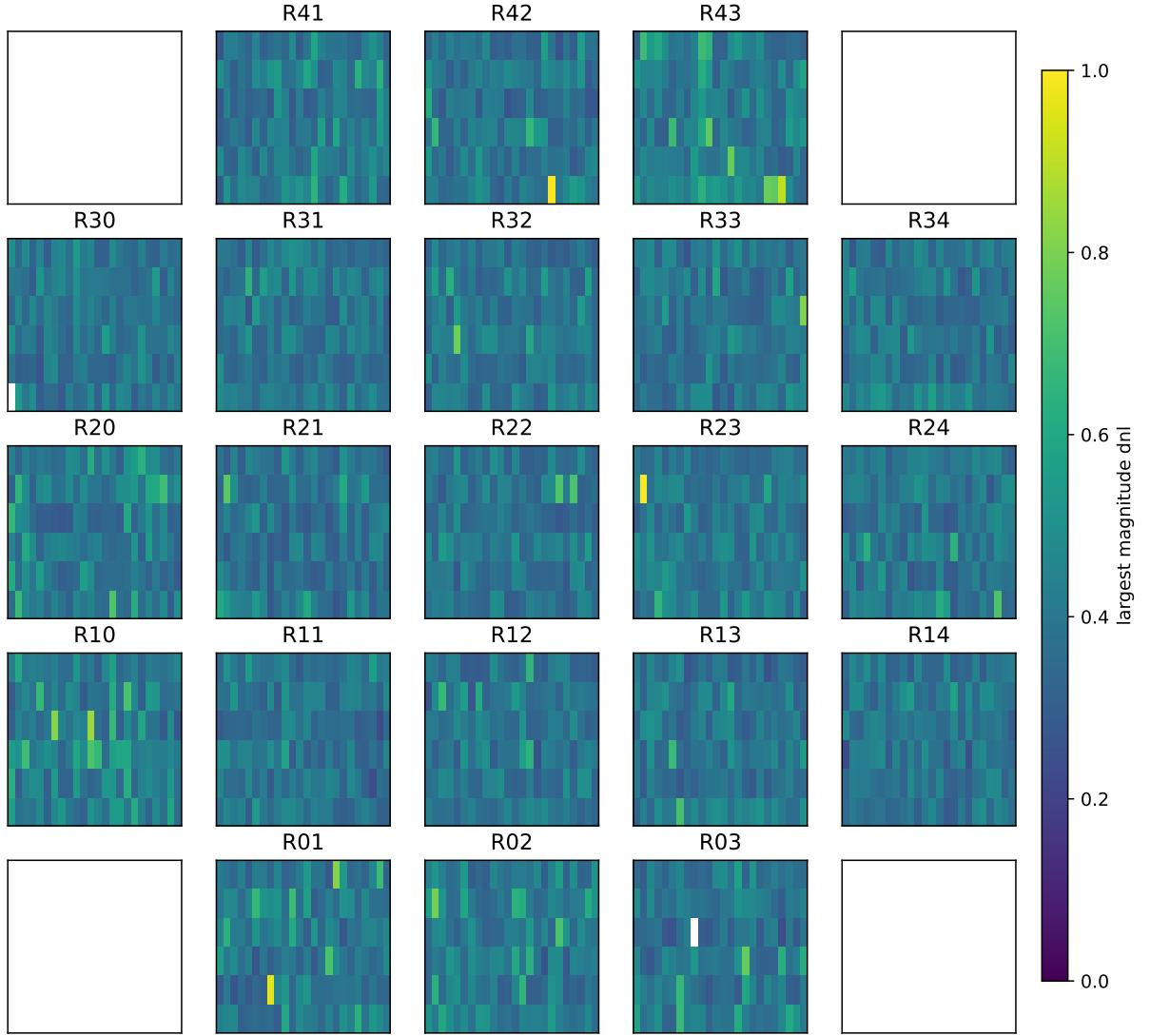


Figure 8: The above figure demonstrates the calibration of the ADC using the differential nonlinearity method we described in section 2.2. This method computes edges along the ADC for each 3024 amplifiers individually. The DNL can then be computed for each bin, in each amplifier, and plotted here is the maximum DNL of all bins in each amplifier. Ranging in magnitude between 0 and 1, the maximally achieved DNL for each amplifier is below the maximally allowed by the specification sheet.

13549 max inl

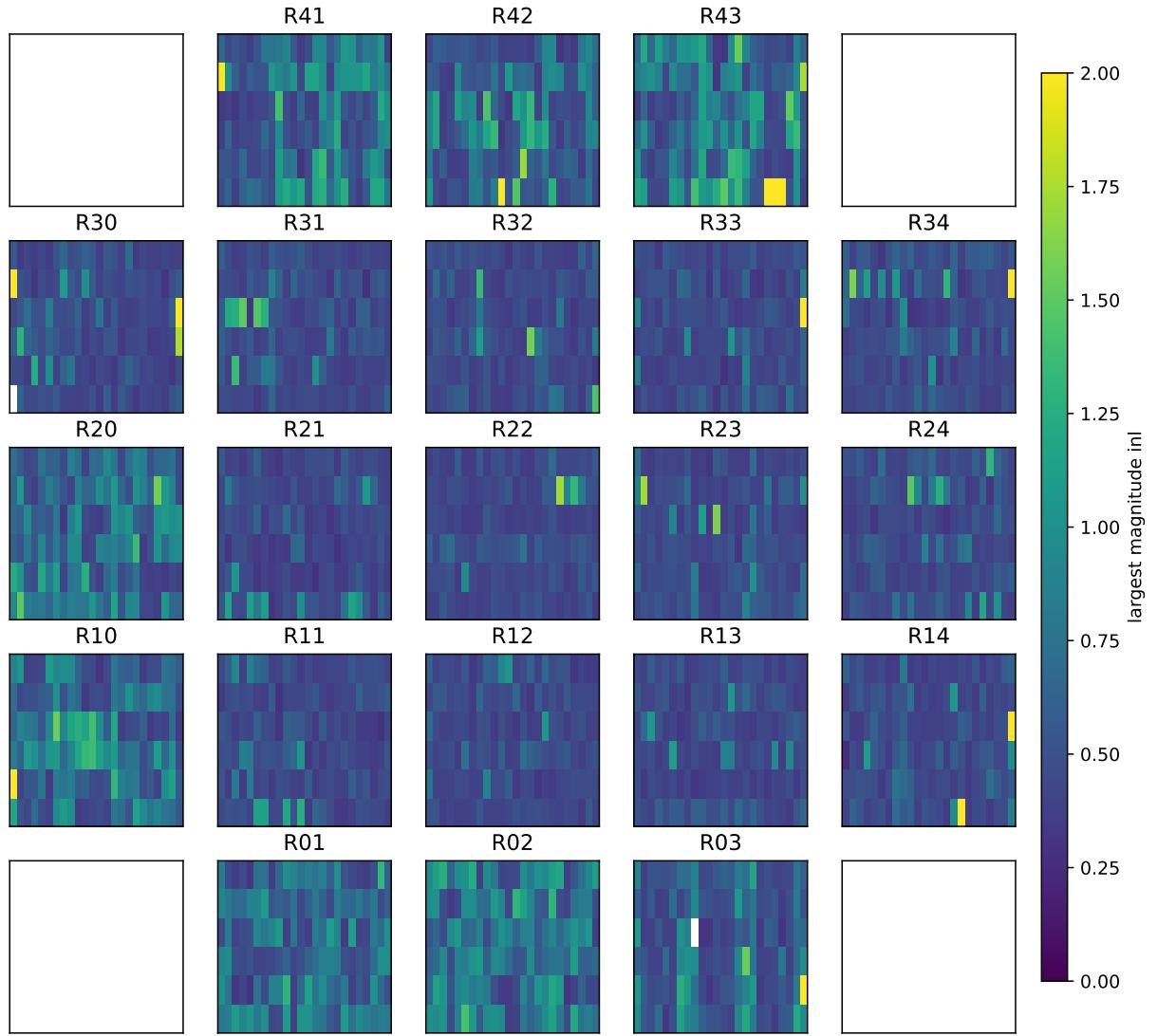


Figure 9: The above figure demonstrates the calibration of the ADC using the differential nonlinearity method we described in section 2.2. This method computes edges along the ADC for each 3024 amplifiers individually. The INL can then be computed for each bin, in each amplifier, and plotted here is the maximum INL of all bins in each amplifier. Ranging between 0 and 2, the maximally achieved INL for each amplifier is below the maximally allowed by the specification sheet.

## 4 Future Work

In future work, this the entire process could be updated for speed, including parallel processing for both pulling the exposures themselves, and creating the edges along the adcs. There is also various applications which one could be able to see the potential impacts of an incorrect or poor calibration may have on science outcomes such as weak lensing shear among other interests.

## 5 Appendix

In this section we detail more methods and processes attempted during the development of the calibration process discussed in this work.

We comment further on motivating a flat distribution being necessary for calibration. Using a unscaled 13144 but the method described in Section 2.2, cumulative binning, the ADC was calibrated depending on a starting bin. This bin normally was taken to be the first ADU in the distribution from the run itself. Yet, one would believe that the DNL and INL would be independent of the location of the starting bin as it should be characterizing the nature of the ADC. The analysis demonstrated this was not true of the INL, and rather the INL would behave very differently depending upon the choice of the first bin. We were able to force the INL to reach a maximum value of anywhere between 2 and 200 over the range of the ADC depending on the choice of starting bin. Work as also done to look for a possible binary correlation between the starting bin and INL and found none was present. We also examined the relative values of the filter and the distribution itself for the starting bin, as well as their change, and found no correlation between the two.

Extensive work was also done in order to choose the filter used in the cumulative sum binning method described in Section 2.2. We examined multiple window sizes, between 25 and 100, and different orders of the filter, 2nd and 3rd for the Savitzky-Golay filter. We also utilized another type of filter, Butterworth with high-pass, lowpass, and bandpass. We utilized Chi-Squared Tests in order to determine which filter best represented the input data and distributions, using both the real data used in this paper, but also fabricated datasets as well. The results demonstrated that the Savitzky-Golay filter with window size 33 and order 3, represented the data used in the analysis and the fabricated data very well, and was therefore chosen for use in the method.

Work was also done to examine the convergence interval method that was used in the cumulative sum binning method. The method uses a convergence interval of  $\pm$  specified value to converge on the right edge of the ADC bin. Through many iterations, it was found that while an increased precision in the value takes much more computation time, it does not positively impact the values of the DNL and INL seen in the distribution. Thus, the convergence interval of 1 count optimizes the computation time while still yielding a 0.2 percent difference between the observed counts and that achieved by the method. Another effort went into looking over the residual amount, either by the method going over or under the goal amount per bin, and whether some bias was present due to it. It was determined that the residuals make no noticeable pattern over the range of the ADC, and rather also suggests that the bins were not larger or smaller than expected. In addition, a test was conducted to add the residuals to the next bin throughout the method. This demonstrated no noticeable changes in the DNL, but a new structure emerged in the INL. These changes, coupled with 0.2 percent difference suggests the residuals were of a negligible amount and no more robust treatment of them is necessary.

Lastly there were various other attempts of calibration methods that were tried. Some of these include terminating the dataset and doing a DNL mean correction in the original method in Section 1.3. These methods both failed due to the other issues discussed throughout this work. Before creating the pre-scale method, we also tried adding a correction to the filtered output in order to more closely match the distribution. This proved to not fix the INL or the DNL problem, and diminished much of the filter's predictive power of acting as a true ADC distribution. The last method we tried involved PDF as used in the cumulative cum binning method, but it did not utilize the cumulative sum. Rather it fit a spline to each of the filtered datapoints, which mathematically would require to deconvolve the PDF in order to recover the edges. This was replaced

by the cumulative sum aspect of the method due to the complexities necessary finding the edge solutions in this method.