

Differential Nonlinearity Analog to Digital Converter Calibration: draft

2.1

Renee Nichols

March 15, 2024

This work seeks to fill in how to calibrate the analog to digital converters (ADC) 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.

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 this 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. A range of analog signals is to be converted into the bins of the ADC. In a perfect ADC, the bins would be integer spaced apart, with integer analog signal mapped directly to integer digital signal. In reality, there are deviations of these edges from expectation, resulting in the digital edges of the ADC being non-integer values.

In order to understand the output signal, we want to be able to determine the location of these edges and look for any trends that may be present in them. (Define what a binwidth here is as well !)

1.2 Original Method of ADC DNL calibration

In the method developed by A. S., the edges were computed using a DNL informed method. For each bin, the DNL was first computed, then edges were made utilizing width = 1- DNL. More formally, a distribution of flat pairs was used in order to create a combined distribution of the entire range of the ADC. The distribution heavily favors the bias end of the ADC, with many more flat pairs in that region than further along the ADC. As discussed further later, there are some issues with using a run with an inconsistent representation of ADC bins, but this was not known at the time. The distribution detailed "number of pixels with a specified ADU

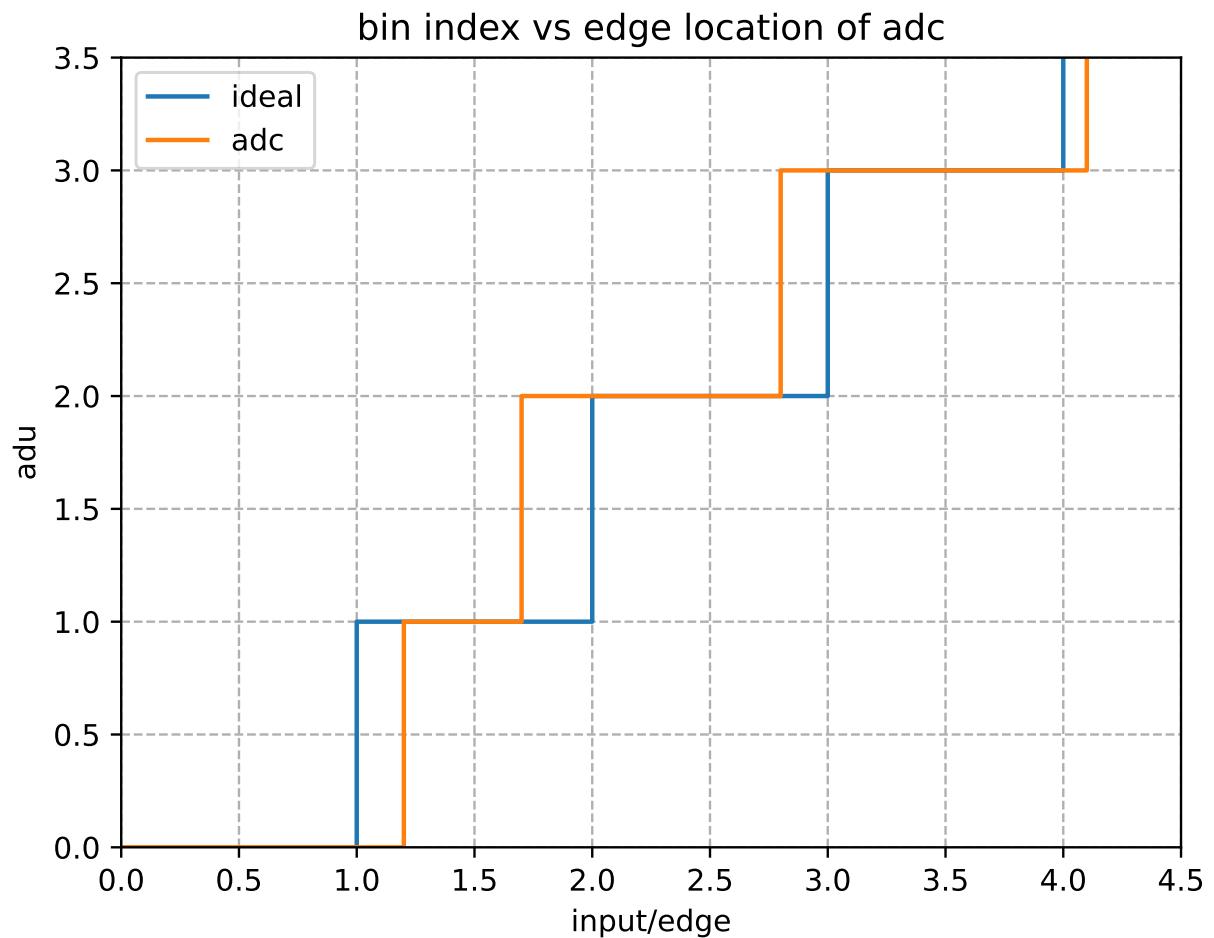


Figure 1: Above is a schematic demonstrating how the analog to digital converter ideally works, and how it truly works. Seen here, the ideal ADC maps a range analog signals into a single bin with integer end points. A true ADC has some variations in the resulting bin, which causes non integer bin edges. These non-integer bin edges then correlate to 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.

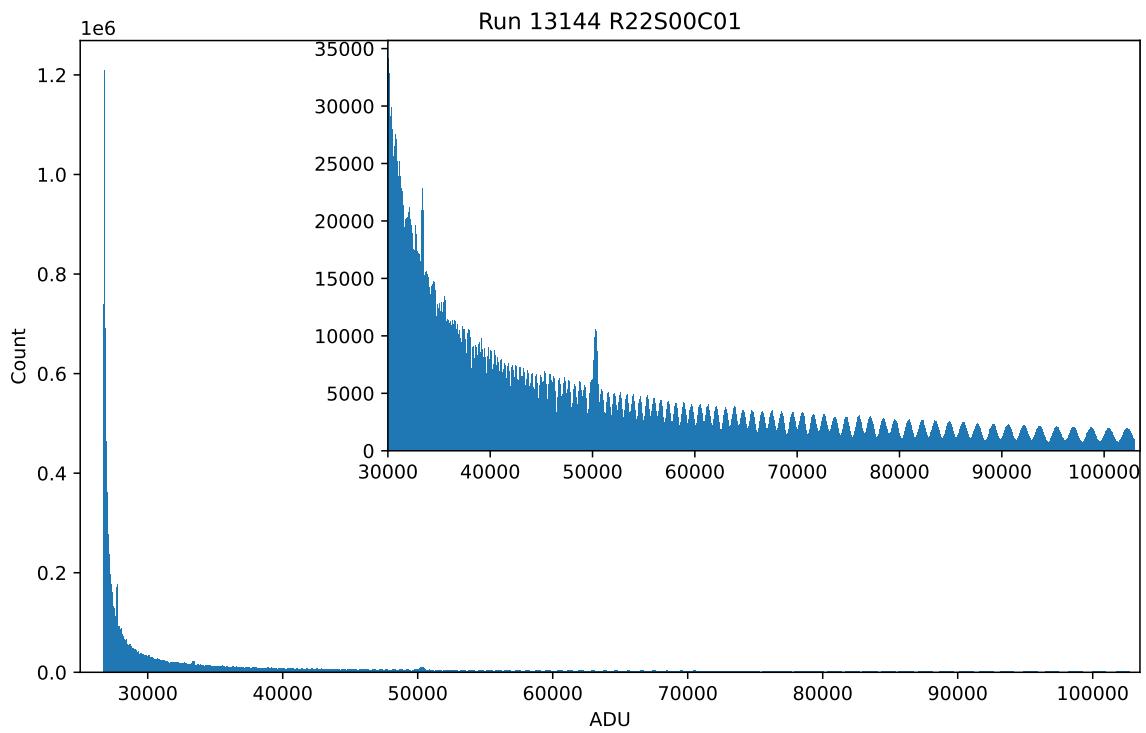


Figure 2: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144. The distribution demonstrates that sub 30000 ADU is more greatly represented than other regions of the ADC, 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.

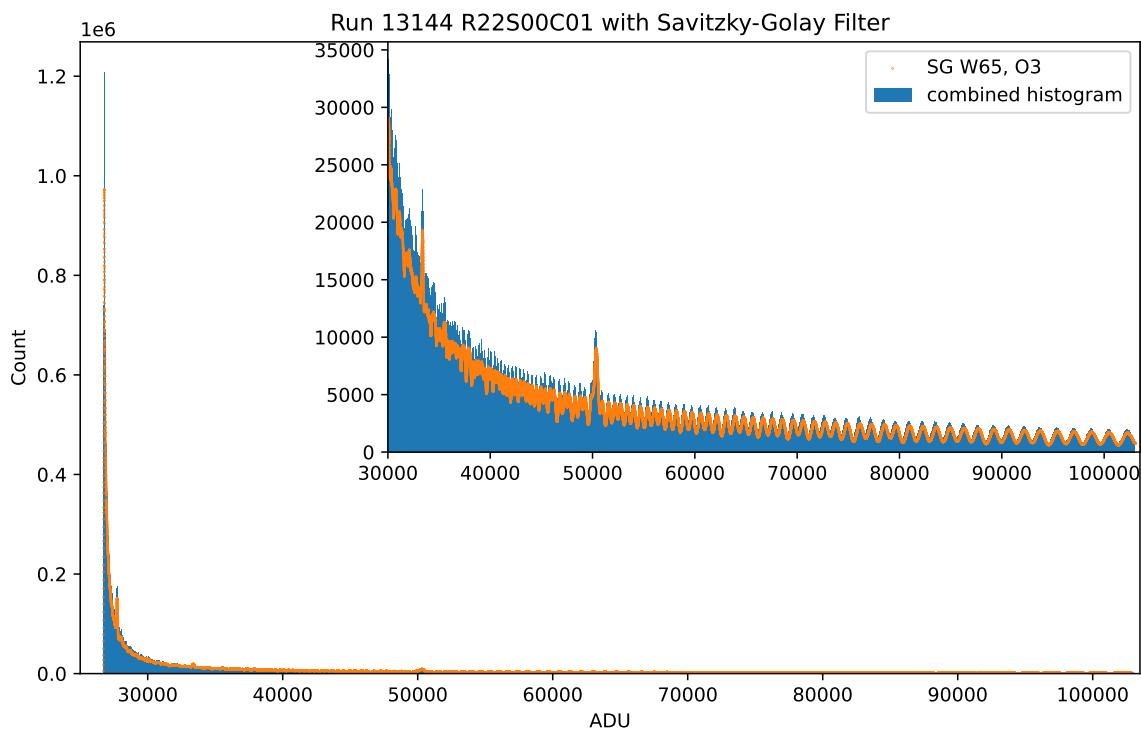


Figure 3: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144. Overlayed in orange is the output from a Savitzky-Golay filter with window size 65, and order 3. The deviations between the filter and the histogram directly demonstrates presence of differential nonlinearity in the ADC.

value” vs ”ADU value”, where ADU is analog to digital unit. This is distribution, or histogram, is noisy with long range behavior which reflects the behavior of the ADC. We utilized a filter to smooth over bin to bin deviations, but preserving the long range structure which we seek to characterize. A Savitzky-Golay filter with a window of 65 and order 3 was used to smooth the distribution for this purpose. This filter type was chosen since it has a good reputation for electronic noise smoothing. The resulting distribution from the filter was assumed to be the expected count distribution if the ADC behaved as an ideal ADC. Therefore, one could take the definition of the differential nonlinearity to be:

$$DNL = \frac{c_e}{c_o}$$

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. Then it was utilized that the width of ADC bins:

$$width = 1 - DNL$$

and the bins were placed accordingly. Starting with a left edge at the lowest ADU recorded in the distribution from the exposures. This allowed the calibration of the ADC over a range of approximately 26-100k ADU, which was deemed sufficient for our science use.

Note: This is not exactly how Adrian originally did this, but it the way we recreated the INL problem, so maybe include this moving forward.

1.3 Differential and Integral Nonlinearity blow up

There are two quantities that we are particularly interested in when looking at the bins and their edges, called the integral nonlinearity (INL) and the differential nonlinearity (DNL). The INL is the deviation of the middle of the true bin from the middle of an ideal middle. Thus if the bins begin to deviate significantly from an ideal ADC, we will see our components of the INL steadily take off. The specifications sheet determines that for any ADC, there should be an INL maximum of +/- 2 ADU for any bin, with a typical value being +/- 1 ADU [CITE]. The DNL is the difference of a bin from having a width of 1, the ideal width, along the ADC. Thus, if there are small or large on average bins, this will show up as having some non zero DNL. The specifications sheet deems that any bin should have a typical +/-0.5, with the maximum being -0.85 or +1.5 ADU [CITE], meaning there should be no skipped bins along the ADC.

The original method, as discussed in Section 1.1, yields both, INL and DNL with deviations from specification. As explained earlier, 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 order of magnitudes 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 region. These large DNL lead 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 was a direct result of how well the filter does with a skewed distribution and not a real characteristic of the ADC. This problem could be avoided when binning began on bins with more consistent data, meaning that there was a DNL dependence on the starting point of the method.

Another issue arose when examining the INL from this method. We found an unbounded INL over the range of the ADC, which resulted in a deviation of over 70 ADU from an ideal ADC. Suggesting that this ADC was extremely out of specification. Further 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 rather than a smoking gun that something was severely wrong with the ADC.

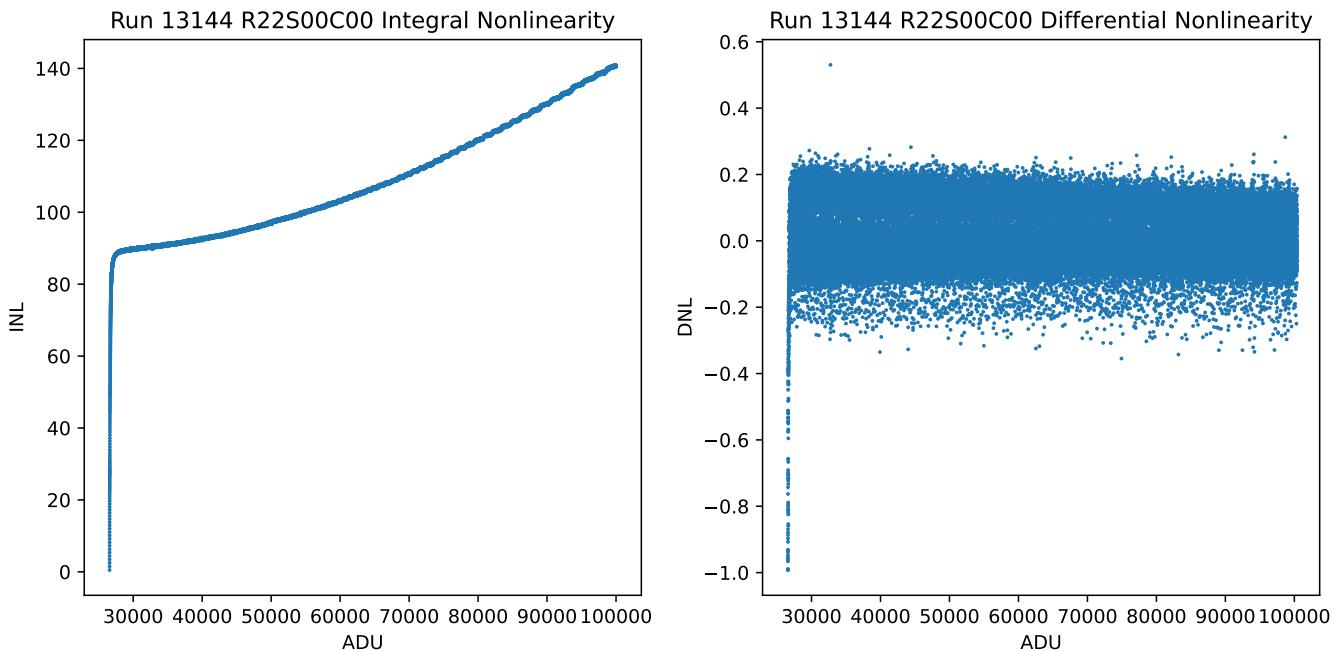


Figure 4: Above on the left demonstrates the contributions to the integral nonlinearity using the method described in Section 1.1. Clearly this demonstrates an unbounded integral nonlinearity that grows progressively throughout the range of the CCD. On the right, the differential nonlinearity is shown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

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.

We also detail important aspects of this calibration, including removing bins that are occupied only by noise and protecting the distribution from any non-occupied bins. Which are all embedded in our code, with the included GitHub repository link included in this section.

2.1 Choosing runs

Calibration of the ADC involves using signal from the CCD over its range, from bias through saturation, which is mapped to roughly 100k so called ADC “codes” with unit ADU. There is some distribution of how the signal is expected to fall on along the ADC, which is called the probability distribution function, which is the amount of signal expected to fall in each bin. We expect each bin to have a different amount, but be relatively consistent over the full range, and is hence a distribution. 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.

Another aspect to consider when determining which data to use in the calibrating distribution, is how

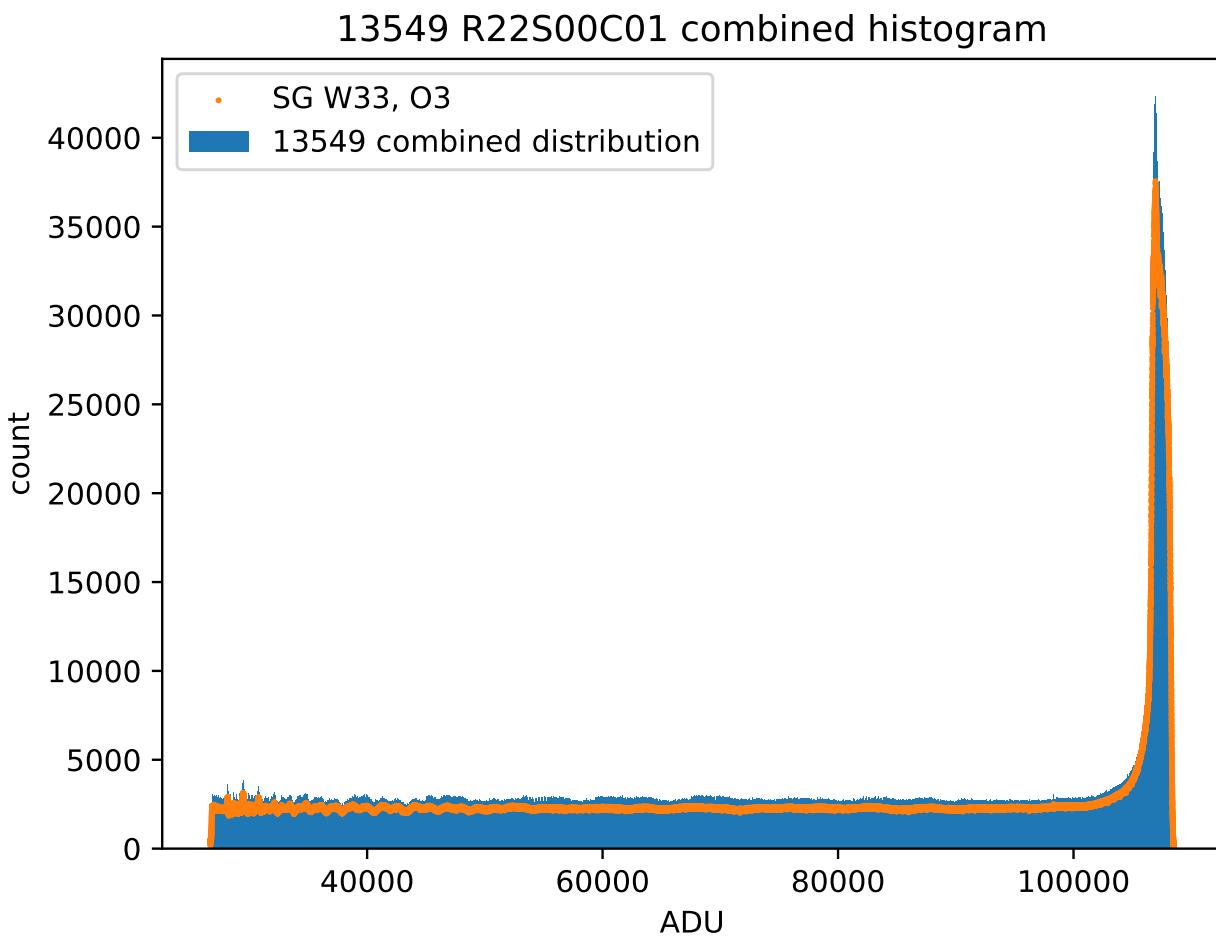


Figure 5: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 234 ramp exposures taken in run electro-optical testing run 13549. Overlayed in orange is the output from a Savitzky-Golay filter with window size 65, and order 3. The deviations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

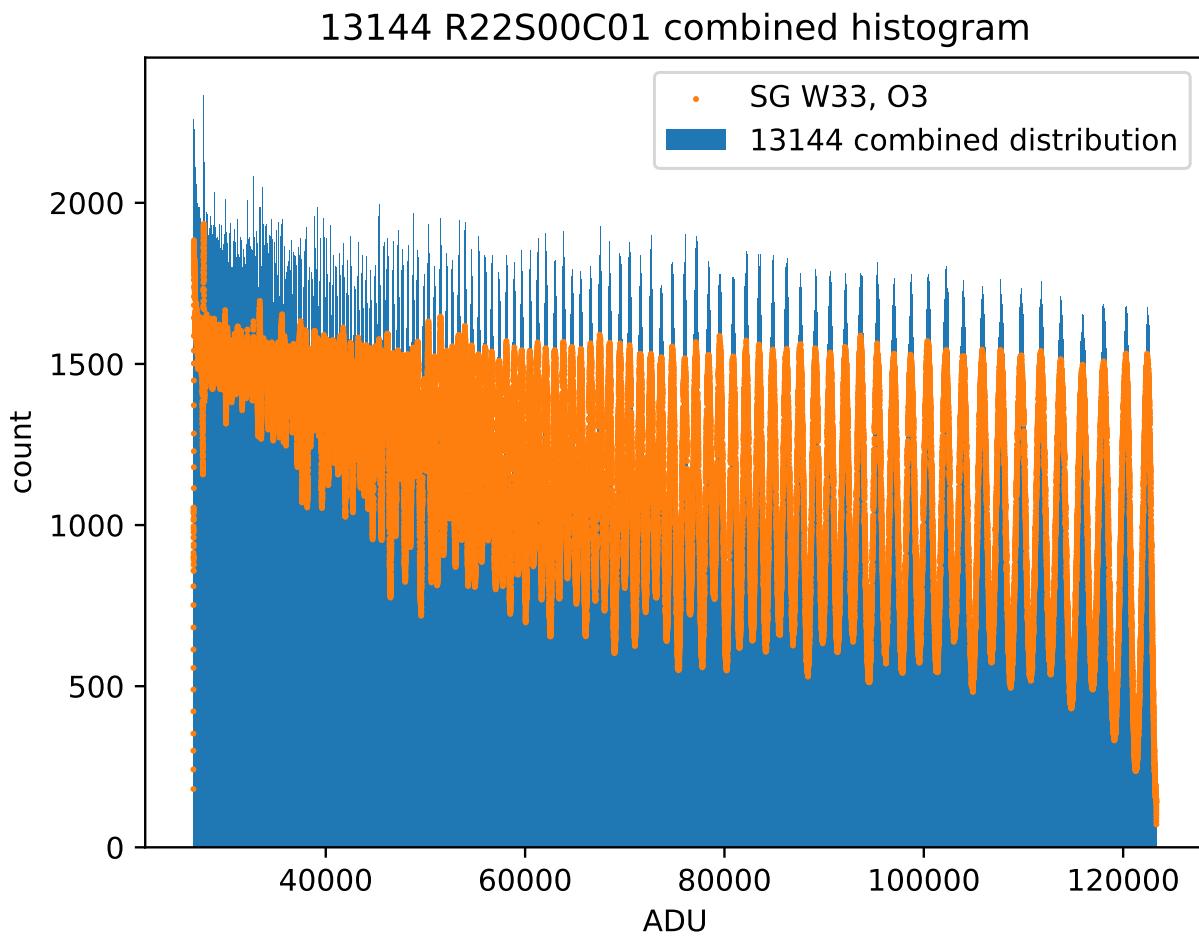


Figure 6: Above is the combined distribution of the readout from all pixels on a single amplifier, R22 S00 C01, along all 687 flat pair exposures taken in run electro-optical testing run 13144. It has then been rescaled, utilizing the pre-scaling process detailed in Section 2.1, in order to reveal this 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.

any features present in the distribution may impact parameters such as the INL or DNL. Through extensive work, we have been able to determine that this calibration method does an excellent job to represent the input data, including periodic features along the ADC and any over or under represented regions of the ADC. Therefore, when 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. See Section 1.2 for more details. Yet when we equally represent the full range of the ADC, either using a flat distribution or trimming our distribution, we yield DNL and INL within the specifications from the manufacturer (See Section 3 for more details). This behavior is expected, as an over or under representing of a region of the ADC will affect the model of our probability distribution function and hence change how our calibration understands these bins. Thus it is very important to have 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.

These two requirements - flatness and the full range of the ADC must be combined to successfully characterize the ADC. Since each amplifier is composed for over 1 million pixels and each pixel is responding and yielding a signal that is proportional to the amount of light incident on the sensor, trying to characterize the ADC over the entire range is functionally impossible, as only a single subregion of the ADC would be represented in any single exposure. 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. 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, 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. This distribution will be extremely right skewed, as explained in Section 1.2, with a majority of exposures being in under 30k ADU. 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 bin. We then compute a pre-scale factor, χ_i , for each subregion of the ADC, with each subregion to include 25 codes, where:

$$\chi_i = \frac{5000}{c_o}$$

where c_o is the observed counts from the raw histogram. 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 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 η , is determined via:

$$\eta = \chi_i * n$$

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. As seen in Section 3, when comparing the results from a pre-scaled run and the ramp run, we note no significant deviations in the calibration bins, suggesting this does a sufficient job in repurposing other types of EO runs for this calibration.

2.2 Determination of ADC edges

Calibration of the ADC requires understanding of the distribution of number of pixels which yield ADU vs ADU. Each bin of the ADC will have counts which are the integral of the pdf, see Section 2, from the left edge of the bin to the right edge of the bin.

$$c_o = \int_l^r \varrho dx$$

where l is the left edge of the ADC, and r is the right edge and dx is the direction along the ADC. We only know the distribution of counts (of pixels), and neither the PDF nor the location of the edges, although for calibration sake we want to find the edges. This leaves looking at the filtered distribution for more clues. There is a relationship between the filtered distribution values and the pdf as well,

$$f_i = \int_i^{i+1} \varrho dx$$

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. That is the filter yields the integral of the PDF if the ADC were ideal, and had all bins with exactly width = 1. The filter still smooths over small bin to bin deviations throughout the ADC, while preserving its characteristic long range behavior.

In this calibration we change the window of the Savitzky-Golay filter to being 33 instead of 65. This was motivated by extensive work described in the appendix. With this, 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.

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

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, ζ , 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:

$$c_o = \zeta(right) - \zeta(left)$$

Rearranging:

$$c_o + \zeta(left) = \zeta(right)$$

Given we choose the first left edge to be our earliest populated bin, this choice preserves as much of the ADC available in order to calibrate as much of the range as possible. 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 +/- 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.

Note: we can (and should probably) include the integral forms of both the filter and distribution for clarity.

2.3 Details

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. 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 1.2. 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.

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 [rejniccho/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. For each individual amplifier's ADC we can see the substructure of the ADC, 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. We can further look at the entire focal plane, and the edges made for each of the adcs. We can see that from both of the runs used for calibration, we have all the adcs within the specifications set by the manufacturer. The INL contributions do not exceed that of +/- 2, and the DNL falls within the 0.85-1.5 range as expected. We no longer have the issues from the previous method, see Section 1.2 and Appendix, and instead there is stability in quantities of the INL and DNL regardless of

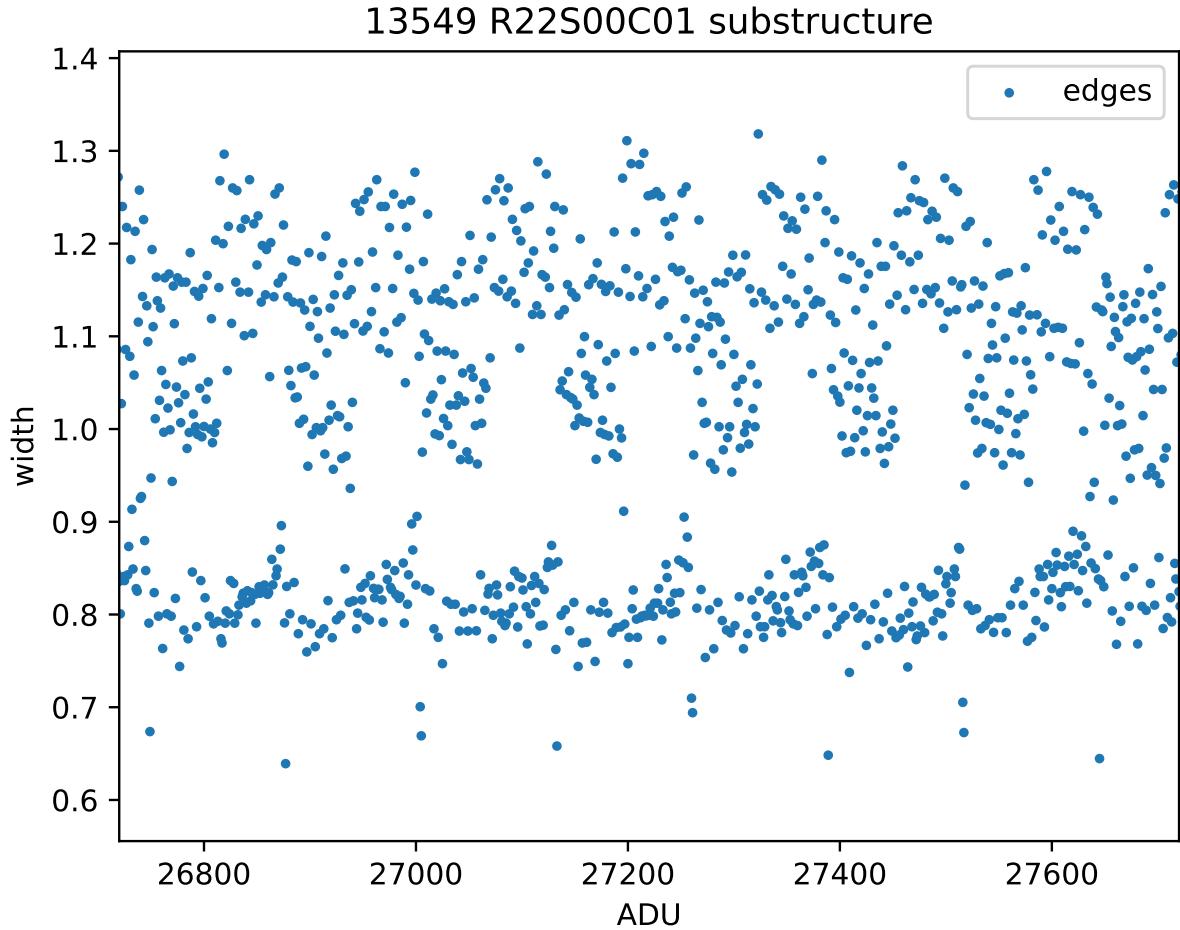


Figure 7: The above figure demonstrates the calibrated bin widths of the the amplifier R22S00C01, near the lowest ADU region. At this zoom level, the substructure of the ADC can be seen clearly, as well as the characteristic macroscopic grouping of DNL ($DNL = 1 - \text{width}$), that was present in the earlier plots.

the commencing bin. We also observe no statistically significant deviations in the widths made from the two runs. 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. In addition, it is noted that there are no artificial structure to the INL, that is the contributions are both positive and negative, whereas with the original method we have unbounded negative deviating INL over the range of the ADC. 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.

Note: Seems fine for now, we need to add plots and explain further which will expand on many of these results I have briefly included here.

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 a incorrect or poor calibration may have on science outcomes such as weak lensing shear among other interests.

13549 max inl

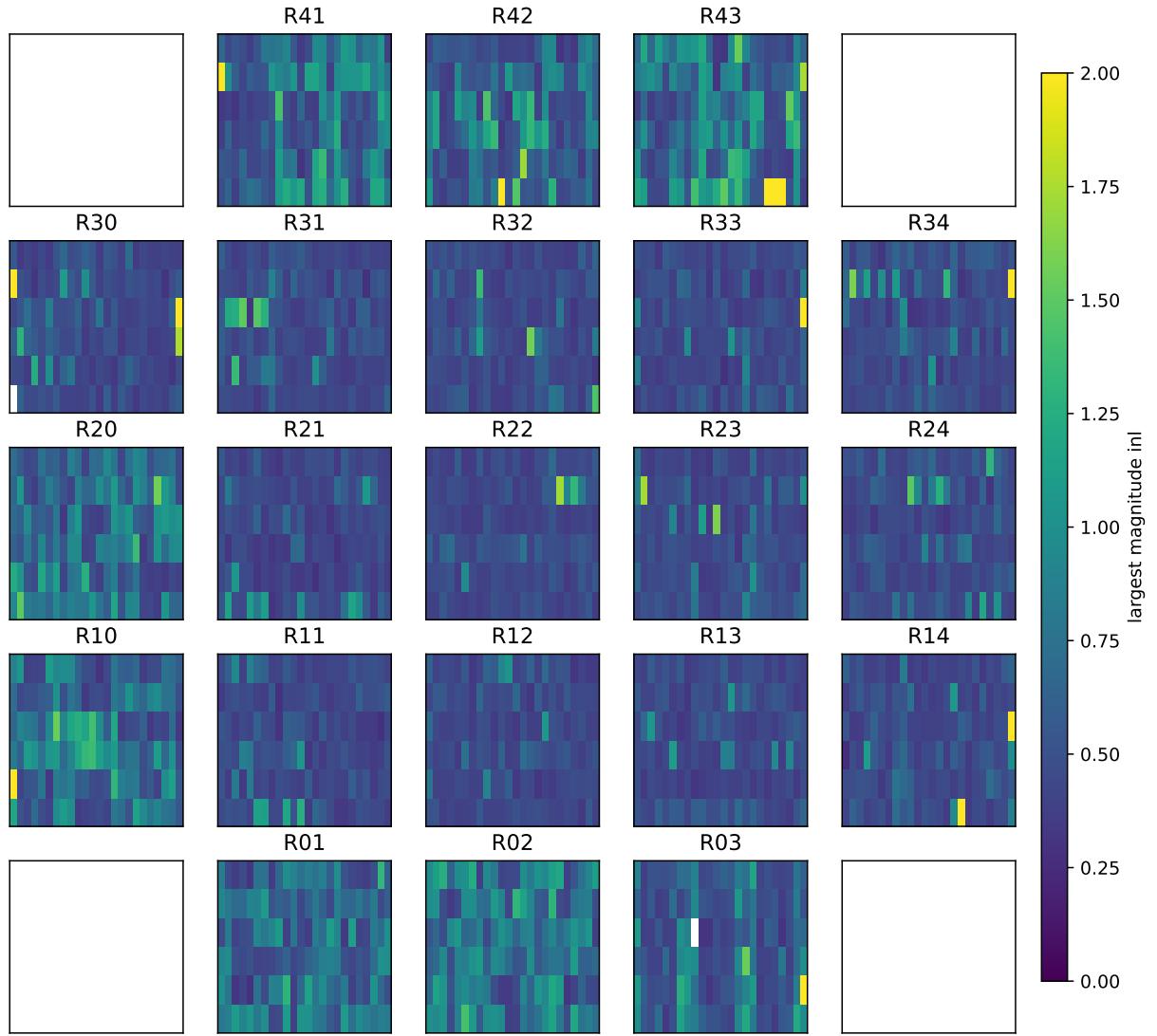


Figure 8: The above figure demonstrates the calibration of the ADC using the differential nonlinearity method we described in section 2.2. This method computes edges along the ADC for each 3024 amplifiers individually. The INL can then be computed for each bin, in each amplifier, and plotted here is 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.

13549 max dnl

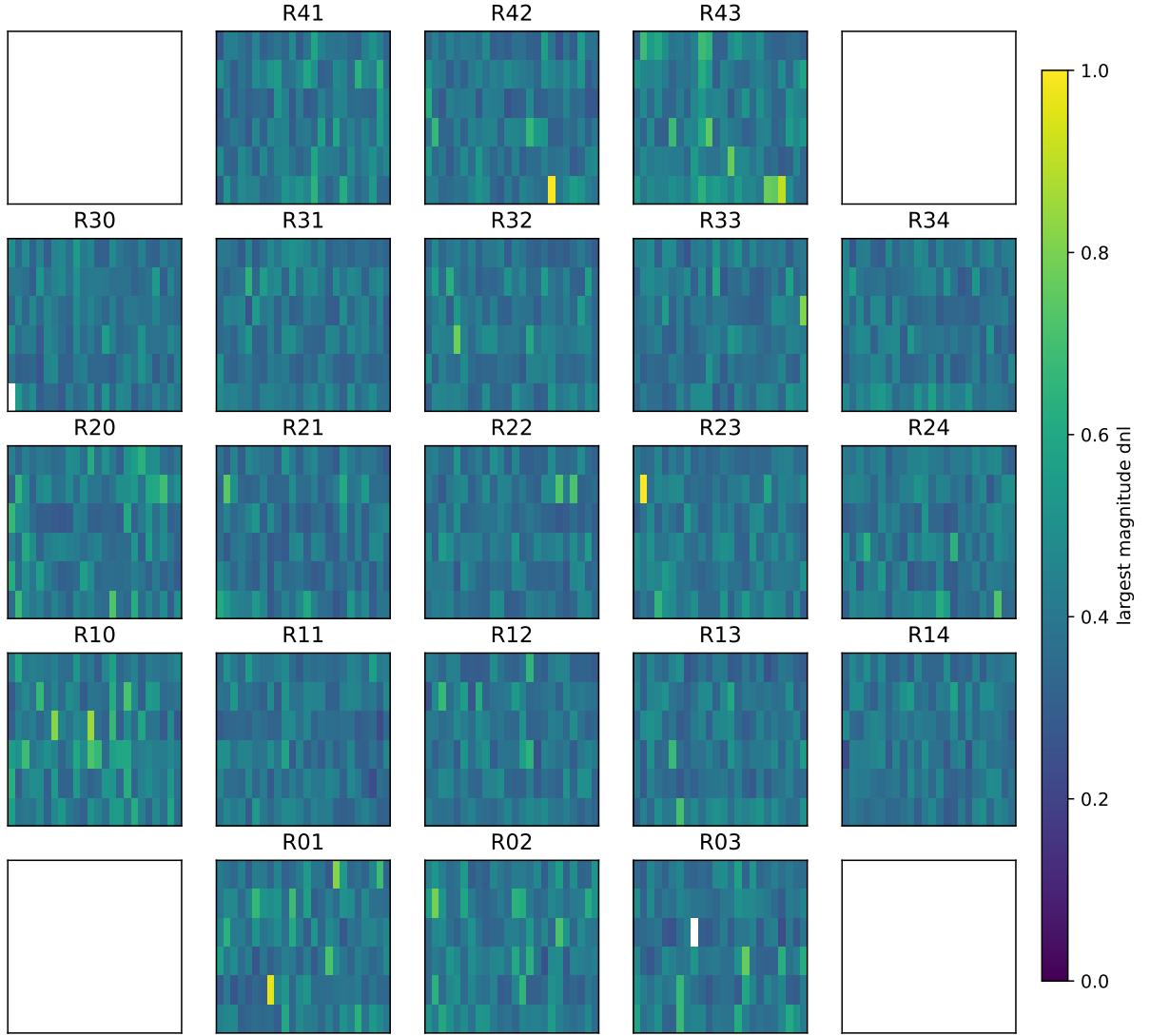


Figure 9: The above figure demonstrates the calibration of the ADC using the differential nonlinearity method we described in section 2.2. This method computes edges along the ADC for each 3024 amplifiers individually. The DNL can then be computed for each bin, in each amplifier, and plotted here is 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.

5 Appendix

This section is aggressively under construction.

In this section we list downfalls encountered along the way of the calibration of the adc. We note them here for the curious reader who may want to do a similar calibration on their own.

a. more about unstable INL - you just verified it existed, but talk about how each binning location made a different inl which was correlated where on the bad distribution binning started. (When starting the calibration method in any given region of the adc, there would be wildly different responses in the inl.) using the cumulative summing on a skewed distribution we still have the inl unbounded (magnitude 15), it is truly the flat distribution with this issue is eliminated

b. There is some notable difference in the locations of the left edges, when you take the histogram the av difference is 0.2 adu, but it is systemic over the entire run. Where as the width difference is centered about 0.

c. work on choosing the filter window

d. convergence interval: time for running and the 0.2 percent difference between the observed counts and that fulfilled by the method on average. Also the timing being about 8s/ 100k edges made. For a bin with an observed depth of over 5000, this is less than 0.2 of a percent difference, which from many iterations and trials of varying convergence intervals this one yields sufficient results in a reasonable timeframe. (rate of edge making I believe it is 8s/100k bins).

e. residuals: Another concern arising from the method is the “residual” or the convergence interval to determine the right edge. As discussed earlier, we allow for a custom convergence interval, which for this work we take to be 1 count. We state that while you could require this convergence to be more strict, we are wary to dictate that the converge shouldn’t be more precise than the rest of the measurements, and thus having a more precise measurement than this (1count/5000), would imply the measurements are more accurate than they truly are. An attempt to reduce residuals by adding/removing the last count from the next bin yielded dnl and inl values not different than ignoring the uncertainty, which suggests such a treatment is unnecessary.