Persistence characterization and data calibration scheme for the RSS-NIR H2RG detector on SALT

Gregory Mosby, Jr.[a], Nathan Eggen[a] and Marsha Wolf[a]

[a]University of Wisconsin - Madison, 475 N. Charter St., Madison, WI, U.S.A.;

ABSTRACT

The University of Wisconsin Madison is building a NIR spectrograph (RSS-NIR) for the Southern African Large Telescope. The detector system uses a H2RG HdCdTe 1.7 μm cutoff array. We performed tests to measure and characterize the persistence of the detector to inform strategies to mitigate this effect. These tests use up-the-ramp group samples to get finer time resolution of the release of persistence. We share these test results. We also present preliminary results of the dependence of persistence on detector temperature. We conclude with an outline and assessment of a persistence calibration scheme.

Keywords: near infrared detectors, infrared arrays, persistence

1. INTRODUCTION

The 11-m Southern African Large Telescope (SALT) is currently the largest optical telescope in the southern hemisphere. SALT hosts a suite of instruments including an imaging spectrograph with several modes, the Robert Stobie Spectrograph[1,2] (RSS) , and a high resolution spectrograph,[3] HRS. These instruments permit the study of a broad range of science topics. To expand SALT's science capabilities further, the University of Wisconsin-Madison is building a NIR spectrograph—RSS-NIR.[4] The spectrograph will be fed by a fiber integral field unit of 300 fibers in a hexagonal configuration. The spectrograph uses volume phase holographic gratings and an articulating six element refractive camera. At the back-end of the camera, RSS-NIR uses a HAWAII-2RG (H2RG) HgCdTe focal plane array with a 1.7 μm cutoff. Teledyne's H2RG HgCdTe detectors have become a staple choice for near infrared instruments in the last 2 decades being used on several instruments. For example, MOSFIRE[5] and the James Webb Space Telescope's NIRCAM, NIRISS, and NIRSpec instruments[6] both utilize H2RGs. However, each H2RG detector system must be optimized for its specified operating conditions. We have optimized the RSS-NIR detector for low read noise and large full well, but, we must also provide a strategy to mitigate persistence.

Persistence is the release of trapped photo-generated charges from a previous exposure in a subsequent exposure. Persistence is commonly seen in NIR detector arrays (e.g. InSb,[7] HgCdTe[8]), but is also seen in some CCD detectors.[9] The presence of persistence in these devices complicates data analysis, especially when precise spectrophotometry or line fluxes are needed. In order to take full advantage of these NIR focal plane arrays, on SALT as well as other facilities, we must find ways to address persistence.

The anticipated varied type of science observations in close succession make RSS-NIR data vulnerable to persistence. SALT observes a wide range of science targets utilizing a queue each night. Persistence using RSS-NIR on SALT might be a problem for routine operation if a user requests arc lamp exposures before a faint observation, if a user requests a grating angle change (bright night skylines would move to different locations), or if going on off target to observe blank sky for objects that fill the slit. More broadly, if NIR observers employ the common nodding technique for sky subtraction, persistence removal will also be important. To minimize this persistence, we currently employ one mitigation technique proposed by Finger et al[11]—global reset detrapping where the detector is continually reset while idle. This resetting has been shown to help detrap persistent charges left over from previous exposures. But, we must still develop a complete strategy to remove persistence in acquired data.

Further author information: (Send correspondence to Gregory Mosby. Jr.)
G.M.: E-mail: mosby@astro.wisc.edu

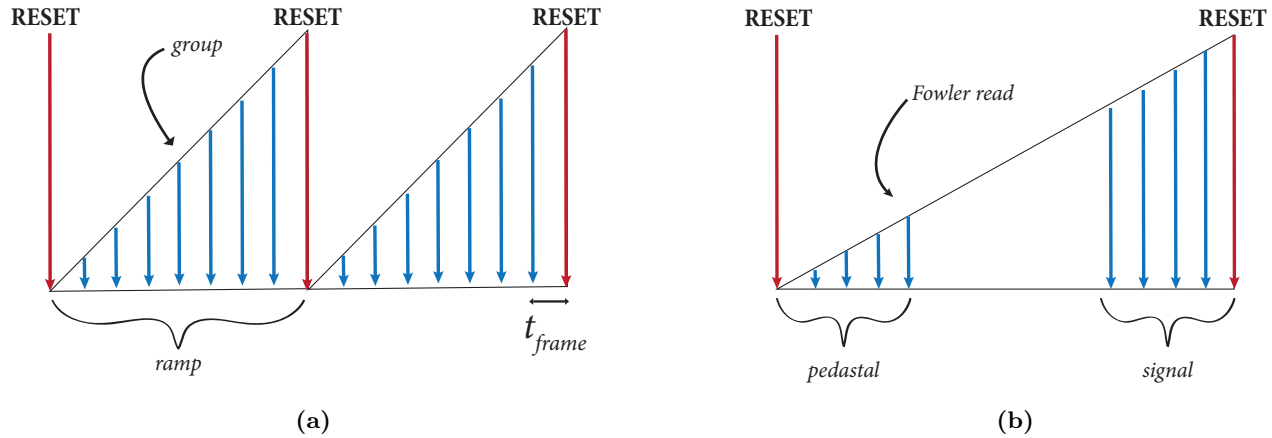


Figure 1: Graphic showing up-the-ramp group sampling (*left*) and Fowler sampling (*right*).

Persistence has been mitigated in other H2RG arrays in additional ways. In Ref. 10, they propose calibrating the persistence and removing it in post-processing. For the JWST mission, persistence is minimized by operating the detectors at very low temperatures. In the low temperature regime (∼40 K for NIRCAM), the persistence decays at very slow rate, but it is also very low amplitude. Another way to mitigate persistence is to expose on different areas of the array. By nodding the telescope, observers can avoid placing science targets on regions that were previously exposed to a large signal (e.g. a foreground star) that will release a large amount of persistence. A less efficient method of minimizing the effects of persistence is to wait for trapped charges to be released before continuing observations. However, the overhead that must be added before new observations reduces observing efficiency. In this paper, we aim to test whether we can correct for persistence using calibration data. In §2 we describe the calibration method and data analysis. In §3 we summarize persistence correction results using persistence decay curves from the calibration data. In §4, we offer some recommendations for future persistence removal methods and characterization.

2. METHOD

We examine how effective a simple persistence correction from calibration data analogous to Ref. 10 is in removing the persistence in up-the-ramp group (URG) and Fowler sampled data (see Fig. 1). However, our study differs in a few key ways. We do not use Fowler sample data to derive the relevant calibration curves, we use URG data. As a result, we can look at how the persistence changes at a higher time cadence. Another key difference is that Ref. 10 looks at how persistence changes up to 7 Fowler frames after an exposure with a signal. We look at how persistence changes within only 1 URG exposure after an exposure with a signal. It is expected for the exposure immediately following an exposure with signal to need the most correction, and using a single frame saves time on data collection. The persistence is first characterized using up-the-ramp group data. The resulting data is then used in a correction algorithm to remove the persistence from collected data. The removal of the persistence is then assessed by comparing the persistence in corrected and uncorrected data.

2.1 Persistence characterization

We focus on two characteristics of the persistence in our H2RG array: the decay time scales as a function of signal and the suppression of the persistence as a function of temporally adjacent exposure signal ratios. The decay properties of the persistence are measured by looking at the integrated persistence in up-the-ramp group sampled data. For all of the data, we run our detector system cooled to 120 K in its faint object mode: gain of 18 dB, $D_{\rm sub} - V_{\rm reset} = 0.499$ V. The CDS read noise of this mode during these tests was estimated at 18.3 e- and the full well in this mode with a 940 nm LED illumination is 73,000 e-. A key variable in the magnitude of persistence is temperature. We show the decay of persistence for three detector temperatures in Fig. 2 derived

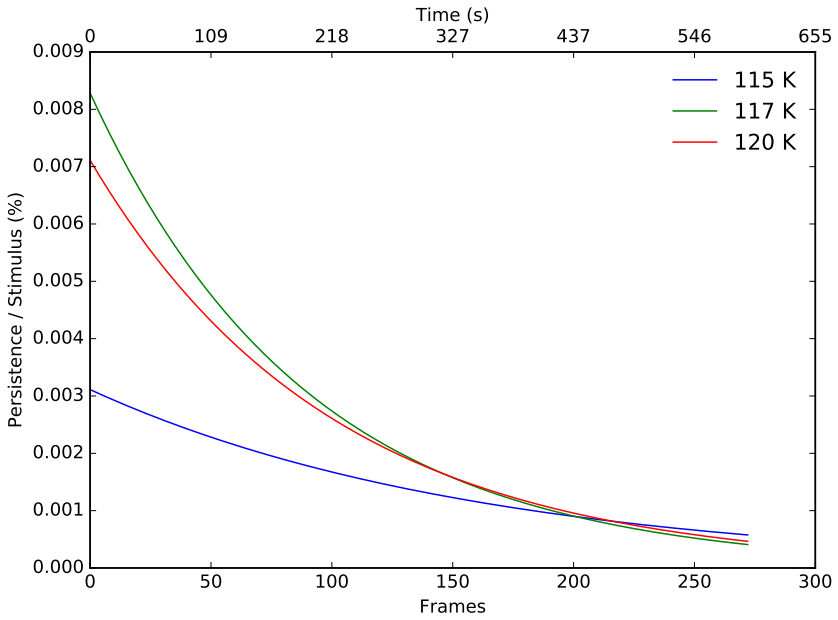


Figure 2: The persistence decay for three detector temperatures 115, 117, 120 K for a mean 31,000 e- 940 nm stimulus. Our system is designed to operate at 120 K.

using the method below to illustrate this, but as previously mentioned we keep detector temperature fixed at 120 K.

In this data, the detector array is illuminated with an NIR LED for a given exposure time to reach a given stimulus signal. At the end of the exposure, the LED is turned off, and the un-illuminated detector is continuously read out for 600 s at a rate of 1 readout per 2.18325 s. An exposure time of 600 s is what is expected to be a typical maximum exposure time of RSS-NIR at SALT set by the sky variability in Sutherland. The cycle of illumination and read out with the LED turned off is repeated for several signal levels with several (typically 10) cycles per signal level (see Fig. 3). The global mean across the detector is measured for each cycle for each group up the ramp in the dark frame. These means are then averaged over all cycles for each signal level. The integrated signal level captured by the averaged means of the dark up-the-ramp group samples is a measure of the integrated persistence released.

The high time cadence of the up-the-ramp group samples allow us to fit precise curves of the integrated persistence as a function of time. We fit these integrated persistence curves with double or single exponential models similar to Ref. 10. See Eq. 1 for the functional form used to fit the integral curves with two exponentials. When the fit of two exponentials is not well matched (by eye) and unconstrained (large covariance values), we reduce the fit to a single exponential*. By taking the time derivative of the fits to these integrated persistence curves, we can look at the persistence decay (see Figs. 4a and 4b).

$$\int_0^{t'} p(t)dt = c + a_1 \exp\left[\frac{-t'}{\tau_1}\right] + a_2 \exp\left[\frac{-t'}{\tau_2}\right], \tag{1}$$

where $p(t)$ is the persistence decay that could be found by differentiating the fit in Eq. 1 yielding:

$$p(t') = -\left(\frac{a_1}{\tau_1}\right) \exp\left[\frac{-t'}{\tau_1}\right] - \left(\frac{a_2}{\tau_2}\right) \exp\left[\frac{-t'}{\tau_2}\right], \tag{2}$$

*Ref. 10 found evidence of traps with varying time constants which might physically represent the different behavior of electrons and holes when they are released.

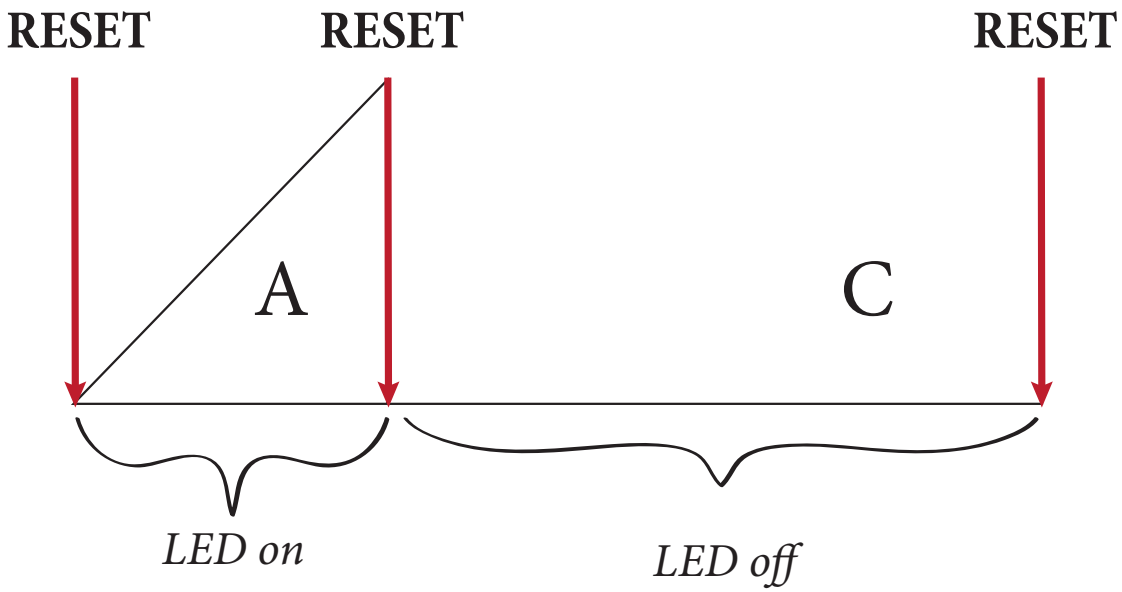
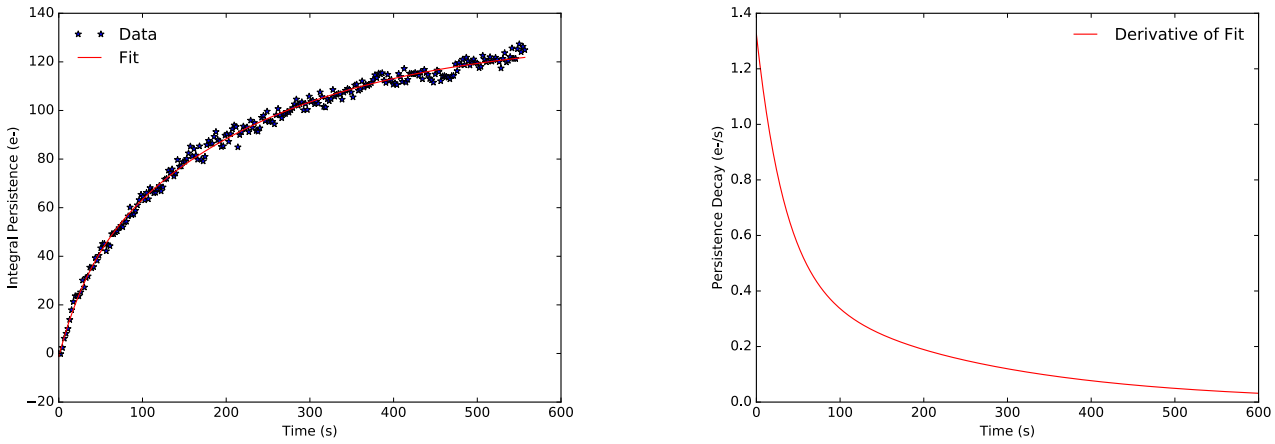


Figure 3: The graphic scheme of the persistence decay tests. One cycle is shown above. About 10 cycles are performed for 9 stimulus levels (i.e. exposure times with constant LED setting).



(a) **(b)**

Figure 4: Example of integral persistence fit (*left*) and derived persistence decay (*right*) for an exposure of 225 s.

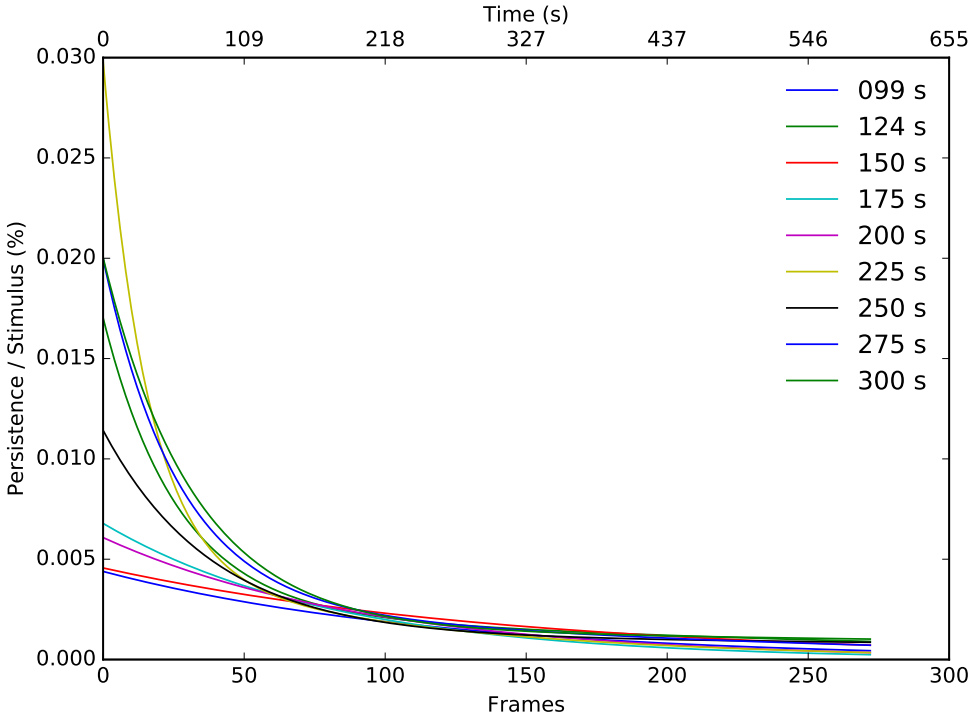


Figure 5: The percentage of stimulus signal that the integrated persistence reaches in a frame time for up to 256 frames after the initial stimulus. For the majority of exposure times, the decay curves have similar shape. For most of the data with lower signal the decay curve has a shallower shape.

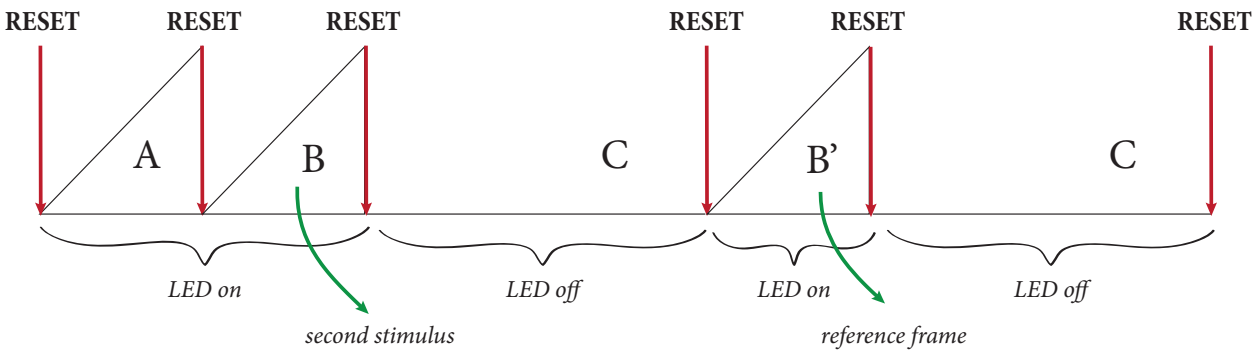


Figure 6: The graphic scheme of the persistence suppression tests. One cycle is shown above. About 10 cycles are performed for 10 stimulus levels (i.e. exposure times with constant LED setting).

Given the persistence decay, we can normalize the derived decay curves by the mean stimulus that produced the persistence and calculate fractional persistence curves. In Fig. 5, we show the fractional decay curves measured for the RSS-NIR detector system.

In addition to the decay properties of the persistence, we must also characterize the suppression of persistence. Ref. 10 found that persistence from a previous exposure can be suppressed in a subsequent frame by the presence of a signal. Physically, while traps left un-illuminated in the subsequent frame (those remaining in the depletion region) can emit their charge, traps currently illuminated (past the depletion region edge) release their charge

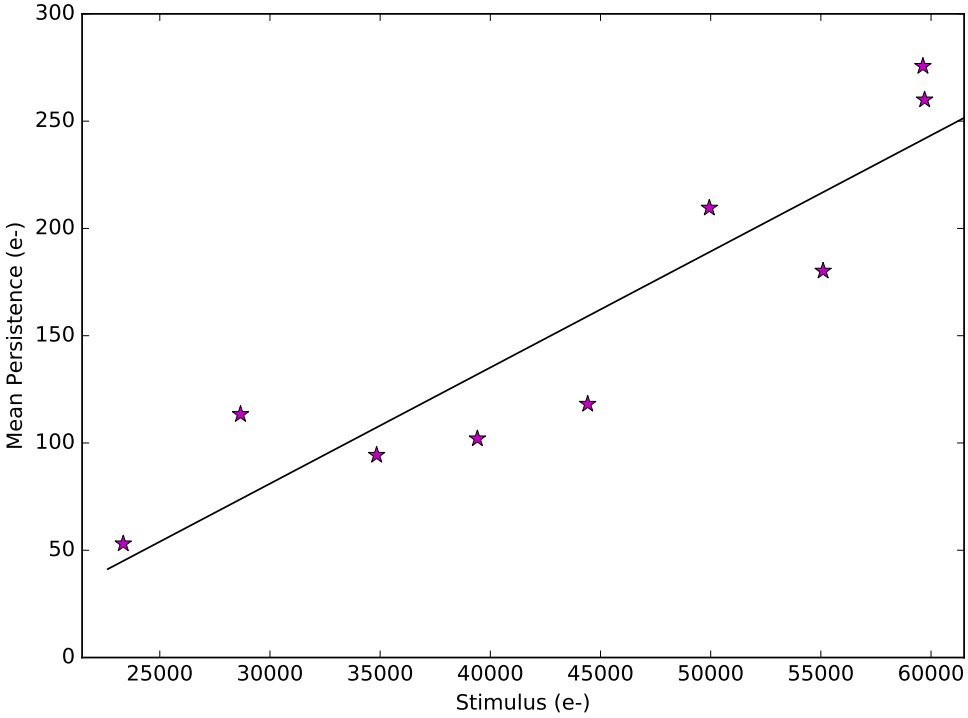


Figure 7: The linearity of persistence from persistence in a 600 s exposure following the stimuli tested to derive decay curves in Fig. 5. The last two data points do have the same stimulus level (exposure time). While errorbars for this fit would be a welcome diagnostic, we omit them from this plot. Because each data point here represents the evaluation of a fit to the average of about 10 cycles of persistence decay tests, deriving a physical and meaningful errorbar would require duplicate cycles at each signal level which is a considerable time investment.

but quickly absorb more photo-generated charge. The net effect is a reduction in persistence. This reduction would depend on the ratio of the intensities of the two consecutive stimuli. To characterize the suppression in the RSS-NIR detector system we perform a series of tests in which we illuminate the detector with several initial stimuli, followed by illumination with a reference stimulus. This double stimulus is followed by a series of dark exposures, and those dark exposures are followed by the reference exposure again (see Fig. 6). Because we use a fixed LED intensity, we achieve the different stimuli by changing the exposure time. We use a 75 s reference exposure and first stimuli ranging from $100 - 400$ s.

By looking at the difference between the reference exposure following the initial stimulus (frame B in Fig. 6) and the reference exposure (frame B') following the darks, we can estimate the persistence from the first stimulus (frame $B - B'$). The persistence estimated for the first stimulus is then compared to the expected persistence estimated using the integrated persistence calculated earlier. We examine the ratio of the estimated persistence from the first stimulus in the presence of a second signal and the persistence from an isolated equivalent signal as a function of the relative strength of the first and second stimulus. This comparison allows us to estimate the suppression of persistence as a function of relative intensity. For our comparisons, since the exposure time of the reference differs from the exposure time of our dark frames used for the single stimulus persistence measurements, we integrate the derived persistence decay curves for frame $B - B'$ to equivalent exposure times. For signal strengths that we do not test in the original decay curve tests, we estimate the persistence from the linear relation between total persistence in a 600 s exposure versus signal derived from the decay tests (see Fig. 7).

In Fig. 8, we show the suppression as a function of relative intensity for the values tested. Data points with

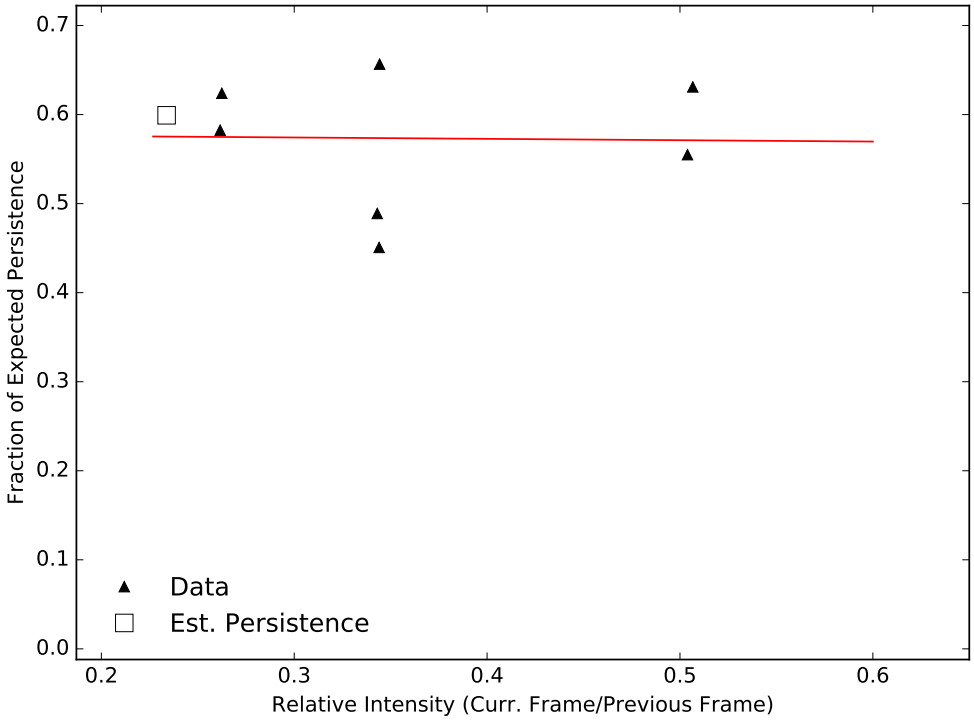


Figure 8: The suppression of persistence as a function of relative intensity of an initial and subsequent illuminated frame. The fraction of the expected persistence observed in an illuminated frame following another is surprisingly flat. To derive the persistence in the shorter frame, we integrate our derived decay curve to an exposure time equivalent to that of the longer exposure time.

persistence estimated from the linear approximation are represented by squares. We fit a linear relation to the suppression curve in Fig. 8. Three estimated points are excluded from this fit as they have estimated persistence close to the read noise and signals strengths that we would expect to have negligible persistence. The linear fit in Fig. 8 ideally can be used to estimate the suppression of persistence when correcting our data. However, our persistence suppression plot is flat. It differs from the decaying plot in Ref. 10 (their Fig. 17). Our results would suggest that there is less persistence in the frames that follow a stimulus, but that this suppression is insensitive to relative intensity. We discuss this result more in the discussion section.

2.2 Persistence correction

Using the characteristics of the persistence in the RSS-NIR H2RG found in §2.1, we can assemble an algorithm to predict the amount of persistence in a frame given the pixel's observation history. For each pixel in an exposure, given its previous integrated signal, its final integration time, and the decay curves in Fig. 5, we can estimate the amount of persistence expected. This estimate amounts to multiplying the previous pixel signal by an appropriate decay curve integrated from the starting time of the current frame after the stimulus and the current frame's end. This estimate of the persistence is then modified by the suppression curve in agreement with the intensity ratio between the previous and current signal. Finally, the estimated persistence is scaled by a normalized dark to account for spatial variation in the persistence after which it is subtracted from the current frame. This is summarized in the following equation:

$$P(x,y) = I_0(x,y)S(x,y;\eta)F(x,y)\int_0^{t'} f(t)dt, \tag{3}$$

where $P(x, y)$ is the derived persistence image, $I_0(x, y)$ is the previous exposure, $S(x, y; \eta)$ is the suppression as a function of η, current signal divided by previous signal, $F(x, y)$ is the normalized dark, and $f(t)$ is the fractional amount of persistence released per second with t' representing the current image's exposure time. We describe these components of the persistence correction in more detail below.

2.2.1 Persistence decay curves

As can be seen in Fig. 5, the decay shape as a function of time varies with exposure time. This is unsurprising in light of Smith et al.'s theory of persistence that would suggest longer exposures allow potential traps to be exposed to charge for longer periods of time. But these results raise the question, which decay curve should be used for correction? The RSS-NIR instrument will observe at a variety of exposure times. It's expected that these exposure times will be limited to 10 minutes, but the exposure cadence is otherwise undefined. Thus, a priori, there is no preferred decay curve as function of exposure time. We test two options for decay curve choice: a weighted sum of decay curves and the nearest decay curve. For the weighted sum decay curve there are two options: an inverse exposure time weighted sum decay curve and an equally weighted sum decay curve. For the inverse exposure time weighted curve we sum each decay curve multiplied by the reciprocal of its exposure time and then divide the sum by the sum of the weights. The result is fit to attain an analytic decay curve. For the equally weighted decay curve we simply take the fit of the mean of the decay curves. These decay curves are shown against the backdrop of the individual decay curves in Fig. 9.

2.2.2 Persistence suppression curve

Ideally in the persistence correction to find the fraction of the persistence from a previous signal that is suppressed, we would evaluate the fit to the suppression as a function of relative intensity (current stimulus by previous stimulus). We would pin any points that evaluate to a suppression factor greater than 1 at 1.0 and any points that evaluate to negative values as 0.0 as these values would be inconsistent with suppression and not physically motivated in the model for persistence that we adopt from Ref. 8. However, as our suppression curve is flat, we do not use suppression in the correction for the following tests. This is equivalent to setting $S(x, y; \eta)$ to one in Eq. 3.

2.2.3 Persistence spatial variation

Ref. 10 found that the spatial variation of persistence differed from the spatial variation of the illumination and quantum efficiency, so this variation is important to account for in the persistence correction. Similar to Ref. 10 we use the average of the first dark frames following a stimulus that have been divided by the stimulus. The resulting average persistence ratio map reveals the structure of persistence for our detector. We normalize this average image by its mean to use this image as a "persistence flat" in the persistence correction. We note that Ref. 10 found that the spatial structure of persistence does change in subsequent frames. However, Ref. 10 found that calibration errors are relatively small ($8.8 - 12\%$) when using only the first dark frame.

2.3 Correction assessment

We can test how well using persistence calibration data to remove persistence according to our algorithm performs by comparing the amount of persistence remaining after correction to the amount of persistence present originally. This is measured by taking the mean of persistence corrected dark frames taken after a stimulus and dividing by the mean of the uncorrected dark frames taken after a stimulus. This ratio is the fraction of persistence remaining after correction. The results of these comparisons on up-the-ramp group (URG) and Fowler data are detailed in §3.

3. RESULTS

We assess the persistence correction technique outlined in §2.2 using data taken in URG and Fowler mode. The URG data comes from individual cycles that were averaged to form the decay curves. The Fowler data was taken independent of the persistence decay measurements. We compare persistence corrected single images to their uncorrected counterparts for both modes below.

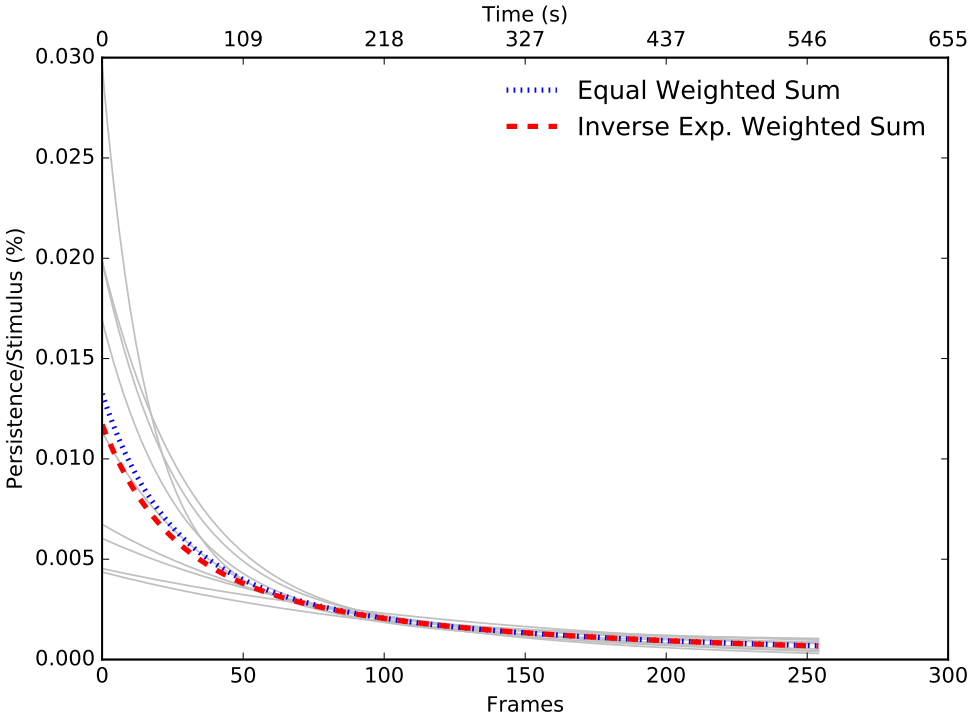


Figure 9: The inverse exposure time weighted decay curve (*red dashed*) and the equal weighted sum (mean) decay curve (*blue dotted*). For the inverse exposure time weighted decay curve we fit the sum over all measured decay curves weighting each decay curve by the reciprocal of its exposure time and then divide the sum by the sum of the weights. The decay curve derived from fitting the mean value as a function time for all decay curves is shown as well. These decay curves are shown above the backdrop of the individual decay curves (*grey*).

3.1 URG tests

We examine the efficiency of our persistence correction technique for decay data that was taken using 3 exposure times 175 s, 200 s, and 300 s. Fig. 10 shows a plot of the median percentage of persistence remaining after correction using the described decay curves (exposure time weighted, equally weighted, and closest). The error bars show the 16th and 84th percentiles, an equivalent 1 sigma range for a normally distributed sample. We find that for the 175 s and 200 s exposures, very little $\sim -10 - 10\%$ of the original persistence remains after correction with the exposure time weighted decay curve. This amounts to a decrease of nearly $90 - 110\%$ of the persistence. To compare the corrected and uncorrected frames, it is also useful to look at their pixel histograms across the full array (see Fig. 11). In Fig. 11, we can see that the correction moves the median value of one 200 s image to approximately the numerical value of the read noise for this mode, representing a good quality correction. The quality of the correction is seen more sharply in comparing the image of the persistence and the corrected frame. This comparison is shown for the 200 s exposure in Fig. 12. In contrast to the 175 s and 200 s results, using the weighted sum decay curve on the 300 s data leaves about 20-30% of the persistence. Using the weighted sum decay curve, we can eliminate much of the persistence, but it does appear that this single curve might not be widely applicable. We find similar results if we use the equally weighted decay curve (Fig. 10). However, more than 80% of the persistence for the 300 s exposure is able to be removed when we use the closest in exposure time decay curve (Fig. 10) with a propensity for over-subtraction. Notwithstanding, the results for the lower exposure times stay promising with 80-90% of the persistence being removed.

We also investigate the efficiency of our persistence correction technique using the persistence suppression data taken. These data are slightly different than the decay data above. A measure of the persistence can be easily obtained from the decay data by comparing the signal in the dark frame following the stimulus (B in

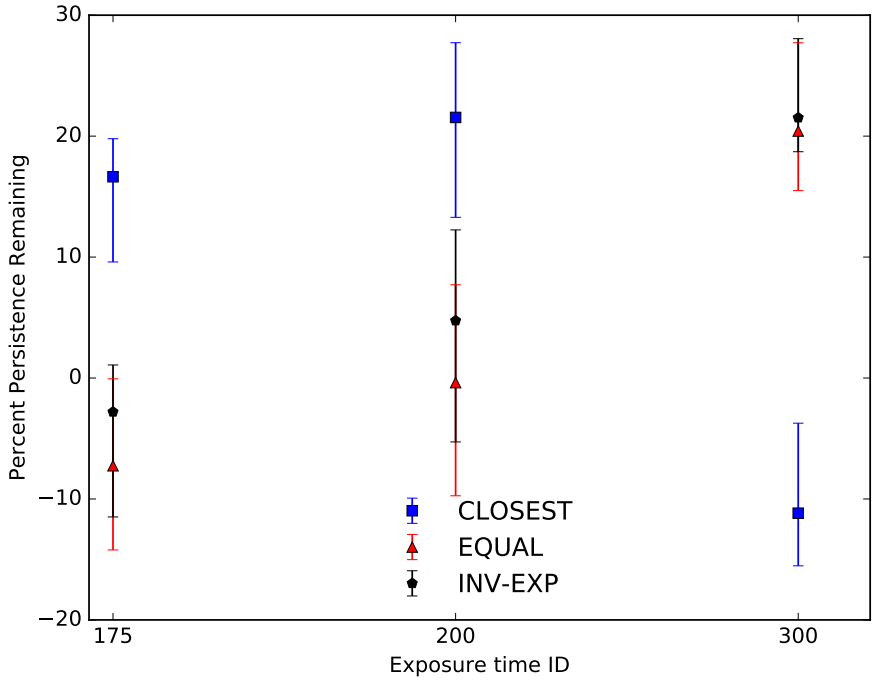


Figure 10: The median percent of persistence remaining in corrected images by URG exposure time using the inverse exposure time weighted decay curve, the equally weighted sum decay curve, and the decay curve with an exposure time closest to the first stimulus image's exposure time. For the 175 s and 200 s trials much of the persistence can be removed, though there is some over-subtracting in the 175 s data. The single weighted sum decay curve performs worse on the 300 s exposure data. However, using the weighted sum decay curve, we can eliminate more than 60% of the persistence in URG images. The results using the equally weighted sum curve mirror those of the inverse exposure time weighted sum decay curve. It is noticeable though that the 200 s median percent persistence remaining falls closer to 0. The results of the correction using the decay curve of the same exposure time show the most diverse results. It alone removes more than 80% of the persistence in the higher flux frame, though there is over subtraction.

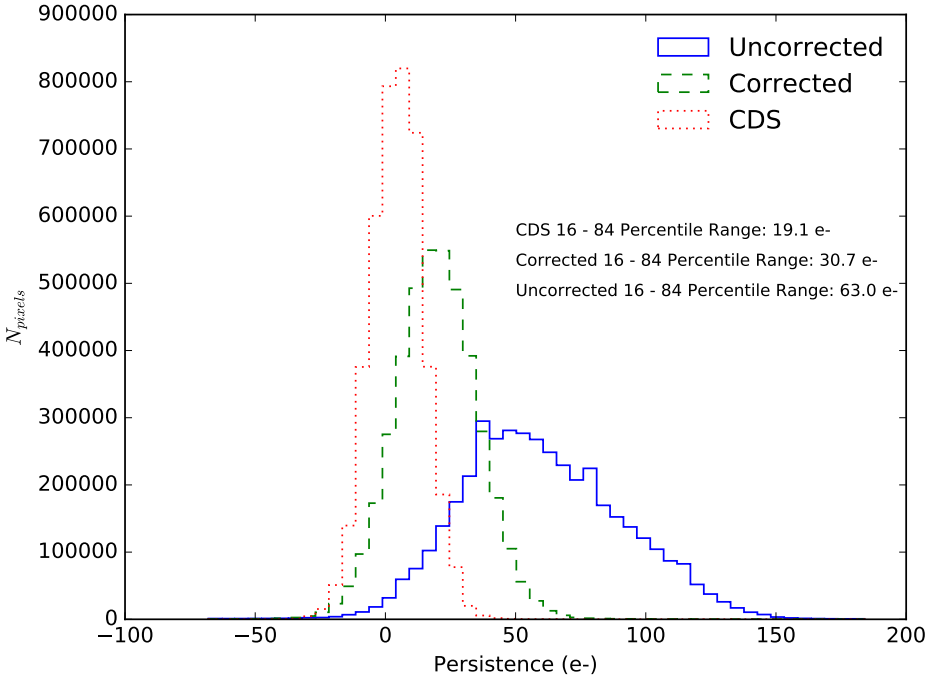


Figure 11: Comparison of the image histograms for persistence corrected (*green dashes*) and uncorrected frame (*blue line*) for the URG 200 s exposure. The persistence correction gets rid of excess variation in the uncorrected frame. A correlated double sample (CDS) frame image histogram (*red dotted*) with no signal is shown for reference.

Fig. 6) before and after persistence correction. For the suppression data, we must compare the difference of the corrected and un-corrected frames (B) with the reference frame (B' in Fig. 6). The excess signal in frame B above B' is from persistence from frame A. This excess signal should be minimized if the persistence correction is efficient. Fig. 13 shows the median percentage of persistence remaining after correction for our suppression data.

We find in Fig. 13 that the persistence correction using the exposure time weighted decay curve is only able to remove a mean of 57% of the persistence, across all exposure times tested with the exception of the low signal case (150). The results are similar for the equally weighted decay curve. For the closest decay curve method, we see a similar pattern of the correction being more efficient at higher signal levels than the previous decay curves tested, however, the amount of persistence removed is still an average of 62%. Note that for the 225 trial, the data derived decay curve is more steep than the other derived decay curves which likely explains its more efficient correction. Setting aside the exceptions, the persistence correction is worse than in the simple decay case. This would imply that we are underestimating the persistence in these stimulus frames. This is even more surprising since we are not using any suppression for this data correction.

3.2 Fowler data

In addition to URG tests, we test the generality of our persistence correction scheme using limited Fowler data taken in two configurations similar to the single (decay data) and double stimulus (suppression) tests for URG sampling (see Fig. 3 and Fig. 6). We asses the effectiveness of our persistence correction on Fowler data in the single stimulus test by comparing the residual image of the corrected and uncorrected first Fowler sampled dark after the stimulus. The exposure times used in these trials are 175 s and 300 s. Fig. 14 shows the percent of persistence remaining after applying a persistence correction according to the exposure time and previous

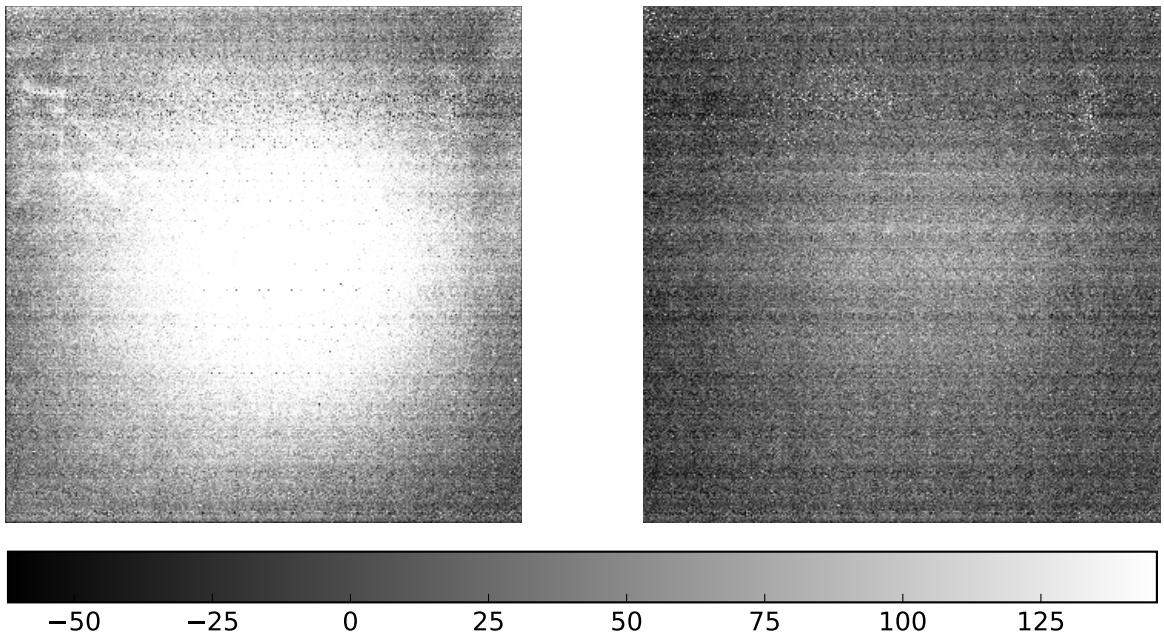


Figure 12: Comparison of the un-corrected image (*left*) and corrected image (*right*) for the URG 200 s exposure.

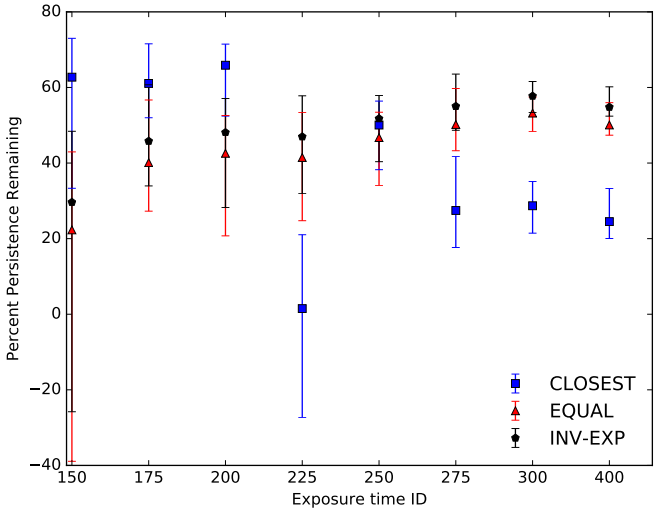


Figure 13: The median percent of persistence remaining in corrected double stimulus URG images using the inverse exposure time weighted decay curve, the equally weighted sum decay curve, and the decay curve with an exposure time closest to the first stimulus image's exposure time. Medians and 16-84 percentile ranges are derived from the sample of cycles per exposure.

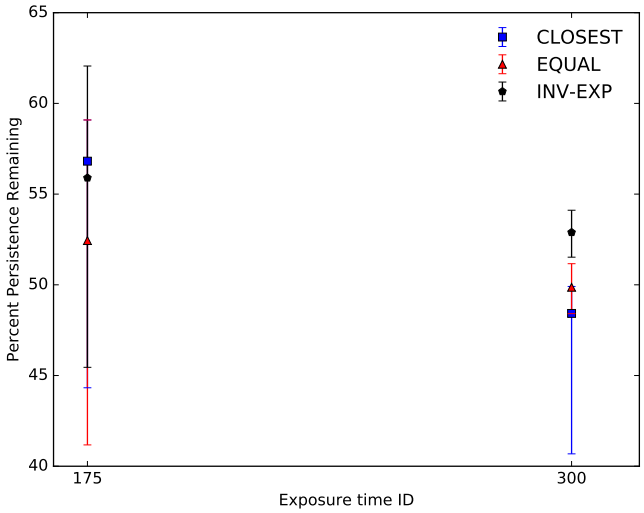


Figure 14: The median percent of persistence remaining in corrected FS images following a single stimulus using the inverse exposure time weighted decay curve, the equally weighted sum decay curve, and the decay curve with an exposure time closest to the first stimulus image's exposure time. Medians and 16-84 percentile ranges are derived from the sample of cycles and Fowler samples (4, 8, 16, 32) per exposure.

stimulus image using our 3 decay curve methods. The correction with the exposure time weighted curve removes 40 - 60% of the persistence from the 175 s data and only about 48% persistence for the 300 s trial. Using the equally weighted decay curve the median persistence remaining is under 55%, and for the 300s data, the median amount of persistence remaining dropped to 50%. When we use the closest decay curves, the median amount of persistence remaining for the 175 trials are worse than the equally weighted decay curve and similar to the exposure time weighted decay curve. The median persistence remaining in the 300s trial, however, decreases below 50%. These results are worse than the URG decay data results. It seems our URG decay curves under-predict the persistence in the dark Fowler frames following a stimulus.

We present results from two double stimulus Fowler tests. In the first trial, we follow a 175 s Fowler readout with a 300 s Fowler readout. In the second trial, we decrease the LED intensity for the second stimulus holding the exposure time constant at 175 s. Fig. 15 shows the median percent of persistence remaining in these trials using the same decay curves from Fig. 9. For the first trial (175s exposure followed by the 300 s exposure, called 175 in Fig. 15), the exposure time weighted decay curve corrected image contains less than 20% of persistence left across the 4 Fowler types. Moreover, the persistence is corrected to a numerical value below the value of the Fowler read noise (Fowler-4: 16.1 e-, Fowler-8: 12.7 e-, Fowler-16: 10.5 e-, Fowler-32: 9.1 e-) for each of the 4 samples tested. For the second trial, the correction removes about 75% of the initial persistence. The persistence does not decrease below the Fowler RN in this trial. The equally weighted decay curve correction improves on the previous results, with a median less than 20% of the persistence remaining in the second trial and median 12% persistence remaining in the first trial. The closest decay curve achieves the minimum persistence remaining for the first trial, but for the second trial. In the second trial, the closest decay curve correction does comparable to the other decay curves with a wider spread across the different Fowler pairs.

Overall, these results are better than the Fowler decay results. In the first trial, we would expect to see results similar to the 175 s FS decay tests. Since the second stimulus is greater than the first, we would expect no suppression. However, it seems we do much better than the 175 s FS decay test. For the second trial, we would expect that there should be some suppression, but the data suggest that a no suppression correction on Fowler double stimulus data is just as efficient as the correction on URG decay data. But, this result does not show an excess of persistence seen in the URG double stimulus data. This is likely influenced by the differences

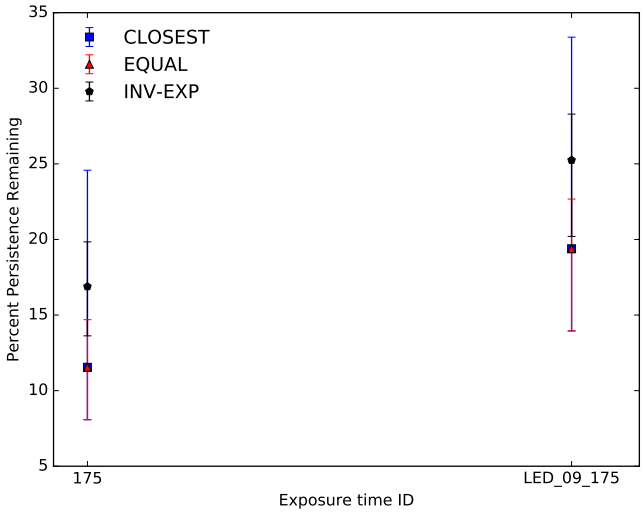


Figure 15: The median percent of persistence remaining in corrected double stimulus FS images using the inverse exposure time weighted decay curve, the equally weighted sum decay curve, and the decay curve with an exposure time closest to the first stimulus image's exposure time. Medians and 16-84 percentile ranges are derived from the sample of cycles and Fowler samples (4, 8, 16, 32) per exposure.

in the two sampling methods. While we get a single measurement of the integral signal in URG sampling. For Fowler sampling, we take several displaced in time measurements of the integral signal that are then equally weighted in a mean calculation. For primarily linear signals, this is a valid approximation to the integral signal with improved signal-to-noise. For a nonlinear signal, such as the persistence decay, Fowler sampling might underestimate the integral signal.

4. DISCUSSION

Persistence is a common issue with NIR detectors that is a confounding factor in signal measurement. In this paper we test whether we can correct for persistence using calibration data. We characterize our persistence with URG data and use those data to subtract predicted persistence off from individual images, and we compare the corrected images with the uncorrected frames. We find that we can correct for persistence using the URG calibration, but two critical points emerge. There must be a wise selection of the decay curve used in the persistence correction, and we must explain an underestimation of persistence in the presence of signal.

4.1 Which decay curve?

Our results for the decay data, not surprisingly suggest that using a single decay curve for all images is worse than using decay curves derived from similar exposures. Persistence does depend on stimulus intensity and stimulus duration. For our tests, stimulus intensity was held fixed, but the duration was allowed to vary, and this did produce different decay curves. In the future, to correct the widest range of type of data, it might be wise to derive an empirical decay curve that transforms the decay time constants and amplitudes with stimulus duration.

4.2 Why no suppression?

The performance of our persistence correction on the suppression data is surprising. We find that we are not removing a substantial amount of persistence in the URG data even though we run the removal algorithm on the suppression data without the suppression factor (see Fig. 13). This hints at an underestimation of the persistence being released. The effectiveness of our simple algorithm on the URG single stimulus decay test data and similar

data that was not used in deriving the fractional decay curves suggests that there is additional persistence present in the second stimulus of our suppression data that is not present at a similar time after a single stimulus. This contributes to the flat suppression curve in Fig 8.

Two key factors at play in the formation of the suppression curve are the different exposure times between the two stimuli and the speed at which persistence is released. Since our exposure times for the second stimulus are shorter than our decay data, to compare the persistence expected in the suppression trial to that observed, we must integrate the second stimulus' derived decay curve to the matched exposure time of the decay data. We find that on this longer timescale that the persistence that would be released in the second stimulus had it the same exposure time would be less. However, if we instead integrate the decay curve of the decay data for the shorter exposure time of the second stimulus, we find that the suppression curve becomes positive—more persistence is released in the second stimulus frame in the same amount of time. This suggests that the persistence during the second stimulus is released much quicker than it is released in a dark frame. It seems the accelerated release of persistence leads to the our underestimation of the estimated persistence in the persistence correction in frames with signal. The magnitude of this mismatch is median 51% (43% mean), and this mismatch seems relatively flat as a function of intensity ratio, at least for the single decay curve URG tests. It would make sense for the ratio of estimated to expected persistence to be constant at different intensity ratios if the persistence in both frames (dark decay data frames and second stimulus frames with signal) used to estimate the ratio scale with their respective stimuli—a reasonable assumption.

We rule out a number of simple explanations for the increase in persistence for the URG double stimulus data. Persistence depends on detector temperature, but for both the decay and suppression tests, the detector temperature does not change more $\sim$0.5 K. Persistence also depends on accrued signal. In these suppression tests, however, the first stimuli (presumed to be responsible for the persistence in the second stimuli) do not have significantly different exposure times (signal) than the stimuli in the decay test data. Moreover, the LED intensity was fixed at the same level for all of these trials. Only the exposure times were changed.

One plausible explanation for the excess persistence we see in the suppression tests is an underestimation of the persistence present. Our decay data used to derive the decay curves have a pause before the un-illuminated detector data is collected for the experimentalist to switch off the LED. In our current setup, the detector continuously resets in IDLE mode which has been shown to decrease persistence.[11] In the time between the decay data collection, if the detector goes into IDLE mode, our measurements of the persistence would be depressed lower than they would be in the suppression tests where there is no significant time delay between the first and second stimulus. As a result, we would see an enhancement of persistence in the suppression tests. However, the decrease seen in Ref. 11 were from using global resetting. We employ resetting by line as we saw better noise performance with this setting. To test this hypothesis, we would need to look at the persistence characteristics without resetting in IDLE mode.

The addition of persistence in the second frame might also be an accelerated release of charges from traps whose closing energy match the energy of the 940 nm LED (1.32 eV). Persistence is thought to be the trapping of charge carriers at intermediate energy levels between the valence band energy and the conduction band energy of the detecting semi-conductor. Sites in the semi-conductor that become traps are called so because charge carriers are not released until energy is provided to excite the trap to release the carrier. One proposed mitigation strategy for decreasing persistence proposed by Ref 10 is the"flashy reset". In this scheme, the detector is illuminated with light (potentially above and below the detector cutoff) to accelerate the releasing of persistence during a reset frame. The broad illumination is used to provide energy to close traps that might require energies larger than the band gap energy. In this method, it's crucial that this illumination happens in a reset frame so that any photo-generated charge is not detected. In our suppression test setup, we illuminate the array while persistence from the previous stimulus frame is being released. If there is a trap population sensitive to 940 nm light there might be accelerated release of persistence observed. In order to test this hypothesis, we can switch the wavelength of the LED from 940 nm to 1550 nm and see if the additional persistence in the second of two stimuli remains. If we are indeed seeing the accelerated closing of traps resonant with 940 nm light, we would expect the amount of additional persistence to vary with wavelength.

4.3 Persistence characterization recommendations

In our trials of characterizing the persistence in our setup we followed these general steps.

1. Characterize persistence decay after several stimulus levels using URG data with multiple cycles to average integral persistence measurements.
2. Characterize persistence suppression using double stimuli of varying intensity ratios also using URG cycles.
3. Fit integral persistence from decay data with single or double exponential laws. The time derivative of this integral persistence is the rate of persistence emission per second, the persistence decay curve.
4. Divide derived persistence decay curves by the mean stimulus of the respective decay data to form fractional decay curves.
5. Use matched integrations of the persistence decay in referenced subtracted second stimulus data and the dark images following a single stimulus to compare the estimated to expected persistence ratio at various intensity ratios. This forms the suppression relation.
6. Use Eq. 3 and derived curves to correct images directly following illuminated frames.

Beyond the data and methods listed, however, we would recommend additional data on the persistence at different wavelengths. This can be useful in diagnosing any dependence of the persistence on incoming photon energy. We would also recommend using a tunable, automated light source (or shutter). In our trials for the decay tests, we have small but measurable gaps between when the LED is turned off and data taking resumes. This is and can be taken into account in the decay curves derived here, but with a synchronized shutter or LED, we could have measured more of the initial persistence as it was released. Additionally, a finely tunable light source would make performing suppression tests with different intensity ratios easier, and with enough flexibility in the light source, one could eliminate the need to change exposure time in suppression tests.

5. CONCLUSIONS

The calibration with URG data seems to work well when decay curves are used that more closely match those derived from similar stimuli, though at low signals a single derived decay curve can remove ∼80% of persistence. However, removing persistence in frames with signal needs more exploration. We observe what appears to be an accelerated release of persistence in those frames not accounted for by the model of persistence adopted. Additionally, the URG calibration data did not consistently perform well on Fowler sampled data. There was better correction of persistence in the presence of a signal than in un-illuminated frames.

ACKNOWLEDGMENTS

This work was carried out under Washburn Astronomical Laboratories, the instrumentation group within the University of Wisconsin-Madison Astronomy Department that is partially funded by the UW College of Letters and Science.

REFERENCES

[1] Kobulnicky, H. A., Nordsieck, K. H., Burgh, E. B., Smith, M. P., Percival, J. W., Williams, T. B., and O'Donoghue, D., "Prime focus imaging spectrograph for the Southern African large telescope: operational modes," in [*Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*], Iye, M. and Moorwood, A. F. M., eds., **4841**, 1634–1644 (Mar. 2003).

[2] Wolf, M. J., Mulligan, M. P., Smith, M. P., Adler, D. P., Bartosz, C. M., Bershady, M. A., Buckley, D. A. H., Burse, M. P., Chordia, P. A., Clemens, J. C., Epps, H. W., Garot, K., Indahl, B. L., Jaehnig, K. P., Koch, R. J., Mason, W. P., Mosby, G., Nordsieck, K. H., Percival, J. W., Punnadi, S., Ramaprakash, A. N., Schier, J. A., Sheinis, A. I., Smee, S. A., Thielman, D. J., Werner, M. W., Williams, T. B., and Wong, J. P., "Project status of the Robert Stobie spectrograph near infrared instrument (RSS-NIR) for SALT," in [*Ground-based and Airborne Instrumentation for Astronomy V*], **9147**, 91470B (July 2014).

[3] Bramall, D. G., Sharples, R., Tyas, L., Schmoll, J., Clark, P., Luke, P., Looker, N., Dipper, N. A., Ryan, S., Buckley, D. A. H., Brink, J., and Barnes, S. I., "The SALT HRS spectrograph: final design, instrument capabilities, and operational modes," in [*Ground-based and Airborne Instrumentation for Astronomy III*], **7735**, 77354F (July 2010).

[4] Sheinis, A. I., Wolf, M. J., Bershady, M. A., Buckley, D. A. H., Nordsieck, K. H., and Williams, T. B., "The NIR upgrade to the SALT Robert Stobie Spectrograph," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], **6269**, 62694T (June 2006).

[5] Kulas, K. R., McLean, I. S., and Steidel, C. C., "Performance of the HgCdTe detector for MOSFIRE, an imager and multi-object spectrometer for Keck Observatory," in [*High Energy, Optical, and Infrared Detectors for Astronomy V*], **8453**, 84531S (July 2012).

[6] Gardner, J. P., "The James Webb Space Telescope," in [*AKARI, a Light to Illuminate the Misty Universe*], Onaka, T., White, G. J., Nakagawa, T., and Yamamura, I., eds., *Astronomical Society of the Pacific Conference Series* **418**, 365 (Dec. 2009).

[7] Campbell, R. D. and Thompson, D. J., [*Scientific detectors for astronomy 2005: Explorers of the Photon Odyssey*], ch. The Effects of Charge Persistence in Aladdin III InSb Detectors on Scientific Observations, 507–514, Springer Netherlands, Dordrecht (2006).

[8] Smith, R. M., Zavodny, M., Rahmer, G., and Bonati, M., "A theory for image persistence in HgCdTe photodiodes," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7021**, 0 (July 2008).

[9] Baggett, S., Biretta, J., and Hsu, J. C., "Update on Charge Trapping and CTE Residual Images in WFPC2," tech. rep. (Sept. 2000).

[10] Smith, R. M., Zavodny, M., Rahmer, G., and Bonati, M., "Calibration of image persistence in HgCdTe photodiodes," in [*Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*], *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* **7021**, 0 (July 2008).

[11] Finger, G., Dorn, R. J., Eschbaumer, S., Ives, Mehrgan, L., Meyer, M., and Stegmeier, J., "Recent Performance Improvements, Calibration Techniques and Mitigation Strategies for Large-Format HgCdTe Arrays," in [*ESO Detectors for Astronomy Conference*], (2009).