Nancy Grace Roman Space Telescope (Roman)
Technical Report Roman-STScI-000394

Determining the best-fitting slope and its uncertainty for up-the-ramp sampled images with unevenly distributed resultants

Stefano Casertano

June 15, 2022

Abstract

We address the problem of determining a best-fitting slope and its uncertainty for near-infrared detectors capable of multiple non-destructive reads "up-the-ramp", focusing in particular on the case of unevenly spaced individual or combined reads. We extend the results previously obtained for evenly-spaced combinations to the more general case, relevant to the Nancy Grace Roman Space Telescope, of the combination of a variable number of reads during an exposure. These expressions allow close-to-optimal slope-fitting for a variety of signal levels, and yield a mathematically accurate estimate of the uncertainty of the slope thus obtained; they are suitable for efficient, array-based implementation in modern interpreted languages. We also provide numerical tests of the expressions we obtain.

1. Introduction

The focal plane of the Wide Field Instrument (WFI) aboard the Nancy Grace Roman Space Telescope (Roman) will have eighteen H4RG detectors, each 4096x4096 pixels in size, for a total focal plane count of over 300 million science pixels (for more information on these detectors, see Mosby et al. 2020). Each detector will be read continuously at a cadence of approximately 3 seconds per frame in normal imaging mode; other modes may require up to 6 seconds, depending on guiding. At 3 seconds per frame, the full focal plane generates. $\sim$ 200 MB of raw data per second. The rate of data generation significantly exceeds Roman's very large download capability (1.4 TB/day) even after on board compression; therefore downloading all individual readouts is currently not feasible.[1] As with most infrared detectors, the H4RG are capable of non-destructive readouts, which allow the accumulated signal to be read without changing it; as a result, the *read noise*, caused by the readout process, can be effectively reduced by reading each detector pixel multiple time while the signal accumulates. Using as many readouts as possible is beneficial for most Roman science, as it improves the signal-to-noise ratio (S/N) that can be achieved in each exposure, especially for faint sources.

In order to realize most of the benefit of frequent non-destructive readouts while remaining within Roman's transmission capabilities, the WFI will adopt a procedure whereby several individual readouts can be averaged on-board, and only their averages, called *Resultant frames* (or resultants, in brief), are transmitted to the ground. On average, at most one read or resultant every 15–20s, or roughly one per six reads, can be sent to the ground; this number depends on the compression ratio that can be achieved by on-board algorithms, the overhead and margin of the transmission, and the details of the specific observing program. The Mission has on-board storage capacity sufficient to store multiple days' worth of data; thus under special circumstances, all reads can be saved as separate resultants and sent to the ground for a limited number of exposures within a day.

A similar averaging method is adopted by instruments on board the James Webb Space Telescope (JWST), where *Groups* are created by averaging consecutive reads, with the possibility of skipping some reads between groups if desired; in the following we will focus on the patterns for the NIRCam imager. The Roman WFI method is more flexible than the JWST method, as the latter requires that all groups/resultants be evenly spaced and contain the same number of reads. In the following, we use either "group" or "resultant" to refer to a data frame constructed by averaging multiple reads; we prefer "group" when these are evenly spaced, "resultant" for the more general case when they may not be spaced uniformly. More details and examples are provided in Section 2.

At the most basic level, the counts accumulated in a pixel throughout an exposure increase roughly linearly with time (see Section 4.1 for more details). The rate at which counts accumulate

[1]For general information on the Roman telescope and its science mission, see, e.g., `https://roman.gsfc.nasa.gov/science/Roman_Reference_Information.html`

provides a measurement of the impinging light, thus a critical goal of processing detector data is to determine the *slope* of the counts vs. time relation, which measures directly the source brightness at that location, and its *variance*, which relates to the uncertainty of that measurement. This topic has been frequently addressed in the literature; see Section 3.2 for a discussion of existing results. The present work largely builds on previous treatments, but it addresses specifically some of the readout patterns proposed for Roman WFI, including the possibility of unevenly-spaced resultants, and it discusses how such data can be handled in the WFI calibration pipeline.

Most processes in use adopt a weight vector, in which the cost function is the weighted sum of the squares of the residuals (measurement minus model). Here we consider the more general case of a weight *matrix*, in which the cost function is a quadratic form that combines the residuals from different measurements; see Equation 2 for its definition. The weight vector corresponds to the special case in which the weight matrix is diagonal. The expressions and treatment for the two cases are very similar; we develop the treatment for the weight matrix case, and provide simplified equations for the weight vector case where appropriate. The advantage of the weight matrix case is that it includes the *optimal* linear algorithm for the slope determination; as we discuss in Section 4, this method is not necessarily of practical use, but it proves very useful in evaluating the effectiveness of more practical methods.

The contents of this manuscript are as follows. In Section 2 we describe the general properties of both JWST/NIRCam and WFI readout patterns, and provide several examples that will be used to compare the two. In Section 3 we introduce some general concepts underlying least-squares estimators for ramp fitting, and briefly describe previous results. In Section 4 we obtain the general equations for weighted least-squares estimators of the slope, and formal expressions for their expected variance, for the general case of combined, unevenly-spaced resultants and of a two-dimensional weight matrix. Section 5 provides all the steps needed to encode the calculation of both the slope and its variance; readers especially interested in the implementation of the present method may wish to start there. Finally, in Section 6, we provide direct comparisons of the methods discussed in this paper via numerical calculations and Monte Carlo simulations, quantify their effectiveness, and indicate possible biases that might occur if estimated variances are used without proper care.

2. Nomenclature and examples: JWST/NIRCam and Roman/WFI

The basic unit of observation we consider is the interval between detector resets. After the initial reset, signal accumulates on the detector and is read out (non-destructively) at various times, accumulating in a linear fashion until the next reset. This element is usually called "exposure", although JWST calls it "integration"; for JWST, an exposure consists of multiple integrations chained together. Roman does not chain integrations, and the term "exposure" appears to be more commonly used to refer to this unit; therefore we will use "exposure" in the following, and always assume that the exposure under consideration consists of a single integration (in JWST

parlance).

The sequence of readouts, their timing, and how they are processed in flight differ to some extent between JWST and Roman. In the following we give a basic description of JWST/NIRCam and Roman/WFI readout patterns.

2.1. NIRCam readout patterns

The readout patterns for the Near Infrared Camera (NIRCam) aboard JWST, illustrated in Figure 1, assume that there is an underlying sequence of equally spaced reads, and that these reads are combined (averaged) into equally sized, equally spaced groups to be sent to the ground.[2] The frame time, or time between the starts of consecutive reads, depends on the instrument, the subarray mode, and the settings of the control electronics, but is generally constant within one exposure. (This differs form some HST infrared instruments, for which reads can be commanded with variable timing.) The number and combination of reads for an exposure is fully characterized by four numbers: NFRAMES, NSKIP, NGROUPS, and NINT. NFRAMES is the number of consecutive reads averaged into each group; NSKIP is the number of frames skipped between the end of one group and the beginning of the next; NGROUPS is the number of groups constituting an integration; and NINT is the number of (identical) integrations constituting an exposure. Since WFI does not combine integrations into exposures, we will only consider in our discussion NIRCam exposures consisting of one integration, i.e., NINT = 1, and will use the Roman term "exposure" as equivalent to a NIRCam integration. Since the last NSKIP frames are not taken—the integration ends after the last set of NFRAMES reads—an integration consists of NGROUP $\cdot$ (NFRAMES + NSKIP) $-$ NSKIP reads, and the total exposure length is that number multiplied by the frame time, plus any additional time needed at the beginning or at the end of the exposure.

In the following, we will consider five readout patterns defined via the JWST parameters NFRAMES, NSKIP, and NGROUPS, and chosen to be similar to the proposed WFI patterns (with uneven resultants) described in the following Section. These patterns are named SHALLOW4-10, EVEN 3-6, EVEN 5-10, EVEN 6-16, and EVEN 10-20; details are given in Table 1. These even patterns provide a good comparison for corresponding WFI patterns in terms of exposure time and number of downloaded frames, and help illustrate the impact of using even vs. uneven grouping. It should be noted that of the five NIRCam-like patterns used here, only one, SHALLOW4-10, is actually available to NIRCam users; the others are similarly constructed, but have NSKIP=0, unlike most of the patterns in the NIRCam list. We will see later that skipping frames can improve dynamic range, but reduces S/N for faint sources; with the ability

[2]See `https://jwst-docs.stsci.edu/jwst-near-infrared-camera/nircam-instrumentation/nircam-detector-overview/nircam-detector-readout-patterns` for more information on the readout patterns available for NIRCam.

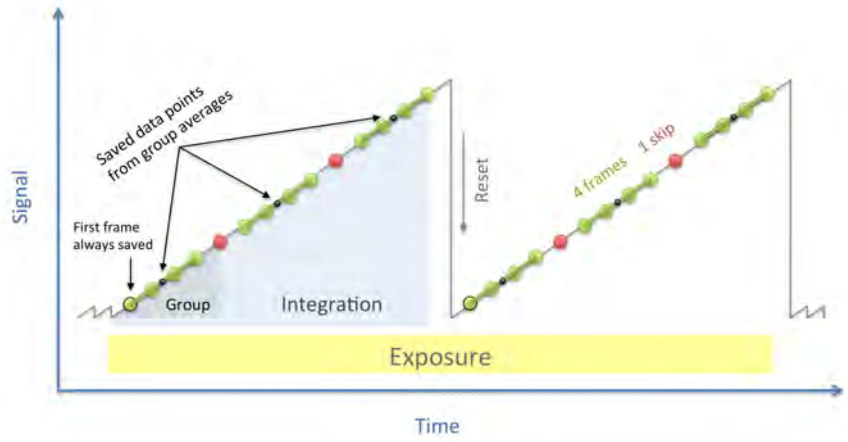

Fig. 1.— Graphic illustration of a NIRCam readout pattern, from `https://jwst-docs.stsci.edu/jwst-near-infrared-camera/nircam-instrumentation/nircam-detector-overview/nircam-detector-readout-patterns`. In our analysis, we will always assume a single integration, and we will use the term "exposure" to refer to the sequence of operations between resets.

of using uneven resultants, the dynamic range improvement provided by skipping frames is not necessary.

A few details on timing are in order. We assume that each exposure begins with a read/reset frame, not included in the frame count, which may be recorded and subtracted from all subsequent frames; this frame will generally not be available for flight data. We will refer to this as the "zero" read, and mark its start as $t = 0$; subsequent reads are numbered beginning with one.[3] Available laboratory data indicate that there are significant differences between the zero read and all subsequent ones; the subtraction is carried out with the goal of increasing compressibility of other frames, but there remain significant offsets that need to be taken into account in the analysis. Consequently, one cannot assume that the residual signal in following frames begins at zero at the time of the zero frame, and an intercept (also called "pedestal") needs to be determined for each pixel; in terms of the linear fit, that means that both the slope and the intercept need to be determined by the fit, separately for each pixel. Here we focus on the value and uncertainty of the slope, which is relevant for scientific analysis; however, the value of the intercept can be useful in understanding the behavior of the detector and control electronics. A corollary is that the first read does not, by itself, provide useful photometry for bright sources; at least two valid, unsaturated measurements (resultants) after the zero read are needed for a meaningful estimate of the slope of the ramp.

[3]Because many modern languages, including python, use zero-based numbering, the read numbered "1" is indicated with the index "0" in our formulae; see, e.g., Eq. 2. This is purely a matter of convenience in translating formulae to code.

We assign to each read a time corresponding to the start of that read after the start of the zero read. For simplicity, we assume the interval, called the "frame time", to be equal for each read. In reality, the frame time of the first read may be slightly different from the frame time of all subsequent reads, as the read/reset frame may take slightly longer than all following reads; however, the difference—which depends on the timing patterns ultimately adopted for WFI detectors—is inconsequential to our analysis, since it would only affect the overall pedestal of each ramp, which is not used further. Similarly, the times at which pixels are read will be different; as the detector readout progresses, pixels with higher row index (Y) will be read later than pixels with lower Y values. Within each row, the readout proceeds from left to right in odd-numbered amplifiers, and from right to left in even-numbered amplifiers, so the X (column) coordinate also affects the exact readout time of a pixel. In addition, the readout time does not increase linearly with pixel index; there can be extra time at the end of each row readout, and larger gaps occur every 256 rows because of the switch to guide star readout.

The calculation of the slope is not affected by the absolute origin of time, but only the relative temporal separations between adjacent/consecutive reads. Therefore, many of the previously mentioned details will only become useful for time domain science, where the absolute time start and separations between reads is required to high-precision.

In Table 1, the exposure time $t_{\rm exp}$ is defined as the difference between the midpoint of the first resultant and the midpoint of the last resultant; it is the quantity that drives the total signal collected from a source, as recorded in the resultants (see Equation 10), and it also enters the Poisson noise calculation. On the other hand, the total time $t_{\rm total}$ is the time between the *beginning* of the read/reset frame and the *end* of the last frame in the last resultant. This is the time span during which the instrument is engaged in obtaining the exposure, and thus it represents the total encumbrance that exposure imposes on the instrument. Of course, the total time needed to obtain the exposure may also need to account for appropriate overheads for activities before and after the exposure, such as filter changes and data transfers, if applicable. The quantity NRES is the total number of resultants for that pattern, and NComp is the number of component reads for each resultant; this is a constant (and equal to NFRAMES) for the NIRCam readout model, but it is expressed as a range for the Roman patterns. The Sequence column indicates how reads are averaged into resultants, and is explained in Section 2.2.

In Section 6.2 we provide more details about the performance of these readout patterns as a function of signal level; in Section 6.3 we discuss their expected saturation level and performance on very bright sources.

2.2. Multi-Accum Tables and the WFI readout patterns

The readout and combination patterns for WFI are similar to those for NIRCam, with the major difference that the combined reads, called "resultants", can differ in size and spacing throughout the exposure. An example is shown in Figure 2.

Table 1: Sample readout patterns considered in this study

Name	Even?	Nreads	$t_{\rm exp}$	$t_{\rm total}$	NRES	NComp	Sequence
ML	No	19	48.64	60.80	6	1–5	1, 2–3, 4–6, 7–10, 11–14, 15–19
EVEN 3-6	Yes	18	45.60	57.76	6	3	1–3, 4–6, 7–9, 10–12, 13–15, 16–18
HiLat	No	47	129.20	145.92	9	1–8	1, 2–3, 4–6, 7–10, 11–15, 16–23, 24–31, 32–39, 40–47
SHALLOW4-10	Yes	49	136.80	152.00	10	4	1–4, 6–9, 11–14, 16–19, 21–24, 26–29, 31–34, 36–39, 41–44, 46–49
HDR-150	No	50	138.32	155.04	10	1–8	1, 2, 3–4, 5–6, 7–10, 11–18, 19–26, 27–34, 35–42, 43–50
EVEN 5-10	Yes	50	136.80	155.04	10	5	1–5, 6–10, 11–15, 16–20, 21–25, 26–30, 31–35, 36–40, 41–45, 46–50
HDR-300	No	98	284.24	300.96	16	1–8	1, 2, 3–4, 5–6, 7–10, 11–18, 19–26, 27–34, 35–42, 43–50, 51–58, 59–66, 67–74, 75–82, 83–90, 91–98
EVEN 6-16	Yes	96	273.60	155.04	16	6	1–6, 7–12, 13–18, 19–24, 25–30, . . . , 91–96
HDR-600	No	197	573.04	601.92	20	1–16	1, 2, 3–4, 5–6, 7–10, 11–18, 19–26, 27–34, 35–42, 43–50, 51–58, 59–69, 70–85, . . . , 182–197
EVEN 10-20	Yes	200	577.60	611.04	20	10	1–10, 11–20, 21–30, 31–40, 41–50, . . . , 191–200

Note: readout patterns considered in this work. Each pair of lines includes a proposed WFI pattern, with unevenly spaced resultants, and a NIRCam-style pattern with even resultants and a similar overall exposure time and number of resultants. Only SHALLOW4-10 is an actual NIRCam pattern, however.

The patterns are defined by a set of entities called "Multi-Accum Tables". Each WFI detector is read out constantly at a rate that depends on the size of the guide windows and other timing parameters; currently, normal imaging data are expected to have a frame time of 3.04 s, while the frame time for spectroscopy observations, which use a larger guide window, is just over 4 s; special guiding modes may have different frame times. In the following, we adopt a frame time of 3.04 s, but most of our calculations scale simply if a different frame me is adopted. Each Multi-Accum Table determines which of these reads are combined into resultant frames for transmission to the ground, and which are discarded. A Multi-Accum Table contains additional information for use by the on-board software; for our purposes, its main content is the list of reads (by number or by time) included in each resultant frame, along with the total number of resultant frames included in the exposure as a whole. Thus a sufficient representation of a Multi-Accum Table that directs, say, one, two, four, and six reads to be combined into resultants, while two reads are skipped between frames, would have the form: $\{[1], [4, 5], [8, 9, 10, 11], [14, 15, 16, 17, 18, 19]\}$. In the following, we will only consider cases in which no reads are skipped *within* a resultant, so that we can use the more compact form $\{1, 4–5, 8–11, 14–19\}$. This description is used in the "Sequence" column of Table 1.

Note, however, that the assumption that there are no skipped reads within a resultant is only a convenience of representation. All of the analysis presented here applies to the more general form in which any number of reads can be skipped within or between resultants. The only requirement is that reads used within different resultants not be interleaved; otherwise the

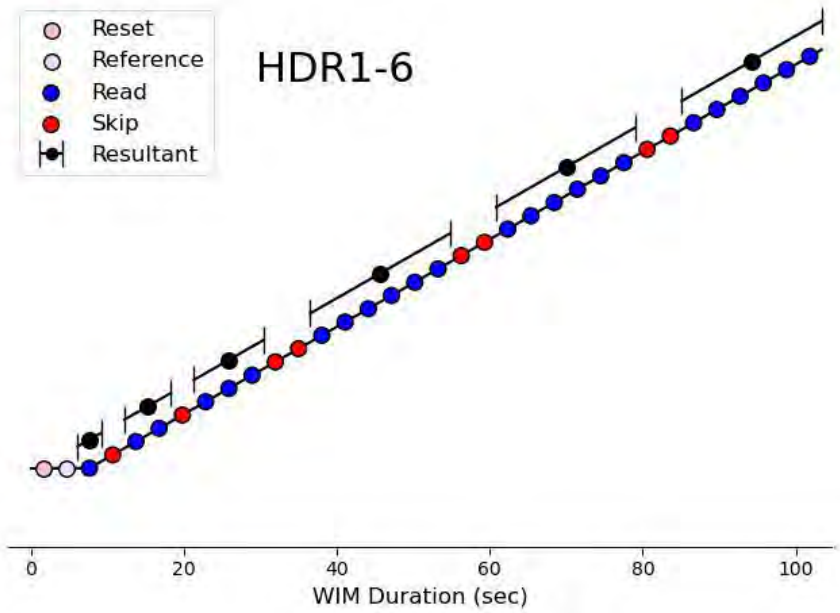

Fig. 2.— Graphic illustration of a WFI readout pattern, courtesy of Tyler Desjardins. Unlike the NIRCam pattern shown in Fig. 1, exposures always consist of a single integration, and the number of reads combined into a "resultant" can vary throughout the exposure.

assumptions used to derive Equation 19 would be violated.

The set of Multi-Accum Tables available for science observations has not been defined yet, and it can be changed in the future, even after launch. However, an obvious difference between exposure combinations driven by Multi-Accum Tables and those for NIRCam, defined by the parameters NFRAMES, NSKIP, NGROUP, and NINT, is that the latter impose equal spacing of the averaged reads (called groups), while the former allow for different averaged reads (called resultants) to be spaced unevenly, and to include different number of component reads. One of the advantages of the WFI method is that resultants can be more frequent early in the exposure, thus improving the dynamic range of the data. This will be discussed further in Section 6.3.

We will consider here five options for Multi-Accum tables that may be representative of different types of observations. The first two, called ML (for MicroLensing) and HiLat, are among those suggested by the Roman Project Science Office (Jeff Kruk, internal communication) as candidates for the Galactic Bulge Time Domain Survey and the High Latitude Wide Area Survey, respectively. The ML Multi-Accum table has a total time of about 60s and six resultants; HiLat has a total time of about 150s and nine resultants. The remaining three have early resultants with few components, designed to optimize the ability to observe bright sources, and have total times of roughly 150, 300, and 600s; they are called HDR-150 ((HDR stands for High Dynamic Range), HDR-300, and HDR-600, and have 10, 16, and 20 resultants, respectively. Details on these patterns are given in Table 1, together with NIRCam-style patterns of similar length. Corresponding patterns with uneven (WFI) and even (NIRCam-style) grouping are given on

consecutive lines. The performance and bright-source limits for these patterns are discussed in Sections 6.2 and 6.3, and summarized in Table 3.

3. Determining the count rate: slope-fitting a sequence of resultants

In this Section we present the general concepts underlying the use of least-square fitting to obtain a linear estimator for the signal in an exposure, including previous results and some notes on the implementation used in existing pipelines. The full treatment for unevenly-spaced resultant frames is presented in Section 4.

The key information we want to obtain from a sequence of up-the-ramp reads is the *slope* of the ramp, i.e., the counts per second registered in that pixel, and its uncertainty. [4] For an ideal detector under ideal circumstances, that slope is a direct representation of the signal impinging on the detector at that location. This signal will generally include contributions from detector background (e.g., dark current), from other instrumental backgrounds (e.g., thermal contribution from the environment), and from the region of the sky projected onto that pixel. Here we focus simply on determining that slope; its interpretation in terms of its various contributors (dark current, sky background, sources, etc) is beyond the scope of the current analysis.

It is generally assumed that the signal in a given pixel increases linearly with time, from a *pedestal* value at the time defined as $t = t_0$ to larger and larger values as signal accumulates. The linearity of this accumulation is only valid under certain assumptions; at a minimum, 1) the impinging light and other backgrounds must remain constant during the exposure, and 2) the conversion of light and background into charge and then measured counts must happen at constant efficiency. Assumption 1) could be violated because the source itself varies or moves (e.g., asteroids), but also if the telescope moves (resulting in a different part of the sky projecting onto the pixel under consideration), the instrument temperature fluctuates (resulting in variable background or dark current), or the attitude of the telescope with respect to sources of background (such as the Earth and the Moon) changes during the exposure. A prime example with the WFC3 instrument on HST is the variable background caused by the atmospheric He I line at 1083 nm, which causes a rapidly increasing background as the telescope points closer to the Earth's limb (Brammer et al. 2014; Brammer 2016; Pirzkal & Ryan 2017). Because Roman will be around the Lagrangian point L2, and not in low Earth orbit, such rapidly variable conditions are less likely to obtain, but they cannot be ruled out a priori, especially at the level of stability desired by Roman science investigations. Other sources of variability, due, e.g., to internal temperature variations, are also less likely to be significant; however, jitter will exist and affect the count rate at a level that depends on the astronomical scene under observation. We do not consider these complications here, but they may be important for certain conditions and science investigations.

[4] Although the *value* of the intercept is not used, simultaneously fitting for both intercept and slope affects the values and uncertainties of both.

The conversion of signal into electrons and then into pixel voltage—and thus Data Numbers, the units of the measured signal—can also deviate from linearity, thus violating Assumption 2. One reason is the so-called "classic non-linearity", present in most astronomical detectors and especially significant for infrared detectors. Classic non-linearity causes the recorded voltage to lag below the accumulated electron counts as the latter increase; the effect becomes especially significant as the number of accumulated electrons approaches the full-well capacity of the pixel. Mosby et al. (2020) test the linearity correction up to 80,000 e^-, beyond which—for the detector they consider—most pixels approach full well, and the linearity correction becomes large and uncertain. Below 80,000 e^-, they report that, after applying a polynomial linearity correction, the residual deviations from linearity are typically below 1%.

Classic nonlinearity not only causes the measured signal to be a non-linear representation of the accumulated counts, but also changes to some extent the apparent noise properties of the signal. Henceforth we assume that a classic nonlinearity correction is applied to the measured signal, so that it correctly represents the accumulated counts. The Poisson noise in this signal is then also correctly represented, since the nonlinear transformation occurs after individual events (which drive the Poisson noise) have been recorded. However, a small inaccuracy enters here because of read noise, which applies to the *measured* signal; thus if that signal is scaled up to correct for classic nonlinearity, the read noise should be scaled by the same factor. While it is in principle possible to carry out all the calculations that follow taking into account the scaled read noise, this would significantly complicate both the analysis and the pipeline calculations. Laboratory measurements of the flight candidate H4RG devices suggest that the classic non-linearity is very modest ($< 1\%$) for counts below several thousand. The typical read noise is approximately 10 e^-; therefore whenever linearity corrections are significant, the total noise is dominated by Poisson noise, which needs no adjustment for the linearity correction.

In addition to these deviations from linearity, other effects can affect the measured signal. For example, time-correlated artifacts due to electronic signatures in the signal chain are likely to affect an exposure in ways that differ from resultant to resultant, possibly leading to 1/f patterns. We assume that steps taken in the calibration pipeline will remove such effects. If signatures of imperfect calibration remain, those need to be studied carefully; their impact on slope fitting and its variance will depend on their properties, and will need to be determined separately.

3.1. Definitions

As in previous treatments, we consider the slope estimation problem in the form of a linear least-squares optimization. We assume that the signal expected after time t can be expressed as:

$$S(t) = f_0 + f \cdot t \tag{1}$$

Assume that we have N measurements S_i of the signal at times t_i. In its general form, the least-squares approach defines the estimated values of f_0 and f, designated $\hat{f}_0$ and $\hat{f}$, respectively,

as the values that minimize the weighted cost function Q, defined as a quadratic form in the residuals $S - f_0 - f \cdot t$:

$$Q = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij} \, (S_i - f_0 - f \cdot t_i)(S_j - f_0 - f \cdot t_j) \tag{2}$$

where the weight matrix W_{ij} is symmetric and positive-semidefinite. [5]

A simplified version of the least-squares approach assumes that the weight matrix is diagonal, with non-negative diagonal elements $W_{ii} = w_i$; in that case the definition of Q takes the more common form:

$$Q = \sum_{i=0}^{N-1} w_i (S_i - f_0 - f \cdot t_i)^2 \ . \tag{3}$$

The frequently-used χ^2 solution is a special case of least-squares optimization. If the covariance matrix C_{ij} of the data is known, then the least-squares formulation reduces to the χ^2 solution if we take $W_{ij} = (C^{-1})_{ij}$, where C^{-1} is the matrix inverse of the covariance matrix. This reduces to the diagonal case if the data are uncorrelated (thus the covariance matrix is diagonal). However, for the infrared detectors under consideration, different reads (or resultants) in up-the-ramp sampling are significantly correlated, as discussed, e.g., in Robberto (2007); see Robberto (2014) for its application to the problem at hand, and Gennaro & Khandrika (2021) for a more rigorous approach to the ramp-fitting issue. Following Robberto (2014), it can be shown (see Section 4.4) that the χ^2 solution is optimal within our current assumptions, in that it has the smallest variance of any least squares solution with arbitrary weights; however, the covariance matrix is not known a priori, even in the idealized noise model we have adopted, since it depends on the *true* count rate. A simplistic signal-based noise model, e.g., assigning each resultant a weight inversely proportional to its variance, is not desirable, as it fails to take into account the substantial covariance between resultants; see Section 6.1.1 for further discussion.

We will show in Section 4.6 that the optimization of Q leads to an *unbiased* slope estimate, both for one-dimensional and two-dimensional weights, regardless of their values. This, however, is only valid provided the weights are *fixed*; if the weights are defined from the measured data, biases are possible. We will discuss this further in Section 6.4.

3.2. Previous results: Slope and variance estimates for evenly spaced resultants

The question of how to determine the slope and its variance for infrared detectors, with their non-destructive reads, has been frequently addressed in the literature. In most treatments (Rauscher et al. 2007; Regan 2007; Robberto 2009, 2014), a linear least-squares approach is used

[5] The assumption of symmetry is not restrictive; if Q were defined with a non-symmetric weight matrix Z, it is trivial to show that the weight matrix $W = (Z + Z^T)/2$, which is symmetric, yields an identical quadratic form Q.

to determine the slope and estimate the variance. Fixsen et al. (2000) formulate their approach in terms of differences between successive reads, which is mathematically equivalent to a least-squares fit up the ramp. An alternative, more rigorous treatment is provided by Gennaro & Khandrika (2021), who use a maximum-likelihood optimization that enables them to include a number of non-linear effects. Most existing treatments focus primarily on individual reads, used for HST NICMos and WFC3, or evenly spaced resultants, used for JWST instruments such as NIRCam.

The treatment of the signal measured in successive reads needs to consider at least two sources of measurement uncertainty: the read noise, due to the readout process itself, and the Poisson noise, inherent in the process of detecting discrete photon events. Both terms contribute to the noise, and thus to the variance in each measurement (see Eq. 16 and 19). Determining the uncertainty in the slope measured for noisy data requires consideration of both sources of error, and proper accounting for the variance and covariance of resultants. Several authors (Rauscher et al. 2007; Regan 2007; Robberto 2009) have computed the slope variance for the case of evenly spaced resultants, with specific focus on the readout and grouping patterns adopted for JWST. Those variance estimates take an especially convenient closed form in the case of uniform weights, and form the basis for the computation of the *error array* provided by the NIRCam pipeline.

Although theoretically appealing, the χ^2 solution obtained by Robberto (2014) is not currently in use in the NIRCam pipeline. As discussed in the previous Section, in that formalism the weight matrix is the inverse of the covariance matrix, which has significant off-diagonal elements. Furthermore, the covariance matrix depends on the signal itself, which is not known *a priori*. Thus, the χ^2 solution requires a matrix inversion for each pixel, as well as an iterative approach to converge to the correct covariance matrix. For these reasons, this method is impractical in the case of large data volumes, and will likely not be adopted in the Roman WFI pipeline. However, it is very useful in simulations—when the true flux is known—e.g., as a way of quantifying the performance of other methods.

3.3. The NIRCam pipeline approach: A practical, quasi-optimal weighting scheme

As noted above, real data do not usually lend themselves to a full χ^2 optimization, even if the simple noise model (Gaussian read noise plus Poisson noise from the signal) is correct. However, the NIRCam pipeline adopts a useful approximation suggested by the work of Fixsen et al. (2000). They consider that the weight setting based on the Poisson+read noise model has two limiting cases. In the Poisson-dominated regime, where the total signal is much larger than the square of the read noise (i.e., $f \cdot \max(t_i) \gg (rn)^2$), the optimal weighting is achieved by giving the most weight to the first and last data points, and zero weight to all points in between. In the read-noise-dominated regime ($f \cdot \max(t_i) \ll (rn)^2$), the optimal weighting scheme gives all points the same weight. Instead of attempting to produce a weighting scheme that exactly matches the (unknown) value of f, they suggest using a simplified scheme in which the weights are set as

a simple function of time, namely power laws with an exponent that ranges from 0 (read noise dominant, uniform weights) to 10 (signal noise dominant, first and last point have almost all the weight). The value of the exponent is chosen for each pixel on the basis of the estimated S/N in that pixel; only six discrete values are used. The exact algorithm is described in Section 6.1; as the range of S/N that corresponds to each discrete exponent value is broad, the weights are not generally a function of the measured signal, except in boundary cases. (Recall that any weighting scheme with noiseless weights is unbiased; the difference is in the variance of the result.) Their formulation has the advantage of being simple to implement and efficient, in that it does not require either matrix inversions or an expensive iterative processes to determine the signal slope in each pixel, at a modest cost in final S/N. This is the algorithm currently used in the NIRCam pipeline.[6]

The algorithm as used in NIRCam is not optimally suited for uneven resultants, since it does not account for the smaller effective read noise in resultants that include more reads. Simply scaling the weights by the number of reads is correct in the read-noise dominated case, but it leads to a poorer solution at high signal. In Section 6.1 we present a modification of the NIRCam weighting scheme that takes into account the number of components per resultant; we show in Section 6.2 that this modified scheme works well for a broad range of readout patterns and signal levels.

4. Slope fitting: The general case of uneven resultants

In this Section we derive the general equations for the least-square fitting presented in Section 3, applied to the case in which individual reads are averaged into resultant frames. We obtain expressions for the expectation value and (co)variance of resultant frames for arbitrary one- or two-dimensional weights, and we obtain the key equations for the slope and its variance under fairly general assumptions, detailed in Section 4.1.

Throughout this Section, the signal is expressed in electrons, i.e., the electronic gain is unity. The value of the electronic gain affects the signal variance, as the individual events driving its Poisson noise are photon detections, not counts. Including the effect of non-unit gain in the equations is trivial—it amounts to appropriately scaling the relation between signal-variance and count rate. However, other effects related to digitization and integer arithmetic, both for individual reads and when reads are combined into resultants, can affect the mean and variance of the measured data. These effects are generally very small, and will be included in a future treatment.

[6]See the documentation at `https://jwst-pipeline.readthedocs.io/en/latest/jwst/ramp_fitting/description.html#slope-and-variance-calculations`

4.1. Assumptions

In this calculation we make several simplifying assumptions regarding the properties and statistics of the signal S detected in individual reads. Definitions and properties for resultant frames are given in Section 4.2.

1. The expected value S of the accumulated charge at time t is linear in time, i.e.,

$$S = f_0 + f \cdot t \tag{4}$$

where f is the count rate for that pixel, assumed constant throughout, and including detector and instrumental signals ("background"). Note that the constancy of f implies no variable sources of signal, such as variable backgrounds, motions of the target field, or other effects. We assume the pedestal f_0 to be an unknown to be determined as part of the slope fitting process. Typically, f_0 is not subject to Poisson noise, as it is determined primarily from electronic effects; it may be possible to determine f_0 from external information (superbias, reference pixels, reference output, etc), but for our purposes, it is a free parameter to be determined empirically. At the moment, existing ground data suggest that determining f_0 externally is very challenging, and substantial pixel-to-pixel variations (of several hundred counts) are seen that cannot be readily predicted from bias or reference pixel information; if the determination of f_0 becomes more deterministic, it may be necessary to revisit this assumption.

2. The accumulation of signal is subject to ordinary Poisson noise, i.e., the signal detected after time t obeys the usual Poisson statistics associated with independent discrete events, with variance $f \cdot t$. While it is common to assume that the Poisson process can be approximated with a Gaussian with the appropriate mean and dispersion, we will not use the explicit distribution, but only its first and second moment, which do not require the Gaussian approximation.

3. The measurement of S at time t is subject to a random Gaussian error, called "read noise," with zero mean and dispersion (rn). Therefore the measured signal $S(t)$ has variance:

$$\mathrm{Var}(S) = (rn)^2 + f \cdot t. \tag{5}$$

4. In the present analysis, we treat the signal as measured and processed in real numbers, and neglect effects introduced by digitization and integer arithmetic. Such effects include both additional ("truncation") noise, and small biases due to the process of convertung the signal into digital counts and to the on-board averaging. Robberto (2014) and Gennaro & Khandrika (2021) include the discretization noise in their analysis; we plan to include both noise and bias in a follow-up analysis, together with other real-world effects such as the detection and impact of cosmic rays.

5. We assume that the read noise for each pixel is uncorrelated between frames, i.e., its value between reads has zero covariance. As a consequence, the read noise term averages down as $1/\sqrt{N}$ as multiple reads are averaged together (Eq. 13). This assumption is commonplace practice for other missions. At the time of this report, the Roman mission does not have test data in hand to determine if the read noise is correlated between frames. Future low-illuminations tests are needed to determine if this assumption is valid, or to characterize read noise correlations if they exist.

6. The Poisson term of the variance *is* correlated, because the signal accumulates up the ramp; the signal received at the earlier readout is in common with the later readout, thus they have the same realization of the Poisson variance term for the time in common.

   Therefore the covariance of the measurements S_i, S_j at different times t_i, t_j is:

$$C_{ij} \equiv \mathrm{Covar}(S_i, S_j) = f \cdot \min(t_i, t_j) \qquad (i \neq j) \tag{6}$$

7. The noise model includes only the Poisson noise, due to the randomness of electron detection, and the Gaussian read noise. Other sources of error, such as variation in the impinging flux, changes in detector and electronic properties, linearity corrections, and so on, are not included. Nonlinear effects, while perhaps straightforward to implement in some cases, would add complexity not considered in our current approach.

8. The signal is assumed to have been corrected for various detector effects, including normal dark current, additive electronic terms, and sources of non-linearity, such as classic non-linearity, brighter-fatter effect, and so on. Some of these effects do not change the noise properties of the signal—e.g., the classic non-linearity correction restores the number of photons detected in the pixel, thus the Poisson noise is essentially correctly described by our model after the correction. Others may have a subtle noise impact after correction, such as the noise correlation across pixels that can be introduced by correctiong for the brighter-fatter effect; these terms are beyond the scope of the current treatment.

9. The contribution of dark current to the signal variance needs to be addressed. The subtracted dark signal is usually defined to include both true dark signal and electronic effects. The true dark signal includes counts generated in the detector substrate by thermal effects, as well as internal background generated in the enclosure; both are subject to ordinary Poisson noise. Electronic effects can include residual bias and reference pixel corrections, both of which may be subject to random noise that does not necessarily scale as Poisson noise. If the dark signal is low, it may be safe to neglect its contribution to the Poisson noise, and fitting a ramp to dark-current subtracted resultants is acceptable. (We assume that the noise in the dark signal to be subtracted is negligible, since it is determined from dozens of dark images.) However, if the dark signal contributes significantly to the Poisson noise, the correct approach would be to subtract off only the non-Poisson term of the dark signal, and

fit the ramp with the linearly growing, Poisson-noise-affected signal. Alternatively, the dark signal could be added to the fitted ramp for the purpose of determining the slope variance (see Section 4.5). This level of detailed information on the dark signal may not always be available. The treatment presented here assumes that the dark current either contributes a negligible amount to the noise, or it can be estimated and included in the slope variance calculation.

4.2. Resultant frames: variance and covariance

Resultant frames are the product of on-board averaging of different reads. Formally, the value R_i of a resultant frame is the average of N_i component reads, designated as $S_{i,k}$, each obtained at times $t_{i,k}$, where $k = 0, 1, \ldots, N_i - 1$).

It is not required that the component reads $S_{i,k}$ be consecutive; indeed, an arbitrary number of reads can be skipped between component reads. However, we shall assume that the resultants are *strictly ordered*, i.e., if $\bar{t}_i < \bar{t}_j$, then $t_{i,k} < t_{j,l}$ for all values of k and l. In other words, reads in different resultants are *not* interleaved.

We assume that the average of multiple reads into a resultant is unweighted, as in the current on board data processing plan for WFI; if some form of weighted averaging is desired, the necessary changes can be introduced as needed. Thus, by definition:

$$R_i \equiv \sum_{k=0}^{N_i-1} S_{i,k}/N_i \,. \tag{7}$$

Under the above assumptions, we can now proceed to compute the expectation value and variance of a resultant frame and the covariance between resultant frames. Using the expectation value of each component, $E(S) = f_0 + f \cdot t$, we obtain:

$$E(R_i) = f_0 + \sum_{k=0}^{N_i-1} f \cdot t_{i,k}/N_i \,. \tag{8}$$

If we define the *mean resultant time* $\bar{t}_i$ as:

$$\bar{t}_i = \sum_{k=0}^{N_i-1} t_{i,k}/N_i \,, \tag{9}$$

then we have the simple expression:

$$E(R_i) = f_0 + f \cdot \bar{t}_i \,. \tag{10}$$

Thus the expectation value of the resultant is the same as for a read obtained at the mean time of the component reads.

This is not true for the variance of the resultant. One obvious difference is that the read noise averages down, since it is assumed to be uncorrelated across reads. A more subtle effect arises when taking properly into account the covariance of the various component reads.

Using the general expression for the variance of a linear combination in terms of the known (co)variance of the individual components, we find:

$$\mathrm{Var}(R_i) = \sum_{k=0}^{N_i-1} \sum_{l=0}^{N_i-1} \frac{d\,R_i}{d\,S_{i,k}} \frac{d\,R_i}{d\,S_{i,l}} \mathrm{Covar}(S_{i,k}, S_{i,l}). \tag{11}$$

Now Equations 5 and 6 yield:

$$\mathrm{Covar}(S_{i,k}, S_{i,l}) = \delta_{kl}(rn)^2 + f \cdot \min(t_{i,k}, t_{i,l}) \tag{12}$$

where δ_{kl} is the Kronecker delta in k and l: $\delta_{kl} = 1$ if $k = l$, 0 otherwise.

Substituting into Equation 11 and noting that $dR_i/dS_{i,k} = 1/N_i$, the expression for the variance of the resultant becomes:

$$\mathrm{Var}(R_i) = (rn)^2/N_i + (1/N_i^2) \sum_{k=0}^{N_i-1} \sum_{l=0}^{N_i-1} f \cdot \min(t_{i,k}, t_{i,l}) \tag{13}$$

As anticipated, the variance due to read noise (the first term on the right hand side) is reduced by a factor $1/N_i$ when N_i reads with uncorrelated read noise are averaged. The second, signal-based term, contains N_i^2 terms, with each time $t_{i,k}$ appearing $2\,(N_i - k) - 1$ times; the last time, with $k = N_i - 1$, appears only once, the previous one ($k = N_i - 2$) appears three times, and so on.

We can simplify its expression by defining the *variance-based resultant time* τ_i as:

$$\tau_i = \sum_{k=0}^{N_i-1} (2\,(N_i - k) - 1) \cdot t_{i,k}/N_i^2; \tag{14}$$

in general, $\tau_i < \bar{t}_i$ if $N_i > 1$, as in the variance expression earlier times appear more often.

The variance-based resultant time can be easily computed (and stored for future use) from the individual read times. Note that if all times are shifted by the same amount, τ_i is similarly shifted; therefore changing the origin of time (which might come into play, e.g., to account for the different zero read times at different pixels) does not change the difference between the mean resultant time $\bar{t}_i$ and the variance-based resultant time τ_i.

For the special case of n evenly spaced reads separated by Δt, we have:

$$\tau = \bar{t} - (n-1)(n+1)\Delta t/(6n)\,. \tag{15}$$

With the definition of τ, Equation 13 becomes:

$$\mathrm{Var}(R_i) = (rn)^2/N_i + f \cdot \tau_i. \tag{16}$$

We now determine the expected covariance between two resultants, R_i and R_j, obtained by averaging N_i and N_j reads obtained at times $t_{i,k}$ and $t_{j,l}$, respectively. We assume that the resultants are strictly ordered, i.e., that either all $t_{i,k}$ are strictly smaller than all $t_{j,l}$ for all values of k and l, or all $t_{i,k}$ are strictly larger than $t_{j,l}$, for all values of k and l.

Applying the same rule as earlier, we find that

$$\mathrm{Covar}(R_i, R_j) = \sum_{k=0}^{N_i\ -1} \sum_{l=0}^{N_j-1} \frac{d\,R_i}{d\,S_{i,k}} \frac{d\,R_j}{d\,S_{j,l}} \cdot \mathrm{Covar}(S_{i,k}, S_{j,l})\,. \tag{17}$$

The derivatives equal $1/N_i$, $1/N_j$ respectively. We now use Equation 6; considering for definiteness $i < j$ (otherwise we swap i and j), the strict ordering assumption guarantees that $\min(t_{i,k}, t_{j,l}) = t_{i,k}$ for all k, l. Substituting yields:

$$\mathrm{Covar}(R_i, R_j) = 1/(N_i\,N_j) \cdot N_j \sum_{k=0}^{N_j-1} f \cdot t_{i,k} = f \cdot \bar{t}_i \qquad (i < j). \tag{18}$$

Considering both $i < j$ and $i > j$, we obtain the general expression for the covariance of disjoint, strictly ordered resultants:

$$\mathrm{Covar}(R_i, R_j) = f \cdot \min(\bar{t}_i, \bar{t}_j) \qquad (i \neq j). \tag{19}$$

4.3. Slope-fitting expression for resultant frames

As discussed in Section 3, we determine the best-fitting slope and intercept by minimizing the quadratic form in Equation 2 for matrix weights, or in Equation 3 for diagonal weights. Here we provide the more general formulae for the case of matrix weights; the expressions for the diagonal case are easily obtained by setting $W_{ij} = w_i\,\delta_{ij}$, and in most cases the double summations in the equations that follow reduce to single sums if the weight matrix is diagonal.

Let's rewrite the quadratic form Q from Equation 2 in terms of the resultants R_i, for which the linear model predicts $R_i \sim f \cdot \bar{t}_i + f_0$:

$$Q = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}(R_i - f \cdot \bar{t}_i - f_0)(R_j - f \cdot \bar{t}_j - f_0) \tag{20}$$

The values $\hat{f}, \hat{f}_0$ of f, f_0 that minimize Q can be found by differentiating with respect to f and f_0 and setting the derivatives to zero:

$$\begin{aligned} \hat{f} \cdot \Big(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}\,\bar{t}_i\,\bar{t}_j \Big) + \hat{f}_0 \cdot \Big(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}\,\bar{t}_i \Big) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}\,\bar{t}_i\,R_j \\ \hat{f} \cdot \Big(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}\,\bar{t}_i \Big) + \hat{f}_0 \cdot \Big(\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij} \Big) &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} W_{ij}\,R_i \end{aligned} \tag{21}$$

It is convenient to define the auxiliary quantities:

$$\begin{aligned} F0 &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} W_{ij} \\ F1 &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} W_{ij}\bar{t}_i \\ F2 &= \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} W_{ij}\bar{t}_i\bar{t}_j \\ D &= F0 \cdot F2 - F1^2 \end{aligned} \tag{22}$$

We also define the array of coefficients $K_i\,(i = 0, 1, \ldots, N-1)$ as:

$$K_i = \Big(F0 \cdot \sum_{j=0}^{N-1} W_{ij}\bar{t}_j - F1 \cdot \sum_{j=0}^{N-1} W_{ij}\Big)\Big/ D \tag{23}$$

We note that

$$\sum_{i=0}^{N-1} K_i = (F0\,F1 - F1\,F0)/D = 0\,, \tag{24}$$

and

$$\sum_{i=0}^{N-1} K_i\bar{t}_i = (F0\,F2 - F1^2)/D = 1\,, \tag{25}$$

regardless of the definition of the weights, as long as $D \neq 0$.

Then the solution for $\hat{f}$ of the linear system in Equation 21 can be written as:

$$\hat{f} = \sum_{i=0}^{N-1} K_i\,R_i \tag{26}$$

4.3.1. Alternate formulae without double sums

The quantities $F0$, $F1$, and $F2$, as defined in Equation 22, involve double sums over the indices i, j, therefore they require $3N^2$ array operations to be evaluated as written. If the number of resultants N is large, this evaluation can be computationally onerous. Alternatively, it is possible to reduce their evaluation to five single sums, thus $5N$ array operations, by computing and storing two new array quantities G_i and H_i, with N components each, as follows:

$$
\begin{aligned}
G_i &= \sum_{j=0}^{N-1} W_{ij} \\
H_i &= \sum_{j=0}^{N-1} W_{ij} \bar{t}_j \\
F0 &= \sum_{i=0}^{N-1} G_i \\
F1 &= \sum_{i=0}^{N-1} H_i \\
F2 &= \sum_{i=0}^{N-1} H_i \bar{t}_i \,.
\end{aligned}
\tag{27}
$$

Whether the formulation of Equation 27 is advantageous in comparison to Equation 22 depends on the computational constraints—execution time vs. memory footprint—as well as on the number of resultants. This issue only arises for non-diagonal weights; for diagonal weights, as employed in the current HST and JWST pipelines, all expressions involve only single sums, thus there is no need to introduce G_i and H_i.

4.4. Ideal case: weighting by the inverse of the covariance matrix

A special case of the least-squares fitting described in Section 4.3 is obtained when the weight matrix W_{ij} is defined as the inverse of the covariance matrix C_{ij}. The Gauss-Markov theorem, extended to the case of a full covariance matrix (Aitken 1935), proves that for a linear optimization problem with known covariance matrix, the generalized least-squares optimizer for which the weights are defined as the inverse of the covariance matrix is the best linear unbiased estimator (BLUE). Robberto (2014) provided a direct calculation of this estimator for a sample case. We refer to this case as the *optimal* estimator. As a consequence of the expressions for the covariance of individual reads (Eq. 6) and resultants (Eq. 19), the covariance matrix is in general non-diagonal and it depends on the flux rate f, thus the optimal weight matrix is also non-diagonal and it depends on f as well. While the dependence of the covariance matrix on f is linear, the dependence of the weight matrix will be much more complicated.

In practice, the optimal estimator is not widely used in the analysis of infrared detector data, for a number of reasons. First, most pipeline treatments, including those in use for HST and JWST data, assume diagonal weights; several existing pipelines would need to be rewritten for non-diagonal weights, and both the execution time and memory footprint could increase. Second, obtaining the weight matrix from the covariance matrix requires a matrix inversion for every pixel; this can be a significant performance impact. Third, and most important, the flux f is a significant component of the covariance matrix, and it is not known a priori. While it is possible to develop an iterative process that uses the estimated flux to successively refine the weight matrix at each pixel, this would be a complicated process, and the possibility of introducing subtle biases

in the resulting slope is not negligible (see, e.g., Garnett & Forrest 1993); see also Section 6.4 for more discussion of possible photometric biases resulting from the use of signal-based variance. Of course, the method *can* be used on simulated data, for which the true covariance matrix is known, and thus provides a useful point of reference for the performance of other, more practical methods (see Section 6).

4.5. Variance of the fitted slope

The variance of the value $\hat{f}$ obtained in Equation 26 can be readily obtained from the expression for the variance and covariance of the resultants, Equations 16 and 19.[7]

Applying the general expression for the variance of a linear combination, we have:

$$\mathrm{Var}(\hat{f}) = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} K_i \, K_j \, \mathrm{Covar}(R_i, R_j) \tag{28}$$

Substituting and noting the symmetry of the expression in i and j, we obtain:

$$\mathrm{Var}(\hat{f}) = \sum_{i=0}^{N-1} K_i^2((rn)^2/N_i + f \cdot \tau_i) + 2 \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} K_i \, K_j f \cdot \bar{t}_i \tag{29}$$

Most terms in this equation can be computed from the known values of $\bar{t}_i$ and τ_i, the weights W_{ij}, and the read noise (rn), resulting in a linear expression in the value of f (generally not known a priori). As in the expression for the variance of a resultant, we find that the variance of the fitted slope is the sum of two components, one that scales with the read noise and one that scales with the (true) signal. We find it convenient to define the variance V_r due to read noise and the shot-noise variance V_s per unit signal as follows:

$$V_r = (rn)^2 \sum_{i=0}^{N-1} K_i^2/N_i \tag{30}$$

$$V_s = \sum_{i=0}^{N-1} K_i^2 \, \tau_i + 2 \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} K_i \, K_j \, \bar{t}_i \tag{31}$$

where all the terms depend only on the times of the resultants, the read noise, and the weights used. The shot-noise term scales with the signal rate f, and the term V_s is the *coefficient* by

[7]Unless otherwise stated, we assume throughout this paper that the weights W_{ij}, while arbitrary, are independent of the measurements R_i. If this is not the case, e.g., if the weights are set on the basis of estimated slope or S/N, the results presented here are not strictly applicable, and biased results are possible. The impact of any dependence of the weights on the measured signal is discussed in Section 6.4; see also the discussion in Regan (2007).

which the signal rate f is multiplied. Thus the predicted variance in the estimated slope resulting from weighted least-square fitting is:

$$\mathrm{Var}(\hat{f}) = V_r + V_s \cdot f. \tag{32}$$

The quantities V_r and V_s can be computed as arrays over the image from sums over the coefficients K_i and the relevant times. The expression for V_r involves a double sum over resultants even in the case of diagonal weights, requiring N^2 array operations; as shown in Section 4.3.1 for the case of the slope calculation with matrix weights, the double sum can be avoided by defining an additional vector quantity (see Eq. 27), resulting in $2N$ array operations at the cost of requiring storage of an additional N arrays.

As discussed in the last item in Section 4.1, the signal rate f used in Equation 32 may need to include the signal rate due to dark current; even if the dark current has been subtracted, its linear growth will still contribute to the Poisson noise of the signal in exactly the same way as photons from the sky.

4.6. The least-squares solution is unbiased

It is straightforward to prove that, under the assumptions in Section 4.1, the slope solution expressed in Equation 26 is unbiased, i.e., its expectation value equals the true value of the slope.

Indeed, assume that the resultant values R_i have expectation value $E(R_i) = f_0 + f \cdot \bar{t}_i$. Then the expectation value of the slope $\hat{f}$ given by Equation 26 is:

$$E(\hat{f}) = \sum_{i=0}^{N-1} K_i \, E(R_i) = \sum_{i=0}^{N-1} (f_0 \, K_i + f \, K_i \bar{t}_i) \tag{33}$$

Applying Equations 24 and 25, we obtain:

$$E(\hat{f}) = f. \tag{34}$$

This expression is valid regardless of the weights W_{ij} chosen, as long as they are fixed, i.e., independent of the measured values R_i. In practice, the slope estimate for a single exposure is unbiased—within measurable limits—even when this assumption is violated, e.g., by choosing weights on the basis of the measured signal; this is discussed further in Section 6.2.

5. Steps for a practical implementation of slope-fitting calculation and variance estimate in the case of unevenly-spaced resultants

The expressions given in the previous Sections can be readily computed in the calibration pipeline. Here we provide a possible sequence of computations that could be carried out efficiently

in array form as part of a pipeline implementation. We limit the discussion to diagonal weights, since this is the case currently considered in the HST and JWST pipelines.[8]

We assume that the read noise for each pixel has already been determined and stored in a relevant reference file, and that the Multi-Accum table in use has been ingested, so that the values of N_i, $\bar{t}_i$, and τ_i for each resultant frame are known. We assume that the weights w_i have been chosen, possibly with a different value for each pixel. For clarity, quantities that vary from pixel to pixel are designated by calligraphic letters, such as $\mathcal{R}_i$ for the resultants, $\mathcal{W}_i$ for the weights, $(\mathcal{RN})$ for the read noise, $\mathcal{D}$ for the denominator of Equation 23, $\mathcal{V}$ for the variance of the slope, and so on. Although only the resultant index is shown explicitly (if needed), each of these quantities is really a two- or three-dimensional array, with the pixel index assumed implicitly.

We carry out the calculation using a single value of each time quantity for the whole array. Individual pixels are read at different absolute times, but the time offset between each read/resultant and the zero read is expected to be the same for all pixels. It is then convenient to define the time origin for each pixel as the time at which the pixel is read in the zero-read; then all times are the same for the whole array. If for any reason this assumption is found to be inapplicable, then the actual timing for each pixel needs to be used to define $\bar{t}_i$ and τ_i, which then become pixel-dependent quantities. All other calculations remain the same; note that a change in the origin of time only changes the interpretation of the pedestal f_0 (itself a pixel-dependent quantity), which we are essentially ignoring in our treatment, but leaves the interpretation of the slope $\hat{f}$ unchanged.

We also assume that bad pixels have already been identified and flagged, with their weight set to zero. The question of identifying cosmic ray hits is not handled explicitly here. The method currently in use identifies cosmic rays from significant deviations from uniformity of the differences between consecutive resultants; this method will need to be modified for unevenly-spaced resultants. Once cosmic rays are identified, affected ramps are split into two or more segments, an independent slope is fitted to each segment, and the slopes are combined according to their inverse variance. This method can probably be improved, but we will leave its treatment to a future discussion.

With this starting information, here are the steps to be carried out:

Auxiliary quantities The auxiliary quantities $\mathcal{F}0$, $\mathcal{F}1$, and $\mathcal{F}2$, defined in Equation 22, are computed as:

[8]Generalizing the computations to matrix weights is straightforward, but it requires making a choice between number of operations and memory footprint, as discussed in Section 4.3.

$$\begin{aligned} \mathcal{F}0 &= \sum_{i=0}^{N-1} \mathcal{W}_i \\ \mathcal{F}1 &= \sum_{i=0}^{N-1} \mathcal{W}_i \bar{t}_i \\ \mathcal{F}2 &= \sum_{i=0}^{N-1} \mathcal{W}_i \bar{t}_i^2 \end{aligned} \tag{35}$$

These quantities depend only on weights and times. If the weights are fixed (uniform weighting) or only assume a limited set of values (e.g., by using the NIRCam-like weighting scheme suggested in Section 6.1), then their calculation can be greatly simplified. The same consideration applies to the resultant coefficients and the variance terms, Equations 37, 39, and 40.

However, individual pixels can be rejected only for some resultants, e.g., because of data transmission errors or other anomalies. This can result in a more general form of the weight array, requiring the full array computation for those pixels.

Denominator The quantity $\mathcal{D}$, also defined in Equation 22, is computed as a single two-dimensional array:

$$\mathcal{D} = \mathcal{F}2 \cdot \mathcal{F}0 - \mathcal{F}1^2 \tag{36}$$

This quantity vanishes only when most weights vanish, i.e., a pixel is flagged as bad for all but at most one resultant. Such pixels are excluded from further computations, and their slope estimate is flagged as invalid.

Resultant coefficients The coefficients $\mathcal{K}_i$ defined in Equation 23 are computed as N two-dimensional arrays:

$$\mathcal{K}_i = \Big(\mathcal{F}0 \cdot \bar{t}_i - \mathcal{F}1\Big) \cdot \mathcal{W}_i / \mathcal{D} \tag{37}$$

This calculation is skipped wherever $\mathcal{D}$ is zero. As for Equation 35, the values of $\mathcal{K}_i$ depend only on weights and times, thus their computation can be greatly simplified if the weights are uniform or take a small number of values.

Slope determination The estimated slope $\hat{\mathcal{F}}$ is computed as a sum over the resultants $\mathcal{R}_i$ and the coefficients $\mathcal{K}_i$:

$$\hat{\mathcal{F}} = \sum_i \mathcal{K}_i \mathcal{R}_i \tag{38}$$

The calculation is skipped for pixels that have $\mathcal{D} = 0$. Note that the coefficient $\mathcal{K}_i$ vanishes for each resultant that has a bad pixel, as a consequence of $\mathcal{W}_i$ vanishing.

Read-noise variance The read-noise component $\mathcal{V}_R$ of the slope variance is computed as:

$$\mathcal{V}_R = \sum_{i=0}^{N-1} \mathcal{K}_i^2 \cdot (\mathcal{R}\mathcal{N})^2 / N_i \tag{39}$$

Signal variance The *coefficient* $\mathcal{V}_S$ of the count rate in the signal-based component of the slope variance is computed as:

$$\mathcal{V}_S = \sum_{i=0}^{N-1} \mathcal{K}_i^2 \tau_i + \sum_{i<j} 2\mathcal{K}_i \mathcal{K}_j \cdot \bar{t}_i \tag{40}$$

Total variance If desired, a (biased) estimate of the total slope variance $\mathcal{V}$ can be computed by adopting $\hat{\mathcal{F}}$ as the estimate of the slope:

$$\mathcal{V} = \mathcal{V}_R + \mathcal{V}_S \cdot \hat{\mathcal{F}}. \tag{41}$$

Other signal estimates can be used in place of $\hat{\mathcal{F}}$. However, please note the caveats on possible biases due to the correlation between slope error and variance estimate; more discussion can be found in Section 6.4.

6. Numerical Estimates

In this Section we provide numerical estimates, based on Monte Carlo realizations of ramps with the appropriate properties, for the expected noise and possible bias in the signal derived with various choices of readout patterns and weights. We consider the ten readout patterns defined in Sections 2.1 and 2.2, with exposure times ranging from ~ 50 s to ~ 600 s. We also consider briefly the ideal case in which all reads are retrieved. We will discuss different weighting schemes, as well as the impact of the choice of readout scheme on saturation effects; and we provide some considerations on biases that might be incurred unless care is exercised in the treatment of the fit results.

6.1. Weighting schemes

The weighting scheme used by the JWST pipeline for NIRCam is based on a "robust" estimate of the signal to noise ratio in each pixel. The scheme, which follows the recommendations ofFixsen et al. (2000), weighs early and late frames more for high signal pixels, while it uses uniform weighting for low signal pixels.

Specifically, the weight for the frame (read or resultant) with index i is

$$w_i^{\rm NIRCam} = |i - i_{mid}|^P \tag{42}$$

where i_{mid} corresponds to the midpoint of the exposure, and the exponent P varies from 0 (uniform weights) to 10 (early and late frames heavily weighted) according to Table 2.[9] Equation 42 explicitly assumes that frame index maps linearly to time; a simple modification that accounts for the different length of time between resultants would be:

$$w_i^{\rm NIRCam,mod} = |\bar{t}_i - \bar{t}_{mid}|^P \tag{43}$$

where $\bar{t}_{mid}$ is the midpoint between the values of $\bar{t}_i$ of the first and last resultant.

As we will see in the following, the NIRCam weighting scheme works quite well for evenly spaced resultants, but it needs modification for uneven resultants. For low signal pixels, resultants consisting of more reads enjoy better read noise suppression, thus carry more information and should be preferentially weighted more. In the limit of dominant read noise, each resultant should be weighted by N_i, the number of contributing reads; we call this NComp-weighting. On the other hand, for high signal pixels, read noise is irrelevant, and better results are obtained by weighting early and late frames the most, independent of the number of reads that go into each. (In the absence of read noise, the difference between last and first read carries all the information needed.)

Therefore we propose a weighting scheme that is consistent with the NIRCam scheme of Equation 43 when resultants are evenly spaced, but accounts for the number of reads in each resultant, and reduces to N_i in the case of negligible signal (i.e., dominant read noise). Following the NIRCam practice, we define the signal-to-noise parameter s for a given pixel as

$$s = S_{max}/\sqrt{(rn)^2 + S_{max}} \tag{44}$$

where S_{max} is the maximum signal in electrons, as estimated from the last available read or resultant, and rn is the read noise, also in electrons. Accordingly, we define the WFI weights $\tilde{w}$

[9]See also the online documentation at `https://jwst-pipeline.readthedocs.io/en/latest/jwst/ramp_fitting/description.html#slope-and-variance-calculations`.

Table 2: Choice of Exponent P as a function of signal-to-noise parameter s for NIRCam, defined in Eq. 44.

Condition	Value of P
$s < 5$	0
$5 \leq s < 10$	0.4
$10 \leq s < 20$	1
$20 \leq s < 50$	3
$50 \leq s < 100$	6
$s \geq 100$	10

as follows:

$$\tilde{w}_i = \frac{(1+P)\cdot N_i}{1+P\cdot N_i}\ |\bar{t}_i - \bar{t}_{mid}|^P \tag{45}$$

with the exponent P defined in the same way as for NIRCam.

For even resultants, all values of N_i are equal, and $\bar{t}_i \propto i$; in this case, the expression reduces to the NIRCam form, aside from a multiplicative factor independent of i, which cancels out in the weight definition. For low signal, when the read noise is dominant, the exponent $P = 0$ and the expression reduces to $w_i \propto N_i$; at high signal, when P is large, this scheme weights the first and last read the most, nearly independent of N_i. We will test this expression numerically, comparing it to the NIRCam formula, unweighted fits, and the optimal weights defined in Section 4.4.

In the following comparison, we consider five choices of weights:

Unweighted or Uniform: All resultants are given equal weight.

NIRCam: Use the NIRCam weighting scheme based on time, rather than resultant index, per Equation 43.

NComp: Weight each point by the number of components, NComp, i.e., the number of frames that were averaged for that resultant. This method is expected to be close to optimal in the limit of very low signal, when the read noise contributes most of the noise.

WFI-proposed: Use the NIRCam weighting scheme adjusted for the variable number of reads included in each resultant, according to Equation 45.

Optimal: Use the inverse covariance matrix to weight the linear fit. This is not usable in practice, as it requires knowledge of the actual count rate, which is not known in a realistic scenario, but it provides a useful limiting case. We will discuss in Section 6.4 the possibility of using an estimated count rate for weighting purposes, and will show that the method has potential pitfalls.

6.1.1. Signal-Based weight

One weighting scheme that is occasionally suggested is to use the inverse variance of each data point, based on Poisson (total counts) and read noise, as the relative weight of that point. We will refer to this as "Signal-Based" weighting. Specifically, one assigns each measurement a variance equal to the total number of counts (Poisson) and the square of the read noise, and assigns that point weight equal to the inverse of this variance in the linear fit. If desired, this can be modified to account for the noise model for resultants (Eq. 16).

This method has very poor performance, especially in the limit of high signal. From the work of Fixsen et al. (2000), we know that in the high signal regime, it is best to weight early and late points the most, and down-weight intermediate measurements; in contrast, signal-based weighting gives less weight to late measurements, since they have the most signal (and thus

the most variance). This is a direct consequence of neglecting the off-diagonal elements of the covariance matrix, which carry the information that much of the Poisson variance of late points is actually in common with previous measurements, with which they share most of the counts.

In addition, signal-based weighting use the signal itself to determine the weight of each point, thus violating the assumption of Section 4.6; in principle, this could result in a biased estimate of the signal. However, our numerical tests show that any bias related to this choice is very small, less than 0.1% of the slope itself, for a broad range of signal levels. More details are given in Section 6.2.

6.2. Performance comparison

In this Section we characterize the relative performance of different choices of Multi-Accum tables and weighting schemes, focusing on the S/N that can be achieved for different count rates and readout schemes, and as a function of the weighting scheme used. In the next Section we discuss saturation and dynamic range.

Consistent with our treatment throughout, we do not include non-ideal effects such as linearity corrections, variations in input signal over time, residual electronic terms after reference pixel correction, correlated noise, and so on. The focus here is on the performance under nominal conditions. Non-nominal behavior will impact all weight choices, but possibly in different ways, and likely necessitate a reassessment of the relative merits of the possible options.

Also note that in this treatment, we do not distinguish the source of "signal", be it source, sky background, dark current, or other backgrounds. Under the assumption that each of these terms is constant and they are all additive with each other, the best measurement of the total will also yield the best measurement of the source; other terms can usually be subtracted to higher accuracy (e.g., darks from dark reference frames, and backgrounds from spatial averages in the neighborhood of a source).

Under these assumptions, the expected S/N of exposures of given length will depend to some extent on how individual reads are used: transmitted to the ground, averaged, or skipped. This behavior is controlled by the Multi-Accum table used, and thus the S/N, compared to the (non-realistic) ideal case in which all reads are used individually, is a good gauge of their relative value.

The Multi-Accum tables considered here are listed in Table 1.

Figures 3 and 4 show the performance of the different weighting schemes listed in Section 6.1 as a function of the signal level and for different Multi-Accum tables. All S/N estimates are scaled to the "optimal" weighting for that same Multi-Accum Table at the same signal (the horizontal line at 1.0), and are for a single exposure, assuming a read noise of 10 e^-. The purple line at values below 1.0 is the relative performance of the ideal case when all reads are used individually with optimal weight; the amount by which the purple line falls below the reference line at 1.0 indicates the net loss in S/N due to the averaging or skipping of reads, not including additional effects such

as cosmic ray hits. Note that for the five JWST-like schemes shown in the right-hand panels, the number of components is the same for every group (resultant); therefore the Unweighted and NComp weighting schemes are equivalent, and so are the NIRCam and WFI-Proposed schemes.

A few general patterns can easily be seen. For all choices of readout patterns, an unweighted fit yields a good performance at low signal levels, especially for even patterns, but becomes increasingly worse at high signal levels. (Here and in the following, we compare the performance of different weighting schemes on the basis of the effective S/N at the same signal level.) NComp weighting outperforms Unweighted for all uneven patterns, approaching the performance of optimal weighting at low S/N. The NIRCam weighting performs generally better, i.e., offers lower variance than NComp weighting at high S/N, but it becomes substantially worse (and close, by construction, to Unweighted) at low signal levels, as it assigns the same weight to resultants with different number of components, despite the fact that resultants with more components have lower read noise. The WFI-Proposed weight performs well throughout the range of assumed signal, always within 1% of optimal weighting, for exposure times up to 150s (Fig. 3); as expected, it fixes the issue with NIRCam weighting for uneven resultants at low signal, where it performs similarly to NComp weighting and close to optimal weighing. It is almost always better than NIRCam weighting, only appearing very slightly worse for narrow ranges of signal. However, for longer exposure times (Fig. 4), both the NIRCam and the proposed weighting scheme perform significantly worse, with a penalty of up to 4% for certain ranges of signal. The problem appears to be primarily in the choice of exponent P; at signal levels comparable to the sky background, about $0.3e^-\,\mathrm{s}^{-1}\mathrm{pix}^{-1}$, it appears that the choice of $P = 0$ is not optimal. However, changing the choice of weight away from the NIRCam pattern may also introduce complications. If the 2–4% penalty below $1e^-\,\mathrm{s}^{-1}\mathrm{pix}^{-1}$ is seen as excessive, further optimization of the weight choice can be conducted.

One final note concerns the possibility of a biased estimation of the slope for an individual exposure. In Section 4.6, we proved that any linear slope estimator with fixed weights is unbiased; however, some of the weighting schemes considered here *do* depend on the measured signal, thus violating the assumptions of Section 4.6. Our numerical tests indicate that any biases that might exist due to this effect are less than 0.1%, the estimated precision of our experiments, and thus are negligible even at the level of accuracy required of Roman.

6.2.1. A note on cosmic rays

Cosmic rays can be a significant consideration in the choice of exposure times and readout schemes for science observations, and their impact has been included in the analysis of expected JWST observations, e.g., by Giardino et al. (2019). The frequency of cosmic rays, outside of solar events, is expected to be roughly 2 to 4 times higher at L2 than in low Earth orbit; Giardino et al. (2019) predict a rate between 2 and $5\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. Each event can impact several pixels, depending on the angle of incidence of the particle on the detector. Events can have different

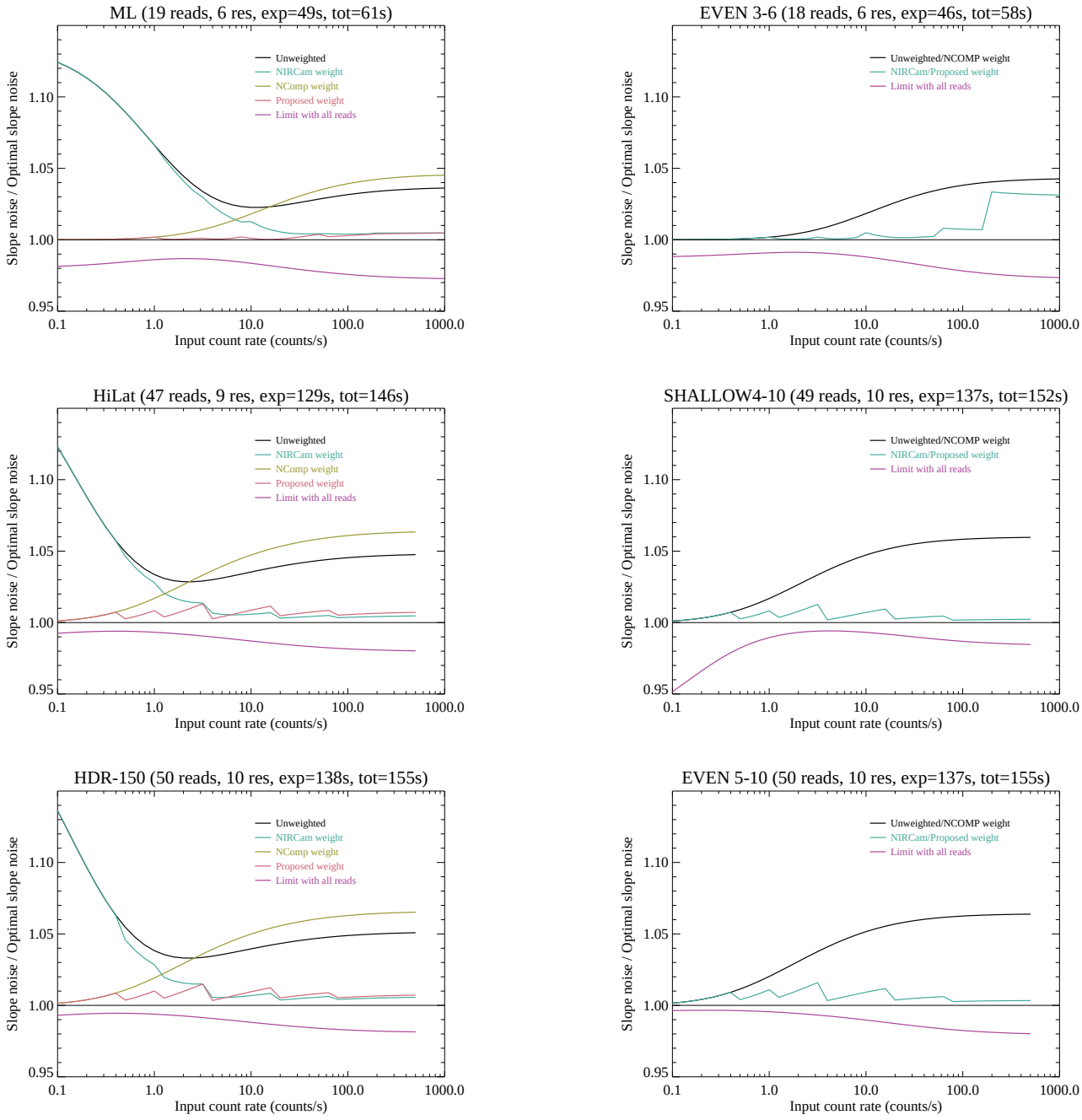

Fig. 3.— Expected S/N vs. count rate for different choices of weighting. Each panel refers to a different readout pattern, with characteristics shown in Table 1. The panels on the left are for patterns with unevenly spaced resultants, chosen to improve dynamic range; the corresponding right-hand panels are for JWST-style readout patterns with evenly spaced resultants and comparable overall characteristics. The expected S/N is scaled to the optimal weight case (horizontal line at 1.0). The proposed weight scheme (Eq. 45) performs within 1% of optimal weight for all count rates; however, see Fig. 4 for longer exposures. For reference, the purple curve below the horizontal line shows the relative performance that could be achieved if all reads were downloaded.

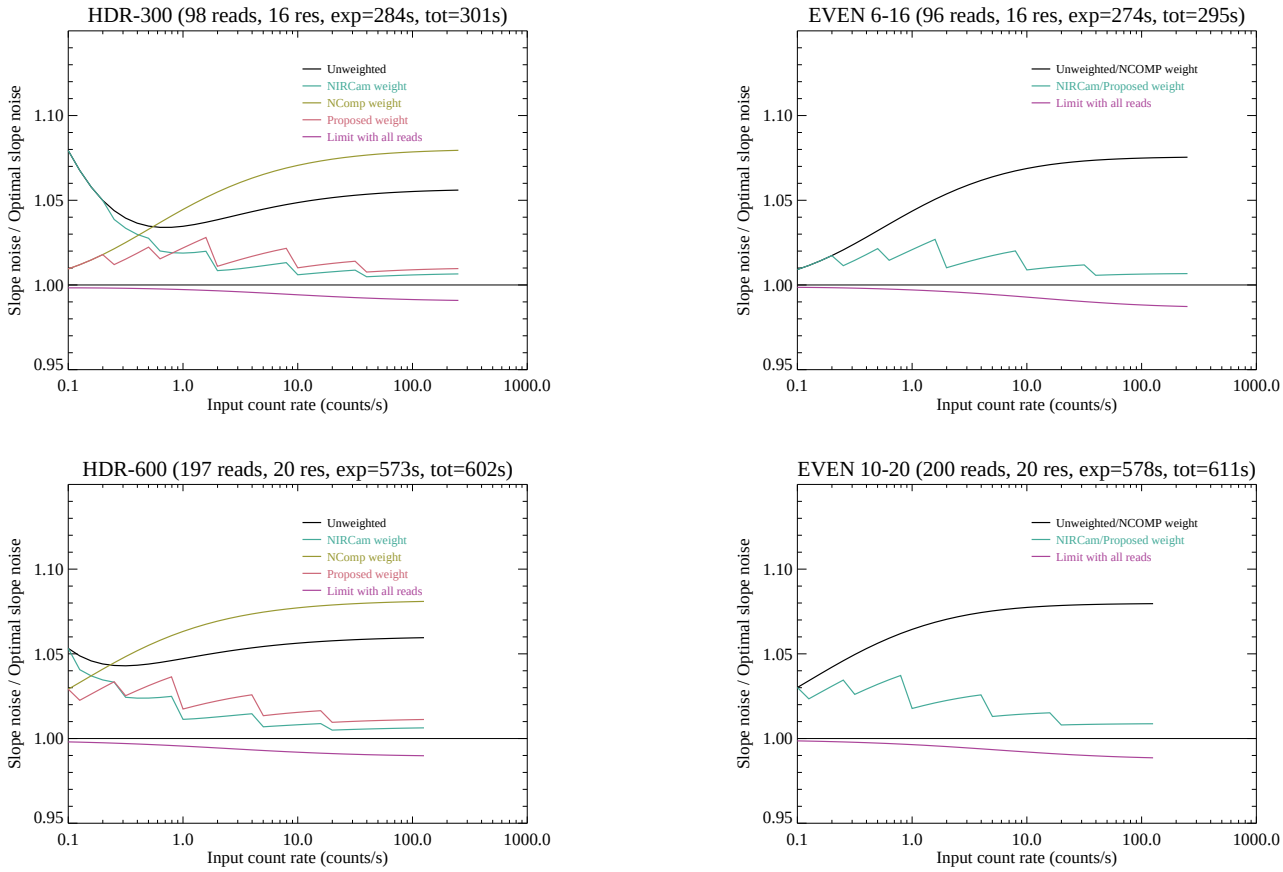

Fig. 4.— Same as Fig. 3, but for the readout schemes with exposure times 300s or longer. Unlike shorter exposure times, the proposed weight scheme (and the NIRCam weight scheme for evenly spaced resultants) are about 2–4% less efficient than optimal weighting at low signal.

effects, depending, e.g., on the impinging particle (most often protons, but heavier nuclei will also be present) and their energy (low-energy particles can generate larger showers). JWST will soon provide very detailed information on the frequency, effect, and other characteristics of cosmic ray events at L2, with detectors that are very similar to Roman's H4RG.

However, some impact estimates are possible based on our current expectations. If the events behave as those in LEO, each will deposit a significant amount of charge near-instantaneously in one or more pixels, after which the behavior of that pixel will return to normal (i.e., accumulating charge linearly with time). Cosmic ray events can be detected by comparing the amount of charge increase between resultants with the expectation from a simple linear rise; if the deviation is large enough, a cosmic ray is assumed to have taken place between the affected resultants, and the interval is discarded. In practice, this can be achieved by fitting two constant-slope segments before and after the event; the slopes can be averaged, possibly with weights, after the fact, or the fit can be conducted as a linear slope with an unknown jump (a third parameter). In rare cases, an event might affect two resultants, or happen early or late enough that there is not enough information to fit a slope to the shorter segment. In any of these cases, the effective S/N

Table 3: Performance and saturation for sample Multi-Accum tables

Name	Reads[2]	Time[2] (s)	Max Count Rate (e/s)	Δm (mag)	$m_{\rm sat}$ (AB mag)	SNR (1 exp.) 0.3 e/s	SNR (1 exp.) 10 e/s	SNR/$\sqrt{t_{\rm total}}$ 0.3 e/s	SNR/$\sqrt{t_{\rm total}}$ 10 e/s
ML	3	9.12	8771.93	1.31	16.38	1.86	21.15	0.24	2.71
EVEN 3-6	6	18.24	4385.96	0.55	17.13	1.74	20.45	0.23	2.69
HiLat	3	9.12	8771.93	1.31	16.38	4.86	35.48	0.40	2.94
SHALLOW4-10	9	27.36	2923.98	0.11	17.57	4.93	36.57	0.40	2.97
HDR-150	2	6.08	13157.89	1.75	15.94	5.12	36.70	0.41	2.95
EVEN 5-10	10	30.40	2631.58	0.00	17.69	5.12	36.77	0.41	2.95
HDR-300	2	6.08	13157.89	1.75	15.94	8.25	52.68	0.48	3.04
EVEN 6-16	12	36.48	2192.98	-0.20	17.88	8.14	52.11	0.47	3.03
HDR-600	2	6.08	13157.89	1.75	15.94	12.33	74.87	0.50	3.05
EVEN 10-20	20	60.80	1315.79	-0.75	18.44	12.44	75.53	0.50	3.06

Notes: "Reads[2]" is the total number of reads at the end of the second resultant. "Time[2]" is the time of the last read in the second resultant; a pixel that is not saturated at that time has at least two unsaturated resultants. "Max Count Rate" is the largest count rate that can be accommodated before a pixel saturates at Time[2]. "Δm" is the gain in saturation magnitude with respect to "EVEN 5/10". $m_{\rm sat}$ is the AB magnitude corresponding to saturation in F146 for a star landing in the center of the pixel, averaged over the field of view. "SNR (1 exp)" is the S/N that can be achieved with that Multi-Accum table for a pixel with count rate of 0.3 and 10 e/s, respectively, in one exposure under ideal conditions, assuming a read noise of 10 e. "SNR/$\sqrt{t_{\rm total}}$" is that S/N scaled by the inverse square root of the total on-target time, an indicator of the relative efficiency of observations of different length; higher is better.

for the affected pixel is likely decreased, possibly by 10–15% depending on when the event takes place and on the readout scheme used.

Assuming the high "normal" rate from Giardino et al. (2019) of 5 events $\rm cm^{-2}\,s^{-1}$, and given the H4RG pixel size of 10μm, a 1000x1000 region of the detector will be hit roughly 5 times every 1000 seconds. If each event affects an average of 4 pixels, then 2% of the detector pixels are affected every 1000s, or 1.2% in the longest exposure considered here (600s). Therefore the net impact on efficiency for 600s exposures (or shorter) is likely a small fraction of 1%. These estimates will need to be revisited if 1) JWST shows a substantially higher rate of events or number of affected pixels per event than our current estimates, or 2) a fraction of the events have additional impacts on pixel response, such as destabilization of the affected pixel or enough charge to cause saturation; under those circumstances, a pixel affected by a cosmic ray would become unusable for some time after the hit, and a larger fraction of the exposure would be lost for that pixel.

We also note that the current JWST pipeline does not have the capability to identify and flag cosmic ray events for uneven resultants. The jump-detection algorithm used for JWST is

predicated on the assumption that all resultants are evenly spaced, as it compares the differences between subsequent resultants to one another. Therefore a new algorithm will need to be devised; it will be useful to revisit the impact of cosmic rays, including detection efficiency, when the new algorithm is in place and more information on cosmic ray events at L2 is available.

6.3. Saturation effects

One of the advantages of unevenly spaced resultants is the improved dynamic range they provide for exposures of any length. The performance of even vs. uneven resultants for faint sources is extremely similar (Table 3), with a very slight advantage (less than 1%) for unevenly spaced resultants in terms of S/N achieved per unit time. However, bright sources have the potential of saturating (exceeding the maximum correctable count rate) before useful data can be accrued. Infrared detectors, including these H4RG as tested in the lab, usually suffer from a "soft" saturation: as the charge collected approaches the full-well limit for the pixel, the voltage response becomes increasingly non-linear, and eventually approaches a maximum value where it is insensitive to the signal deposed. At count levels less than half the well depth, this phenomenon, often described as classic (or signal-dependent) non-linearity, can be readily corrected in calibration. However, as the deposited charge increases, the correction becomes increasingly less accurate, and at some point the pixel is considered saturated. For the purpose of this discussion, we adopt a somewhat arbitrary saturation level of 80,000 electrons; in reality, the saturation level will vary form detector to detector and even from pixel to pixel. A different saturation level does not change the relative merits of different readout schemes.

As discussed in Section 2.1, current tests indicate that the initial read-reset does not generally provide a valid starting point to determine the count rate in a pixel; therefore it is usually necessary to obtain two valid count values, *after* the initial read-reset, in order to fit a slope for that pixel. [10] Therefore we consider saturated any source for which the total counts in the brightest (central) pixel, *after the last read of the second resultant,* exceed the nominal saturation level of 80,000 electrons. The column "Time[2]" in Table 3 reports this time for each choice of readout, and the column "Max Count Rate" shows the largest count rate that leads to a measurable slope under those circumstances.

The gain offered by the WFI readout patterns, and especially by those labeled "HDR", is now obvious. For example, HDR-300 allows the observation of a source a factor 6 (1.95 mag) brighter than its equivalent even-exposure pattern, EVEN 6/16; HDR-600 offers a full factor 10 (2.5 mag) over EVEN 10/20. The column Δm in Table 3 gives the difference in saturation magnitude compared with the EVEN 5/10 case; a positive number means that brighter stars are observable with that readout pattern. The column $m_{\rm sat}$ gives the saturation limit in F146,

[10] This is slightly simplistic, as there is always the chance that a cosmic ray hit affects one of the resultants. With only two valid (unsaturated) resultants, it is not possible to identify a cosmic ray event. If desired, similar considerations can be applied to the third resultant instead.

expressed in AB magnitudes, with each readout pattern; the WFI-like patterns have saturation limits between 0.8 and 2.5 mag fainter than the corresponding even patterns. The saturation magnitude has been obtained assuming that 29.6% of the light falls on the central pixel (for a star centered in a pixel), and that the zero point for F146 is 27.558; these values, and thus the saturation magnitude, will likely evolve as we learn more about the instrument, but the relative benefit in dynamic range across readout patterns will remain the same.

6.4. Signal-based weighting and photometry bias

The contribution of each measurement to a photometry estimate is subject to the Poisson noise inherent to discrete events (detection of photons) happening at a constant rate. The Poisson variance in the number of events is the expected number of events itself.

In analogy to common implementations of least-square fitting, it is tempting to use the intrinsic variance of each measurement (i.e., resultant) as weight for that point. With this signal-based weighting, defined in Section 6.1.1, the weight of each resultant would be its variance, defined in Equation 16, i.e., the combination of Poisson and read noise. However, this simplistic approach is not recommended for the case we are considering, primarily because of the covariance of the resultants—i.e., the off-diagonal elements in Equation 16; proper consideration of the full covariance requires that the weight matrix be set to the inverse of the covariance matrix, as seen in Section 4.4 (see also Robberto 2014). The results of adopting the signal-based weighting are shown in Figure 5, and show that it pays a penalty of more than 10% at high signal levels. The reason is obvious: signal-based weighting down-weights late resultants, which have the largest signal and thus the largest nominal contribution from Poisson noise, while we know from previous discussions that early and late resultants carry the most information at high signal levels. This Figure is specific to the HiLat readout pattern, but similar results apply to all other readout patterns, whether the resultants are evenly spaced or not.

Although signal-based weighting is less efficient (i.e., noisier) than other forms of weighting, it does not result in a bias in the slope determined from a single exposure. The question arises because signal-based weighting grossly violates the assumptions of Section 4.6: not only is the weighting scheme dependent on the signal, but the weight given to each point depends on the value at that point. However, numerical experiments show that any bias due to this effect is less than 0.1%, and thus negligible.

The issue of photometry bias *across multiple observations* is subtler but more general, as it applies to *any* weighting scheme. The problem arises from combining multiple measurements with weights that depend on the respective signal; it has long been known (see, e.g., Garnett & Forrest 1993) that using the measured signal to estimate its uncertainty can lead to a bias in the resulting measurement, i.e., the photometry of the source. As shown in Equation 32, the estimated variance of the slope obtained with any weighting scheme is the sum of two terms, one independent of the signal and the other proportional to the signal. In the simplest terms, data

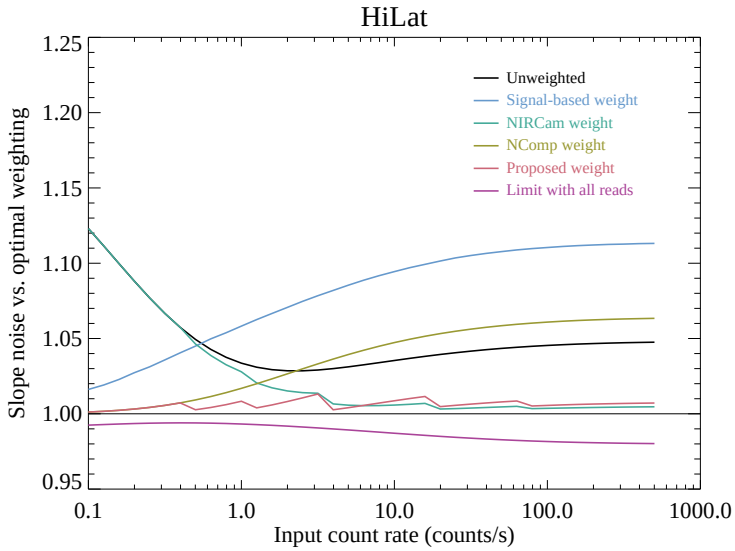

Fig. 5.— Expected S/N achieved with different forms of weighing for the HiLat readout pattern, as in Fig. 3, but with the inclusion of signal-based weighting (light blue line). The naive signal-based weighting, which down-weights late resultants, pays a significant penalty in the achieved S/N at high count levels. Similar results apply to all readout patterns.

points with a positive deviation (more than expected counts) would be under-weighted because their estimated variance is larger than the truth; data points with a negative deviation (fewer than expected counts) will be overweight-ed because their estimated variance is smaller than the true variance. The result is underestimating the source flux in that pixel, by an amount that increases at low S/N.

Figure 6 shows the bias resulting from averaging multiple measurements when a weighted average is taken, with the weight of each measurement based on its measured slope. This bias is small in comparison with the single-measurement variance; however, as this is a *systematic* effect, it can become significant when multiple measurements are combined, reducing their individual scatter. As a significant part of the Roman science mission is based on very large surveys, systematic effects can easily become very significant. An example is the case of deep images resulting from combining many individual exposures; if each is assigned a weight based on Equation 32, using the measured slope for that exposure, the combined image will be biased low, by an amount that depends on the signal in each pixel. This obviously applies to deep fields, but also to any other case in which the final image is obtained by stacking a large number of individual exposures, such as might be obtained by some versions of the Galactic Bulge Time Domain Program. More subtly, the same effect could apply to the photometry of a collection of sources, each with its own estimated uncertainty, if those uncertainties are used to determine the average signal.

In order to avoid such biases, it is possible to use weighting schemes that do *not* rely on the signal measured for each exposure. Examples include weights based on exposure time, on sky estimated from a very large number of pixels, or on the source brightness *as determined from all*

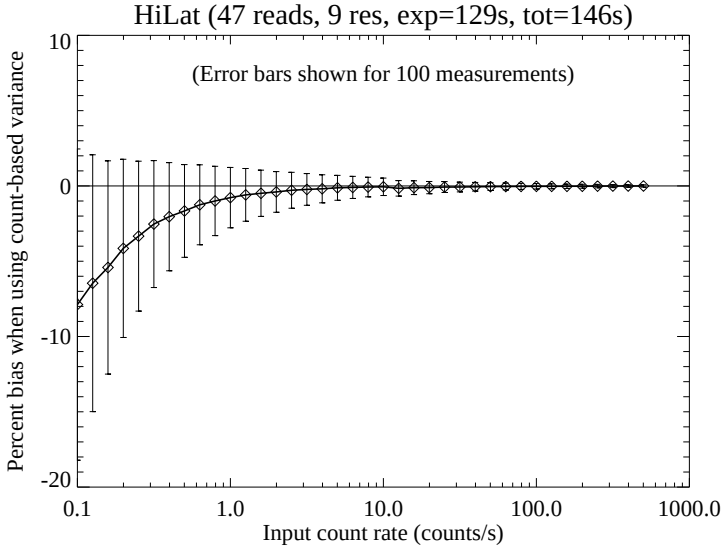

Fig. 6.— Photometry bias resulting from inverse-variance weighting, if the variance is based on the estimated signal via Eq. 32. The error bars reflect taking the average of 100 measurements; thus the bias, while negligible for an individual measurement, can quickly become significant as multiple measurements of the same source, or across different sources of comparable brightness, are averaged together. This figure is for the HiLat readout pattern, but similar results apply to all readout patterns, whether even or uneven, and for any weighting scheme.

contributing exposures together. In practice, this process would entail keeping track of the V_r and V_s terms in Equation 32 separately, and only carrying out the final combination once the signal in each pixel is computed from all exposures. We are not aware that such a capability exists in current image combination methods; therefore less efficient, but safer, weighting methods could be used until such tools are readily available. It is also possible that for certain types of analysis, depending on the type of data and typical S/N per exposure of the sources under consideration, these effects are negligible; as the answer depends on the science scenario under consideration, we are unable to offer more specifics on this point.

7. Conclusions

We have evaluated the modifications required in the handling of up-the-ramp sampled data if individual reads are grouped and sampled unevenly in time, as proposed for the operation of the Wide Field Imager on the Roman Space Telescope. We find that the changes in the slope-fitting process are modest, and can be efficiently coded via array operations; we predict that any impacts on the execution time of the slope-fitting step will likely be minor. However, changes will be needed in the weighting scheme used for the slope fit, and the computation of the variance of the resulting slope requires additional computations of scale similar to the slope fit itself. We consider several weighting schemes and provide a numerical evaluation of their performance for a

range of possible readout patterns. We also identify possible photometry biases that might occur if the detected signal is used to define the relative weights of multiple measurements.

Finally, we note that the identification of cosmic rays events will also need to be modified from the existing NIRCam algorithm; this will be the topic of a separate study.

I am grateful to my colleagues in the Roman Telescope Branch, and especially to Andrea Bellini, Rick Cosentino, Tyler Desjardins, Andreea Petric, and Russell Ryan, for their feedback on this study and for reading and commenting on earlier versions of the manuscript, and to Sebastian Gomez, Rick Cosentino, Harry Ferguson, Megan Sosey, and Cristina Oliveira for a very careful review of the manuscript. People in the DCL group at Goddard Space Flight Center, and especially Bob Hill, Greg Mosby, Bernie Rauscher, and Maxime Rizzo, have greatly helped me understand some of the nuances of the analysis of H4RG data and the behavior of the detectors. Dale Fixsen, Bernie Rauscher, Massimo Robberto, and their coauthors have developed the methodology for the rigorous analysis of infrared data; the work presented here is a small increment on their pioneering results. The encouragement received from Gisella de Rosa and Jonathan Hargis is gratefully acknowledged. I am also grateful to my colleagues in the Data Management System group, and especially Nadia Dencheva and Megan Sosey, for assistance in understanding the properties of the NIRCam pipeline and for motivating additional analysis on Roman-specific issues.

REFERENCES

Aitken, A. C. 1935, Proc. Royal Soc. Edinburgh, 55, 42

Brammer, G. 2016, Reprocessing WFC3/IR Exposures Affected by Time-Variable Backgrounds, Space Telescope Instrument Science Report WFC3 2016–16

Brammer, G., Pirzkal, N., McCullough, P., & MacKenty, J. 2014, Time-varying Excess Earthglow Backgrounds in the WFC3/IR Channel, Space Telescope Instrument Science Report WFC3 2014–03

Fixsen, D. J., Offenberg, J. D., Hanisch, R. J., et al. 2000, PASP, 112, 1350

Garnett, J. D., & Forrest, W. J. 1993, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 1946, Infrared Detectors and Instrumentation, ed. A. M. Fowler, 395–404

Gennaro, M., & Khandrika, H. 2021, A maximum likelihood approach to estimating the flux in infrared detectors non-destructive ramps, Space Telescope Instrument Science Report WFC3 2021–11

Giardino, G., Birkmann, S., Robberto, M., et al. 2019, PASP, 131, 094503

Mosby, G., Rauscher, B. J., Bennett, C., et al. 2020, Journal of Astronomical Telescopes, Instruments, and Systems, 6, 046001

Pirzkal, N., & Ryan, R. 2017, Variable He I emission in grism data, Space Telescope Instrument Science Report WFC3 2017–05

Rauscher, B. J., Fox, O., Ferruit, P., et al. 2007, PASP, 119, 768

Regan, M. 2007, Optimum weighting of up-the-ramp readouts and how to handle cosmic rays., Technical Report JWST-STScI-001212

Robberto, M. 2007, Analysis of the sampling schemes for WFC3-IR, Space Telescope Instrument Science Report WFC3–2007–12

—. 2009, Derivation of the correct noise equation for general MULTIACCUM readout, Technical Report JWST-STScI-001853

Robberto, M. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, ed. J. Oschmann, Jacobus M., M. Clampin, G. G. Fazio, & H. A. MacEwen, 91433Z