Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing scholarworks-group@umbc.edu and telling us what having access to this work means to you and why it's important to you. Thank you.

Principal component analysis of up-the-ramp sampled infrared array data

Bernard J. Rauscher
Richard G. Arendt
Dale J. Fixsen
Alexander Kutyrev
Gregory Mosby
Samuel H. Moseley

Principal component analysis of up-the-ramp sampled infrared array data

Bernard J. Rauscher,[a,*] Richard G. Arendt,[b] Dale J. Fixsen,[c] Alexander Kutyrev,[c] Gregory Mosby,[a] and Samuel H. Moseley[a]

[a]NASA Goddard Space Flight Center, Observational Cosmology Laboratory, Greenbelt, Maryland, United States
[b]CRESST II/UMBC/GSFC, Greenbelt, Maryland, United States
[c]CRESST/UMd/GSFC, Greenbelt, Maryland, United States

Abstract. We describe the results of principal component analysis (PCA) of up-the-ramp sampled infrared (IR) array data from the Hubble Space Telescope wide field camera 3 (WFC3 IR), James Webb Space Telescope NIRSpec, and prototype Wide Field Infrared Survey Telescope's wide field instrument detectors. These systems use, respectively, Teledyne H1R, H2RG, and H4RG-10 near-IR detector arrays with a variety of IR array controllers. The PCA shows that the Legendre polynomials approximate the principal components of these systems (i.e., they roughly diagonalize the covariance matrix). In contrast to the monomial basis that is widely used for polynomial fitting and linearization today, the Legendre polynomials are an orthonormal basis. They provide a quantifiable, compact, and (nearly) linearly uncorrelated representation of the information content of the data. By fitting a few Legendre polynomials, nearly all of the meaningful information in representative WFC3 astronomical datacubes can be condensed from 15 up-the-ramp samples down to 6 compressible Legendre coefficients per pixel. The higher order coefficients contain time domain information that is lost when one projects up-the-ramp sampled datacubes onto two-dimensional images by fitting a straight line, even if the data are linearized before fitting the line. Going forward, we believe that this time domain information is potentially important for disentangling the various nonlinearities that can affect IR array observations, i.e., inherent pixel nonlinearity, persistence, burn in, brighter-fatter effect, (potentially) nonlinear interpixel capacitance, and perhaps others. © *2019 Society of Photo-Optical Instrumentation Engineers (SPIE)* [DOI: 10.1117/1.JATIS.5.2.028001]

Keywords: methods; statistical instrumentation; detectors.

Paper 18025 received Apr. 23, 2018; accepted for publication Feb. 4, 2019; published online Apr. 9, 2019.

1 Introduction

The Wide Field Infrared Survey Telescope's (WFIRST)[1] cosmology and microlensing surveys require excellent control of detector systematics. For example, the WFIRST High Galactic Latitude Weak Lensing Survey requires knowledge of the size and ellipticity of the point-spread function (PSF) to <0.1% to avoid biasing measurements of dark energy properties and other cosmological parameters.[2] However, the PSF's shape is measured using reference stars that are typically much brighter than weakly lensed field galaxies, necessitating large linearity corrections.

When large linearity corrections are required, nonideal instrument signatures—including inherent pixel nonlinearity, the brighter-fatter effect (caused by charge integrating in neighboring pixels as pixels fill up),[3] residual persistence from prior exposures, (potentially) nonlinear interpixel capacitance,[4] and telescope pointing jitter (this list is almost certainly incomplete)—have the potential to compromise PSF knowledge and thereby WFIRST science. Within the WFIRST IR Detector Working Group, understanding nonlinear pixel response is particularly a high priority. We therefore began our study by trying to better understand modern linearity correction techniques and the linearity properties of prototype WFIRST infrared (IR) arrays.

Today's IR calibration pipelines for missions—including the Hubble Space Telescope's (HST) wide field camera 3 IR (WFC3)[5–8] and the James Webb Space Telescope (JWST)[9]—linearize the up-the-ramp sampled data from each pixel before fitting a straight line to infer brightness from the fitted slope. (For a description of up-the-ramp sampling, please see Sec. 9. The WFC3 IR calibration pipeline is described in Sec. 3.3 of The WFC3 Data Handbook.[10] A similar pipeline is in development for JWST and is planned for WFIRST.) The planned pipeline calibration sequence for JWST is typical. It includes: (1) bias correction, (2) reference pixel correction, (3) linearization, (4) dark subtraction, (5) cosmic ray detection, and (6) slope fitting. Linearization is based on fitting a low degree polynomial to the up-the-ramp samples in calibration flats. As described by Vacca et al.[11] and Hilbert,[12] the resulting polynomial fit coefficients are used to linearize astronomical exposures before fitting a straight line to each pixel to infer the flux. Typically, linearization assumes that the charge integration rate at the beginning of the exposure is the "true" one and uses the calibration polynomial fit coefficients to make multiplicative corrections to later samples for which nonlinearity is significant.

Fitting a "deg" degree polynomial to n up-the-ramp samples projects the data from the Cartesian n-space in which they were acquired into a monomial space of deg+1 dimensions (the monomial basis vectors are $\{z^0, z^1, z^2, \ldots, z^{\text{deg}}\}$, where z is the time index for equally spaced samples). The co-ordinates in monomial space, the polynomial fit coefficients, provide a representation of a pixel's response that is optimal in a least squares sense to the specified fit degree. However, the

monomials are not an orthogonal basis, and they do not offer insight into how high the fit degree needs to be. Moreover, as shown in Sec. 2.5, there often exist significant correlations between the monomial fit coefficients. For these reasons and others, we decided to use principal component analysis (PCA) to see if there might exist a better basis for modeling up-the-ramp sampled IR array data than the monomials.

The input data were provided by the NASA Goddard Space Flight Center Detector Characterization Laboratory as part of the WFIRST project. WFIRST's wide field instrument (WFI) uses 18 Teledyne H4RG-10 near-IR detector arrays. The 4K × 4K pixel H4RG-10[13] is the most recent member of Teledyne's HxRG family of HgCdTe near-IR detector arrays. The underlying HxRG architecture is an outgrowth of the 2K × 2K pixel H2RG[14] that was introduced in about 2003 for JWST. The "H" in HxRG stands for Hawaii, $x \in \{1, 2, 4\}$ refers to the number of kilopixels in the vertical and horizontal directions, "R" indicates the presence of reference pixels, and "G" indicates the availability of guide mode. Compared to the earlier 1K × 1K pixel H1R detector that was used by WFC3, the H2RG and H4RG include a built-in guide mode. Two versions of the H4RG are available. The H4RG-10 has 10 μm pixels and is used by WFIRST because mass and volume are at a premium. The physically larger H4RG-15[15] offers 15 μm pixels. It may be better matched to the optics of large, ground-based telescopes.

At first, we studied laboratory flatfield data from two different WFIRST H4RG-10s controlled by third generation "Leach" controllers from Astronomical Research Cameras, Inc., and both gave identical results. The PCA immediately suggested a dramatically better basis than the monomials: the Legendre polynomials. If z is an index that runs over frame number and $s(z)$ is integrated signal, then today's monomial approach fits as follows:

$$s(z) = \sum_{i=0}^{\deg} a_i z^i, \tag{1}$$

where the a_i terms are the fit coefficients. Typically, deg = 3 is used for linearization while deg = 1 is used for slope fitting (see Ref. 12 and references therein). When fitting Legendre polynomials, z is uniformly mapped to the interval $-1 \le x \le +1$, and signal integrates according to

$$s(x) = \sum_{i=0}^{\deg} \lambda_i P_i(x). \tag{2}$$

In this expression, λ_i are the fit coefficients and the Legendre basis vectors are as follows:

$$\{P_0(x), P_1(x), P_2(x), \ldots, P_{\deg}\}.$$

Unlike the monomials, the Legendre polynomials are an orthonormal basis. Moreover, because the Legendre polynomials approximately diagonalize the covariance matrix, they provide a representation that is significantly less correlated than the monomials. Finally, because the covariance matrix is roughly diagonal in Legendre space, the Legendre fit coefficients provide a useful way to quantify how the information content falls off by fit degree.

Although we began the PCA using WFIRST H4RG-10 laboratory flats, we have since shown that the results are general,

applying equally to other systems including the JWST near-IR spectrograph and HST WFC3. The findings do not depend on the type of IR detector or the IR array controller (H1R+Ball electronics; H2RG+SIDECAR ASIC; H4RG-10 + Gen. III Leach controller). All systems have given similar results. These are the findings.

1. The Legendre polynomials generally provide a good basis for the time dimension of up-the-ramp sampled IR array data. The monomial basis that is used in today's pipelines is a comparatively poor choice because the basis vectors are not orthogonal and monomial fit coefficients are correlated.

2. Up-the-ramp sampled IR array data contain time domain information that is not utilized by fitting a straight line to the linearized up-the-ramp samples. As a corollary, for archival research, it is important to save more information than is contained in only the fitted bias and slope images.

3. Legendre fits are highly compressible. If it is not possible or practical to save all samples, then Legendre fitting offers a simple way to significantly reduce the data volume while retaining nearly all of the scientific information.

Legendre fitting naturally produces datacubes rather than two-dimensional (2-D) images. Two kinds of datacubes are commonly encountered in near-IR astronomy today. Often, the unprocessed up-the-ramp samples are saved in a three-dimensional (3-D) array, with the axes being time (or equivalently time index for equally space samples) and the two angular dimensions on the sky, (t, α, δ). Another common astronomy datacube has wavelength running along the zeroth dimension. Here, we introduce the idea of a "Legendre cube." A Legendre cube differs from a time-ordered 3-D array in that the zeroth dimension contains the fit coefficients, λ_i, for the Legendre polynomial $P_i(x)$. The axes of a Legendre cube are (i, α, δ).

This article is structured as follows. In Sec. 2, we explain the PCA in the context of WFC3's frontier field observations of the Abell 370 strong lensing field. We selected Abell 370 because the field is information rich, and it contains a wide range of source brightness ranging from sky background to hard saturation. This is followed by Sec. 4, where we explore how astrophysical information manifests in Legendre cubes. The higher Legendre orders are information poor, leading to Sec. 5, where we discuss data compression in Legendre space. Finally, in Sec. 6, we begin to describe how these findings can potentially be applied to astrophysical observations. This section is necessarily incomplete as we are only beginning to work with Legendre cubes today. We close with a summary.

2 Principal Component Analysis

The purpose of this section is to describe the PCA in the context of real WFC3 observations. We previously obtained similar results with WFIRST and JWST lab data (including both flat field images and unilluminated darks). We begin with a short introduction to the WFC3 data. This is followed in Sec. 2.2 by a PCA refresher, and then by the PCA itself.

2.1 Abell 370 Frontier Field

The Frontier fields were an HST Director's Discretionary Program that aimed to exploit the amplification of light by strong gravitational lensing to image faint, high-z galaxy populations.[16] We selected one of the Frontier fields, Abell 370 (Fig. 1), to test the ideas that we had conceived earlier using WFIRST lab data.

The data were acquired between August and September, 2016, as part of proposal ID 14038 (J. Lotz PI). We selected all available 16 frame SPARS100 exposures taken with the F160W filter ($\lambda_{\text{peak}} = 1.545 \ \mu$m and FWHM $= 0.29 \ \mu$m). (We also looked at other filters. The choice of filter does not affect the PCA results.) In SPARS100 mode, WFC3 acquires a "reset frame" followed by 15 uniformly spaced nondestructive samples at ~100 s intervals. Because the reset frame did not follow the same approximately linear trend as later samples, it was discarded (it is also not used for fitting in the WFC3 pipeline). All told, we were able to use 36 EXPTIME ≈ 1403 s dithered SPARS100 exposures, each with 15 up-the-ramp reads, resulting in a total exposure time of about 14 h.

Our focus in this paper is the PCA. For more information about WFC3 readout modes and data products, the interested reader is referred to The WFC3 Instrument Handbook[8] and The WFC3 Data Handbook.[10]

2.2 PCA Refresher

Up-the-ramp sampled data are typically stored in datacubes or in the case of WFC3 as FITS files with one extension per frame. For these WFC3 data, the unprocessed files unpacked to $15 \times 1024 \times 1024$ datacubes after discarding the reset frames. Although time-ordered datacubes are an intuitively obvious way to store data, they are neither the most compact nor the most useful way to represent the information. PCA is a
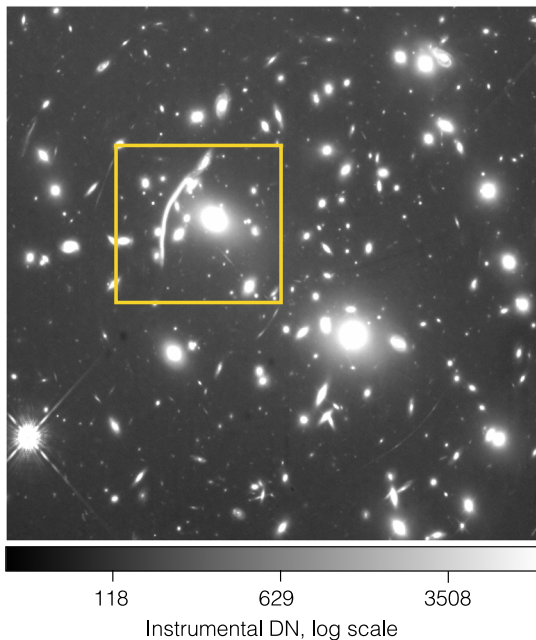
mathematical tool for finding a more compact representation that reduces the number of variables that are needed to represent the dataset while retaining much of the information.[17] Throughout this paper, we use lowercase boldface letters to represent vectors and uppercase boldface letters to represent matrices.

PCA is built on the concept of the covariance matrix, Ω, which is a generalization of the variance that allows for the possibility of correlations. Consider a univariate statistical process, **d**. For present purposes, **d** is a vector containing 15 up-the-ramp samples. If the expectation value of **d** is known to be $\langle \mathbf{d} \rangle$, then the residuals for any one realization can be represented by a column vector:

$$\delta = \mathbf{d} - \langle \mathbf{d} \rangle. \tag{3}$$

If the experiment is run n times (or we have n pixels), we can put all of the different realizations of δ into a matrix, Δ. The covariance matrix is then defined as follows:

$$\Omega = \frac{1}{n} \Delta \Delta^{\text{T}}. \tag{4}$$

Ω is by definition square, symmetric, and positive definite. If there is any correlation in the data (there is with up-the-ramp sampled detector data), then Ω will have off-diagonal elements. In PCA, we seek a basis in which the covariance matrix is diagonal.

Since Ω is square, by the Eigen decomposition theorem, it can be factored as follows:

$$\Omega = \mathbf{VWV}^{-1}. \tag{5}$$

In this expression, **W** is a diagonal matrix containing the positive eigenvalues of Ω. **V** is a square matrix containing the corresponding eigenvectors (one eigenvector per column). Because Ω is symmetric, the eigenvectors form an orthogonal basis. In the PCA lexicon, each orthogonal basis vector corresponds to a "component."

Because each column of **V** is a basis vector, applying $\mathbf{V}^{\text{T}}$ to the data, $\mathbf{d}' = \mathbf{V}^{\text{T}} \mathbf{d}$, accomplishes a change of basis yielding, $\mathbf{d}'$, which contains the same information but with orthogonal deviates (i.e., the covariance matrix is diagonal when computed in this basis).

This is a more convenient representation because each component can now be considered independently. It can save computation (ignoring the computation to convert it into this form) since weighting can be done component by component instead of multiplying by a matrix. The data can also be stored more compactly. This is because each component only needs to be stored with enough precision to encompass the noise of that component. Moreover, in many physical systems, most of the information lies in the first few components.

118   629   3508

Instrumental DN, log scale

Fig. 1 This image is the median λ_1 parameter (slope) from stacking the available WFC3 F160W data. The integrated exposure time is about 14 h. Other than reference correction (Sec. 9) and fitting Legendre polynomials, the image is uncalibrated. The yellow box is the region of interest (RoI) shown in Fig. 5.

2.3 PCA Method

We used nearly all of the pixels shown in Fig. 1 (about 98.5%) for the PCA. The ~1.5% that were discarded were found by doing a simple reference correction and preliminary fifth degree Legendre fit. We then rejected high rms pixels based on a histogram. In practice, this is very loose quality control. Nevertheless, the PCA gave consistent results in spite of there being cosmic ray hits and other artifacts in the data.

After discarding the reset frame, reference correction, cropping off reference pixels, let the data from the i'th of $n \approx 1014^2$ pixels be represented by a column vector:

$$\mathbf{d}_i = \begin{pmatrix} d_{0\ i} \\ \vdots \\ d_{14\ i} \end{pmatrix}. \tag{6}$$

By flattening the two spatial dimensions on the sky, the data can be represented by a matrix as follows:

$$\mathbf{D} = \begin{pmatrix} d_{0\ 0} & \cdots & d_{0\ n-1} \\ \vdots & \cdots & \vdots \\ d_{14\ 0} & \cdots & d_{14\ n-1} \end{pmatrix}. \tag{7}$$

Each column in $\mathbf{D}$ contains the up-the-ramp samples from one pixel.

Define the matrix, $\langle \mathbf{D} \rangle$, to be a matrix that has the same shape as $\mathbf{D}$ but with every column equal to a column vector that is the expectation value of $\mathbf{D}$ averaged over all columns. All columns of $\langle \mathbf{D} \rangle$ are therefore identical. The covariance matrix is as follows:

$$\Omega = \frac{1}{n}(\mathbf{D} - \langle \mathbf{D} \rangle)(\mathbf{D} - \langle \mathbf{D} \rangle)^{\mathrm{T}}. \tag{8}$$

To complete the PCA, following Sec. 2.2, we computed Ω's eigensystem. Figure 2(a) shows the eigenvectors $\mathbf{v}$. By inspection, the eigenvectors are similar to the Legendre polynomials [Fig. 2(b)]. For clarity, we show only the first few here, although the striking similarity continues as more are plotted. PCA of WFIRST and JWST flats (Sec. 10) produced similar results. However, we find this demonstration using real astronomical data to be particularly compelling. Moreover, the result is not particular to Abell 370. We did the same analysis for WFC3 observations of the 47 Tucanae globular cluster and got the same results.

Figure 3 shows the eigenvalues, w_i, plotted by PCA component. They provide a quantitative measure of the variance in each component. Throughout this paper, we equate variance with "information." Not all information is scientifically useful because some is just noise. We define the "meaningful information" to be the remaining variance after subtracting off the noise.
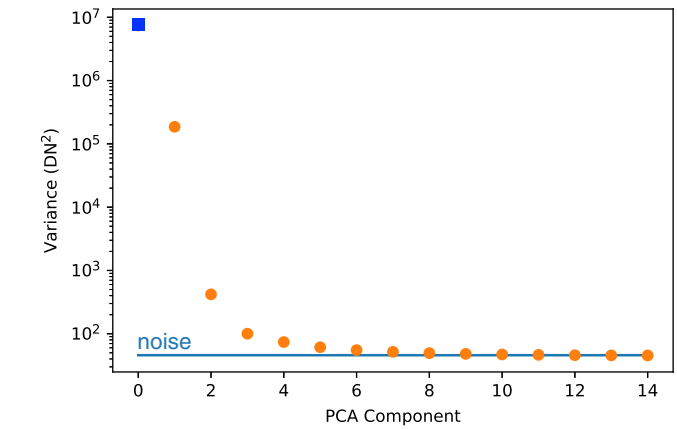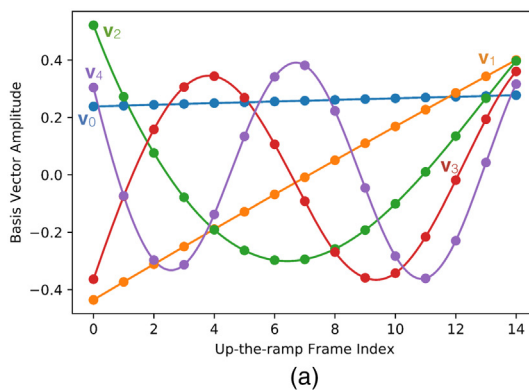

Fig. 3 This plot of eigenvalue versus PCA component (λ index) shows that essentially all of the meaningful information is contained in $\lambda_0 - \lambda_5$. The symbol for λ_0 is different because it is dominated by the instrument signature (detector bias pattern and hot pixels etc.). Excluding $>3.5\sigma$ statistical outliers, the noise is essentially Gauss-normal for λ_6 and higher (Sec. 11). For reference, WFC3's conversion gain is about $g_c \sim 2.25\ e^- \mathrm{DN}^{-1}$. According to the WFC3 Instrument Handbook, the read noise is between $20.2 - 21.4\ e^-$ per correlated double sample.[8] The read noise per sample is therefore about 15 e^-, which corresponds to the variance of the blue noise line that is overlaid on the plot.

Consistent with this plot, a variety of statistical tests (Sec. 11) show that the noise is essentially Gauss-normal for λ_6 and higher once $>3.5\sigma$ outliers are excluded. Later, in Sec. 4, we will discuss images (Fig. 5) of information content by PCA component. For these observations, the evidence is consistent with nearly all of the scientific information being contained in $\lambda_0 - \lambda_5$.

2.4 Why Legendre Polynomials?

It is reasonable to ask, is there a physical explanation for why the Legendre polynomials emerge so clearly? If one views the problem of fitting the up-the-ramp samples as both a physics problem and as a linear algebra problem, then one can plausibly argue that the first few Legendre polynomials ought to be a good basis for modeling a pixel's response to light.

Whenever a pixel is reset, there is a statistical uncertainty in the number of charges on the pixel's capacitance equal to,
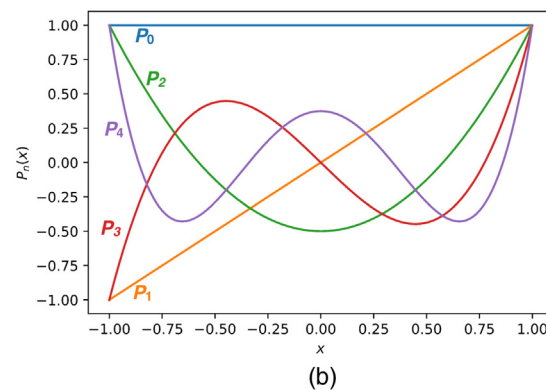

Fig. 2 Panel (a) shows the first five eigenvectors computed from the WFC3 Abell 370 data. They are very similar to the (b) Legendre polynomials. For comparison purposes, we show the actual Legendre polynomials in (b). If we were to change the normalization to match (a), then the two plots would look even more similar.

$\sigma_{\mathrm{kTC}} = q_e^{-1}\sqrt{k_B T C}$.[18] In up-the-ramp sampled data, this "kTC noise" appears as a constant offset affecting all samples. In a linear model parameterized by up-the-ramp frame index, z, kTC noise appears as a constant time z^0. After the pixel is reset, by design, the ideal pixel starts to integrate the charge in a highly linear manner. We therefore expect a term that looks like a constant time z^1. As the charge integrates, we know from looking at data that most pixels gradually roll over and lose response in a manner that can be roughly approximated with a quadratic, i.e., a constant times z^2, before steepening upon entering harder saturation as a constant times z^3 (and so on). We can use these observations as the starting point for a linear model of the pixel's response that includes the basis set, $B = \{z^0, z^1, z^2, z^3\}$.

Without loss of generality, we can map z to the interval $-1 \leq x \leq +1$, and furthermore, define an inner product for this interval:

$$\langle a, b \rangle = \int_{-1}^{+1} a(x)b(x)\mathrm{d}x. \tag{9}$$

If one uses Eq. (9) to orthogonalize B, the result is the first four Legendre polynomials, $B' = \{P_0(t), P_1(t), P_2(t), P_3(t)\}$.[19] One can extend this argument to a higher degree, although the justification for the higher order terms becomes less physical the higher one goes.

Viewed in this way, the Legendre polynomials emerge as a natural orthogonal basis for pixels that have kTC noise, respond to light in a roughly linear matter at low signal levels, and for which the response gradually rolls off before steepening upon entering saturation. Empirically, it so happens that they also approximately diagonalize the covariance matrix.

2.5 *Advantages of Legendre Polynomials*

One advantage of the Legendre polynomials compared to the monomials is that they provide a less correlated representation of the data. To quantify this, we looked at the Pearson correlation matrix, $\mathbf{P}$ (also known as Pearson's r; see Ref. 20, Sec. 14.5). To compute $\mathbf{P}$, define the matrix Δ' that has the same shape as $\mathbf{D}$. The i'th column of Δ' is equal to the column vector:

$$\delta_i' = \begin{pmatrix} \frac{d_{0i} - \langle d_0 \rangle}{\mathrm{std}\, d_0} \\ \vdots \\ \frac{d_{14i} - \langle d_{14} \rangle}{\mathrm{std}\, d_{14}} \end{pmatrix}, \tag{10}$$

where std is the usual standard deviation. With these definitions, the Pearson correlation matrix becomes

$$\mathbf{P} = \frac{1}{n}\Delta'\Delta'^{\mathrm{T}}. \tag{11}$$

To compare the two representations, we fitted $\mathbf{D}$ to fifth degree using Legendre polynomials and the monomials. The nearly diagonal correlation matrix of the Legendre polynomials [Fig. 4(a)] is preferable to the checkerboard of the monomials [Fig. 4(b)].

There are at least two advantages to a diagonal covariance matrix. One advantage is that in order to do calculations, one needs to only know and deal with six components (the diagonals) instead of 21 components (including the off diagonals) for
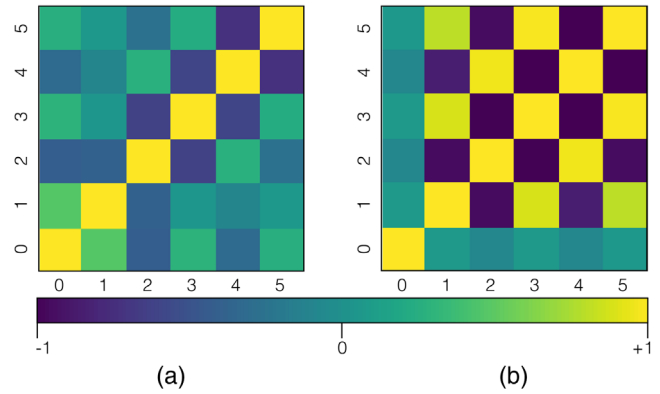

Fig. 4 We fitted the same Abell 370 data to fifth degree (six free parameters) using (a) Legendre polynomials and (b) monomials and computed the Pearson correlation matrices. The Legendre representation is strikingly less correlated than the monomials. As described in the text, the nearly diagonal correlation matrix of the Legendre polynomials is preferable to the checkerboard that the monomials produced.

these observations. This reduces the number of parameters that one needs to know a priori and saves computer space and computer time. The other, arguably more important, advantage is that inverting a large matrix with large off-diagonal components leads to larger errors. For monomials, this leads to manifestly wrong answers with single precision arithmetic at deg = 6 and even with double precision at deg = 11.

3 When Might One do Better?

The PCA shows that the Legendre polynomials are generally a good choice for modeling the response of sampled up-the-ramp pixels to astronomical scenes like Abell 370, the globular cluster 47 Tucanae (not discussed here), laboratory flats, and darks (when sufficient data are available). As discussed in Sec. 2.4, one would expect the Legendre polynomials to emerge as a natural basis in many astronomical situations.

However, in some special cases, it might be possible to characterize and measure the actual eigenvectors. For example, this might be true for some transiting exoplanet observations. In these cases, the Legendre polynomials would not be a bad choice. However, one might expect to do even better by measuring and using the actual eigenvectors when it is practical to do so.

4 Information Content by Legendre λ

For Abell 370, Fig. 3 shows that a fifth degree Legendre fit extracts nearly all of the meaningful information. The only pixels that experience significant information loss are those that strongly saturate, are hit by cosmic rays, or are statistical outliers in some other way. For these rare events, techniques like PCA that rely on ensemble averages are not effective.

Figure 5 provides a visual impression of how information content falls off with increasing i in λ_i. This figure was made by computing the median of all 36 Legendre cubes after offsetting them to account for dithering. The offsets were computed by cross correlating the λ_1 (slope) coefficients. The integrated exposure time on the sky in each panel is about 14 h.

λ_0 is the mean value of the up-the-ramp samples. It contains much of the instrument signature (bias pattern, hot pixels, bad pixels, etc.), although there is also astrophysical information.
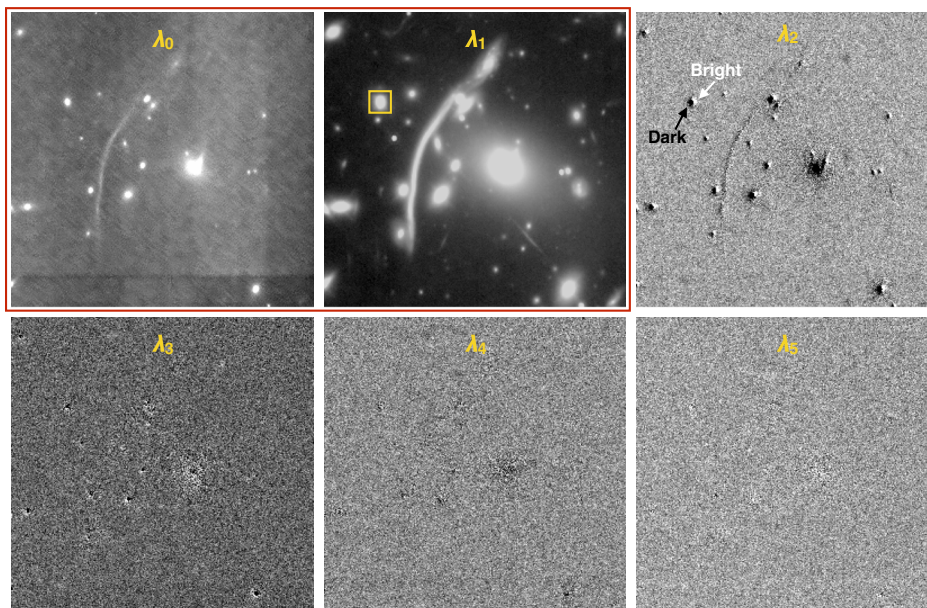
Fig. 5 This figure shows the six Legendre λ coefficients for the RoI of Fig. 1. The red box highlights the information that is used by the WFC3 calibration pipeline. After linearization, the pipeline fits a straight line (λ_0 and λ_1), thereby not utilizing the information that is contained in $\lambda_2 - \lambda_5$ to constrain model parameters. Consistent with Fig. 2(b), λ_5 and higher contain very little astronomical information. For comparison, we fitted the same data with a two-parameter straight line. The yellow box is a photometer aperture. For this source, there is a 0.9% difference in brightness between the fifth degree Legendre fit and the straight line. The images show detector edges from stacking (especially λ_0). The bright-dark artifacts seen especially in λ_2 are more interesting. These may be caused by ≈ 1.5 mas guiding errors (1% of WFC3's pixel pitch) during each exposure. This is discussed more fully in Sec. 12.

Where there is a bright source, the mean value of the ramp is higher, so the sources are visible in λ_0.

λ_1 is proportional to the conventional slope image (the conversion factor is given in Sec. 6). As expected, it is strongly dominated by astronomical sources. As in conventional pipelines, λ_1 can be used directly for noncritical measurements.

λ_2 is the first term to capture curvature [see Fig. 2(b)]. For nominal pixels that slowly lose response as they fill until entering saturation, it should always be negative when it exceeds the noise. λ_2 strongly reflects nonlinearity. We anticipate that it will be extremely sensitive to anything that can impart curvature to the ramp. Some examples include intrinsic nonlinearity for bright sources, nonlinearity in the readout integrated circuit (ROIC) when/if bright sources perturb the ROIC's electrical state, the brighter-fatter effect, and pointing jitter and drifts.

The brighter-fatter effect can make a "bull's-eye" pattern around each bright star, with a dark central hole surrounded by a bright ring in the λ_2 image. The effect of drifting in/out from pixels would appear like the bright-dark artifacts, as shown in Fig. 5. Here, λ_2 seems to have been sensitive to 1.5 mas drifts (1% of pixel pitch). For more information on the bright-dark pattern, please see Sec. 12.

The higher order λs contain diminishing information. But, as can be seen in Fig. 5, information is still present in these Abell 370 data until at least λ_5. For disentangling time-dependent instrument signatures, we believe that the time domain information contained in all six terms may be important.

5 Information Compression

Legendre cubes are highly compressible. Looking at the Abell 370 data, consisting of the reset frame and 15 up-the-ramp

samples, a complete data dump would be $16 \times 2 = 32$ bytes/pixel. However, essentially all of the information can be retained in six Legendre coefficients. If these are stored as 4 byte floats, this is 24 bytes per pixel, a modest 25% compression. Compressing a typical astronomical data file with "gzip −best" results in 26.7 bytes/pixel, a 17% compression.

However, floating point numbers are unnecessary and difficult to compress. Keeping 8 digits of precision on a number that has an uncertainty at the second digit is clearly excessive. The ancient instruction of keeping a single digit into the uncertainty is good advice. To update this to binary, one should keep 2 or 3 bits into the noise. One way to do that is to convert the Legendre coefficients from floating point to integer (or calculate them that way to begin with), and then multiply them with a prearranged constant to make the 2 or 3 least significant bits be in the band of uncertainty. For astronomical data dominated by Poisson noise, the uncertainty is proportional to the square root of the signal. Taking the square root halves the number of bits required even before compression. If, as often the case, there is additional readout uncertainty, the proper treatment is to add a constant before taking the square root (see Ref. 21, Fig. 1).

Here, the first two coefficients, λ_0 and λ_1, are offset to make them positive and the square root results in data that can fit into 12 bits each per pixel, including three bits of noise. The λ_2 coefficient fits into 10 bits and the remaining coefficients fit into 9 bits each including three bits of noise. Thus, a full ramp can be packed into 60 bits or 7.5 bytes, a compression of 75%, or a factor of 4. This file can furthermore be gzipped, resulting in a compression ratio of approximately $5 - 6\times$ relative to the starting data while retaining nearly all of the scientific information.

Table 1 Compression trades.

	Data in various forms			
Data form	Item size	Number	Byte per pixel	Compression
Original data	2 bytes	16 sample	32	—
Legendre Trans.	4 bytes	6 Coefs	24	25%
Compress to int	10 bits avg	6	7.5	77%
Zipped file			5.6	83%
Keep 2 bits of noise			4.6	87%

The final gzip compression of about 25% is the best that can be expected since much of the "data" are the three bits of noise in each coefficient. If only two bits of noise are kept, the data compresses to 4.6 bytes/pixel, for a total compression ratio of 7. Higher compression ratios can be expected for longer ramps. This is because even a 64 sample ramp on the WFIRST H4RG-10 requires only ≈ 8 Legendre coefficients, and these will require only a few extra bits for the coefficients. Table 1 provides a summary of the compression that can be achieved after the different steps described above.

6 Astrophysics in Legendre Space

For many years, the standard practice in IR astronomy has been to linearize up-the-ramp sampled data and then fit straight lines to make a slope image.[11,12] Although the details of the implementations differ, the focus has been on collapsing 3-D datacubes to 2-D images. The 2-D image, or more commonly averaged 2-D images, has almost always been the unit of data that is used for science analysis.

The PCA clearly shows that this approach does not optimally use the time domain information that is contained in the higher order terms. Using calibration files to linearize the data before fitting a straight line seeks to calibrate this information out, rather than use it to constrain model parameters. To the extent that the linearity correction is always imperfect, some time domain information is always left behind. To utilize all of the meaningful information (both astrophysical information and instrument signatures), one must retain more than just 2-D slope images. In this regard, Legendre cubes are helpful because they approximate the eigenvectors.

As of today, however, we are only just starting to learn how to work with Legendre cubes. As place holders for proper algorithms, we have found the following to be handy although imperfect tools for collapsing 3-D Legendre cubes to familiar 2-D images for debugging and comparison with legacy data products. Going forward, our aims include developing techniques for performing common analysis algorithms directly on Legendre cubes. We think that this will probably rely more on forward modeling than is common today.

6.1 Slope Estimators: Two Handy but Imperfect Tools

If a 2-D image is desired, then λ_1 is proportional to the conventional slope image. If there are n up-the-ramp samples and the frame readout time is t_f s, then the slope in units of data numbers (DN) per second is as follows:

$$\text{Slope} = \frac{2\lambda_1}{(n-1)} t_f^{-1}. \tag{12}$$

If used as the only flux estimator, Eq. (12) has the undesirable attribute that it does not optimally use the time domain information that is contained in the higher order terms.

Another handy way to get a 2-D image that captures more of the information is to compute the integrated signal in DN using the Legendre fit:

$$\text{integrated signal} = \sum_{i=0}^{\text{deg}} \lambda_i (P_i(+1) - P_i(-1)). \tag{13}$$

Only the odd numbered terms contribute to the sum in Eq. (13) because the even numbered Legendre polynomials are even functions. As such, some information is not optimally used. Compared to Eq. (12), Eq. (13) has the advantage that it utilizes more of the available information and it is not much more difficult to compute. Equation (13) can straightforwardly be converted to the mean photocurrent during the exposure (units: DN per second):

$$\text{mean photocurrent} = \sum_{i=0}^{\text{deg}} \frac{\lambda_i (P_i(+1) - P_i(-1))}{(n-1)} t_f^{-1}. \tag{14}$$

These image estimators are far from perfect. The first problem is that they collapse the 3-D Legendre cube to a 2-D image before performing any scientific analysis. In doing this, resolution on the time domain information is lost. However, the time domain may hold the keys to disentangling source flux from persistence, burn in, and other detector artifacts mediated by charge trapping and release.

A second well-known problem is that the detectors are inherently nonlinear. Nonlinearity must be accounted for when mapping instrumental brightness in DN to astronomical source brightness. As has already been discussed, today's pipelines use calibration data to correct for nonlinearity before slope fitting (or vice versa). These operations generally do not commute: that is to say that if slope fitting is done before linearization, one gets a slightly different result. Our eventual aim is to account for nonlinearity simultaneously with measuring source brightness as part of fitting (or forward modeling) a multiparameter model to the pixel's response as recorded in the Legendre cube. However, this is work for the future.

6.2 Plans to Study Linearity

Given the lack of a satisfactory network of faint standard stars spanning a range of brightness and color, we understand that we must learn how to model nonlinearity in Legendre cubes. To address this, we are pursuing two parallel tracks.

One track is purely experimental. We are upgrading a laboratory test setup at Goddard so that we can apply light of known wavelength and relative brightness over a wide, $10^{4-5}\times$ dynamic range. The aim is to use this to study the linearity properties of Legendre fitted pixels. If an opportunity were to come up to fly a similarly calibrated detector in a very well understood telescope, then that would be invaluable in establishing a network of faint astronomical calibrators with known relative brightnesses.

The second track combines theory and experiment. By experimentally studying the properties of individual pixels and building physical models of them, we hope that it may be possible to apply valid linearity corrections using less calibration data. However, work is just beginning on the theoretical modeling. The test setup described above is still under construction. We look forward to saying more about these topics as things mature.

7 Future Work

Our hypothesis going forward is that for WFIRST's most demanding surveys, it may be beneficial to replace linearization followed by line fitting with fitting (or forward modeling), a multiparameter model to Legendre cubes. Linearization would notionally be part of this process, as would measuring source brightness, and accounting for artifacts such as persistence.

For space, we must also understand how to recognize cosmic rays in Legendre cubes. Work on this is just beginning. However, it may be that the Legendre cubes themselves contain sufficient information to differentiate cosmic ray hits from normal pixel response. In other words, cosmic ray hits may leave a distinct signature in Legendre space that differs from normal pixel response. This is something that we can begin to explore with existing laboratory data from JWST and WFIRST and flight data from WFC3.

8 Summary

In this paper, we reported the results of PCA of laboratory flats and real astronomical data from a variety of IR instruments. The detectors included WFC3's Teledyne H1R, JWST's H2RGs, and WFIRST H4RG-10s. The IR array controllers included WFC3's flight electronics, JWST SIDECAR ASICs, and Gen-III Leach controllers from Astronomical Research Cameras, Inc., for WFIRST. In all cases, the Legendre polynomials emerged as a better basis (in the linear algebra sense) for modeling up-the-ramp sampled pixels than the monomial basis that is used for polynomial fitting and linearization in today's calibration pipelines. [Some colleagues have suggested that the Fourier basis might be a good choice (after projecting out the offset and slope terms). While preparing this manuscript, we tried doing this. Compared to the Legendre polynomials, the Fourier basis is also a poor choice.] The Legendre polynomials are an orthogonal basis, whereas the monomials are not. Moreover, the PCA shows that the Legendre polynomials are often a reasonable proxy for the eigenvectors that diagonalize the covariance matrix, whereas the monomials are not. As such, the Legendre polynomials provide a less correlated representation of the data than the conventional monomials.

Building on the PCA, in Sec. 5, we described how the information content in astronomical scenes falls off with fit degree. For the HST WFC3 Frontier Field observations of Abell 370, we showed that nearly all of the meaningful information could be compressed from 15 up-the-ramp samples down to 6 Legendre coefficients per pixel. Moreover, Legendre coefficients are compressible. In Sec. 5, we describe concepts for compressing data like these down to 14% of the raw data volume while retaining nearly all of the meaningful information.

Looking to the future, we believe that a change of basis from the monomials to the Legendre polynomials should benefit not only compression but also IR array calibration. The PCA shows that up-the-ramp sampled data contain more information than is contained in the conventional offset and slope images. The additional information exists in the time domain, and we believe that

it may be important for disentangling time-dependent artifacts like inherent nonlinearity, brighter-fatter effect, persistence, and pointing drifts, etc. Linearizing the up-the-ramp samples before fitting slopes, as is done today, does not use the time domain information to constrain model parameters. Rather, it seeks to remove it. Our hypothesis going forward is that it may be possible to do better by directly fitting, or forward modeling, multiparameter models to Legendre cubes.

9 Appendix A: WFC3 H1R Detector, Up-the-Ramp Sampling, and Reference Correction

9.1 *WFC3 H1R Detector*

WFC3 IR uses a Teledyne H1R near-IR detector array [Fig. 6(a)]. The H1R is a 1024×1024 pixel HgCdTe detector array. Pixels are read out in four 512×512 pixel "quadrants." The HgCdTe is tuned to have a cutoff wavelength of about 1.7 μm. The "R" in its name denotes reference pixels.

The H1R was the first astronomical near-IR detector to provide engineered reference pixels embedded in the video outputs. In the H1R, the reference pixels appear in a 5 pixel wide band framing the 1014×1014 photosensitive pixel area [Fig. 6(b)]. Although the reference pixels are designed to electronically mimic a regular pixel, they do not respond to light. During calibration, they are used to remove electronic drifts, as described in Sec. 9.3.

9.2 *IR Array Readout and Up-the-Ramp Sampling*

Unlike in a CCD, it is possible to nondestructively sample the pixels in H1R and HxRG (and similar IR array) detectors many times without resetting them. At first, "multiple nondestructive reads" were used to enable "Fowler sampling," which quickly samples the detector several times at the beginning and end of each integration to "average down" the read noise. Fowler and Gatley[22] still provide a good introduction to Fowler sampling and other uses of multiple nondestructive reads.

Building on these ideas (and briefly mentioned in Fowler and Gatley), it is possible to nondestructively sample a detector uniformly throughout an exposure. Today, this is known as up-the-ramp sampling. Up-the-ramp sampling is widely used in space because cosmic ray disturbance is easily recognized and can sometimes be corrected for. Figure 7 shows an example of up-the-ramp sampling as it is implemented in JWST NIRSpec.

In all IR array readout modes that use multiple nondestructive samples, there are correlations between the samples within an exposure that affect the uncertainty in the measured flux. Garnett and Forrest[24] provide a good introduction to correlated noise trades as a function of readout mode. Vacca et al.[11] provide an updated discussion that includes the effects of detector nonlinearity. Rauscher [Eq. (1)][25] and Rauscher et al.[26] provide an update to Vacca et al.'s noise model for the special case of JWST readout.

9.3 *Reference Correction*

WFC3's H1R detector is read out in four "quadrants" [Fig. 6(a)]. The quadrants can (and do) have DC offsets with respect to each other. The reference pixels are used to take these out in postprocessing. The outermost 5 pixels on all sides are reference pixels. Of these, the outmost rows/columns contain reference pixel design variations. The inner four rows/columns of reference
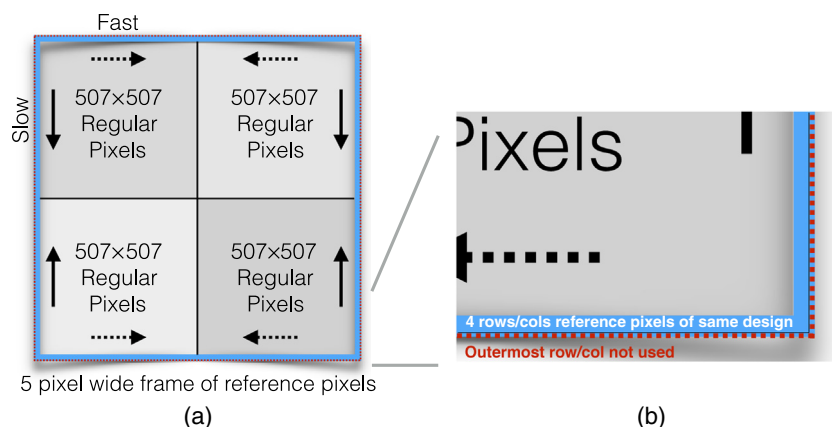
Fig. 6 (a) WFC3's H1R detector is read out in quadrants. The full array has 1024×1024 pixels of which 1014×1014 are regular photosensitive pixels. These are surrounded on all sides by a 5 pixel wide border of reference pixels. (b) The outermost row/column of reference pixels was not used because it contains several different reference pixel designs. The inner four rows/columns were used because these reference pixels are all of the same design.

pixels all use the same design and these are the ones that we used for this study.

We used a very simple reference correction scheme for this study. The reference correction was applied to each frame individually as the first processing step. We treated each quadrant separately and computed the median of all reference pixels in each quadrant excluding the outermost rows/columns. This median was then subtracted from every pixel in that quadrant. The end result was a DC reference correction that was applied frame-by-frame and quadrant-by-quadrant. After making the reference pixel correction, we cropped all reference pixels from each frame.
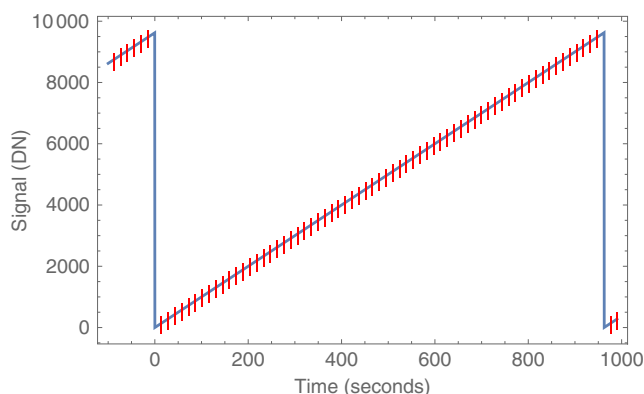
More sophisticated reference correction is possible given the right hardware, clocking patterns, and calibration software. For example, some instruments treat the reference rows and columns separately. With careful tuning, it is sometimes possible to use the reference columns to remove noise within the outputs (not just DC). For JWST NIRspec, we developed improved reference sampling and subtraction (IRS^2; pronounced "IRS-square").[27] IRS^2 uses a specialized clocking pattern to interleave many more reference pixels into the data than is otherwise possible, resulting in cosmetically cleaner images with less correlated noise.

In any case, the conclusions of this paper appear to be robust against changes in how the reference pixels are used. In addition to the very simple DC correction described above, we also did PCA on IRS^2 sampled flats (Sec. 10). In both cases, PCA showed that the Legendre polynomials were a good approximation to the eigensystem.

10 Appendix B: PCA of JWST and WFIRST Flats

The Legendre polynomials are a good choice of basis for both astronomical scenes and flats. Figure 8 shows the eigenvectors
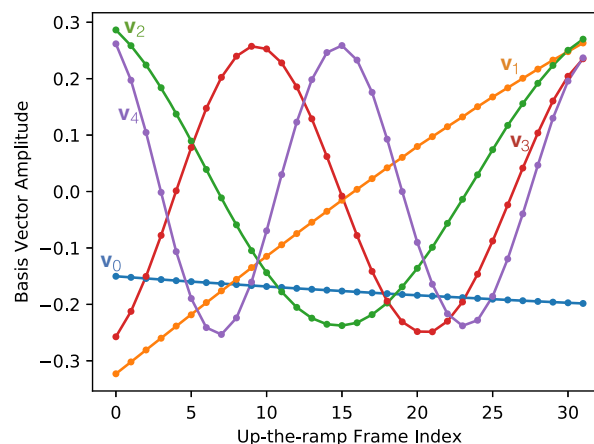

Fig. 7 This figure shows up-the-ramp sampling for one pixel as it is implemented in the JWST NIRSpec. There are 65 nondestructive samples per $EXPTIME = 933$ s exposure. To maintain constant power dissipation, the detector is clocked and pixels are digitized at a constant cadence. For the NIRSpec mode that is shown here, the cadence is about 14.6 s/frame. Charge integration begins upon the completion of reset for the previous exposure ($t = 0$ s). Each vertical red line indicates a sample. Note that some signal has integrated by the time the pixel is digitized for the first time. The intent of this figure is to show the timing from the perspective of one NIRSpec pixel. JWST observers should see the NIRSpec information pages[23] for technical information about actual JWST data products. Although the details differ slightly for WFC3 and WFIRST, the underlying concept of acquiring multiple nondestructive samples is the same.

Fig. 8 The eigenvectors, **v**, of a JWST NIRSpec channel 491 flatfield exposure to nearly full well are very similar to the Legendre polynomials. WFIRST flats show similar behavior.

for a JWST NIRSpec flatfield image exposed to nearly full well (6×10^4 e^-). The procedure was as described in Sec. 2. By inspection, the eigenvectors are similar to the Legendre polynomials [Fig. 2(b)]. We have repeated this analysis for flats taken with WFIRST H4RG-10s and gotten similar results.

11 Appendix C: Read Noise Distribution at High λ_i

If all of the information were captured by $\lambda_0 - \lambda_5$, then the distribution in λ_6 and higher would be Gauss-normal noise. To test this hypothesis, we used three standard statistical tests: (1) the Kolmogorov–Smirnov (KS) test, (2) $Q - Q$ plots, and (3) the Anderson–Darling test. The KS test shows that by λ_6, the noise is statistically indistinguishable from Gauss-normal. The $Q - Q$ plots also showed this to be true, except for statistical outliers. The Anderson–Darling test weights statistical outliers heavily, so it shows that although the distribution is trending toward Gauss-normal, there are nevertheless too many statistical outliers to be well represented by an underlying Gauss-normal distribution. However, if we clip statistical outliers, then the Anderson–Darlington test, like the $Q - Q$ plots and KS test, becomes consistent with an underlying Gauss-normal distribution. In a statistical sense, for these HST data, essentially all of the meaningful information is captured by $\lambda_0 - \lambda_5$ once a very small sample of outliers has been eliminated.

11.1 KS Tests

As the information content (i.e., variance) decreases in the higher order λ_i (Fig. 3), we also expect that the distributions of λ_i values should become increasingly well represented by a Gauss-normal distribution. Thus, a one-sided KS test (see Ref. 20, Sec. 14.3) was applied to compare the λ_i values to a Gaussian function for each of 111 ramps (36 at F160W, the rest using other filters). The KS test calculates the maximum deviation (D) between the cumulative distribution functions for the data and for the Gaussian model. Then, one calculates the probability for finding a value of $\geq D$ under the assumption that the data are drawn from a Gauss-normal distribution. Low probabilities indicate that D is improbably large for data with a truly Gauss-normal distribution. Over many repeated tests of Gauss-normal data (with known mean and σ values), the probabilities should fall in a uniform distribution over the range [0,1]. Because of the detector readout pattern and the reference correction (Sec. 9), the four detector quadrants were tested independently.

Figure 9 shows the KS probabilities for each of 111 λ_i values, in cases where $n = 7$ Legendre components were fit. There is zero chance that λ_0 through λ_2 are Gauss-normal distributions. However, the probability of Gauss-normal distributions increases from λ_3 through λ_5. For λ_6 and λ_7, the distributions are indistinguishable from Gauss-normal. The apparent bias toward high probabilities is a result of fitting the mean and σ for each of the tests, as these parameters are not known a priori. In summary, for these Abell 370 data, the KS test suggests little or no information is contained beyond λ_5.

11.2 Q–Q Plots

A quantile–quantile ($Q - Q$) plot is a graphical diagnostic to test whether two samples come from the same distribution. They plot one sample's quantiles against another sample's quantiles. If the samples come from the same distribution, the points lie
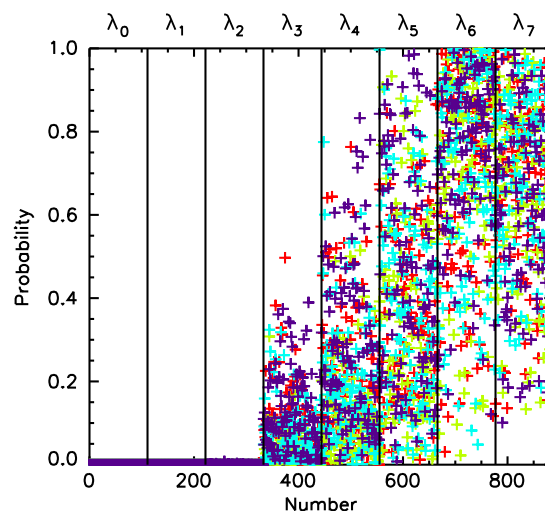

Fig. 9 One-sided KS probabilities for testing the Gauss-normal distribution of λ_i. The horizontal axis simply disperses the different λ_i and 111 ramps. The lowest λ_i carry information and are strongly non-Gaussian. λ_6 and λ_7 are consistent with Gauss-normal distributions, accounting for the bias produced by fitting the mean and σ of each distribution rather than comparing to a priori fixed values. Symbol colors distinguish the independently tested four quadrants for each λ_i.

along a 1:1 line. Here, what is meant by a quantile is simply a data point below which a certain proportion of the data falls. To test whether the noise distributions in the higher λ_i terms are Gauss-normal, we can compare the sample quantiles of the λ_i data by quadrant to the theoretical quantiles of a Gaussian normal distribution.

Figure 10(a) shows the resulting $Q - Q$ plots for λ_5 with each quadrant represented by a different color of points. The λ_5 data has had the mean value subtracted and been divided by its sample standard deviation. A 1:1 line is shown in red. In this plot, it is clear that data in this higher λ_i term are mostly aligned with the 1:1 line suggesting that these data approach Gaussian normality at higher order. However, the data deviate from normality at the extremes. These deviations are the results of heavy tails in the distribution of the λ_i values, which could be due to cosmic rays or other uncorrected effects. Examining the higher order $\lambda_5 - \lambda_7$ $Q - Q$ plots, we see the significant deviations from normality occur at about 3.5 standard deviations from the mean. Dashed lines show this limit, and if we remove these data, we see that the data at higher order become more consistent with a Gauss-normal distribution as measured quantitatively with the Anderson–Darling tests described in Sec. 11.3.

11.3 Anderson–Darling Test

The Anderson–Darling test tests the null hypothesis that the given sample data is drawn from a specific distribution.[28] The Anderson–Darling test returns the Anderson–Darling statistic that can be compared to tabulated critical values (given a chosen significance level) to assess the validity of the null hypothesis.[29] The Anderson–Darling test, like the KS test, compares data distributions, however, the Anderson–Darling statistic is calculated applying weights that make the Anderson–Darling test sensitive to differences in the beginning and ends of distributions, i.e., the tails. We select a significance level of 1% corresponding to a 1 in 100 chance of rejecting
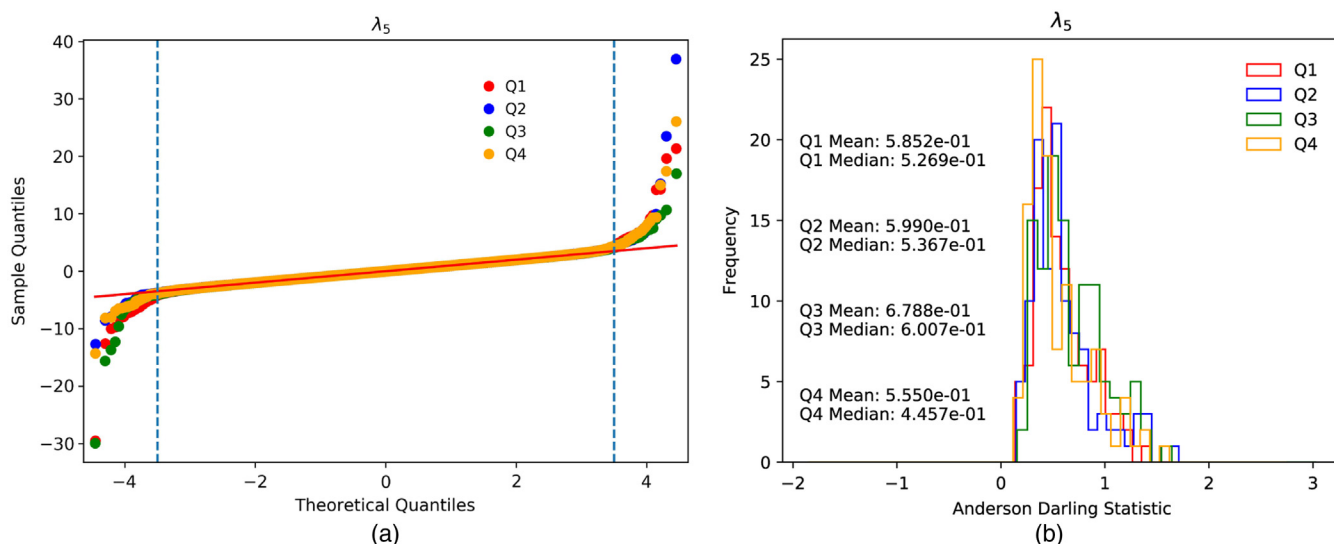
Fig. 10 (a) $Q - Q$ plot comparing λ_5 data with a Gauss-normal distribution with each quadrant represented by a different color of points. The λ_5 data have had the mean value subtracted and been divided by the sample standard deviation. A 1:1 line is shown in red. Higher order terms tend to approach the 1:1 line implying increasing normality, but heavy tails do exist at the extrema. Dashed lines mark 3.5 standard deviations from the mean where it seems the data start to deviate from normality in the higher order terms. (b) Histogram of the Anderson–Darling statistic for λ_5 excluding the >3.5 standard deviation wings (about 1% of pixels) with different quadrants represented by different colors. The mean and median of the statistic for each quadrant are printed in the plot. All means are less than 1.092 suggesting the clipped data are consistent with a Gauss-normal distribution.

Gaussian normality when the data are actually Gaussian normal. The corresponding critical value for the Anderson–Darling statistic when comparing to the normal distribution for a large sample size is 1.092.

We calculate the Anderson–Darling statistic for the λ_i data for each quadrant across 111 images for data fit to six Legendre polynomials. Without removing the outliers in the tails of the data distribution marked by the dashed lines in Fig. 10(a), the Anderson–Darling statistics for all images are greater than 1.092 suggesting a non-Gaussian normal distribution. However, if we exclude the data in the wings of the distribution, on average the Anderson–Darling statistics become less than one suggesting the data is consistent with a Gauss-normal distribution. Figure 10(b) shows the histogram of the Anderson–Darling statistic for λ_5 excluding the >3.5 standard deviation wings with different quadrants represented by different colors. The mean and median of the statistic for each quadrant are printed in the plot. Accounting for the outliers in the tails of the data distribution, the data suggest that at higher order, the coefficients in the Legendre fit become Gauss-normal.

12 Appendix D: Bright-Dark Artifacts in λ_2

As $P_2(x)$ is the first nonlinear Legendre polynomial, we expected that λ_2 would be strongly imprinted with the nonlinear detector response found at very bright sources. As such, we expected λ_2 to be negative and strongly biased toward the brightest sources. However, Fig. 5 reveals a bright-dark artifact such that λ_2 has both negative and positive values on opposite sides of bright sources.

We hypothesize that these artifacts may be caused by a drift in the WFC3 pointing during the course of the observations. Where a source is drifting into a pixel, the accumulating up-the-ramp signal will steepen with time, requiring a positive value of λ_2. Conversely, as a source drifts away from a pixel,

the ramp will flatten (as it would from a standard nonlinear response), and λ_2 will be negative. For small shifts, the induced values of λ_2 should be the product of the gradient of the brightness and the amount of the drift.

To test this, we approximated the brightness gradient by differencing images of the median λ_1 shifted by $(-\delta x, -\delta y)$ and by $(+\delta x, +\delta y)$. To match the apparent direction of the shift in the λ_2 image, the actual gradient image used here is the average of gradient images made with $(\delta x, \delta y) = (1,1)$ and $(\delta x, \delta y) = (1,0)$. This approximates a gradient image of λ_1 in a direction that is 30 deg from the x axis ($d\lambda_1/dz$, where $\hat{z} = (2\hat{x} + \hat{y})/\sqrt{5}$). Next, we performed a linear correlation between λ_2 and $d\lambda_1/dz$ for 6728 pixels with significant but not saturated signals. The slope of this correlation provides a measure of the fractional pixel shift. This shift translates to 1.54 mas given the 130 mas pixel size of for the WFC3 IR detector. The subtraction of the scaled λ_1 gradient leaves a λ_2 residual, where the imprints of astronomical sources are almost entirely negative, and are tightly (but not linearly) correlated with λ_1, as would be expected for standard nonlinear response.

The amount of drift during each of the 36 ramps was also estimated from examination of the RA and DEC data in the associated jitter files. The systematic drifts during each ramp are quite small (comparable to the standard deviation) and averaged across 36 ramps also had a value of 1.54 mas. The directions of the drifts were similar for all ramps, which was likely essential to their systematic imprint in the median λ_2. However, the direction of the drift implied by the most straightforward interpretation of the jitter files is nearly 90 deg different from the apparent drift in the data. When asked, former members of the WFC3 IR Development Team at Goddard told us that the recorded jitter was not necessarily for the specific line-of-sight to our field, and that known (or suspected) flexures and offsets in the observatory could potentially explain the

discrepancy. The status is therefore that the drift amplitude implied by the bright-dark artifacts is consistent with the amplitude recorded in the jitter files. However, we were not able to reconcile the apparent drift directions with the jitter files.

If one accepts that the bright-dark artifacts are consistent with pointing drifts, then a small drift as implied by λ_2 would necessarily distort the shapes of sources measured in that data. Use of the information in λ_2 to try to deconvolve the image to get the true source shapes is likely to be unstable. A better use of this information is to incorporate the drift into models that are compared to the data so that both the data and model are drifted and comparisons between the two will be unbiased by the effect.

References

1. D. Spergel et al., "Wide-field infrared survey telescope-astrophysics focused telescope assets WFIRST-AFTA 2015 report," arXiv:1503.03757 (2015).
2. C. Shapiro et al., "Weak gravitational lensing systematics from image combination," *Publ. Astron. Soc. Pac.* **125**, 1496–1513 (2013).
3. A. A. Plazas et al., "Laboratory measurement of the brighter-fatter effect in an H2RG infrared detector," *Publ. Astron. Soc. Pac.* **130**, 065004 (2018).
4. A. C. Moore, Z. Ninkov, and W. J. Forrest, "Quantum efficiency overestimation and deterministic cross talk resulting from interpixel capacitance," *Opt. Eng.* **45**(7), 076402 (2006).
5. J. W. MacKenty et al., "Wide field camera 3: science capabilities and plans for flight operation," *Proc. SPIE* **7010**, 70101F (2008).
6. J. W. MacKenty et al., "On-orbit performance of HST wide field camera 3," *Proc. SPIE* **7731**, 77310Z (2010).
7. J. W. MacKenty, "Performance and calibration of the HST wide field camera 3," *Proc. SPIE* **8442**, 84421V (2012).
8. L. Dressel, *The WFC3 Instrument Handbook*, Version 10.0, 10th ed., STScI, Baltimore (2018).
9. J. P. Gardner et al., "The James Webb Space Telescope," *Space Sci. Rev.* **123**, 485–606 (2006).
10. M. Gennaro, *WFC3 Data Handbook*, Version 4.0, 4th ed., STScI, Baltimore (2018).
11. W. D. Vacca, M. C. Cushing, and J. T. Rayner, "Nonlinearity corrections and statistical uncertainties associated with near-infrared arrays," *Publ. Astron. Soc. Pac.* **116**(818), 352–361 (2004).
12. B. Hilbert, "Updated non-linearity calibration method for WFC/IR," STScI Instrument Science Report 2014-17, pp. 1–32 (2014).
13. E. C. Piquette et al., "Progress in development of H4RG-10 infrared focal plane arrays for WFIRST-AFTA," *Proc. SPIE* **9154**, 91542H (2014).
14. M. Loose et al., "HAWAII-2RG: a 2k x 2k CMOS multiplexer for low and high background astronomy applications," *Proc. SPIE* **4850**, 867–879 (2003).
15. M. Zandian et al., "Performance of science grade HgCdTe H4RG-15 image sensors," *Proc. SPIE* **9915**, 99150F (2016).
16. J. M. Lotz et al., "The Frontier Fields: survey design and initial results," *Astrophys. J.* **837**, 97–24 (2017).
17. I. Jolliffe, *Principal Component Analysis*, pp. 1094–1096, Springer, Berlin, Heidelberg (2011).
18. G. H. Rieke, "Infrared detector arrays for astronomy," *Annu. Rev. Astron. Astrophys.* **45**, 77–115 (2007).
19. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. **1**, Wiley-VCH, New York (1989).
20. W. H. Press et al., *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn., Cambridge University Press, New York (1992).
21. J. D. Offenberg, D. J. Fixsen, and J. C. Mather, "Memory-efficient up-the-ramp processing with cosmic-ray rejection," *Publ. Astron. Soc. Pac.* **117**, 94–103 (2005).
22. A. M. Fowler and I. Gatley, "Noise reduction strategy for hybrid IR focal-plane arrays," *Proc. SPIE* **1541**, 127–133 (1991).
23. "JWST User Documentation Home," Space Telescope Science Institute, Baltimore, Maryland https://jwst-docs.stsci.edu (2016).
24. J. D. Garnett and W. J. Forrest, "Multiply sampled read-limited and background-limited noise performance," *Proc. SPIE* **1946**, 395–404 (1993).
25. B. J. Rauscher, "Erratum," *Publ. Astron. Soc. Pac.* **122**, 1254 (2010).
26. B. J. Rauscher et al., "Detectors for the James Webb Space Telescope near-infrared spectrograph. I. Readout mode, noise model, and calibration considerations," *Publ. Astron. Soc. Pac.* **119**, 768–786 (2007).
27. B. J. Rauscher et al., "Improved reference sampling and subtraction: a technique for reducing the read noise of near-infrared detector systems," *Publ. Astron. Soc. Pac.* **129**, 105003 (2017).
28. T. W. Anderson and D. A. Darling, "Asymptotic theory of certain goodness of fit criteria based on stochastic processes," *Ann. Math. Stat.* **23**(2), 193–212 (1952).
29. M. A. Stephens, "EDF statistics for goodness of fit and some comparisons," *J. Am. Stat. Assoc.* **69**(347), 730–737 (1974).

Bernard J. Rauscher is an experimental astrophysicist at NASA's Goddard Space Flight Center. His research interests include astronomy instrumentation, extragalactic astronomy, and cosmology, and most recently, the search for life on other worlds. His work about developing detector systems for JWST was recognized by a (shared) Congressional Space Act award and NASA's Exceptional Achievement Medal.

Richard G. Arendt is a senior research scientist at UMBC, working in the Observational Cosmology Laboratory at NASA's Goddard Space Flight Center. He has performed software and data analysis tasks related to the COBE/DIRBE and Spitzer/IRAC instruments while researching dust in the solar system and Galaxy, supernova remnants, and the cosmic IR background.

Dale J. Fixsen has worked at NASA's Goddard Space Flight Center calibrating the FIRAS instrument on the Cosmic Background Explorer (COBE), designing and building the TopHat instrument, and designing the data pipeline for the Infrared Array Camera of the Spitzer Space Telescope. Prior to that, he worked on SQUID magnetic gradiometers for submarine detection and automated data fusion. He holds his PhD in astrophysics from Princeton University and bachelor's degree in mathematics/physics from Pacific Lutheran University.

Alexander S. Kutyrev: Biography is not available.

Gregory Mosby is a NASA postdoctoral program fellow at NASA Goddard Space Flight Center. He received his BS degree in astronomy and physics from Yale University in 2009 and his MS and PhD degrees in astronomy from the University of Wisconsin—Madison in 2011 and 2016, respectively. His current research interests include near-IR detectors, astronomical instrumentation, and applications of machine learning to observational astronomy. He is a member of SPIE.

Samuel H. Moseley is a senior astrophysicist at NASA Goddard Space Flight Center. He has received many awards for his work, including a (shared) Gruber Cosmology Prize, the American Astronomical Society's Joseph Weber Award, and SPIE's George W. Goddard Award.