TECHNICAL REPORT

SPACE TELESCOPE SCIENCE INSTITUTE
Operated for NASA by AURA

Title: NIRCam Dithering Strategies II: Primaries, Secondaries, and Sampling	Doc #: JWST-STScI-002473, SM-12 Date: June 21, 2011 Rev: -
Authors: Jay Anderson Phone: 410 – 338 - 4982	Release Date: 16 April 2014

1 Abstract

A previous Technical Report (Anderson 2010) developed a least-squares approach to combining multiple dithered exposures. This new approach allows us to reconstruct an astronomical scene from a set of dithered images in a way that rigorously respects the sampling of the output frame and also provides an estimate of the covariances among the output pixels. This report will use the algorithm to explore two important issues. The first issue concerns how much sampling is required to fully represent the information delivered by the telescope optics to the detector; this will be a function of wavelength, since blue PSFs are sharper than red PSFs. The second issue involves using the algorithm to determine how well various combinations of primary and secondary dithers are able to produce output images with different levels of oversampling. It is anticipated that the results from this document will eventually feed into the "NIRCam Dithering Handbook".

2 Introduction

When planning a program, NIRCam users will determine from the ETC how much total exposure time they need to spend on their source to achieve the desired S/N ratio. Once this has been determined, they will still have a few decisions to make.

They will have to decide how to divide this time up among multiple exposures, and they will have to decide how to dither the multiple exposures. Dithering will be compulsory for all but a few specialized NIRCam observations (such as coronagraphy and transit studies). The decision about how to break up the total time into individual, dithered exposures will involve a compromise between random sources of noise (such as read-noise), systematic sources (such as L-flat or detector defects), the desired sampling for output-image reconstruction, and overheads (such as slew-times).

Once the general question of how many exposures to take has been answered, the last question to answer concerns the specific choice of dither pattern. Anderson (2009) developed a set of "canned" dither patterns that will be flexible enough to allow users to

pick an optimal pattern for their science, while at the same time consisting of a small enough set that the pipeline can be optimized for the particulars of each pattern.

NIRCam will offer three types of primary dithering: full-field, intra-module, and intra-chip. The choice of which to employ should be based simply on the size of the target. So, all that is really left is the allocation between primary and secondary dithers. The scripts will take N_S secondary dithers at each of the N_P primary-dither pointings, so that the total number of exposures, N_{EXP}, will simply be $N_{EXP} = N_P \times N_S$. Both types of dithering will increase our knowledge of the scene, but some projects may benefit from having more secondaries and fewer primaries, or vice versa.

Primary dithers accomplish several objectives. For the full-field targets and intra-module targets, the primary dithers mitigate the holes in field coverage caused by the gaps between the detectors. Primary dithers move targets around by several arc-seconds, and as such they will ensure that even large detector defects will impact a source in a minimal number of primary pointings. Low-frequency errors in the flat fields can similarly be mitigated by primary dithers. Large dithers end up expanding the entire region covered, while at the same time shrinking the region covered with the maximum depth. The primary-dither offsets will be so large that we should not expect any pixel-phase coherence among them. Each primary dither will essentially place a given object at a random pixel phase.

Secondary dithers will largely be used to improve the sampling of the scene. At wavelengths below 2μm in the SWC (and 4μm in the LWC), the NIRCam detectors will be undersampled. This means that a single exposure cannot contain all the information that is delivered to the detector by the telescope optics. By taking multiple exposures, with each one having the scene slightly shifted on the pixel array, it is possible to recover the high-spatial-frequency information that is missing in a single exposure at a single pointing. The level of undersampling present will depend on which filter is used, thus the optimal number of secondary dithers will naturally be a function of filter as well. Since achieving adequate sampling of the scene is the main goal of the secondary dithers, the secondary dithers cannot be very large, as distortion across the field will cause large dithers to lose pixel-phase coherence[1]. An additional goal of secondary dithers will be to mitigate the effect of bad pixels; each secondary dither will be offset with respect to the others by 2 or 3 pixels to ensure that a bad pixel will affect a star image only once.

Both primary and secondary dithers will provide some mediation for detector/calibration and sub-pixel sampling issues, but in general the primary dithers will provide better mitigation for detector-related issues and secondary dithers will provide better sub-pixel coverage. Users of NIRCam certainly will tailor their dither distribution such that they can optimize the particular aims of their science program. Programs that are focused on

[1] The distortion in NIRCam is required to be less than 2%. This means that the plate scale at the far corner of the detector can differ from that at the center of the detector by at most 2%. If we dither by 10.5 pixels at the center of the chip, then this means the dither at the edge will be at most 10.71 pixels. Thus, instead of having observations at pixel-phase 0.0 and pixel-phase 0.5, we will have them at 0.0 and 0.71. This should not impact image reconstruction too much, but much more pixel-phase phase-drift would start to impact our ability to reconstruct the scene.

point sources may not require as much sub-pixel dithering and thus may prefer to focus on large, primary dithers to better mitigate the L-flat errors. Conversely, programs that are interested in resolving fine-scale structure in faint galaxies or in measuring minor elongations due to weak lensing will depend on good sampling to accomplish their goals, and as such these programs will want to take as many secondary dithers as are necessary[2].

It is not the goal of this document to categorize various projects into what kind of dithers they require, but rather to document what frequency of sub-pixel sampling is needed for various NIRCam filters, and then to demonstrate which dither combinations can achieve adequate sampling.

Since "adequate" sampling will enable a full reconstruction of the scene, there is no need to sample more finely than the Nyquist frequency. Therefore, the goal for most programs will be to use as many secondary dithers as are necessary to achieve adequate sampling, and once that is done, to maximize the number of primary dithers. The amount of sampling necessary will depend on the wavelength of light being studied. The detectors are marginally well sampled at the longest wavelengths for each channel and become progressively more undersampled as the filter wavelengths get shorter. The ideal sampling will depend on the filter used for each observation.

The first half of this report will study models of JWST's PSFs to determine how finely the scene must be sampled in order to recover all the spatial structure that will present in the scenes that NIRCam will image. A scene cannot contain information on a finer scale than the PSF can resolve, so it will be good to know what sampling will fully resolve the structure in images through each of the filters. This will result in a minimal sampling recommendation for each filter.

The second half of this report will study the dither patterns. It will evaluate the sampling achieved by various combinations of secondary and primary dithers in terms how small we can make the output pixels, while still having good constraints on them. The sampling we get from the secondary patterns alone is obvious; but by combining the carefully phased secondary patterns with the (essentially) randomly phased primary patterns, *some* additional sampling is achieved. It will be useful to understand the net result when planning an observation.

3 PSFs and Sampling

Section 3 in the first report discussed how the astronomical scene is represented in the detector pixels and how undersampling results in an incomplete representation in a single exposure. An easy way to conceptualize these ideas is by thinking of the "effective" scene. The effective scene is literally the astronomical scene convolved with the optical PSF and the detector-pixel profile. The effective scene is a continuous, two-dimensional function with as much structure as can be represented by the convolution of the PSF and Π (the pixel-response function).

[2] This may be counter-intuitive, but although point-sources are more under-sampled than marginally resolved sources, there is actually less information in them, since a point source can be entirely specified by three parameters (its x,y position and its flux) but a marginally resolved source can have a structure at *many* spatial frequencies.

The value of the effective scene at any given point in the field represents how much flux would be recorded by a pixel that happened to be centered at that location. As such, each pixel in each exposure represents a single point-sampling of the effective scene. Similarly, the array of pixels in an exposure represent an array of point-samplings of the effective scene. To the extent that there is structure in the effective scene on scales smaller than one pixel, then the scene is considered to be "undersampled" by the pixels, and sub-pixel dithering can help us recover the missing structure.

In this section, we will make use of simulated JWST PSFs in order to investigate how much spatial structure, relative to the native pixel sampling, images at different wavelengths can contain. This will tell us what kind of sampling we will need in the reconstructed image in order to represent all the information.

3.1 PSFs from JWPSF

The Institute has made available to the public a software program, JWPSF, that generates simulated PSFs for JWST, making use of OPD (optical-path-difference) maps that attempt to represent the state of JWST's primary-mirror (PM) segments at typical times between fine-phasing adjustments (see Cox & Hodge 2006). These PSFs are "instrumental" in the sense that they represent the PSF that arrives at the detector, before the detector collects the photons into its pixels. (We note that there is a new, more sophisticated PSF-generating routine available now, WEBB PSF[3], designed by Marshall Perrin, but the structure-scale issues we are investigating here should be the same in both.)

We used JWPSF to generate these "instrumental" PSFs through all of NIRCam's wide-band filters[4] in the SWC and LWC and integrated them over detector pixels in order to arrive at the "effective" versions of the PSFs. It is these "effective" PSFs that will correspond to point sources in the scenes we are attempting to reconstruct. We computed three versions of the PSF for each filter. One comes from the typical OPDs discussed above. A second PSF corresponds to the best PSF that we might expect on orbit, based on correcting the typical OPDs by piston for each primary. The final PSF corresponds to a "perfect" 6.5-m mirror, which has no wave-front error at all. By examining these three PSFs, we should successfully bracket typical, optimal, and ideal observing conditions.

Figure 1 shows an array of PSFs for various SWC filters. The top two rows show the instrumental PSFs for perfect and typical mirror states, the next two rows show the analogous effective PSFs (which include the convolution with the pixel-response function). Finally, the bottom row shows a slice through the center of the PSFs, in order to give a more graphical sense of how sharp the PSF features are.

It is clear that the PSF becomes sharper as the wavelength gets shorter. In addition, the longer wavelengths are less sensitive to wave-front error. The LWC sees almost no

[3] See: http://www.stsci.edu/~mperrin/software/webbpsf.html.

[4] We do not yet have the actual filter curves for all of NIRCam's filters, so the routine simply represents the filter as a top-hat with the specified wavelength zeropoint and width. This simplification should not affect our results significantly.

difference between the well-phased and the typically-phased PSF, but the SWC sees significant differences starting with F150W and getting worse all the way through F070W. We should keep in mind that even though the PSF may have significant sub-structure at short wavelengths with the realistic OPDs, that does not mean there is more high-frequency structure. In general, the amount of sampling needed to resolve whatever PSF information is present is a separate issue from the distribution of that information. In other words, wave-front error may prevent all of the flux from being at the center, but it will not change the scale-size of the PSF features. The sharpness of features is more related to λ/D-type considerations, which is constant for a given filter.

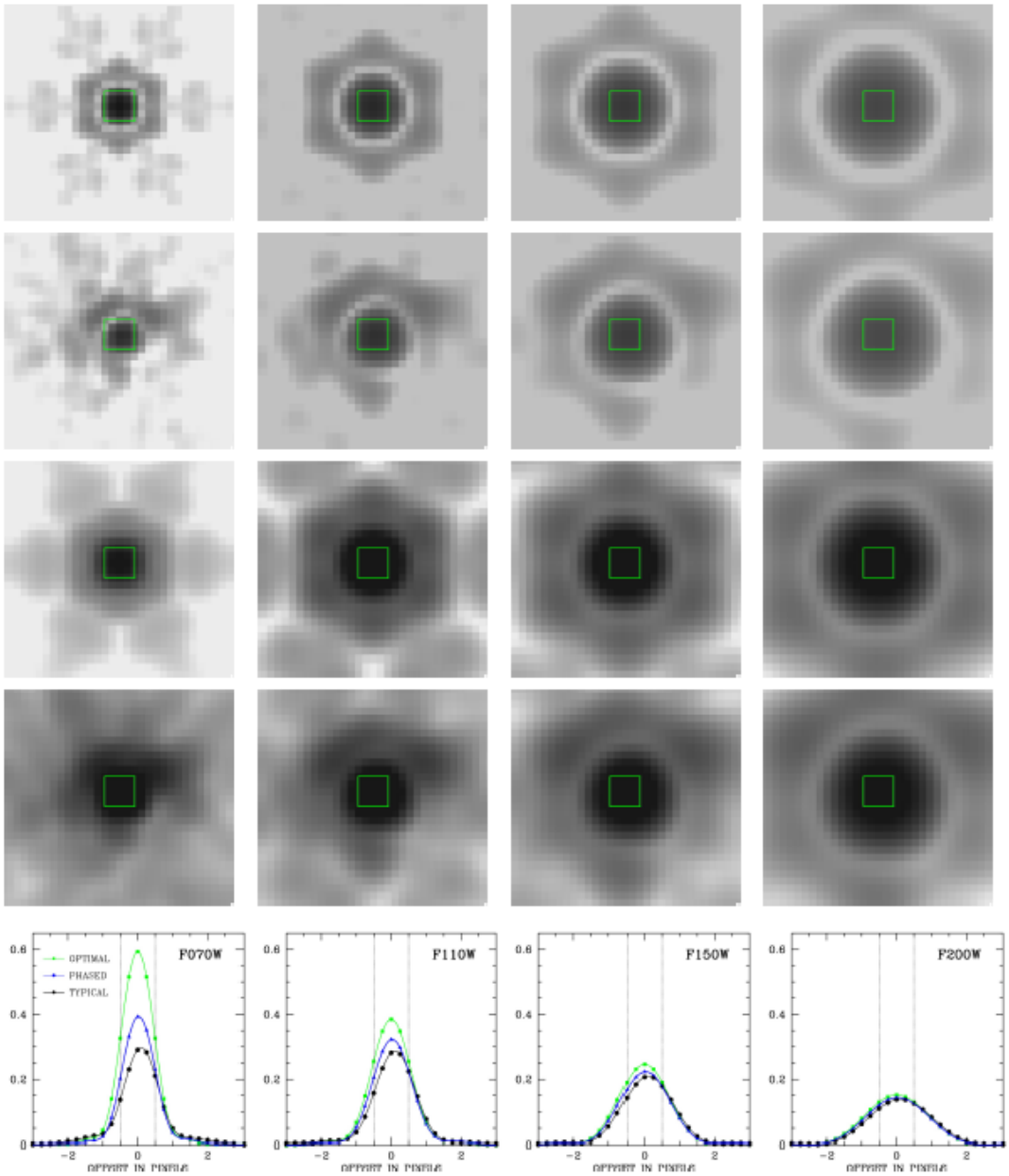

Figure 1: The columns from left to right show information for filters F070W, F110W, F150W, and F220W. (Top row) The *instrumental* PSF for a perfect JWST primary. (Second row) The instrumental PSF for a "typical" mirror state. (Third row) The *effective* PSF for a perfect JWST. (Fourth row) The effective PSF for a typical mirror state. (Bottom row) Plots along the $y = 0$ center of the effective PSFs for the optimal case (green), the typical case (black), and the best-phased case (blue).

3.2 The image-reconstruction algorithm

We will use the least-squares image-reconstruction algorithm developed in the previous TR to probe what kind of sub-pixel sampling will be needed to model the structure in the PSFs. This algorithm was an attempt to address limitations in the multi-drizzle approach to image reconstruction. The multi-drizzle algorithm rigorously preserves the flux in its contributing images. It generates an image that has the flux as close as possible to its true location (as regulated by the "pixfrac" parameter), however it provides no metric to assess how regularly sampled the output image is. In addition, when `pixfrac` is not zero, the output drizzled pixels can be correlated with each other, but information about this correlation is not available to the users. The intention of the previous TR was to develop a framework where both flux *and* sampling could be preserved in image reconstruction, and where the correlations among the output pixels can be quantified.

The assumption in this least-squares approach was that the "effective scene" could be represented as a bicubic spline as constrained by an array of gridpoint-values. There are many ways to specify a bi-cubic spline. A cubic spline naturally assumes that in between gridpoints the behavior of the function can be represented by a cubic function, which is constrained to go through the points. A bi-cubic spline is the extension of this to two dimensions (as described in the previous TR).

There are several flavors of cubic splines. To evaluate a function between two points, a cubic function allows four constraints: the values of the function at the two points and the derivatives of the function at the two points. The values of the function at the two points are given directly by the gridpoints themselves, but the derivatives can be constructed in a variety of ways. The previous TR chose to construct the derivatives that constrained the bi-cubic spline in the easiest way possible, by setting $f_i' = (f_{i+1}-f_{i-1})/2$. This has a very narrow spatial domain and is the reason that the spline-fit over the square region between four pixels depended only on the surrounding 4×4 grid of 16 pixels. It has the downside, though, that although the function and the first derivative are continuous throughout the domain, the second derivative is not continuous.

An alternative way to specify the first derivative of each point in the grid is to fit a quartic function to the surrounding ±2 pixels, and use the value of the derivative to specify the derivative of the cubic spline that goes through that point. This way, the second derivative would also be continuous across the boundary. We decided to explore this and found that it was not to hard to generalize the formalism in the previous TR to do this. Now, instead of each input pixel depending on (and having influence on) the surrounding 4×4 gridpoints (16 total), each input pixel now depends on the surrounding 6×6 gridpoints (36 total).

So, the plots and results shown in this document represent the improved bicubic spline, rather than the one presented in the previous document. The difference is small, but it is worth ensuring the smoothness of as many derivatives as are practical.

3.3 Simulating sampling

As we saw above, the software routine JWPSF gives us a ×4-supersampled version of the "instrumental" (pre-pixelated) PSF. We integrated this over the detector pixels (which we presumed to be square, flush, and uniformly sensitive[5]) to arrive at the "effective" PSF, again super-sampled by a factor of four.

Our goal then is to determine how much sub-pixel sampling is needed to represent this PSF. To answer this, we will plunk down a square array of grid-points with a specified spacing to sample the effective PSF. This square array will represent the sampling of the scene in a reconstructed image. We are free to reconstruct an image with an arbitrary pixel scale relative to the native detector pixels, so we consider that spacing to be a free parameter here.

Figure 2 shows how well we can sample NIRCam's F070W PSF with nodes centered on the PSF with various node spacings. We used the image-reconstruction routine discussed in the previous TR to find the value of the node points such that the bi-cubic spline through them represents the input PSF as well as possible, in a least-squares sense. By spacing the grid-points with different spatial frequencies, we can see how well a given sampling frequency is able to represent the structure in the PSF.

The ×4 reconstruction is essentially perfect. It's not until the sampling gets worse than about ×2.5 that we start to see residuals greater than 1%. The ×1 sampling has severe ringing everywhere, at about the level of 10% of the central value, and even goes significantly negative just outside of the core. While the fractional error at the center is not bad, the fractional error in the outskirts is enormous.

While this test gives us one sense of the sampling issues, it is only a partial sense. In general, the sources we are trying to model will not be centered on the grid that we use to sample them, so it will be important to see how well the reconstruction can be done for an arbitrary offset with respect to the output grid. Figure 3 shows the same plot, but for samples that are offset by a half-node. It is clear that the residuals change drastically. The size of the residuals for samplings finer than ×2 is negligibly small, as above; but when the sampling is coarser than this, the grid really has a hard time following the PSF. There is significant ringing at the 20% level for the ×1.33 case, and at the 30% level for the ×1 case. In general, the pattern of residuals is quite different for the two sampling cases, meaning that it is impossible to draw any inferences about what inadequate sampling will do to a random object, which of course will land at an arbitrary location

[5] The effective PSF treatment does not actually depend on these assumptions and can accommodate almost any pixel-response function. But it is more convenient to deal with flat, flush, pixels.

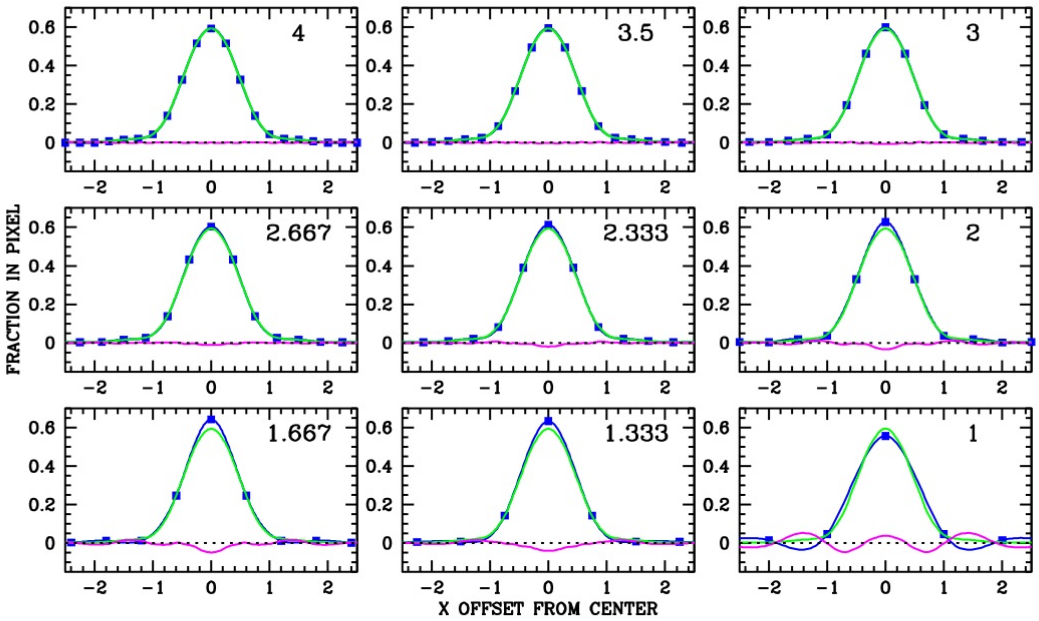

Figure 2: The "optimal" F070W PSF (green). Panels from the upper left to the lower right show the best we can do sampling the function at from 4× to 1× the NIRCam pixel sampling. The blue curve shows the best-fitting spline curve through the allowed sample nodes. Residuals are shown in magenta.

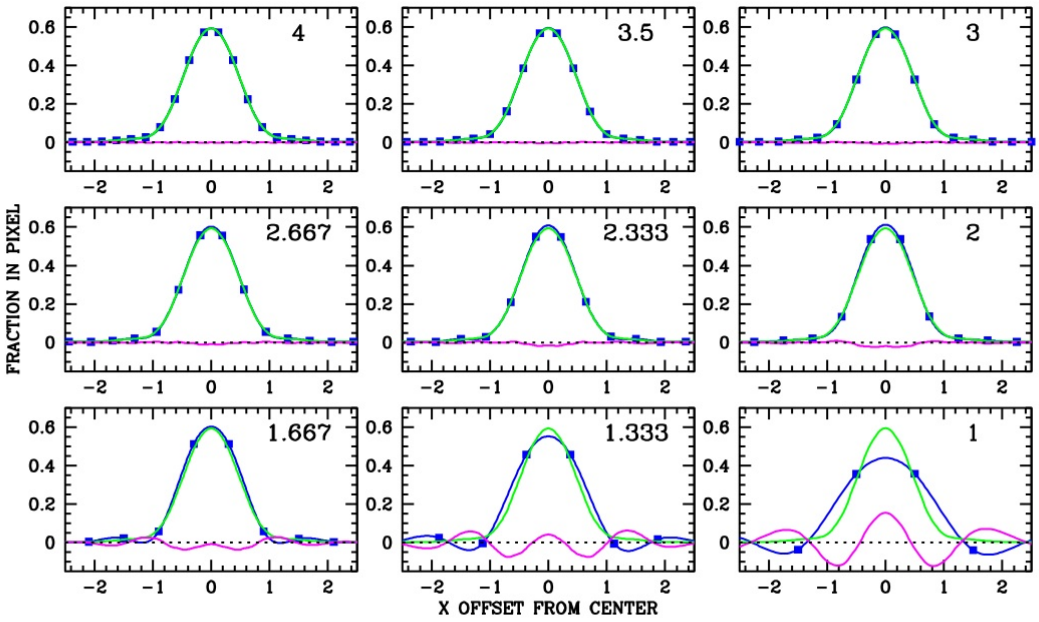

Figure 3: Same as Figure 1, except that here the nodes we use to sample the function are displaced by half a node spacing. It is clear that the reconstructed function is very different from above.

with respect to our sampling nodes. Sometimes the poor sampling ends up increasing the sharpness of the core, other times it broadens the core.

In an effort to include these meta-issues in our sampling analysis, when we evaluate how well a particular sample spacing can represent the PSF, we will perform many trial offsets of the grid-point array relative to the PSF array. We expect observed stars or other astronomical objects to fall at random locations with respect to the reconstructed grid, so this should be a fair test of the realistic situation.

We clearly need to come up with some metrics to assess how the sampling of the dither pattern will impact our ability to reconstruct the scene. The actual test itself involves comparing the true scene, $\Psi(x,y)$, against the model scene, $M(x,y)$, which has been reconstructed using a given sampling frequency. Three quality-of-fit metrics come to mind:

1) the error at the worst-fit point, in terms of the maximum value of the PSF: $\max(|M(x,y)-\Psi(x,y)|] / \max[\Psi(x,y)]$,
2) the RMS error over the entire PSF: $\sqrt{\int [M(x,y)-\Psi(x,y)]^2 \, dx \, dy} / \int \Psi(x,y)$, and
3) the total absolute residual: $\int |M(x,y)-\Psi(x,y)| \, dx \, dy / \int \Psi(x,y)$

We will evaluate each of these three metrics for 19 different sampling frequencies for PSFs for the F070W, F110W, F150W, and F200W filters. We saw above that the particular locations of the grid relative to the center of the PSF can significantly impact the quality of fit, so we investigated an array of 10×10 sub-pixel placements of the grid relative to the PSF center to ensure that there would be no bias in this respect.

Figure 4 shows the results for the simulations. For each sampling/filter combination evaluated, we report the average value of the metric for all the appropriate simulations. The worst-fit metric is shown in green, the RMS error in black, and the absolute error in blue. Each panel shows the trends for a particular PSF. We use the same three PSFs for each filter as in the previous section: one where we assume a perfect mirror, one where we assume a perfectly pistoned mirror, and a final PSF where we assume a typical state of the mirror. These mirror-quality variations are shown from left to right. The filters are shown from top to bottom, with the bluest SWC filter (F070W) on top and the reddest SWC filter (F200W) on the bottom.

In general, we find that the variation of the mirror/alignment quality does not significantly affect sampling-related errors. Evidently, the wavelength of observation determines the size of the finest structure in the PSF and, even if the structure is not phased perfectly coherently, a given wavelength of light still has the same scale of variations.

It is obvious that as we go from the bottom panels, for F200W, where the PSF is essentially Nyquist-sampled, to the top panels, where it is extremely undersampled, we require more and more sampling to achieve a given level of fidelity in the reconstruction. With the F110W PSF, we need better than 2× sampling to get to below 1% RMS and worst-point errors. (We will find in the later sections that a 6-point dither pattern can deliver 2.33× sampling). Finally, the F070W filter is the most undersampled. It requires 3× over-sampling to ensure that the reconstruction artifacts will be below 1%.

These plots tell us what kind of sampling is necessary to represent the structure that is present in an effective image. It is worth noting that even though the F200W PSF is formally Nyquist sampled (its FWHM is 2 pixels), we are not able to fully represent it with simple ×1 sampling. The reason for this is that our reconstructions are based on spatially-limited splines (i.e., splines that depend only on the local scene), which therefore are not formally band-limited, like the PSF. Therefore, we are attempting to represent a band-limited function with a spatially-limited one; this can be done only imperfectly, and the errors in the approximation are related to the sampling frequency.

The reason we chose not to model the entire astronomical scene with a Fourier transform or with sinc interpolation, which would do a higher-fidelity representation of the PSF, is that we will not be dealing with simple point sources in most astronomical scenes. Rather we will be dealing with resolved and marginally resolved objects that, thanks to the tight PSF, have pixels that are almost *completely* uncorrelated over distances of a few pixels. It would therefore introduce unnecessary noise into the reconstruction if the noise in every input pixel contributed to the noise in every output pixel, as is the case in a Fourier reconstruction or a sinc interpolation.

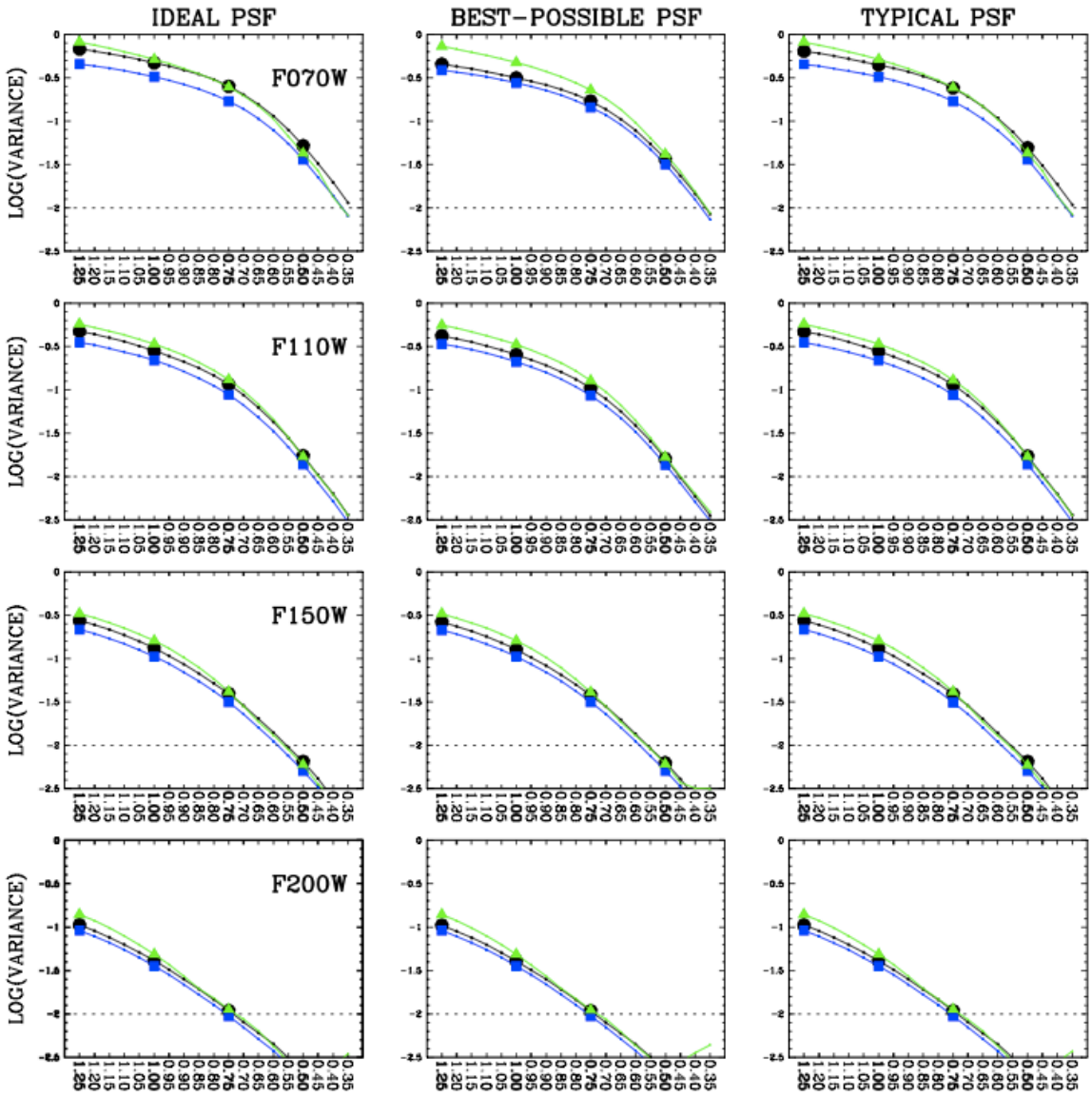

Figure 4: The results of our simulations of what errors to expect when representing various PSFs with grids of various sampling frequencies relative to the NIRCam pixels. The panels from top to bottom show filters from short wavelength (F070W) to long wavelength (F200W). The panels from left to right show how the state of the mirror impacts these results. The leftmost panels show a perfect mirror, the middle panel shows a realistic primary mirror right after fine-phasing, and the right panels show the PSF from a "typical" state of the mirror. The green curves show the worst-case error, the black curves the RMS fractional error, and the blue curve the absolute fractional error. The dotted line at the bottom for LOG(VARIANCE) = −2 corresponds to an error of 1%.

The goal of our approach here is to keep things as local as possible, in an effort both to represent the true structure and to minimize the impact of bad pixels. The downside is that we may need slightly finer sampling to represent the scene with our non-band-limited functions, but the benefit is that we are able to assess more directly the error in reconstructing each pixel, as well as the covariances among pixels. These would be very difficult to estimate from a Fourier-based approach. We note that the pixel-sampling treatment in a drizzled image is even more local than this, so it would be less able to accurately represent the pixel-to-pixel structure in the effective scene.

The plots shown in Figure 4 are for the SWC filters; similar plots could be constructed for the LWC, but we defer that to a future document. In general, the LWC will not be as undersampled as the SWC, since its shortest-wavelength filter is only F277W, which is only half as undersampled as the SWC's bluest filter. The next section will tell us what combination of dither pointings can give us the needed sampling.

3.4 Consistency with previous studies

These conclusions are largely consistent with those arrived at by Koekemoer & Lindsay in their 2005 study of dither strategies for NIRCam and other instruments. Their study focused on photometry and astrometry, not directly on high-fidelity image restoration, so it is not possible to make a direct comparison against these results. Also, their study was considerably broader than this, in that they investigated trade-offs between field coverage, overheads, etc. The present document is focused exclusively on sampling issues.

4 Evaluating Combinations of Secondary and Primary Dithers

In the previous section, we evaluated how well a given regularly sampled grid could represent scenes with various amounts of structure in them. In this section we will turn the question around and determine how the various dither patterns (including combinations of primary and secondary patterns) will be able to constrain output grids with various amounts of oversampling.

We begin with an effective scene represented by an output grid that is sampled with a frequency S. This grid-sampling frequency is assumed to represent all the structure in the scene. The previous section tells us what the sampling frequency S needs to be for a given NIRCam filter.

The goal here, then, is to see how well each dither pattern allows us to reconstruct an output effective scene with a given sampling frequency relative to the detector pixels. The effective scene will be constructed in terms of output-pixel grid-points, so we can evaluate our ability to reconstruct the scene in terms of our ability to accurately measure one representative grid point within it. We could do this for any arbitrary point in the grid, but we will choose the central point for simplicity.

The set of dither patterns evaluated will involve all possible combinations of primary and secondary dithers. Each dither combination provides a set of points that "sample" the effective scene at known locations. Using the least-squares algorithm discussed in the previous TR, we are able to take this arbitrary set of point-samplings and determine from them the regularly spaced grid that best represents them in a least-squares sense. The

covariance matrices from the least-squares fit will then tell us how well the central pixel of the output grid is constrained by the dither pattern. We are particularly interested in the variance of this pixel, in terms of the errors inherent in each of our input samples, but we would also like to know the covariance among neighboring pixels. Both of these diagnostics will tell us how well a given dither pattern can reproduce a super-sampled scene.

If we have N dithers, then the fundamental limit to how fine we can make the output grid is 1 output pixel every $1/\sqrt{N}$ input pixels. This is constitutes a situation that is formally constrained, so long as all of the input pixels are at unique locations. However, we will see below that when the output grid is only marginally constrained, it can result in considerably amplified random errors.

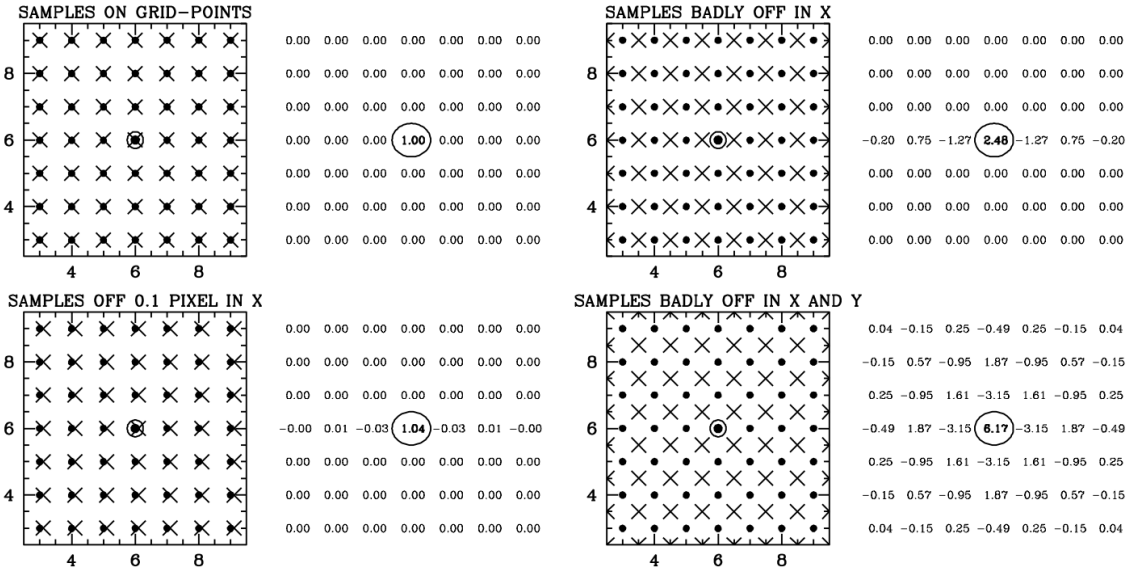

Figure 5: This shows the covariance matrix achieved when we resample a grid of observed samplings (the crosses) at the locations of an output grid (the filled circles). The covariance matrix corresponds to the central grid-point (circled). The circled point in the covariance matrix represents the variance. The variances and covariances shown are with respect to a single observation.

4.1 Resampling

Before we delve into reconstructing an image from a set of dithered observations, it is informative to consider how well we can resample a single image with a grid that has the same sample spacing. In the single-image case, we have no hope of improving the sampling, so the only reason to this would be to remove distortion, at the expense of introducing re-sampling errors.

The upper-left plot shows the trivial case of resampling the grid at the same points as the input constraints. There is no covariance among the output pixels and the variance in the output pixel is the same as that of the input pixels (1.0). The next case, in the lower left, shows what happens when the destination grid is shifted by a tenth of a pixel to the left relative to the input grid. The variance in the output pixel is now greater than 1.0, and

some covariance has been introduced among the pixels to the right and left. Note that there is no covariance among pixels in the vertical direction.

The upper-right plot shows what happens when we resample the grid in the case where the output pixels are placed exactly in-between the input pixels in the horizontal direction. The variance shoots up to 2.48 (the error in the pixel value will be the square root of this, or 1.57 times the typical error in an input pixel), and the covariances are also quite high. This makes sense. In order to estimate the value of the function in between two pixels, we interpolate among the input samplings, which involves taking sums and differences of the samples. Imagine trying to solve for the value at the central point by fitting a cubic function using the surrounding 4 sample points (the two ×'s to the left and the two to the right), and it is clear that there will be considerable covariance, and the error in the resulting function at the output point will be larger that that of a single input grid point.

Finally, we consider the worst possible case: where our output grid is maximally out of phase with our input grid. In the bottom right panel, we see that the variance in the value of the output pixel is 6.17, meaning that the error in our estimate of that output grid-point will be 2.48 times worse than the error in an input grid point. There will also be considerable covariance among the output grid points. This has to be the case, as the output grid-points are all constructed from sums and differences of the same input pixels. One might wonder why we would do this, but when a single distorted image gets resampled onto an geometrically correct output grid, this will happen to a large fraction of the pixels.

4.2 An example with additional secondary dithers

The resampling example above is actually quite a bit more relevant to our problem than one might have originally thought. When we reconstruct an output grid in a real-world situation, we are essentially doing the same thing, except that (1) there can be more than one input sample within the projected area of each input pixel (thanks to multiple dithered exposures) and (2) the output grid can be spaced more finely or less finely than the input grid. The same issues, though, are relevant, since we are still resampling the input data — which comes in the form of point-samplings — onto a different output grid. We will explore the various aspects of this below.

Figure 6 shows what happens when we have 2-point and 3-point dithers and try to reconstruct an output image with the original pixel sampling. The top case for each represents the situation where one of the output dithers is placed exactly on the output grid. In this case, the variance is always less than 1.00, since we have one perfectly placed sample and additional samples that help constrain the output pixel as well. The improvement is naturally better when we have 3 dithers.

The bottom plots show what happens when we now try to reconstruct the output grid with grid-points that are as far as possible away from the input samples, we find that we have a variance of 0.91 in the 2-point case and 0.56 in the 3-point case. The 2-point case is much better than the 1-point case shown in the previous figure. Clearly the additional dithers help us constrain the "worst" point much better than a single dither can. The

"worst" point for the 2-point dither is actually better reconstructed than the "best" point for the 1-point dithers.

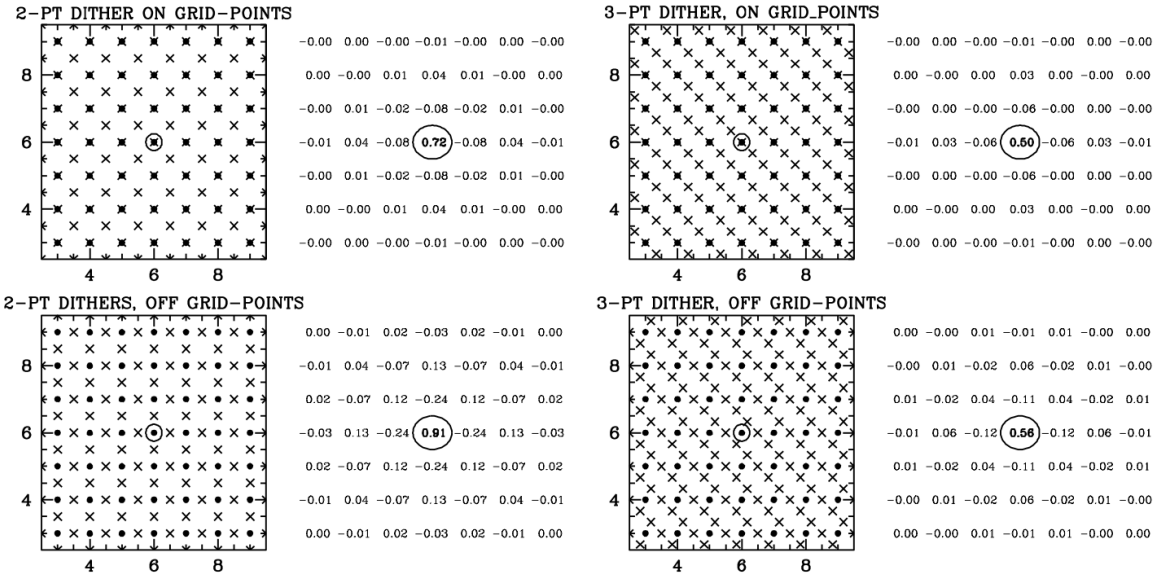

Figure 6: Similar to Figure 5, but this shows the reconstruction of a 1× sampled output image using a dithered data set. On the left, we have a 2-pt dither and on the right we have a 3-point dither. In the top panels, the output grid is aligned with one of the samples and in the bottom grid, the output grid is placed in the worst possible phase to sample the input grid. The variances and covariances shown are with respect to that of single observation.

4.3 Varying the sampling for a given dither pattern

The above examples for 2-point and 3-point dithers illustrate that our ability to reconstruct a grid will depend not only on the input and output samplings, but on the details of exactly where the samples are located relative to the output grid. In general, we will not have control over this, so we must plan for the worst possible case.

In Figure 7 we explore how a given 4-point dither pattern is able to reconstruct an output grid with various output sampling frequencies. If we try to resample the grid with 1:1 output pixel to input pixel spacing, we have essentially 4 constraints for every output pixel. It makes sense that the output variance is quite small, 0.40. In principle, it could be as low as 0.25 if all the input samples were placed at the exact location of the output samples, but that is not practical. We have chosen to space out our dithers in an effort to mitigate the worst-case for grid reconstruction. It is worth noting that we will never remove the covariances between pixels, since input-pixel samplings that fall in between output grid-points will naturally contribute to multiple output grid-points.

The other panels show how our reconstruction is impacted when we reconstruct the output image with finer and finer sampling. The first step is to go to 1:0.90 sampling, which means that the output pixels are 0.9 times the linear size of the input pixels. The quality of reconstruction is only slightly worse than in the 1:1.0 case. We are able to go all the way down to a sampling of 1:0.7 before the variance in the worst output pixel is worse than that in a single input pixel. The 0.7-pixel output pixels have an area of about

0.5 square input pixel, and as such we have roughly two input constraints for every output grid-point. Nevertheless, since these input constraints are not placed at the exact location of the output pixels, the achieved variance ends up closer to 1.0 than to 0.5.

The lower-left panel shows the quality of reconstruction for the worst-case placement of the output grid when we reconstruct with the highest possible resolution allowed by the number of constraints we have. The output grid here is only marginally constrained by the input grid, and the variances and covariances are large. They are, in fact, identical to what we saw in Figure 5 with the worst-case 1:1 resampling of a single observation. This makes sense, since in both cases we have the same array of input and output samplings, modulo an overall scale change. Note that in all these cases, we chose the worst possible centering for our output grid to emphasize the worst-case scenario.

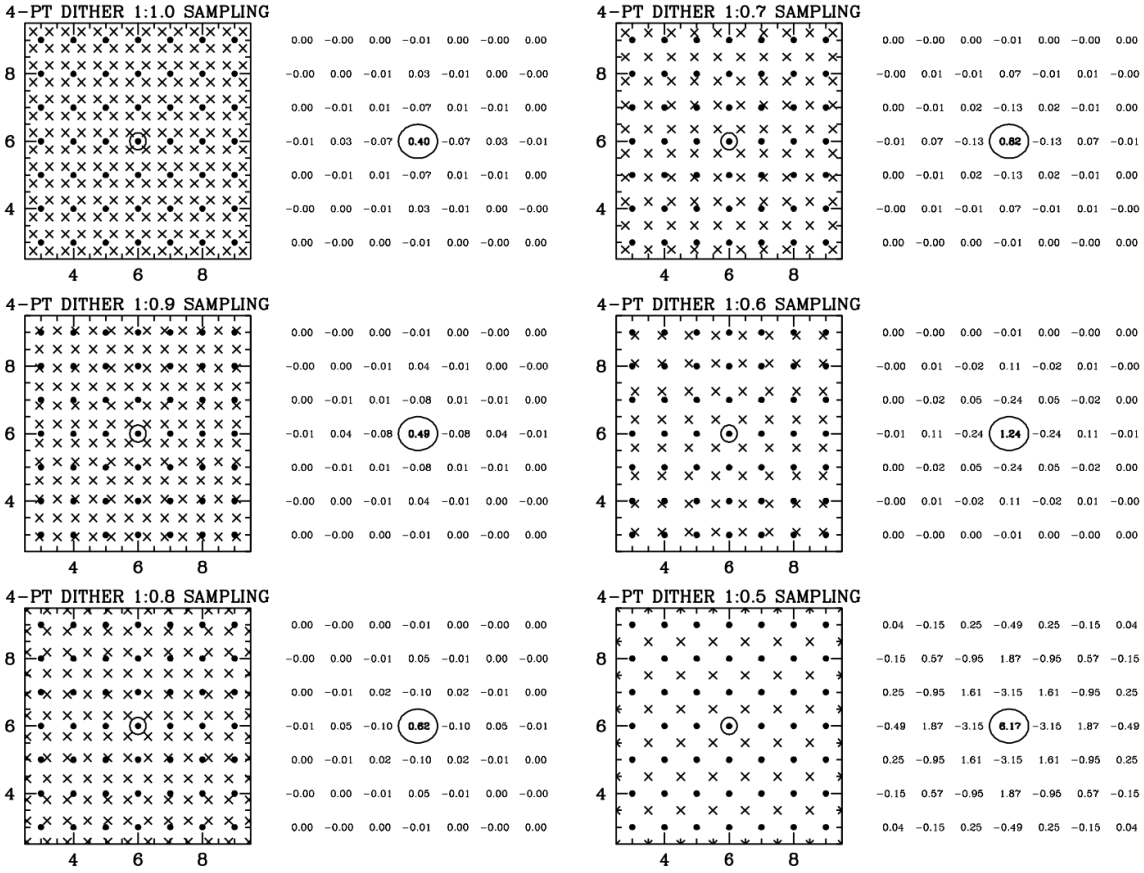

Figure 7: These panels show how well a perfect 2×2 4-point dither pattern is able to reconstruct scenes with a range of output-sampling frequencies. In all cases we have chosen the "worst" location for the output grid relative to the input grid, namely the offset that maximizes the distance between the central point and the input samplings. The variances and co-variances shown here are with respect to those of a single observation.

4.4 Folding in the random element.

Choosing the optimal dither pattern will usually involve a compromise between mitigating sampling errors and mitigating other issues, such as gaps in field coverage, L-flat errors, etc. For this reason, NIRCam users will often want to combine primary and

secondary dithers. In general, the secondary dithers will all have coherent pixel-phase coverage, but the dithers at different primary locations will in general not be pixel-phase coherent with respect to one other, since the offsets will be so large that distortion will randomize the pixel-phase offsets achieved in different parts of the field.

We simulate this by examining the quality of reconstruction at various sampling frequencies (F) for a combination of P × S dithers, were P is the number of primary dithers and S is the number of secondary dithers executed at each primary location. The secondary dithers will all be coherent within a given primary pointing, but will be incoherent among the different primary pointings. We will model this as each primary dither having a random offset with respect relative to the other primaries.

Figure **8** gives one example of this. In the left panels, we show the sampling achieved by four different realizations of a single pair of 2 secondary dithers (a 1×2 pattern) placed randomly within the output grid space. In this case, the dithers are all separated by (0.5,0.5) pixel, and the only difference is the arbitrary shift of the output grid relative to the input pixels. In general, we are able to reconstruct the output grid with a variance between 0.79 and 0.96.

In the right panels, we show four random realizations of a 2×1 pattern. The samples from each primary image have a random dither with respect to the output grid. Sometimes this results in satisfactory sampling to reconstruct the central point (0.67 variance), and sometimes it results in relatively poor constraints on the central point (1.21 variance).

Figure 9 provides one more example of this. Here, we compare issues involved in reconstructing an F=0.7 output image for the worst-case 4-point dither (a 1×4 pattern) and several realizations of two 2-point dithers (2×2) pattern. In the latter, the secondary dithers are coherent within each primary pattern, resulting in a certain level of regular coverage of the pixel phases, but the dithers for the different primaries are not coherent, so the coverage is not completely even. The very worst point of the regular 1×4 pattern (shown in the upper left) gives us a variance of 0.82; the variance for all other points within the output image will be better than this. The five different random realizations of a 2×2 pattern give variances between 0.84 and 28.83. A variance of 28.83 means that the error in that pixel would be more than 5× worse than the error in a typical input pixel.

It is clear that sometimes when we leave some aspect of the sampling to chance, we risk having very poor constraints on some of the output-image pixels. In the worst case (lower left), we see that the two primary pointings just happened to sample nearly the same pixel phases, leading to considerable degeneracy in the resulting grid solution.

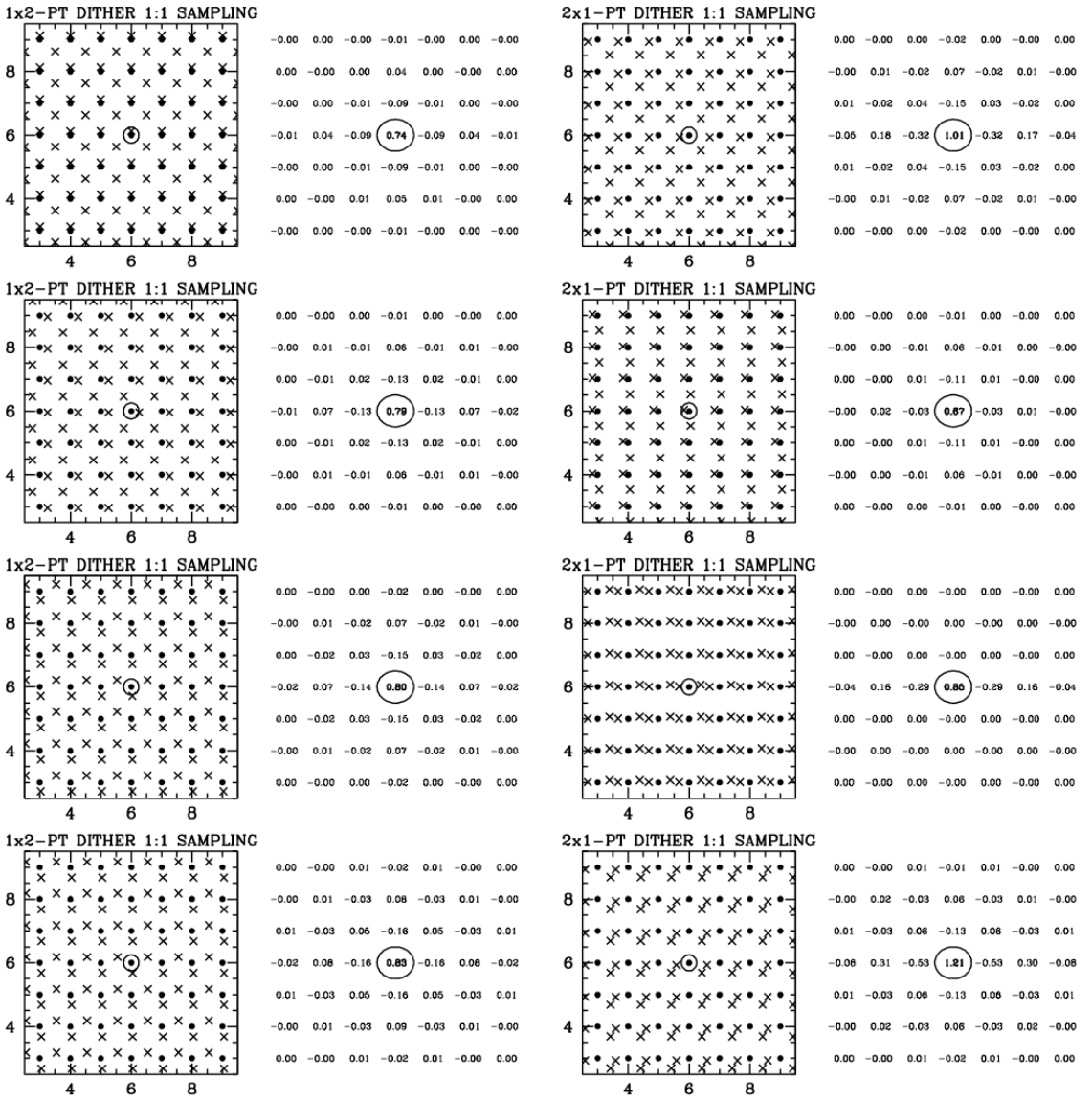

Figure 8: These panels show how out how well we can constrain the ×1 sampled output grid from the case where we have two well spaced dithers and two randomly spaced dithers. The left panels show the resulting sampling quality for four random placements of a 2-point secondary pattern relative to the output grid. Note that the pattern is always the same but its offset from the output grid varies from realization to realization. In all cases the variance is less than 0.83. On the right, we have simulated two completely random, uncorrelated dithers. The best case here (0.67) is better than the best case in the 1x2 example, but the worst case (1.21) is far worse. The variances and covariances shown here are with respect to the error in a single exposure.

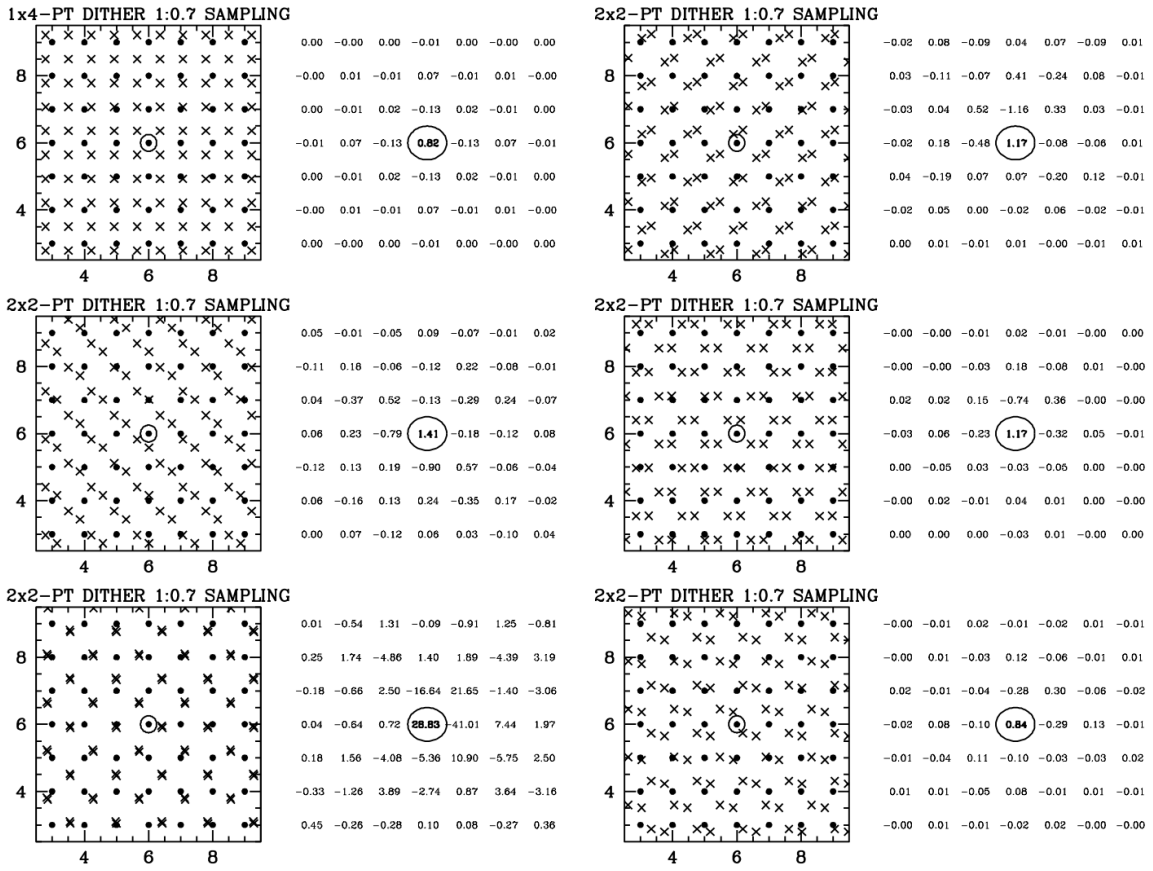

Figure 9: The upper-left panel shows the sampling achieved for the worst field location of a 1×4-point dither pattern, for a S=0.7 output grid. Even though this is the worst point in the field, we still get a variance of 0.96. The other panels show five different realizations the random offset between two 2-point dithers and the output grid. It shows the distribution of coverage we could expect from a 2×2-dither pattern. Note that the realization in the lower-left panel shows a pathological case, where the two primary dithers happen to coincide, pixel-phase wise. The variances and covariances shown here are with respect to the error in a single observation.

4.5 From specific examples to distributions

The specific examples shown in Figures 5 through 9 demonstrate the kinds of reconstruction errors we might expect for particular arrays of samplings. It is useful to average over the "random" aspects of these examples to arrive at overall distributions of the output variance for a given combination of sampling (F) and dither pattern (P×S).

To examine these overall distributions, we simulated 10,000 realizations of the 1×4, 2×2, and 4×1 pattern (the latter just consists of four completely random pointings). We then determined how well these sets of samplings could allow image reconstruction with various output-sampling frequencies (F).

The panels in Figure 10 show the distribution of variances for the sampling frequency F (labeled 1:F), for the three different dither strategies (all secondary, all primary, and half-and-half). The top plot shows the variance distribution for reconstructing a 1:1 output

grid. In general, all dither strategies are able to produce an output image with a variance well below 1.0. This is not surprising, since in all cases, we have four times the number of constraints as unknowns. Note that sometimes the 4×1 pattern (blue histogram) does better than the 1×4 pattern (black histogram). This is because every once in a while, all 4 will fall very close to the output pixel centers. The best possible case for this would have a variance of 0.25. More typically, though, the 4×1 does worse than the 1×4.

The next panel down shows what happens when we try to reconstruct an image with slightly improved sampling. Even when we improve the sampling to 0.9 of the original pixel scale, we find that the 4×1 pattern begins to develop a high-variance tail. As we make the output sampling finer and finer, there is a bigger and bigger chance that a random dither will not allow a good reconstruction for the randomly chosen output gridpoint. By contrast, the 1×4 and 2×2 patterns continue to provide a high-quality reconstruction.

As we aim for finer and finer output sampling, the 4×1 and 2×2 patterns begin to show more and more of a high-variance tail, corresponding to the output grid points that just happen not to have any nearby constraints. The 1×4 pattern, by contrast, continues to provide uniformly good reconstruction quality all the way down to F=0.6. At this fine output sampling frequency, the 4×1 and 2×2 patterns often cannot even produce a reliable output image, due to the pathological sparseness of the samples.

Finally, at F = 0.5, even the 1×4 pattern gives a distribution with variances between 1.0 and higher. Many of the realizations have variances between 1.0 and 2.0, and there is a tail that goes up to the high levels seen in the lower-right panel of Figure 7, as one would expect when the output points find themselves exactly between the input points.

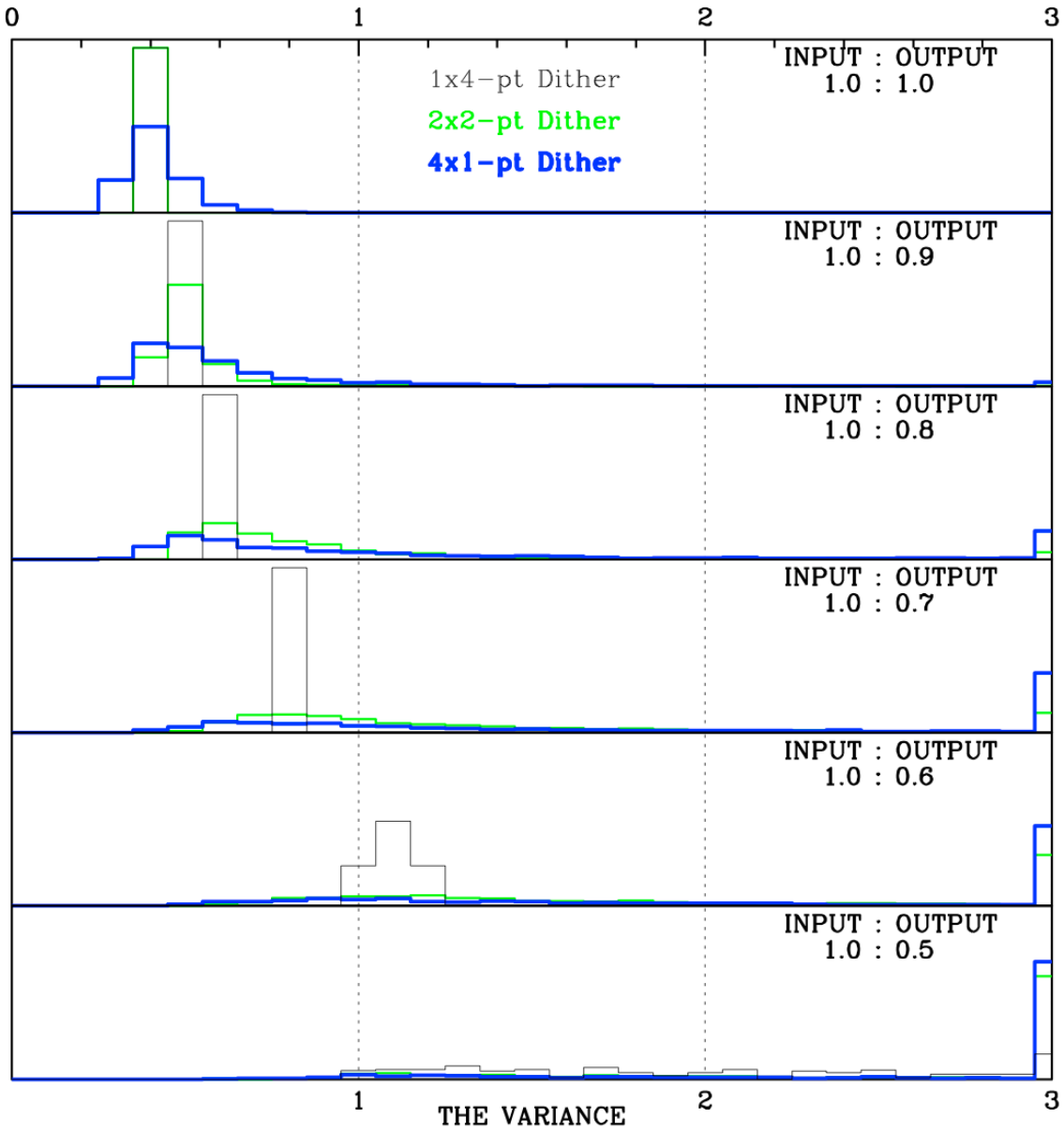

Figure 10: The panels show the distribution of variance achieved in reconstruction of the central pixel of an output image for a variety of the output pixel sizes (1.0, 0.9, to 0.5 input pixels). The histograms come from 10,000 random realizations of three different dither strategies (1×4 in black, 2×2 in green, and 4×1 in blue). The regular pattern clearly provides much more even coverage. The variances shown here are with respect to the error in a single observation.

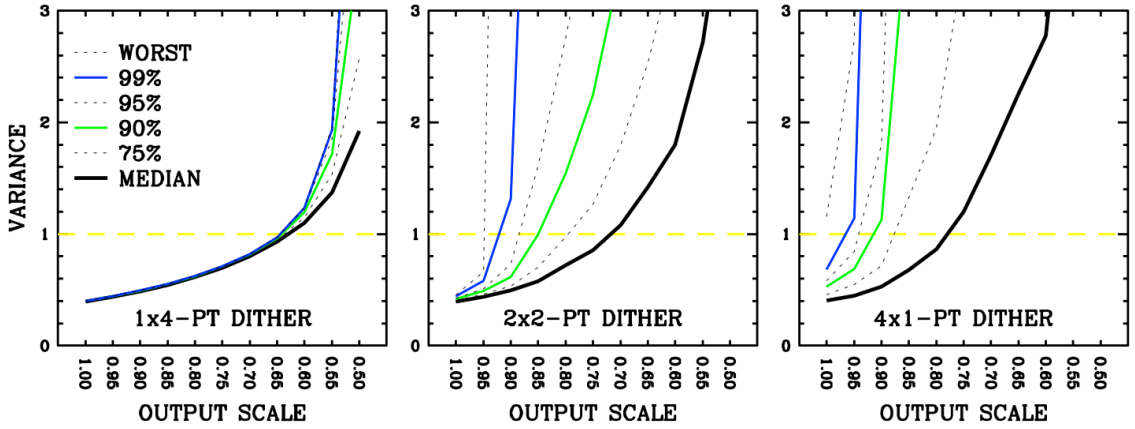

Figure 11: This figure shows the same information in Figure 10, but in a more compact format. Here, we show the percentiles of the variance achieved as a function of output pixel scale for the three dither patterns investigated. The variance is with respect to the error in a single observation.

Figure 11 above shows the same information as was in Figure 10, but in terms of percentiles. The solid black line shows the median reconstruction quality achieved for the 10,000 realizations of the three dithers considered. The other lines show the 75%, 90%, 95%, 99%, and the worst-case percentiles for the 10,000 cases examined.

It is clear that if we wish to sample the image more finely than the input pixels, we really need to have a regular array of samplings. For instance, if we want to always get a variance below 1.0 (namely, if we want the output image to have at least the same signal-to-noise that can be found in a single pixel of the input exposures), then we can at best get 0.65-pixel sampling with a 1×4 dither. If we try to get better sampling than this, we will end up increasing the variance a good number of pixels. For a 2×2 dither, we can only get 0.95 sampling before we start to compromise a few pixels. The 4×1 sampling doesn't even provide a robustly reliable sampling at the original pixel scale.

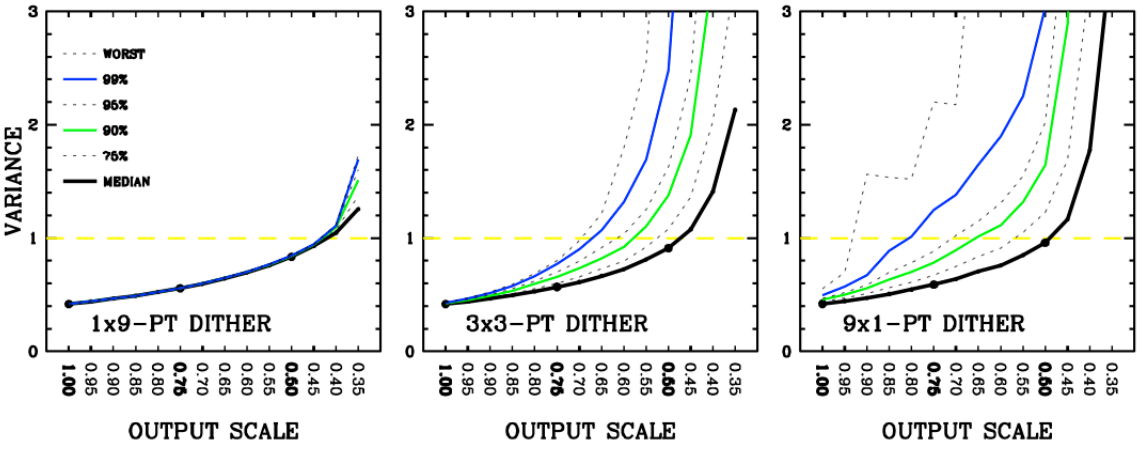

Figure 12: Same as in Figure 11, but investigating the sampling achieved with 9 dithers, laid out with a 1×9, 3×3, and 9×1 dither strategies (P×S). The variance is with respect to the error in a single observation.

Figure 12 below shows a similar plot, but for a set of 9 dithers, distributed again among all secondary, a mix of primary and secondary, and all primary. It is clear that exquisite super-sampling (F=0.45) is possible with a regular dither, and reasonable sampling (F=0.70) can be achieved with a mix of primary and secondary. With the purely random 9-point pattern, 99% of the time we can get a good reconstruction with F=0.80, but even then there is a trail with 1% of the pixels that will have larger variances than 1.0. Sometimes even 9 random samples do not allow us to increase sampling to F=0.8.

At some point, we will come up with plots like Figure 11 and 12 for 6-point patterns with 1×6, 2×3, 3×2, and 6×1 and for 8-point patterns with 1×8, 2×4, 4×2, and 8×1. The goal of the present document is to be more exploratory than comprehensive.

4.6 What can be done with only the standard Secondary patterns?

Given how clearly superior regular dither patterns are at recovering structure in the scene, it is worth investigating the percentile plots for the cases for 1×S, i.e., where there is one primary pointing and S secondary pointings. This will tell us the minimum number of dithers required to achieve a given level of sampling.

Figure 13 shows the variance achieved by percentiles for the 9 standard S-point secondary dithers. The upper left panel shows the reconstruction quality achieved for a 1-point dither — essentially sampling the output image at a random pixel phase. If we reconstruct an image with output pixels that are 1.25-times the input pixel scale (i.e., oversized output pixels), then we can achieve a decent reconstruction. If we try to sample it at the native resolution (F=1.00), then we have some fraction of pixels with very large variances.

The 2-point pattern allows a good F=1.00 resampling, but starts to incur large errors just below this (even at F=0.95). The 3-point pattern can do a reliable reconstruction down to F=0.85 and the 4-point pattern down to 0.65 (as we saw above). The 5-point pattern can go down to F=0.60 and the 6-point pattern down to output pixels that are 0.55-times the size of the input pixels. The 8-point dither can achieve a F=0.50 resampling and the 9-point dither can go down to 0.45 times the input pixel scale.

It is worth noting that none of these is able to achieve the sampling (F=0.40) needed to adequately treat the F070W PSF, which is about 3× undersampled. We have not looked into a finer secondary dither pattern than 3×3 (S=9). That may be worthwhile, but then again, the more sub-sampling one does, the more other systematic issues, such as pixel-response-function regularity and PSF variations, become very difficult to model. At some point, parameterized-model fitting will be superior to blind image reconstruction. It is likely that F=0.40 is close to this limit.

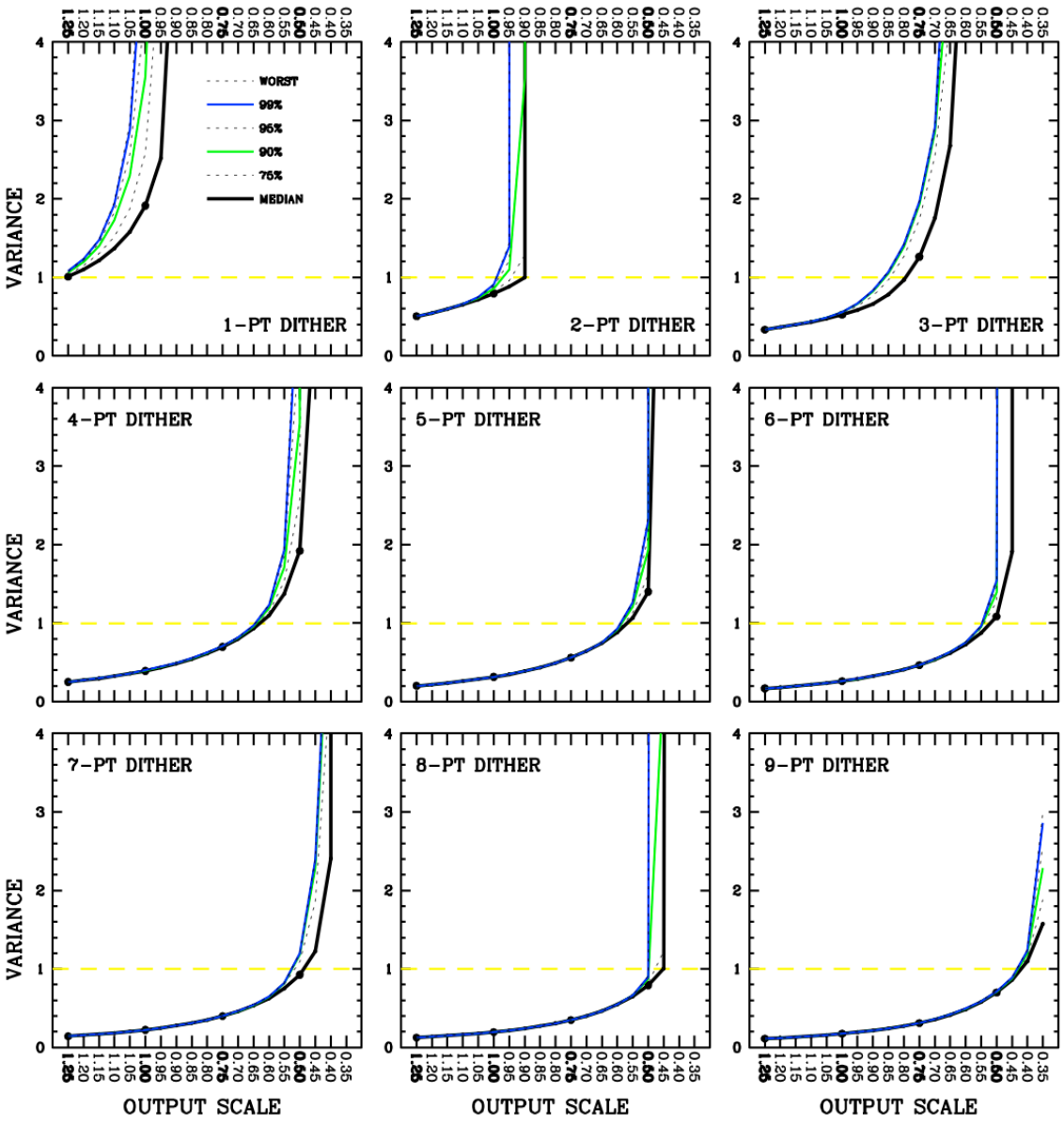

Figure 13: The variance achieved for the 9 standard secondary dither patterns. 10,000 realizations of the dithers have been simulated, and the percentiles of the variance distribution are shown. The take-away from this diagram is at what scale we can super-sample the pixels as a function of the number of regular dithers. In principle, the best-possible output-pixel spacing should be $1/\sqrt{N_{DITH}}$, but in practice to avoid near-degeneracies, we need slightly better sampling than this.

5 Conclusions

This document has discussed two different aspects of sampling: (1) What frequency of sampling is necessary to preserve all the information present in an image? and, (2) What frequency of sampling do we expect to achieve with a given dither pattern? Together, these two investigations will allow users to determine what they need and how to meet those needs.

Of course, not all projects with NIRCam will be sampling-critical. Many, in fact, will have other issues that take priority over sampling. Point-source photometry, for instance, should not depend critically on sampling. Studying well-resolved objects will also not require optimal sampling. There are many examples of projects where considerations other than sampling (such as field-coverage or minimizing overheads) will take precedence.

Nevertheless, many programs *will* want to make the most of NIRCam's resolution potential, and they will want to fold in the considerations presented here to optimize their observing program. We expect that as the operations plans become more and more refined, the Institute will provide supplemental information to this document to help users make wise dithering decisions. Sampling is just one issue among many. It may be worthwhile exploring ways to create stacked images that have regular and improved sampling, but not necessarily at the highest possible resolution.

One issue that has not been addressed here is what happens in the case of missing data. Not all pixels in all dithered exposures will have the same weights, due to dark current variations, CR-hits degrading the up-the-ramp fits, or other issues. Some pixels in some exposures may even be missing, due to defects, pathological CRs, gaps, etc. Whichever image-reconstruction algorithm we adopt should be robust against these missing data. In fact, the final image-reconstruction algorithm will also have to be robust against pixels errors that are larger than anticipated. We may have to iterate the image reconstruction in such a way that we can identify the previously unknown bad pixels and remove them. Least-squares treatments are notoriously sensitive to outliers. We note, though, that the least-squares approach should naturally be able to deal with pixels that have different intrinsic errors.

One issue that has not been discussed is what to do in the case where the PSF varies slightly from exposure to exposure, such that each exposure does not represent the exact same effective scene. In this case, we will have to find some way to regularize the effective PSF. This may involve introducing some kind of convolution kernel into the image-reconstruction procedure.

These issues are, naturally, beyond the scope of the present document, which had the limited aim of determining (1) what sampling is needed for a given PSF, and (2) how to achieve that sampling with a set of dithers. The real-world limitations and complications must be dealt with in turn, but at least here we know what is theoretically possible (and impossible).

It would be interesting someday to compare the results of this technique and Fourier, sync, and drizzle-based techniques to assess where each performs and underperforms in terms of producing a output images with understood sampling properties and covariance properties.

6 References

Anderson, Jay. 2009. JWST-STScI-001738, Dither *Patterns for NIRCam Imaging*

Anderson, Jay. 2010. JWST-STScI-002199, NIRCam Dithering Strategies I: A Least-Squares Approach to Image Combination

Cox, C. and Hodge, P. 2006 SPIE 2625 26, "Point-spread function modeling for the James Webb Space Telescope"

Koekemoer, Anton and Lindsay, Kevin. 2005. JWST-STScI-000647 An Investigation of Optimal Dither Strategies for JWST