Simulating JWST high contrast observations with PanCAKE


Event: SPIE Optical Engineering + Applications, 2021, San Diego, California, United States
Simulating JWST high contrast observations with PanCAKE

Aarynn L. Cartera, Andrew J. I. Skemer a, Camilla Danielski b,c, Jarron Leisenring d, Jason J. Wang e, Kyle Van Gorkom d,f,g, Brian York h, Jea Adams j, Beth Biller i,k, Julien H. Girard h, Sasha Hinkley l, Bryony Nickson h, Marshall Perrin h, and Laurent Pueyo h

a University of California, Santa Cruz, Santa Cruz, CA 95064, USA
b UCL Centre for Space Exochemistry Data, Atlas Building, Fermi Avenue, Harwell Campus, Didcot, OX11 0QR, UK
c Instituto de Astrofísica de Andalucía (IAA-CSIC), Glorieta de la Astronomía s/n, 18008 Granada, Spain
d Steward Observatory, University of Arizona, USA
e Department of Astronomy, California Institute of Technology, Pasadena, CA 91106, USA
f University of Arizona, College of Optical Sciences, Tucson, Arizona, United States

ABSTRACT

The James Webb Space Telescope (JWST) and its suite of instruments will offer significant capabilities towards the high contrast imaging of objects such as exoplanets, protoplanetary disks, and debris disks at short angular separations from their considerably brighter host stars. As the official JWST exposure time calculator (Pandeia) is not capable of simulating a range of observational features that can significantly impact the performance of JWST’s high contrast imaging modes (e.g. target acquisition offsets, temporal wavefront evolution, small grid dithers, and telescope rolls), the development of more advanced software is necessary to produce realistic estimates of contrast performance. Here we present an update to the JWST high contrast simulation tool PanCAKE which dramatically improves on its accessibility and functionality for more general users, whilst maintaining the capabilities of its predecessor in simulating more advanced observational features across both the NIRCam and MIRI instruments. Of particular note is a more streamlined user-interface, and new capabilities to: optimise up-the-ramp readout parameters, estimate and apply wavefront evolution across observations, and perform advanced PSF subtraction strategies. As demonstration of this improved PanCAKE interface and functionality, we simulate a subset of the scheduled NIRCam and MIRI observations of the exoplanet HIP 65426 b as outlined in JWST DD-ERS Program 1386.

Keywords: JWST, high contrast imaging, coronagraphy, exoplanets, NIRCam, MIRI, user support, simulations

1. INTRODUCTION

The James Webb Space Telescope (JWST), currently scheduled for launch in late 2021, offers a variety of observational modes suitable for high contrast measurements. Specifically, the Near-InfraRed Camera (NIRCam) has coronagraphic capabilities from ∼1–5 µm, the Mid-InfraRed Imager (MIRI) has coronagraphic capabilities from ∼10–23 µm, and the Near-InfraRed Imager and Slitless Spectrograph (NIRISS) has aperture masking interferometry capabilities from ∼3–5 µm. The remaining Near-InfraRed Spectrograph (NIRSpec) instrument
does not possess a mode strictly designed for high contrast observations; however, optimised post-processing algorithms (e.g. MEDUSA\textsuperscript{6–8}) may allow for modest contrast measurements to be made with its integral field spectroscopic modes across $\sim$2–5 $\mu$m.

The primary science use case of these high contrast capabilities is to detect and characterise exoplanets, low mass brown dwarfs, and circumstellar disks at the near- to mid-infrared wavelengths available to JWST. JWST will have the largest primary mirror so far for a space-based observatory, and is therefore much more sensitive to these faint objects than past and current observatories such as Hubble and Spitzer. Additionally, due to the cold and atmosphere free environment of the L\textsubscript{2} Lagrange point at which JWST will be located, it will be largely unaffected by limitations of ground-based observatories such as telescope self-emissivity and Earth’s low atmospheric transmittance at particular regions of wavelength space. Such advantages have already been explored and indicate that JWST will offer a unique opportunity to search for and characterise these objects throughout its entire lifetime.$^{9–13}$

Given the highly competitive nature of the JWST proposal process and JWST’s relatively short nominal lifetime of 5–10 years, it is clear that the ability for potential observers to quickly and accurately predict the performance of such observations is critical to not only improving their likelihood of proposal acceptance, but also ensuring that telescope time is not wasted. Here we present an update to the Pandeia Coronagraphy Advanced Kit for Extractions (PanCAKE),\textsuperscript{14–16} a Python-based simulation tool that extends the official JWST exposure time calculator \texttt{Pandeia}\textsuperscript{17} to produce more accurate predictions of JWST coronagraphic performance. This update dramatically improves on its accessibility and functionality for more general users, whilst maintaining and in some cases improving upon the advanced capabilities of its predecessor.

In Section 2 we outline the current available simulation tools relevant to simulations of JWST coronagraphic performance and in Section 3 we describe the performed update to \texttt{PanCAKE}. In Section 4 we provide an example demonstration of the new \texttt{PanCAKE} functionality as applied to the JWST Director’s Discretionary Early Release Science (DD-ERS) observations of the exoplanet HIP 65426 b (PI: S. Hinkley). Finally, in Section 5 we summarise our conclusions and outline future developmental work to be performed.

### 2. SIMULATING JWST OBSERVATIONS

#### 2.1 The JWST Exposure Time Calculator

The most widely used tool for the simulation of JWST observations is the official JWST exposure time calculator, \texttt{Pandeia},\textsuperscript{17} which is accessible through both an online portal\textsuperscript{1}, and a separately distributed Python package\textsuperscript{1}. The primary advantage of \texttt{Pandeia} is its versatility in being able to simulate a broad range of JWST observations, across all available instruments and modes, in a relatively short amount of time. Given the significant diversity in observations that JWST can perform, an all-purpose tool such as \texttt{Pandeia} is highly valuable to the astronomical community as a whole. However, in the case of coronagraphic observations specifically, this versatility comes at the cost of limited simulation accuracy and functionality. In fact, this limitation in accuracy is well understood by the package developers, and the “Known Issues” page as of \texttt{Pandeia} version 1.6.1 states: “This calculation represents the best case scenario for coronagraph observations, and ought to be used with caution...”\textsuperscript{†}. A more precise description of these limitations is provided below:

- **Field Of View** — In order to reduce computational intensity, \texttt{Pandeia} only simulates small “postage stamps” of the true field of view for many of JWST’s instrumental modes. In the case of NIRCam and MIRI coronography, the field of view is reduced to 6.3″ × 6.3″ and 8.8″ × 8.8″ from 20″ × 20″ and 24″ × 24″, respectively. This necessity to use the contrast performance at close separations as approximations for the contrast performance at wider separations will naturally lead to inaccuracies. However, the postage stamps are typically large enough to capture both the contrast- and background-limited regimes of simulated coronagraphic images and significant variations at wider separations are unlikely.

\textsuperscript{‡}https://jwst.etc.stsci.edu/
\textsuperscript{†}https://pypi.org/project/pandeia.engine/
\textsuperscript{†}https://stsci.service-now.com/jwst?id=kb_article&sys_id=b972e4b41b7c7814b8f41d1b24bcb0e (ID: KB0012786)
• **PSF Accuracy** – When generating the constituent point spread functions (PSFs) of a desired coronagraphic simulation, *Pandeia* makes use of a precomputed library of PSFs that are sparsely sampled across the simulated field of view. Although using this library significantly reduces the computational intensity of a simulation, interpolations of the library cannot accurately capture variations in the PSF due to small offsets such as target acquisition errors. Neglecting these variations in the relative PSFs of a target and reference observation, or observations at different roll angles, will lead to over-estimations in the overall contrast performance following the use of subtraction techniques such as reference differential imaging (RDI) and angular differential imaging (ADI).

• **Wavefront Evolution** – Although *JWST* will be placed far beyond the wavefront distorting effects of Earth’s atmosphere, wavefront errors will still arise through effects such as telescope vibrations and the thermal evolution of the telescope.\textsuperscript{15,18} Due to the temporally evolving nature of these effects, the wavefront error experienced between two separate observations (e.g. a target and reference coronagraphic observation) will be different. For coronagraphic observations, this differential wavefront error will lead to variations in the residual stellar PSF, inhibiting our ability to perform an optimal PSF subtraction, and limiting achievable contrast. *Pandeia* offers no functionality to estimate and apply wavefront evolution between observations.

• **Target Acquisition Offsets** – Under true *JWST* operations, the centering of a star behind a coronagraphic mask will be imperfect. This not only leads to sub-optimal masking of the stellar PSF, which introduces more residual stellar flux into the image, but can also drive variations in the measured PSF between separate observations. As PSF subtraction techniques, particularly RDI, rely on observations separate to the target observation to approximate the residual stellar PSF, these variations negatively impact the achievable contrast. At present, the only method of including target acquisition offsets in *Pandeia* is to manually offset all desired sources within a scene by a desired amount, and their influence on the produced PSF will not be simulated accurately due to the use of the precomputed PSF library.

• **Small-Grid Dithers** – In an effort to reduce the impact of target acquisition offsets on *JWST* coronagraphic observations, small-grid dither (SGD) techniques have been developed.\textsuperscript{19,20} In short, SGD makes use of sub-pixel dithers to repeat observations of a reference star and more broadly sample PSF variations resulting from its differential placement behind the coronagraphic mask. This sample of PSFs can then be applied to common subtraction techniques such as KLIP\textsuperscript{21} or LOCI\textsuperscript{22} to optimise the subtraction of the residual stellar PSF. Unfortunately, despite SGD techniques potentially offering up to a factor of $\sim 60$ improvement to the achievable contrast, \textsuperscript{20} *Pandeia* offers no such functionality. However, this is understandable given the previously explained limitations in PSF accuracy and target acquisition offsets.

• **PSF Subtraction Method** – The current version of *Pandeia* can only perform PSF subtraction through simplistic reference differential imaging using a single reference image, there is no functionality for PSF subtraction through ADI or using a collection of reference PSFs. As *Pandeia* presents highly optimistic simulations of coronagraphic performance, more complex subtraction strategies will likely result in marginal increases in the achieved contrast. However, with the inclusion of more complex effects such as wavefront evolution and target acquisition offsets, exploring a broader diversity of subtraction strategies is critical to identifying the limits of the achievable contrast.

2.2 PanCAKE

Given the limitations to *Pandeia* listed above, it is necessary for a simulation tool to exist that can more accurately predict the coronagraphic performance of *JWST*. One such tool is PanCAKE\textsuperscript{14–16} (previously known as *Pandeia–Coronography*), a Python package\textsuperscript{*} which directly extends the capabilities of *Pandeia* to improve upon the accuracy of coronagraphic simulations. Broadly speaking, PanCAKE acts as a structural wrapper to *Pandeia* and offers methods to either build upon or bypass its default calculation procedures. Prior to this update, the construction of a PanCAKE simulation would follow very closely to the construction of a *Pandeia* simulation.

\textsuperscript{*}https://aarynncarter.github.io/PanCAKE

---

Terms of Use: https://www.spiedigitallibrary.org/terms-of-use

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 07 Sep 2021
As PanCAKE is intimately connected to the Pandeia simulation engine, the field of view of all PanCAKE simulations is identical to those simulated by Pandeia, and it is still not possible to directly determine performance at the widest separations. Instead of estimating PSFs solely through the precomputed library of PSFs held within the Pandeia reference data, PanCAKE provides an optional toggle to generate a PSF for a provided observational scene on the fly using the underlying WebbPSF* dependency. Additionally, PanCAKE allows for the injection of a desired input optical path difference map during this process to apply a specific wavefront error to a given observation. Relative offsets between observations, due to differential target acquisition offsets or small-grid dither procedures, can be readily applied using routines based on the current knowledge of JWST performance and operations. Finally, while prior versions of PanCAKE only allowed for a form of RDI PSF subtraction (see Appendix of Carter et al. 2021[13]), the more recent update reported here now includes statistical corrections at small separations, coronagraph throughput corrections, and ADI subtraction procedures.

2.3 Other Available Simulation Tools

Pandeia and PanCAKE are not the only tools available to simulate JWST coronagraphic observations, and a small number of alternative options are currently available or in development. Perhaps the most mature in this regard is pyNRC (Leisenring et. al, in preparation†). pyNRC offers a range of benefits compared to Pandeia such as: simulation of the full available field of view, dynamic and quick PSF computation using polynomial coefficient maps, application and estimation of wavefront evolution between observations, and optimisation of MULTIACCUM ramp parameters. Despite these significant advantages, the most significant downside to pyNRC with respect to JWST coronagraphic observations is that it is a NIRCam specific tool, and cannot simulate any of the MIRI coronagraphic modes. For the more general JWST user, the flexibility in being able to simulate both NIRCam and MIRI observations with PanCAKE may prove particularly advantageous; however, for those considering NIRCam observations specifically, the pyNRC package is a suitable alternative to PanCAKE.

Aside from pyNRC, there are ongoing developmental efforts to produce simulation tools for JWST coronagraphic observations that generate high fidelity data products that closely match those received during true JWST observations. Currently, these efforts are split into separate NIRCam and MIRI specific tools. With respect to NIRCam, these data products are produced through a composite of pyNRC and the Multi-Instrument RAmp GEnerator (MIRAGE) simulations (J. Kammerer, private communication). Conversely, the MIRI data products are produced through a composite of MIRISim and webbpsf_ext simulations (K. Ward-Duong, private communication). Given their early stage of development, it is not clear whether these tools will be suitable alternatives to PanCAKE as an exposure time simulator in terms of ease of use or computational speed. However, they will offer considerable advantages over PanCAKE in producing data products that can be used to test and develop data reduction pipelines.

3. A NEW UPDATE TO PANCAKE

PanCAKE occupies a relatively unique niche in its ability to accurately simulate both NIRCam and MIRI coronagraphic observations for the purpose of assessing observational feasibility and determining appropriate exposure times. However, no tool is perfect, and prior to this update there were a selection of capabilities that, although technically possible to use, were somewhat cumbersome to implement. Additionally, there were a small number of immediately beneficial capabilities that PanCAKE was lacking. It is these deficiencies that the update to PanCAKE presented here seeks to resolve, with more specific details provided below.

3.1 Improved User Interface

As PanCAKE is a direct adaptation of Pandeia, all user inputs need to be provided and executed such that they can be directly interpreted by Pandeia. As Pandeia was designed to simulate all of JWST’s observational modes, simulations for a specific instrumental mode (e.g. NIRCam/MIRI coronagraphy) must be constructed

*https://webbpsf.readthedocs.io/en/latest/
†https://pynrc.readthedocs.io/en/latest/
§https://wiki.miricle.org//bin/view/Public/MIRISim_Public
¶https://github.com/JarronL/webbpsf_ext

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 07 Sep 2021
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
using a range of user-defined input parameters. Furthermore, simulations must be performed independently and more complex sequences of observations can only be performed by manually constructing and executing multiple simulations. For a general, observation non-specific tool, this architecture provides great versatility; however, when considering coronagraphic simulations in particular, it requires a large number of overly specific user inputs. The new iteration of PanCAKE presented here overcomes these issues by providing a user-friendly interface to Pandeia simulation construction and execution in the form of “Scene” and “Sequence” Python classes.

3.1.1 The Scene Class

The construction of a simulation in PanCAKE begins through the creation of astrophysical “scenes” for which synthetic observations can eventually be generated. This nomenclature is identical to that adopted by Pandeia; although PanCAKE restructures the user input framework to add convenience functions and reduce the number of user inputs overall. Separate scenes can be initialised using the Scene class and named as required (e.g. the target and reference scene of a coronagraphic observation), at which point astrophysical objects, or “sources” can be added to those scenes. Currently, PanCAKE only allows for the addition of point sources to these scenes, although methods to add more diffuse emission, such as that from a circumstellar disk, are a desirable future improvement.

The addition of a source to a scene requires a single line of code; however, the precise number of required user inputs depends on how the spectrum of the object will be estimated. In the simplest case, a single string input that corresponds to a valid SIMBAD identifier can be provided. As long as there is an available internet connection, PanCAKE will query SIMBAD directly and then: 1) save a selection of the object’s properties into memory, 2) use the SIMBAD-provided spectral type to identify the closest matching PHOENIX model spectrum included within Pandeia’s data file dependencies, and 3) re-normalise this spectrum to the SIMBAD provided K-band magnitude of the object. Alternatively, the spectral type, normalisation magnitude, and normalisation bandpass for a source can be provided, at which point the closest matching PHOENIX model spectrum will be identified and normalised using the provided normalisation parameters. Finally, the file path of a user input spectrum can be provided, along with the units of the wavelength and flux axes, to directly import any desired spectrum. It is worth noting that for the latter two methods additional properties may also need to be provided (such as the RA and Dec of the object) to fully exploit the more advanced capabilities of PanCAKE.

3.1.2 The Sequence Class

Following the creation of scenes, consecutive observations of these scenes in different observational modes can be defined using the “Sequence” class. Once initialised, observations are added to a sequence according to the chronological order in which they will be executed. In the simplest cases this order will not be important, although if wavefront evolution effects are included the order of observations should be carefully considered. Much like sources in scenes, observations are added to sequences with a single line of code; however, whereas only one source could be added to a scene, multiple observations of a scene using any number of coronagraphic filters across NIRCam and MIRI, and for any number of telescope roll angles, can be added simultaneously. The number of groups, integrations, and readout pattern of JWST’s MULTIACCUM detector readout for each desired observation can either be provided to PanCAKE manually, or estimated using the default optimisation procedures described in Section 3.2. If necessary, specificity in the precise NIRCam and MIRI coronagraphic masks, detector sub-arrays, and small grid dither patterns can also be provided.

Once a sequence has been fully defined, all of the constituent simulations for this sequence can be executed, or “run”, simultaneously within a single line of code. Data products from these simulations are returned directly within the code for further processing, but if desired, they can be saved to a FITS file as they are completed. In the case of large sequences, or sequences using advanced functionality, the computational time is non-negligible and saving prevents unnecessary repetition of simulations. Additionally, in the event of a failed simulation, simulations can be resumed following the last completed and saved simulation. It is also at this point that the more advanced features of PanCAKE, such as on the fly PSF calculations or dynamic wavefront evolution, can be toggled on or off, and their underlying parameters adjusted. Importantly, the running of a sequence does not alter the sequence itself in any way, so multiple runs of a sequence (e.g. with different degrees of application of advanced features) can be performed with relative ease.
3.2 Readout Parameter Optimisation

The readout of JWST detectors can be adjusted by varying the number of groups, number of integrations, or the readout pattern itself (see https://jwst-docs.stsci.edu/understanding-exposure-times for further details). Not only do these properties have a wide range of possible values, but the readout patterns themselves differ between all four of JWST’s instruments. This breadth allows for significant customisation to the observational structure as a whole, but can be slightly mystifying when attempting to identify “optimal” readout parameters for a specific observation. To alleviate this source of confusion for more novice users, PanCAKE now includes rudimentary methods to identify such parameters for both NIRCam and MIRI when provided a desired observation length in seconds.

These optimisations are built upon the assumption that the best overall contrast will be achieved with a combination of readout parameters that results in the largest amount of time between detector resets (i.e. the longest integration times) whilst simultaneously avoiding detector saturation and/or significant cosmic ray contamination. Such constraints maximise the number of photons detected from the astrophysical scene, whilst also minimising the impact of read noise. In reality, it may be that better contrasts can be obtained at wider separations by allowing the innermost regions of the image to saturate, and shorter integrations may be desirable if there is a need to obtain more integration-level images within a defined timescale. Use of this optimisation procedure should be considered carefully; however, in standard circumstances this assumption is unlikely to result in readout parameter estimations that are significantly different from the true optimal values. We describe the steps of this optimisation below, in addition to displaying them visually in Figure 1.

3.2.1 NIRCam

To begin, a simplistic Pandea simulation of a given observation is performed with readout parameters set to one integration, two groups, and the RAPID readout pattern. This corresponds to the shortest possible integration length, and if saturation is present at this stage the overall feasibility of the observation may need to be reconsidered depending on the desired scientific goals. From this simulation, an estimate of the integration length necessary to reach a defined maximum fraction of detector saturation, \( f_{\text{sat,max}} \), is calculated following

\[
t_{\text{sat}} = \frac{2t_{\text{frame}} f_{\text{sat,max}}}{f_{\text{sat}}}, \tag{1}
\]

where \( t_{\text{frame}} \) is the subarray frame time, and \( f_{\text{sat}} \) is the measured level of detector saturation in the simulated image. Importantly, \( f_{\text{sat}} \) does not need to be <1 and partial detector saturation can also be explored.

Prior to assessing the full range of NIRCam readout patterns at reaching this integration time, the BRIGHT1, SHALLOW2, MEDIUM2, and DEEP2 are eliminated from consideration as their BRIGHT2, SHALLOW4, MEDIUM8, and DEEP8 counterparts have fewer skipped frames and in the vast majority of situations optimising for a lower read noise is preferable to optimising for the rejection of cosmic rays (Robberto et al. 2009*). The time for a single integration,

\[
t_{\text{int}} = t_{\text{frame}} \left[ (n_{\text{frame}} + n_{\text{skip}})(n_{\text{group}} - 1) + n_{\text{frame}} \right], \tag{2}
\]

where \( n_{\text{frame}} \) is the number of averaged samples in the readout pattern, \( n_{\text{skip}} \) is the number of skipped samples in the readout pattern, and \( n_{\text{group}} \) is the input number of groups. Starting with the longest readout pattern, the exposure time for a single integration with two groups (i.e. the shortest possible ramp) is calculated following Equation (2). If this time is longer than the original desired exposure time, \( t_{\text{exp}} \), plus a user-defined margin of error, \( t_{\text{margin}} \), or longer than \( t_{\text{sat}} \), no exposure with this readout pattern will be viable and the next longest readout pattern is considered. If this time is shorter than the original desired exposure time, PanCAKE then determines the maximum possible number of groups for this readout pattern and observation,

\[
n_{\text{group,max}} = \min \left( \left[ \frac{t_{\text{sat}} - n_{\text{frame}} t_{\text{frame}}}{t_{\text{frame}} (n_{\text{frame}} + n_{\text{skip}})} + 1 \right], \left[ \frac{t_{\text{exp}} + t_{\text{margin}} - n_{\text{frame}} t_{\text{frame}}}{t_{\text{frame}} (n_{\text{frame}} + n_{\text{skip}})} + 1 \right], n_{\text{group,lim}} \right), \tag{3}
\]

* JWST-STScI-001721: NIRCam Optimal Readout Modes
Figure 1. Block diagram representation of the readout parameter optimisation procedure as described in Section 3.2. Start and end points are marked in purple, failure points are marked in red, NIRCam specific procedures are marked in blue, MIRI specific procedures are marked in orange, and instrument non-specific procedures are marked in green. Directional arrows are not provided for each step for the sake of clarity, although the sequence begins in the top left and broadly flows from top to bottom, and dashed arrows correspond to negative responses to procedures.
where \( n_{\text{group,lim}} \) is the instrumental limit of the number of groups for this readout pattern, and the independent fractions are obtained by rearranging Equation (2). Beginning with this maximum number of groups, the possible number of integrations that correspond to total exposure times between \( t_{\text{exp}} - t_{\text{margin}} \) and \( t_{\text{exp}} + t_{\text{margin}} \) is calculated. If no possible number of integrations fits within this range, this process is repeated with one fewer group per integration, and if no number of groups fits within this range, the next longest readout pattern is considered. If instead there are possible numbers of integrations that fit within the range, that which corresponds to an exposure time closest to \( t_{\text{exp}} \), along with the underlying number of groups and readout pattern, are adopted as the “optimal” readout parameters for this observation. Careful consideration should be made when selecting the value for \( t_{\text{margin}} \), as too narrow a range in acceptable exposure times will bias the calculation towards shorter readout patterns or a smaller number of groups.

3.2.2 MIRI

For MIRI, the procedure for optimisation of readout parameters is much more straightforward than that of NIRCam, as it relies on a number of coronographic and cross-mode recommended strategies*. Specifically, MIRI only has two possible readout patterns, FAST or SLOW, and FAST is ubiquitously recommended for MIRI coronagraphic modes as the primary function of the SLOW readout pattern is to reduce data volume. Data volume considerations are largely irrelevant for MIRI coronagraphic modes as they make extensive use of subarrays and therefore have a comparatively low data output. Following this recommendation, only the FAST readout pattern is considered within PanCAKE.

Optimisation begins in a similar manner to NIRCam, where a simplistic Pandeia simulation of a given observation using one integration, two groups, and the FAST readout pattern is performed. Again, if saturation is present at this stage the overall feasibility of the observation should be reconsidered. As only the FAST readout pattern is considered, the equation to determine the maximum possible number of groups, \( n_{\text{group, max}} \), now follows:

\[
\begin{align*}
    n_{\text{group, max}} &= \min \left( \left\lfloor \frac{2f_{\text{sat, max}}}{f_{\text{sat}}} \right\rfloor, \frac{t_{\text{exp}}}{t_{\text{frame}}} \right) 
\end{align*}
\]

where \( t_{\text{cosmic}} \) is the recommended limit on the exposure time of a single integration to avoid significant contamination due to cosmic rays (~300 s)*. At this stage, the “optimal” readout parameters for observation are obtained in an identical manner to the NIRCam optimisation procedure, where the number of groups is sequentially decreased from the maximum until an acceptable number of integrations is identified that produces an exposure time between \( t_{\text{exp}} - t_{\text{margin}} \) and \( t_{\text{exp}} + t_{\text{margin}} \).

3.3 Dynamic Wavefront Evolution

During JWST operations, the structure of the optical telescope element (OTE) will not be static, and instead will experience distortion resulting from its varying thermal state in response to telescope slwrs, or small vibration-driven tension variations. These effects are very small and do not impact the structural integrity of the telescope; however, they can induce changes in the wavefront error across a variety of different timescales.\(^\text{15}\) Prior to this update it was possible to incorporate these changes into PanCAKE, although the actual implementation was not straightforward to perform. PanCAKE now includes a procedure to dynamically estimate and apply wavefront variations for all observations within a given sequence with a simple toggle, which we describe in more detail below.

The process to apply wavefront evolution to a sequence of observations can only be performed if: a) on the fly PSF calculations are being performed, and b) at least one source for each scene used in the sequence has an assigned RA and Dec. If these conditions are satisfied, PanCAKE is able to calculate variations in the wavefront error between observations using an incorporated version of pyNRC’s OTE\_WFE\_Drift\_Model class. Based on a collection of input observatory pitch angles and elapsed times, this class provides functionality to estimate the expected sources of OTE wavefront drift, specifically: thermal changes in the OTE state, vibrations due to the instrument electronics compartment (IEC) heaters, and vibrations due to the OTE stray light baffle/insulation closeouts (commonly known as the “frill”)*\(^\text{15}\), and apply them to a current estimate of the wavefront. We note

\(^*\text{https://jwst-docs.stsci.edu/mid-infrared-instrument/miri-observing-strategies}\)
that the vibrations due to the electronics heaters are predicted to be an effect of the ground testing support structure and should not be present after launch, although this drift is a useful proxy for potential unknown sources of wavefront drift at short (~minutes) timescales. The underlying WebbPSF dependency provides 10 individual realisations of the wavefront which can be used as a starting point for the evolution; however, varying these realisations does not significantly affect the achieved contrast.\textsuperscript{13} Nevertheless, PanCAKE allows for a specific realisation to be used if desired.

To obtain realistic estimates of the required pitch angle for each scene within a sequence, PanCAKE makes use of the physical locations of these scenes relative to the location of JWST itself. As the distance from Earth to JWST is much smaller than the distance to an arbitrary point on the celestial sphere, the pitch angle required to observe such a point is equivalent to its latitude following a 90° rotation of the ecliptic coordinate system about a line tangent to JWST’s orbit at its current location (see Figure 2a). Under this prescription, the required pitch angles, $\theta_{p,x}$, for a collection of input astrophysical scenes, $x$, are calculated using:

$$\theta_{p,x} = \arcsin \left( \cos(\beta_x) \sin(\lambda_x + \lambda_J + \pi/2) \right),$$

where $\beta_x$ are the ecliptic latitudes of the scenes, $\lambda_x$ are the ecliptic longitudes of the scenes, $\lambda_J$ is the ecliptic longitude of JWST, and the $\pi/2$ accounts for the fact that at a pitch of 0° JWST points perpendicular to the ecliptic longitude at which it is located. Crucially, as $\lambda_J$ varies depending on JWST’s orbital location around the Sun, the required pitch angle to observe an astrophysical scene must also vary correspondingly (see Figure 2b). To identify a single set of pitch angles for all observations within the sequence, PanCAKE first identifies all possible combinations of $\theta_{p,x}$ throughout JWST’s orbit by varying $\lambda_J$ from 0–360° at a resolution of 0.1°. All combinations that include a pitch angle outside JWST’s allowable range of $-5^\circ < \theta_p < 45^\circ$ are impossible to perform, and are therefore rejected. If all combinations of pitch angles are rejected, the structure and overall feasibility of the sequence should be reconsidered. However, if viable combinations remain, the one with the minimum, median, or maximum range in pitch angle values can be adopted to estimate the wavefront evolution.

The elapsed time between observations is a function of the exposure times of these observations and the time necessary to perform slews between them. Importantly, slews are not identical to changes in the observatory pitch angle, and instead correspond to a movement of JWST’s orbit by varying $\lambda_J$ from 0–360° at a resolution of 0.1°. All combinations that include a pitch angle outside JWST’s allowable range of $-5^\circ < \theta_p < 45^\circ$ are impossible to perform, and are therefore rejected. If all combinations of pitch angles are rejected, the structure and overall feasibility of the sequence should be reconsidered. However, if viable combinations remain, the one with the minimum, median, or maximum range in pitch angle values can be adopted to estimate the wavefront evolution.

3.4 pyKLIP PSF Subtraction & Contrast Estimation

The generation of accurate simulated coronagraphic observations alone does not give a complete picture of the overall contrast performance for those observations. Typically, significant image post-processing is necessary to subtract the residual stellar PSF and reach the limiting contrast for a given instrument or mode. In the case of JWST coronagraphy this subtraction will most likely follow an ADI or RDI procedure (or a combination of the two), although ADI procedures alone will be somewhat limited given JWST will at best be able to perform an

\textsuperscript{13}https://jwst-docs.stsci.edu/jppom/visit-overheads-timing-model/slew-times
Figure 2. 2-dimensional representations of the observable field of view of JWST as projected onto the celestial sphere (blue shaded region). The relative sizes and separations of the Sun (orange circle), Earth (blue and green circle), and JWST (orange hexagon) are not to scale and are solely for illustrative purposes. When the ecliptic plane is viewed side on, as in a), it is clear to see that the pitch angle required to observe a particular point on the celestial sphere will correspond to its latitude following a 90° rotation of the ecliptic coordinate system about a line tangent to JWST’s orbit at its current location (i.e. the axis into the page at the location of JWST). In b), where the ecliptic plane is now viewed face on, the orbital paths of the Sun, Earth, and JWST around their common center of mass have been shown. From this perspective it follows naturally that the field of view of JWST will evolve throughout its orbit, which in turns adjusts the pitch angle required to observe a particular point on the celestial sphere. In both examples the ability of JWST to slew a complete 360° around the Earth-Sun-JWST axis is also indicated.

absolute roll of 14°. PanCAKE now includes functionality to perform these PSF subtraction procedures following a Karhunen-Loève Image Processing (KLIP)\textsuperscript{21} methodology using the pyKLIP Python package.\textsuperscript{32,33} Figure 3. As a point of clarity, we note that the ADI+RDI subtraction is achieved by including images at other roll positions in the KLIP modelling and subtraction of the stellar PSF, followed by derotating all of the images and averaging them together.

Following KLIP PSF subtraction, contrast curve metrics of the coronagraphic performance can also be quickly calculated for any given set of observations within a PanCAKE observational sequence. These contrast curves are additionally calibrated for the 2-dimensional throughput of the coronagraphs themselves\textsuperscript{10,34} (MIRI throughputs: C. Cossoi, R. Gastaud, private communication), the intrinsic throughput of the KLIP subtraction process, and the effects of small sample statistics at the shortest separations.\textsuperscript{35} In actuality, PanCAKE will return a selection of contrast curves from different stages in the calibration process; however, particular emphasis should be placed on the final produced contrast and the more naive contrast calculated without a correction for small sample statistics. For a selection of PanCAKE simulations, particularly those performed with MIRI, the residual noise in the simulated images following KLIP PSF subtraction transitions from being dominated by the speckle noise of the residual stellar PSF, to the background noise of the astrophysical scene (e.g. Figure 3). In such a circumstance, the spatial scale of residual noise is closer to \(\sim 1\) pixel and corrections for speckle-driven small sample statistics, which have much larger spatial scales, will result in significant underestimations of the coronagraphic performance. We do not attempt to constrain the observational regimes of this transition within this work; however, at present we do encourage any future users of PanCAKE to assess the structure of the residual noise in their post-subtraction simulated images and decide whether the application of a small sample statistics correction is justified.
4. EXAMPLE WORKFLOW: ERS OBSERVATIONS OF HIP 65426B

To better demonstrate the new functionality of PanCAKE, here we describe in detail how to produce estimated contrast curves for a subset of the JWST DD-ERS coronagraphic observations of the exoplanet HIP 65426 b. As with all Python packages, at first the package must be imported:

```python
[1] import pancake
```

and with this, the user has complete access to the capabilities of PanCAKE. To begin a simulation, we must create the astrophysical scene that we would like to simulate an observation for. Note that at this stage the scene is observation non-specific, and can be used for any possible coronagraphic simulation.

```python
[2] target = pancake.scene.Scene('Target')
[3] target.add_source('HIP 65426', kind='simbad')
[4] target.add_source('HIP 65426b', r=0.83, theta=150, kind='file', filename='model.txt',
                   wave_unit='micron', flux_unit='Jy')
```

Here we have defined the sources within this initial target scene through two different methods. In the first we have passed a string identifier for the star HIP 65426 which can be interpreted by SIMBAD. At this point a query to SIMBAD is performed, and a stellar model spectrum is automatically determined following the description in Section 3.1.1. In the second, we have provided a spectrum for the exoplanet HIP 65426 b directly into PanCAKE, defined the units of its wavelength and flux axes, and also specified its polar coordinates. Sources that do not have their polar coordinates explicitly defined are simply placed at the center of the image. As the DD-ERS program also includes corresponding observations of a reference star, we need to create a scene for this also:

```python
[5] ref = pancake.scene.Scene('Reference')
[6] ref.add_source('HIP 68245', kind='simbad')
```

Figure 3. Example PanCAKE simulations followed by a variety of pyKLIP PSF subtractions for the HIP 65426 system, with the companion HIP 65426 b located at 0.83′′, 150°. Simulations in each filter are performed at 0° and 14° roll angles to enable ADI, 9-point small grid dithers on a separate HIP 68245-like system are also to enable RDI, and the full sequence of observations includes on the fly PSF generation and wavefront evolution following Section 3.3.

Downloaded From: https://www.spiedigitallibrary.org/conference-proceedings-of-spie on 07 Sep 2021
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
where again the source has been added to the scene through a simple query to SIMBAD. The defined names of ‘Target’ and ‘Reference’ for these scenes are arbitrary and are only used to identify their respective observations at different stages in the contrast curve analysis. Whilst not shown here, we also note that there is functionality within PanCAKE to automatically plot the spectra or spatial locations of all sources within a scene to ensure that the scenes have been correctly constructed.

With all necessary scenes defined, and their constituent sources added, an observational sequence can be initialised:

```
[7] seq = pancake.sequence.Sequence()
```

Following this, we can add observations of our previously defined scenes to this sequence in the chronological order in which we want them to occur:

```
[8] seq.add_observation(target, exposures=[('F1550C', 'fast', 250, 60), ('F444W', 'deep8', 15, 2)],
                   nircam_mask='MASK335R', rolls=[0,14])
[9] seq.add_observation(ref, exposures=[('F1550C', 'fast', 100, 19), ('F444W', 'medium8', 4, 2)],
                   nircam_mask='MASK335R', nircam_sgd='9-POINT-CIRCLE', miri_sgd='9-POINT-SMALL-GRID')
```

Here we have constructed exposures in the MIRI F1550C and NIRCam F444W filters using the precise readout parameters from the ERS observation specifications in the format (FILTER, PATTERN, GROUPS, INTEGRATIONS). If instead we did not know these precise parameters, but were interested in what an exposure of a given duration might look like, we could easily redefine the exposure to (FILTER, ‘optimise’, T_EXP) to apply the optimisation schemes described in Section 3.2. In the case of NIRCam, multiple coronagraphic masks can be used for a given filter and therefore the desired mask must be explicitly defined. Separate observations should be defined if the use of multiple NIRCam coronagraphs is desired. Conversely, the MIRI coronographs are tied to the desired coronagraphic filter and a precise definition is not necessary. For the target observation, roll angles relative to the initial scene construction have been explicitly defined. With this simple definition, simulations of all of the defined exposures will be performed for each of these roll angles. Similarly, the small-grid dither patterns defined for the reference observation will ensure that all simulations are repeated for each of the individual small-grid dither positions. Finally, we note that the default behaviour of PanCAKE is to group observations by the coronagraph used, then perform the corresponding rolls. More specifically, all observations for the coronagraph corresponding to the first listed filter will be executed, then their corresponding rolls (if any), then the observations for the second coronagraph, then the second coronographs corresponding rolls (if any), et cetera. In situations where this is not desired, more specific sequences can be defined by separating the exposures into different observations.

Once all of the observations of interest have been added to a given sequence, all of the simulations for this sequence can be performed in a single line of code.

```
[10] results = seq.run(on_the_fly_PSFs=True, wavefront_evolution=True, wave_sampling=11,
                    save_file='results.fits')
```

Here we have opted to enable some of PanCAKE’s advanced functions, specifically: on the fly PSF generation to improve PSF accuracy, and dynamic wavefront evolution following Section 3.3. The wave_sampling argument sets the number of monochromatic PSF’s that will be generated to estimate the overall PSF for a given bandpass. In this circumstance (22 independent simulations) the entire execution took ∼3 hours on a laptop with a 2 Ghz processor and 32 GB of RAM. However, the dominant contribution to this time is the on the fly generation of PSFs and a more rudimentary “quick-check” assessment of observation feasibility can be obtained in a few minutes by setting on_the_fly_PSFs=False.

All output simulations are saved as they are calculated to a user-defined save file, although it is not necessary to do so. Finally, with these simulated images, we can determine the contrast curves for any given subset of observations within the sequence.
To compute the contrast, PanCAKE must be provided contextual information on a “target” scene (from which the residual stellar PSF will be subtracted), and for RDI procedures, a collection of “reference” scenes (from which the residual stellar PSF will be estimated). Additionally, the user must define which subtraction methodology to use. Beyond this, although not shown in the code fragment above, the precise application of the KLIP algorithm (e.g. number of basis vectors, annuli, and subsections) can also be fine tuned, as well as which specific rolls to use for ADI procedures. The execution time of the contrast curve analysis depends on the underlying KLIP parameters, but takes only a few minutes when using the default parameters (25 basis vectors, 1 annulus, 1 subsection). Here we have computed contrasts for both of the filters used for observations of the target scene; however, it is also possible to specify a subset of filters for a given sequence if desired. We show the resultant contrast curves for this calculation (in addition to ADI and RDI subtractions) in Figure 4, and corresponding images are shown in Figure 3.

It is not the intent of this work to formally assess variations in JWST performance, although it is worth describing some of the features observed in the simulated images and contrast curves. Compared to the RDI subtraction, both the ADI and ADI+RDI subtraction have inferior contrasts at shorter separations. This stems from the variation in the residual stellar PSF between different ADI rolls, which is an effect of their different locations underneath the coronagraphic mask, in addition to the evolving wavefront. For the case of ADI+RDI specifically, it may be that including the additional roll image in the PSF library used for the KLIP RDI process is negatively impacting its ability to estimate the residual stellar PSF, although further work is required to explore this possibility in detail. In contrast, at wider separations where the residual stellar PSF is less prominent, the ADI and ADI+RDI subtractions begin to outperform the RDI subtraction. At such separations the contrast is no longer contrast limited, but is instead limited by background noise, and the additional roll angle acts a proxy for increasing the total exposure time, thereby improving the signal to noise.

As might be apparent, the steps involved in performing a PanCAKE simulation are modular. Running a sequence of observations will not alter the underlying scenes within it in any way, and computing contrast curves for a simulated sequence will not directly alter any of its simulated images. Additionally, aside for computational

![Figure 4](https://www.spiedigitallibrary.org/conference-proceedings-of-spie)

Figure 4. Estimated contrasts for observations of the HIP 65426 exoplanetary system as simulated by PanCAKE. Simulations in each filter are performed at 0° and 14° roll angles to enable ADI, 9-point small grid dithers on a separate HIP 68245-like system are also to enable RDI, and the full sequence of observations includes on the fly PSF generation and wavefront evolution following Section 3.3. Dashed lines indicate the annular 5σ standard deviation in the image, whereas solid lines indicate a more formal estimation of contrast that applies small sample statistical corrections for a 3 × 10⁻⁷ false positive fraction. The estimated contrasts of the companion HIP 65426 b compared to the host star are also indicated.
limitations, there is no limit to the number of exposures or scenes that can be included in a given sequence. This framework allows for much more complex sequences of observations than this example case to be simulated and analysed with relative ease.

5. CONCLUSIONS AND FUTURE WORK

We present in this work a new update to the Python package PanCAKE and its application towards simulating high contrast imaging observations of JWST’s coronagraphic modes. For such modes, the default JWST exposure time calculator, Pandeia, lacks in both simulation accuracy and user functionality, and in many cases is not appropriate for obtaining valid estimates of their performance. PanCAKE alleviates these concerns by building a framework around Pandeia that allows for further advanced functionality, such as: on the fly PSF generation, wavefront evolution, small grid dithers, and more rigorous PSF subtraction. Furthermore, PanCAKE now includes a wide range of usability improvements, such as: automatic SIMBAD target querying, readout parameter optimisation, and easily expandable observational sequence construction. In totality, PanCAKE offers a user-friendly and easily accessible way to accurately simulate both NIRCam and MIRI high contrast imaging observations with JWST – a capability of which to date no single tool has been able to provide.

Although this update to PanCAKE is a significant upgrade over its predecessor, there exist a range of additional features that would drastically improve its functionality and increase the fraction of the high contrast imaging community that can apply it to their observations or proposal ideas. Whilst we cannot guarantee that such improvements will be made, we outline those with the highest priority below, and welcome suggestions or collaborations from other members of the community:

- **Complex Astrophysical Sources** – At present, only point sources can be included within PanCAKE simulations. With observations of circumstellar disks constituting a significant portion of high contrast imaging observations as a whole, the inability to include extended emission sources, such as circumstellar disks, is a considerable drawback in PanCAKE’s functionality. The primary limitation in including these sources is that PanCAKE is intimately connected to Pandeia, and at present Pandeia only allows for the inclusion of point sources or non-annular disk emission. Potential avenues of development in this regard include: approximating extended emission through a collection of point sources, creating a new annular emission source type that can be incorporated into a Pandeia simulation, constructing a method to directly convert a 2-dimensional image as an input scene.

- **Faster PSF Calculations** – Without any advanced features, the running of a single PanCAKE simulation is relatively fast (~minutes); however, the process of computing PSFs on the fly using WebbPSF is particularly intensive and can extend the single simulation time dramatically (~hours). A possible improvement in this regard would be to utilise a similar methodology to pyNRC and webpsf_ext, which use polynomial fits to each individual pixel within a collection of initial monochromatic WebbPSF simulations to quickly generate a wide range of subsequent simulations.

- **Non-Coronagraphic Imaging** – The use of coronagraphs will be particularly valuable for a wide range of high contrast observations with JWST, although in some situations in may be beneficial to explore the possibility of imaging without a coronagraph. As Pandeia can readily handle such a calculation, the only limitation towards performing a non-coronagraphic imaging simulation is that the current PanCAKE sequence class as described in Section 3.1 has no functionality to construct observations in these modes. Fortunately, PanCAKE has been designed in such a way that the addition of such functionality should not represent a considerable technical challenge.

**Acknowledgements**

We thank Rebecca Jensen-Clem for useful discussions on image noise regimes following PSF subtraction and their impact on contrast estimation. This project was supported by a grant from STScI (JWST-ERS-01386) under NASA contract NAS5-03127.
REFERENCES


