Publications of the Astronomical Society of Australia (2026), **43**, e075, 34 pages
doi:10.1017/pasa.2026.10194

CAMBRIDGE UNIVERSITY PRESS

Research Article

AMIGO: A data-driven calibration of the JWST interferometer

Louis Desdoigts[1,2], Benjamin Pope[3,4], Max Charles[1], Peter Tuthill[1], Dori Blakely[5,6], Douglas Johnstone[5,6], Shrishmoy Ray[3,4], Anand Sivaramakrishnan[7,8,9], Kevin Volk[7], Jens Kammerer[10], Deepashri Thatte[7] and Rachel Cooper[7]

[1]Sydney Institute for Astronomy, School of Physics, University of Sydney, Camperdown, NSW 2006, Australia, [2]Leiden Observatory, Niels Bohrweg 2, Leiden 2300RA, The Netherlands, [3]School of Mathematical & Physical Sciences, Macquarie University, 12 Wally's Walk, Macquarie Park, NSW 2113, Australia, [4]School of Mathematics & Physics, University of Queensland, St Lucia, QLD 4072, Australia, [5]Department of Physics and Astronomy, University of Victoria, 3800 Finnerty Road, Elliot Building, Victoria, BC V8P 5C2, Canada, [6]NRC Herzberg Astronomy and Astrophysics, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada, [7]Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA, [8]Astrophysics Department, American Museum of Natural History, 79th St at CPW, New York, NY 10024, USA, [9]Department of Physics and Astronomy, Johns Hopkins University, 3701 San Martin Drive, Baltimore, MD 21218, USA and [10]European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany

Abstract

The James Webb Space Telescope (JWST) hosts a non-redundant Aperture Masking Interferometer (AMI) in its Near Infrared Imager and Slitless Spectrograph (NIRISS) instrument, providing the only dedicated interferometric facility aboard – magnitudes more precise than any interferometric experiment previously flown. However, the performance of AMI (and other high resolution approaches such as kernel phase) in recovery of structure at high contrasts has not met design expectations. A major contributing factor has been the presence of uncorrected detector systematics, notably charge migration effects in the H2RG sensor, and insufficiently accurate mask metrology. Here we present AMIGO, a data-driven calibration framework and analysis pipeline that forward-models the full JWST AMI system – including its optics, detector physics, and readout electronics – using an end-to-end differentiable architecture implemented in the JAX framework and in particular exploiting the ∂LUX optical modelling package. AMIGO directly models the generation of up-the-ramp detector reads, using an embedded neural sub-module to capture non-linear charge redistribution effects, enabling the optimal extraction of robust observables, for example kernel amplitudes and phases, while mitigating systematics such as the brighter-fatter effect. We demonstrate AMIGO's capabilities by recovering the AB Dor AC binary from commissioning data with high-precision astrometry, and detecting both HD 206893 B and the inner substellar companion HD 206893 c: a benchmark requiring contrasts approaching 10 mag at separations of only 100 mas. These results exceed outcomes from all published pipelines and re-establish AMI as a viable competitor for imaging at high contrast at the diffraction limit. AMIGO is publicly available as open-source software community resource .

Keywords: Optical Interferometry; Astronomical Detectors; James Webb Space Telescope; Astrostatistics; Neural Networks

(Received 29 June 2025 UTC; revised 31 March 2026 UTC; accepted 6 April 2026 UTC)

1. Introduction

Direct imaging of exoplanets and their environments against the glare of the host star makes extreme demands on precision in calibration of optics, electronics, and computational data analysis. Recovering signals requires simultaneous high contrast and angular resolution with performance levels dictated by the planet's relative faintness and proximity to their host-stars (Follette 2023), being limited mainly by systematics from optical aberrations and detector electronics. Coronagraphic methods can achieve contrast ratios of 10^{-4}–10^{-6}, and are particularly favoured for deployment on modern space observatories like JWST (Gardner et al. 2006; Gardner et al. 2023) that avoid speckle noise arising from atmospheric turbulence. Despite proven performance at high contrasts, coronagraphs remain limited by their Inner Working Angle (IWA, with best performance usually found beyond $> 2\lambda/D$ (Guyon et al. 2006). This inability to study the inner structures of extrasolar systems, regions crucial to the understanding of exoplanetary formation (Wagner, Apai, & Kratter 2019), leaves a glaring observational gap at high angular resolutions.

Aperture masking interferometers provide imaging capabilities at and beyond the classical diffraction limit $\sim \lambda/D$ (Monnier 2003), filling the observational gap left at high angular resolutions. Non-redundant masks (NRMs) propagate interferometric phase information from each mask hole to a unique location in the Fourier plane, enabling precise calibration and subtraction of instrumental effects. Furthermore, closing triangles and squares in the aperture yield closure-phases (Jennison 1958) and amplitudes (Twiss, Carter, & Little 1960; Readhead et al. 1980), respectively, observables that are robust to low-order wavefront and amplitude errors. When combined with short exposures, they enable high angular resolution images through the phase corruption of atmospheric turbulence (Baldwin et al. 1986; Robertson et al. 1991). Further advances have produced images in the Near Infrared (NIR) at the diffraction limit (Monnier, Tuthill, & Danchi 1999; Tuthill et al. 2000), cementing these methods as the only proven

Corresponding author: Louis Desdoigts, Emails: desdoigts@strw.leidenuniv.nl, louis.desdoigts@sydney.edu.au.

Cite this article: Desdoigts L, Pope B, Charles M, Tuthill P, Blakely D, Johnstone D, Ray S, Sivaramakrishnan A, Volk K, Kammerer J, Thatte D and Cooper R. (2026) AMIGO: A data-driven calibration of the JWST interferometer. *Publications of the Astronomical Society of Australia* **43**, e075, 1–34. https://doi.org/10.1017/pasa.2026.10194

© The Author(s), 2026. Published by Cambridge University Press on behalf of Astronomical Society of Australia. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

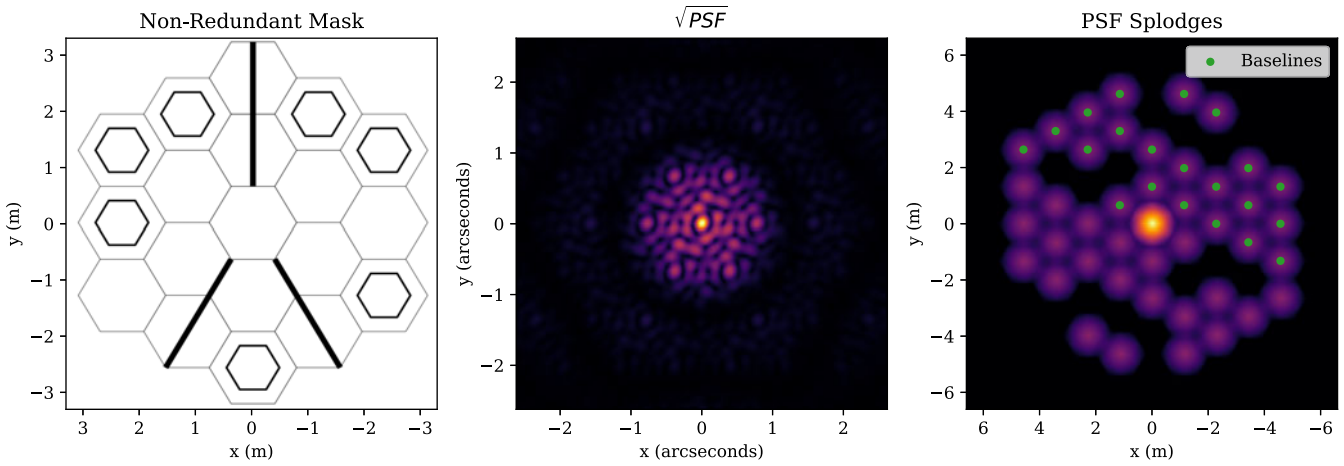


Figure 1. Left panel: Schematic diagram of the 7-hole NRM projected over the primary mirror. Middle panel: The resulting PSF (i.e. interferogram) from the non-redundant mask, visualised on a square-root scale to highlight low-power features. Right panel: The power-spectrum of the PSF featuring baseline-specific regions of fringe power known in the literature as *splodges* that can be conveniently found by Fourier transform of the PSF. The 21 discrete non-redundant baselines are indicated by the overlaid green dots.

way to simultaneously image at both high contrasts and angular resolutions through wavefront phase errors.

Non-redundant and sparse aperture configuration have been used to great success on various ground-based instruments. The Keck observatory (Wizinowich et al. 2000) host a 9-hole NRM that has been used to study planetary formation and disk structures (Willson et al. 2016; Sallum et al. 2023). The Very Large Telescope also has a various NRM modes on the NACO and SPHERE instruments (Lacour et al. 2011; Cheetham et al. 2016; Beuzit, J.-L. et al. 2019) that have been used for similar transition disk and high-angular resolution studies of exoplanetary formation (Lacour et al. 2011; Blakely et al. 2022; Stolker et al. 2024). Many observatories also employ sparse aperture configuration and beam combiners such as the Large Binocular Telescope (Hinz et al. 2016), the Center for High Angular Resolution Astronomy array (ten Brummelaar et al. 2005), and the Very Large Telescope Interferometer (Haguenauer et al. 2010) that allow high angular resolution observations that are resilient to uncorrected residual atmospheric turbulence.

The JWST NIRISS (Doyon et al. 2012; Doyon et al. 2023) also hosts an NRM, providing an AMI observational mode – the first of its kind aboard a space observatory (Sivaramakrishnan et al. 2023; Sivaramakrishnan et al. 2012; Soulain et al. 2020). While the Hubble Space Telescope (HST) hosts an interferometer in each arm of its fine guidance system (Jefferys et al. 1985), AMI on JWST is the first of its kind capable of complex imaging. NRM systems have hitherto been used to mitigate atmospheric turbulence, prompting the question as to their value on a space observatory without a turbulent atmosphere corrupting its wavefronts. However, despite its optical stability, JWSTs segmented aperture demanded precise phasing of the mirrors after launch, a problem not well addressed by in-focus, clear-pupil imagers due to the redundancy of wavefront phase information, but well suited to non-redundant apertures (Cheetham et al. 2012). The NRM configuration allows AMI to act a proven alternate wavefront calibration device that also provides unique scientific capabilities – and was added to NIRISS' suite of observing modes after JWST's preliminary design review. AMI provides a complementary role to the various coronagraphic modes and explores a search space within the IWAs of JWST's coronagraphs, reaching the snow line of nearby exoplanet systems (Ray et al. 2023).

1.1. JWST AMI: A unique space-based interferometer

JWSTs AMI mode employs a 7-hole NRM to form a stable interferometric Point Spread Function (PSF) suitable for classic interferometric analysis methods. Its configuration, shown in Figure 1 provides approximately even sampling in the Fourier plane, and primarily observes in three medium band filters from ∼3–5 μm.

Interferometric data is analysed in the Fourier plane – hereby referred to as the *uv*-plane – by examining its complex Fourier coefficients described as its *complex visibilities*. Adjacent observations of calibrator stars enable the subtraction of both the host star and instrumental signals via a division of these visibilities in complex form, ideally revealing the minute signals of the near-stellar environment.

Various pipelines exist to extract the interferometric observables from calibrated JWST images. AMICAL (Soulain et al. 2020) and SAMpy (Sallum, Ray, & Hinkley 2022) both perform analysis on the Fourier transform of the calibrated JWST images and have been used in the context of ground-based interferometry to great success (Sallum et al. 2023; Vides et al. 2023; Blakely et al. 2022, 2025; Lucas et al. 2024). FOURIEVER (Kammerer et al. 2023) harnesses similar ideas and extends them for kernel phase (Martinache 2010). Other pipelines such as ImPlaneIA (Greenbaum et al. 2015) perform analysis in the image plane, using an analytic forward model of the PSF as a way to better address piston phase errors and pixel-level miscalibrations (Lau et al. 2023; Greenbaum et al. 2019). This work draws much of its intellectual heritage from these earlier pipelines, in particular ImPlaneIA through its use of an analytical forward model for the interferogram, but greatly extended to include the instrument as a whole including the detector and a high-order model of the optical aberrations.

Notwithstanding the rigour and success of these analytical tools, alongside much work from researchers in the field – AMI mode has failed to provide its promised fringe stability and precision, with both found to be up to an order of magnitude worse than expectation in the worst cases (Sivaramakrishnan et al. 2023; Ray et al. 2025; Sallum et al. 2024). Despite this degraded performance, careful calibration has been able to recover images of both the circumstellar disk and companion planets around the PDS 70 system (Blakely et al. 2024), albeit at the cost of discarding very

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

large portions of the data. Similar approaches have been used to resolve the dusty environments around WR 137 (Lau et al. 2023), though visible PSF miscalibration remains persistent. Efforts have been made to fix these problems but follow the approach of simply discarding data without addressing the underlying issues (Goudfrooij et al. 2024).

The most significant factor contributing to the degraded performance of AMI arises from the Fatter Effect (BFE), or charge migration between pixels (Antilogus et al. 2014; Guyonnet et al. 2015; Rowlands, Midwinter, & Warner 2018; Argyriou et al. 2023; Goudfrooij et al. 2024). Electrostatic interactions within the substrate push excited photoelectrons into neighbouring pixels. This results in an effective distortion of the measured PSF, typically seen as a broadening of bright sources that fill neighbouring pixels with more charge than dim sources – hence 'brighter-fatter'. While this effect is troublesome for various imaging modes, it presents a uniquely challenging problem for interferometric analysis dependent on precise inference of the PSF to calibrate observations. To identify why this effect plagues AMI in particular we must examine the JWST data processing pipeline, not the interferometric analysis methods themselves.

These other pipelines adopt *inverse modelling*: a series of transformations of the data to extract summary statistics (a calibrated image, or parameters of that image) that we can fit astrophysical models to. The JWST pipeline applies fixed pixel-wise linearity corrections to the data, and the AMI pipelines assume images are formed linearly downstream of that. This assumption breaks down significantly in the presence of the BFE, which is a local nonlinear convolution – at least when considered in the context of the PSF precision requirements found within interferometry. There is not a known accurate expression or simulator for the BFE, and certainly not an inverse operator to restore the un-blurred ideal pixel response.

This induces a nonlinear change in PSF shape between target and calibrator stars, so that different pipelines return different complex visibilities that do not calibrate in the Fourier domain by simple division. While the BFE was known prior to launch (Rowlands et al. 2018), its seriously harmful effect on interferometric analysis was only realised post launch.

The BFE, together with imperfect gnosis of the AMI metrology, have resulted in AMI under-performing. Until now, fringe stability and precision has not been much better than ground-based observations and AMI proposals have fallen short in the competitive environment of available JWST General Observer (GO) observing time.

In this paper, we present a new approach: Aperture Masking Interferometry Generative Observations (AMIGO), a pipeline in which we jointly train a physical model of the optical system with a hybrid forwards and machine-learned Effective Detector Model (EDM), and extract rich visibility information with a generalisation of kernel phase. We fit the optical and electronic models simultaneously to on-sky calibration data that allow us to separate these effects for the first time, and apply this base instrument model to extract interferometric observables from several science targets, which have hitherto resisted AMI pipelines, achieving near-photon-noise-limited performance in detecting faint companions.

1.2. Differentiable forward models: From pixels to planets

AMIGO uses an end-to-end differentiable forward model of the entire end-to-end chain of physics based on Automatic Differentiation (autodiff; Margossian 2019) – the foundational algorithm of machine learning (LeCun, Bengio, & Hinton 2015). By decomposing functions into a sequence of function primitives and applying the chain rule programmatically, autodiff enables an algorithmic computation of machine-precise derivatives. Importantly, autodiff does not harness finite differences nor symbolic differentiation, instead computing the exact derivative of its input function directly. Its success can be attributed to its computational complexity, scaling with the model itself – even for high-dimensional or nested models – *rather* than the number of parameters being differentiated. Two primary algorithms underpin autodiff: 'forwards' mode (or the tangent method; Siskind & Pearlmutter 2008) and 'reverse' mode autodiff (or backpropagation; Griewank & Walther 2008; Rumelhart et al. 1986). These can be composed to efficiently implement operators such as Jacobians and Hessians through arbitrary computational programs. In particular, such partial derivatives are necessary for optimisation and sampling in high dimensions, for example by stochastic gradient descent (SGD; Ruder 2016) or Hamiltonian Monte Carlo (HMC; Betancourt 2017). This native computational efficiency and accuracy has enabled the training of Machine Learning (ML) models with *billions* of parameters, giving rise to much of the modern world.

In this paper we build on the growing body of work on differentiable modelling in optics. Page & Favaros (2020) demonstrates the design and calibration of microscopes, O'Brien et al. (2026) provides models for cryogenic electron microscopes while improving analysis methodology and Sitzmann et al. (2018) showed the joint design of an optical system with image reconstruction algorithms. In astronomical optics, Wong et al. (2021) explored high-dimensional phase retrieval and mask design, Liaudat et al. (2023) demonstrated super-resolution imaging and recovery of chromatic variations in space telescopes, Desdoigts et al. (2023) showed simultaneous detector and optical calibration while recovering astrophysical scenes, and Desdoigts et al. (2024) showed optical design maximising fisher information.

Differentiable modelling has also proved highly valuable in wider astronomy. Non-exhaustively, Campagne et al. (2023) presents a differentiable theoretical cosmology framework, McDougall et al. (2025) demonstrates lag inference in reverberation mapping, and Horta et al. (2025) provides stellar spectrum models. In exoplanetary astronomy Gully-Santiago & Morley (2022) perform high-dimensional spectroscopic line-profile fitting, Hattori et al. (2024) provides a number of tools for general exoplanetary modelling and analysis, Dholakia et al. (2024) builds ellipsoidal transit models, while Kawahara et al. (2025) and Kawahara et al. (2022) provide exoplanet and brown dwarf spectral modelling and characterisation.

All of these previous works explore how differentiable models can be used within specific sub fields and typically provide a set a open-source tools. However, they all fit within the existing inverse data-reduction paradigm found within astronomical analysis, albeit usually while employing forwards models. AMIGO instead takes the leap to a true end-to-end forwards models, coherently solving for the astronomy and calibrating the instrumental state using the same model, something yet to be comprehensively explored until now. This approach enables the capture of complex non-linear instrumental effects poorly corrected by data-reduction methods in order to recover high-precision astronomical measurements.

Beyond enabling optimisation and sampling with gradients, current autodiff frameworks like JAX (Bradbury et al. 2018) and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

PyTorch (Paszke et al. 2019) offer substantial benefits over standard numerical processing libraries. Built and designed for ML research, almost all autodiff libraries offer many highly optimised tools that ease development and efficiency. Native deployment to hardware accelerators such as GPUs & TPUs, efficient compilers, function vectorisation and parallelisation, and higher order derivatives all give access to a toolbox that can accelerate research and development. Furthermore, many of the bleeding-edge optimisation and inference algorithms rely on the efficient derivatives offered by autodiff. Any new software tools built within autodiff frameworks offer strict benefits over the standard numerical libraries found throughout the astronomical software landscape.

The AMIGO model is built using ∂Lux (Desdoigts et al. 2023) & Zodiax are used as the base framework for differentiable optics and a user-friendly interface for scientific forward models, respectively, which are built on Jax & Equinox (Kidger & Garcia 2021) to ensure it can act as a single end-to-end differentiable system. Jax provides the core autodiff engine with Equinox providing the framework for object-oriented programming as well as the tools required to implement ML models.

The most significant innovation within AMIGO is found in the detector model. Presented with the challenge of forward modelling the charge migration that manifests as the BFE – ultimately governed by differential equations – a novel solution was found by directly integrating a Neural Network (NN) *inside* the detector. This formulation breaks from the three commonly understood modelling paradigms: forwards, inverse and machine-learned, but is part of a emerging body of work known as 'hybrid models' (Akhare, Luo, & Wang 2023; Karniadakis et al. 2021; Willard et al. 2022).

Differentiable forward models show a path towards the next generation of precision calibration and data analysis. For variable nuisance processes that must be inferred directly from science data, which might be affected by heteroskedastic noise and unknown nonlinearities, it is better to fit a forward model to data with Bayesian methods (e.g. discussion in Hogg, Bovy, & Lang 2010). Such models can have very many parameters, necessitating autodiff; and may have to represent noise processes for which a physical simulation may be inadequate or unavailable, but which a neural network can adequately predict (e.g. discussion in Hogg & Villar 2024). In this case, we may not care about interpretability of the model for nuisance processes – in our case, an EDM for the BFE – but only that they perform well and do not damage recovery of the physics we *do* care about (the astrophysical scene).

The hybrid modelling approach presented in this paper connects forward models with machine-learned ones, gaining the best attributes of both approaches. Because the NN is embedded within the overall forward model, it cannot be trained in isolation on a well-curated, diverse dataset, and must instead be evaluated solely through its influence on the end-to-end model predictions. This approach is likely to cause discomfort for many, in particular in its generalisation to datasets very different to those on which it was trained; however its effectiveness in the regime of high contrast imaging will be thoroughly demonstrated in this work.

2. The Amigo model and pipeline

The AMIGO model and pipeline consists of a digital twin of AMI, trained only on high-quality in-flight point source data and flat-field calibration data; and then the application of this pre-trained 'base model' to Bayesian inference from science data of the instrument state (principally wavefront and Fresnel defocus) and astrophysical observables (visibilities, flux, spectrum, though with the option to fit more complex models), holding most of the model's other parameters fixed.

While most data analysis pipelines work on processed and calibrated data, AMIGO is designed to generate predictions of the *uncalibrated* JWST data, fully independent of any other software, including the JWST official pipeline. While the embedded optical model resembles those found in WebbPSF (Perrin et al. 2014; Perrin et al. 2012), the fiducial STScI-supported physical-optics simulator for JWST, its direct integration of a visibility forward model, and a detector model designed to produce the 3-dimensional time-evolving pixels found in uncalibrated data are significant departures from existing models.

AMIGO consists of five distinct and modular sub-models of: the optical system, the complex visibilities, a linear detector, a non-linear ramp, and the read electronics. A flow diagram of the full model is shown in Figure 2, with each component detailed in later sections. While these components are modular and can operate independently, they are trained and behave as a single cohesive system.

2.1. Calibration data

High quality calibration data is of the utmost importance for the ambitious calibration goals of the AMIGO model. Due to a late switch out of the detector hardware on JWST, there is not much publicly available, high-quality, pre-flight calibration data, necessitating an approach exclusively using on-sky data. While a number of calibrator stars have been observed by AMI mode, a requirement for interferometric calibration, most of these programs select bright targets in order to reduce required observation times. Consequently, almost all of the existing data is very coarsely sampled up-the-ramp, making it insufficient for accurate inference of the dynamics of the BFE.

To aid on-going calibration efforts of AMI and to seek a deeper understanding of the BFE, program CAL 4481 (PI: Sivaramakrishnan) was proposed, accepted, and finally observed on 5 May 2024. Designed purely for calibration, it employs a 5-point sub-pixel dither, an uncommon approach for AMI data, with 2×10^9 photons at each dither position. Seeking to better understand the detector systematics, HD 41094 was chosen (with the aid of `SearchCal`; Bonneau et al. 2012) as our calibrator for building the digital twin. It is an appropriately-bright ($W2 = 0.69$ Jy) target for a good number of groups up-the-ramp, with 11, 20, and 30 groups per integration in the F380M, F430M and F480M filters, respectively. As a K0 giant beyond 300 pc, with no known companions, we can be confident that it is a point source in the AMI band. In order to capture the full dynamics of the BFE without dealing with the compounding complexities that arise from deeper exposures, a peak pixel depth of ~ 50 ke^- was chosen – approximately half way to saturation. A summary of the observing program is presented in Table 1.

A future calibration data-set designed to explore the dynamics of the BFE further up-the-ramp and with more complex illuminance patterns, GO 8330 has been already been accepted and observed. Ideally, this program should allow for the AMIGO model to remain performant for brighter targets and deeper observations, expanding the observational capabilities of AMI into the future.

The CAL 4481 proved indispensable to the calibration of AMIGO. Diversity of both wavelengths and dithers, deep exposures with high temporal resolution, with balanced signal across

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 1. Summary of JWST CAL 4481 observations for HD 41094, used for model calibration. All targets are point sources. Pixel well depth values have dimensionless integer units ('digital number' or DN), of which each count corresponds to ~1.6e^-.

Type	Program	Star	Filter	Groups	Integrations	Dithers	Well depth	Photons
Calibrator	CAL 4481	HD 41094	F480M	30	3 800	5	30 000	9.98×10^9
Calibrator	CAL 4481	HD 41094	F430M	20	4 525	5	29 000	9.63×10^9
Calibrator	CAL 4481	HD 41094	F380M	11	5 300	5	30 000	9.39×10^9
Validator	COM 1093	HD 36805	F480M	9	65	1	24 000	0.10×10^9
Validator	COM 1093	HD 36805	F430M	6	82	1	24 000	0.10×10^9
Validator	COM 1093	HD 36805	F380M	3	122	1	22 000	0.10×10^9
Validator	GO 1843	HD 205827	F480M	10	641	1	20 000	0.86×10^9
Validator	GO 1843	HD 205827	F430M	7	1 885	1	20 000	2.15×10^9
Validator	GO 1843	HD 205827	F380M	3	7 800	1	18 000	4.75×10^9
Validator	GTO 1242	HD 18638	F480M	8	4 869	1	16 000	3.93×10^9
Validator	GTO 1242	HD 18638	F430M	5	6 256	1	15 000	3.62×10^9
Validator	ERS 1386	HD 116084	F380M	3	10 000	1	17 000	5.86×10^9
Validator	ERS 1386	HD 116084	F380M	3	6 000	1	17 000	3.52×10^9

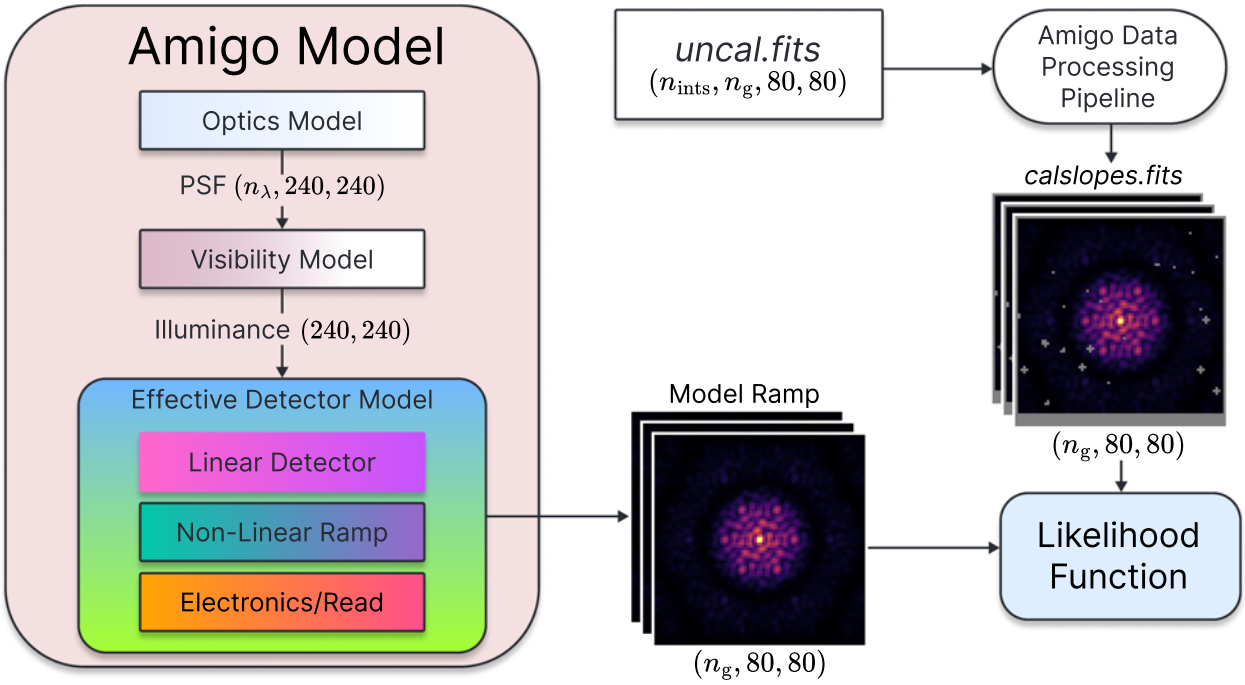


Figure 2. High level flow diagram of the AMIGO model and pipeline, showing the input and output product and shapes passed between each modular component. $n\lambda$ is the number of wavelengths modelled by the optics, n_g is the number of groups in the data, and $n_{\rm ints}$ is the number of integrations. Each of these model and pipeline components are discussed in detail in their own section.

filters all combine to provide an excellent training and testing ground for all AMI pipelines both at present and into the future. All proceeding discussion and presentation of calibration products in this work originate exclusively from this dataset, combined with in-flight flat-fielding data which disentangle the pixel sensitivities and nonlinearity from the spatially-varying effects of the BFE.

In order to ensure model generalisation and avoid over-fitting, a validation dataset was built from existing public programs GO 1843, GTO 1242, ERS 1386, and COM 1093. This set of calibrator star exposures were chosen with the goal of having as much validation data as possible, while keeping a balance of total signal in each filter, detailed in Table 1.

2.2. Data processing pipeline

JWST employs Teledyne HAWAII-2RG(H2RG) near-infrared detectors across various instruments, including NIRISS which hosts AMI mode. These detectors use a non-destructive readout pattern known as 'up-the-ramp' sampling, where the voltage is measured multiple times as it accumulates charge producing a

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

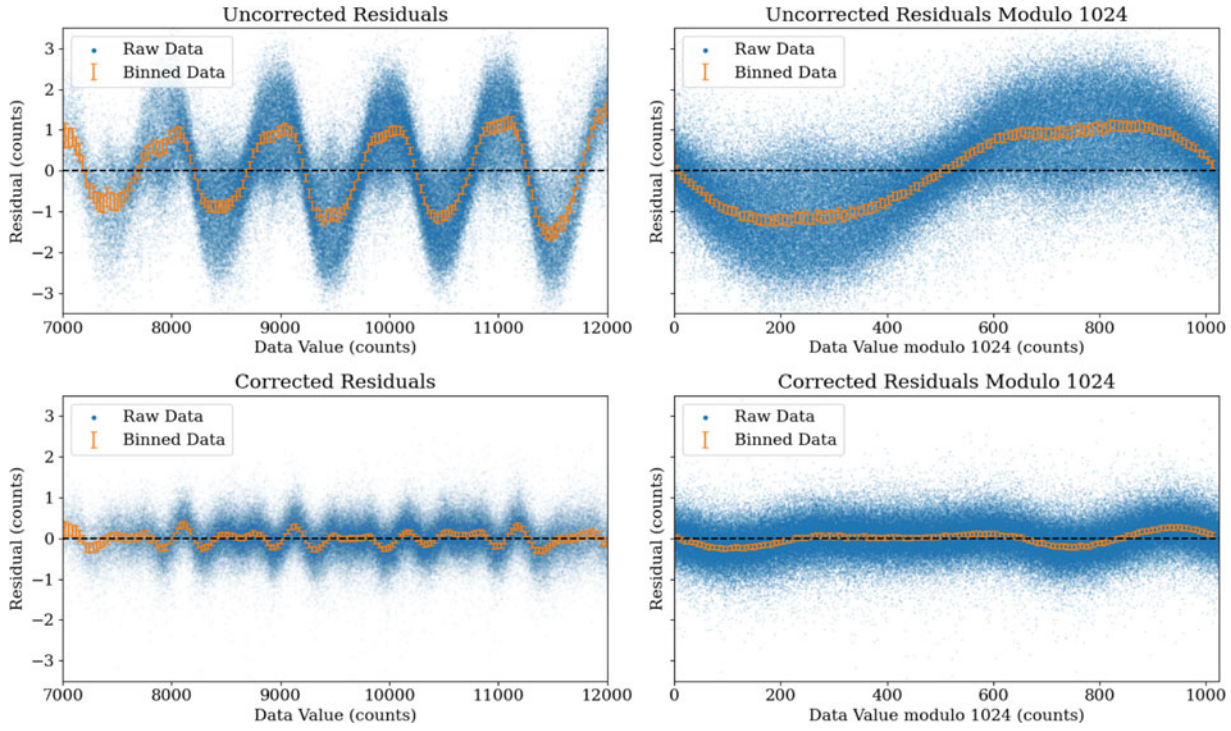


Figure 3. Residuals from second-order polynomial fits to ramp data, shown before (top row) and after (bottom row) applying a sine-wave-based correction for ADC integral non-linearity. The left panels plot residuals as a function of the ramp value, while the right panels show the same residuals folded over a modulo 1024 pattern, revealing periodic structure. Prior to correction, the residuals exhibit a strong sinusoidal modulation. The applied correction consists of a 1024-period sine wave with fixed amplitude, significantly reducing both the overall residual structure and the folded periodicity (bottom panels). Orange points and error bars represent binned data mean and standard error in each bin, highlighting the improved uniformity of residuals post-correction.

time evolving measurement in each pixel. This method provides better read noise characteristics and makes identifying and rejecting cosmic-rays easier (Rauscher et al. 2007). Ideally these pixels can be treated independently with linear fits to the resulting ramp solving for the incident flux.

Standard data calibration approaches based on inverse models seek to subtract these effects sequentially from data, ideally returning a clean signal representative of the input photon distribution. However, non-linear effects like the BFE are self-interacting, non-local, and couple to various properties not addressed at a pixel-level such as PSF shape, sub-pixel curvature, and sub pixel positioning. As a result, these effects do not calibrate straightforwardly by division in the *uv* plane.

The AMIGO data processing pipeline takes a very different approach, seeking to preserve as much of the physics as possible in the output product. The 'up-the-ramp' readout of the H2RG detectors provides 4D uncalibrated data: Two spatial dimensions, one time dimension for the read at each group, and a second time dimension for each integration, or 'image'.

The AMIGO pipeline performs a single calibration to the data, correcting for the Analogue to Digital Converter (ADC) integral non-linearity. There is a strong periodic residual in uncalibrated data that, until we identified this, made interpreting trends in the group-level data challenging as it induces unique per-pixel periodic signals greater than the intrinsic noise. This behaviour is thought to arise from a lack of power being supplied to the amplifiers. It is periodic in *raw counts*, and we estimate it simply by performing a least-squares fit to the average cleaned ramp values. Plotting the residual to this fit against the data value (i.e. including the pixel bias) reveals a strong sinusoidal signal. While we are not able to infer a functional form accurate at all count levels, we find a sufficiently accurate fix for our purposes by subtracting off a sinusoid with period 1024 and amplitude 2. This correction leaves some periodic residuals in the ramp-level data, however the existing correction is currently sufficient. We also found this signal in other observing modes, with a more comprehensive treatment and correction to come in subsequent work (Dholakia in preparation). Figure 3 shows this residual and the post-correction residuals.

Next, simple outlier rejection is performed on the ADC corrected ramps, rejecting the 3σ outliers from each group. This same outlier rejection is performed at 3σ on the data slopes, found by taking the differences between each group read. This outlier detection process is aimed at removing cosmic rays and other spurious jumps in the data.

From this outlier cleaned data the mean and covariance matrix of each pixel is calculated along the integration axis for both the ramp and slope data. This produces the final `calslope` output product. At present the AMIGO model is designed to fit the slope data, although the ramp data is still preserved with the intention of extending the model to predict ramp level, rather than slope level data in order to produce higher fidelity over the pixel gain.

An important consequence of fitting to the slope data is the loss of the first group read. While this had the benefit of avoiding predicting the pixel-to-pixel bias, it also results in a loss of $1/n_{\mathrm{groups}}$ worth of data. As a result, the AMIGO model at present can not predict observations made with a single group read, and loses a

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

higher fraction of the total photons for observations made with a small number of groups.

As most AMI observations deliberately do not go above half pixel depth, in order to avoid BFE, it is not possible with current resources to train the AMIGO model to cope with saturated sources, and we obtain worse fits with increasing well depth. This may be mitigated in future as deeper datasets become available and the model is re-trained to include these.

3. Optical and visibility model

AMIGO combines a diffractive physical optics model with a novel interferometric visibility forwards model. This enables the injection of observed source brightness distributions into PSFs in an optically coherent manner and encoded via its complex visibility.

3.1. Optical model

The first stage of AMIGO uses a differentiable physical optics model based on ∂LUX (Desdoigts et al. 2023; Desdoigts et al. 2024), which has similar features to other Python physical optics packages like WEBBPSF (Perrin et al. 2012; Perrin et al. 2014) and PRYSM (Dube 2019) but provides automatic differentiation and hardware acceleration and parallelisation through JAX.

To circumvent these issues the AMIGO optical model has a strong focus towards flexibility, an approach facilitated by autodiff. With direct calibration from on-sky data, it emerges as the most precise physical optics model of AMI, with dynamically generated, non-linearly distorted aperture geometries, persistent wavefront sensing, broadband PSFs, and Fresnel diffraction.

The optical model can be broken into four components:

1. The spectral model
2. The aperture model
3. The wavefront model
4. The propagation model.

Each component enables the coherent flow of gradients from residuals of the predicted PSF back through to the underlying parameters, enabling precise calibration and resulting in a highly accurate and physically principled PSF describing the instrumental response to incoming wavefronts.

3.1.1. Chromaticity and spectral model

AMI observations can be taken with four different filters, detailed in Table 2, although most observations avoid F277W due to these wavelengths being significantly under-sampled by the NIRISS pixels. Other potential problems arise with the F277W filter, since the BFE is also believed to have wavelength dependence, generalising the implementation of the AMIGO model would require greater complexity and more calibration data. For these reasons, the CAL 4481 program (used for training AMIGO) did not observe using the F277W filter, performance in this band is not well characterised and not discussed further.

All three of the primary AMI filters (F380M, F430M, and F480M) are relatively narrow in comparison to the size of the generally expected spectral features, allowing a simple linear spectral energy distribution model to be used as the default via

$$F(\lambda) = 1 + m(\lambda - \lambda_c) \quad (1)$$

Table 2. Allowed Filters for AMI observations. Values taken from the JWST documentation. Full tabulated curves used in propagation.

Filter	λ central (μm)	Δλ (μm)
F480M	4.815	0.289
F430M	4.285	0.203
F380M	3.825	0.205
F277W	2.771	0.717

where λ_c is the filter's central wavelength and m is the spectral slope parameter. This default model is unlikely to capture the complexities present in some science cases. However, the forward modelling framework enables the inclusion of user-defined differentiable spectral models, tailored to the specifics of any observation.

By default AMIGO propagates 9 monochromatic wavelengths through the optical systems to capture the appropriate spectral diversity. Each wavelength is weighted by the integrated filter bandpass and the spectral weights and summed to a broadband illuminance pattern.

3.1.2. Aperture model

The AMIGO aperture model is similar to, but significantly more flexible than existing aperture models used throughout the field. Each aperture mask hole is modelled as a soft-edged hexagon rendered dynamically on the aperture coordinates. The soft-edges of the aperture enable stable gradient propagation through to the coordinate grid, despite the output array being dominated by the binary ones and zeros. The mask is parameterised by two sets of 2D polynomial distortion coefficients: one that controls relative positions of each aperture mask hole, and one for each of the 7 aperture mask holes, applied to the hexagon coordinate array.

These aperture polynomial distortions provide the flexibility required to model the astrometric distortion seen in real optical systems. Typically, these distortions are modelled by applying a polynomial distortion to the PSF in the image plane, as opposed to the aperture in a pupil plane. The choice to apply these distortions in the pupil plane enables the prediction of a PSF that remains governed by diffractive physics. Additionally, this approach accounts for any relative shears/rotations between the wavefront and the aperture mask when encountered in the filter wheel.

The use of 2D polynomial distortions is important as it enables the removal of two problematic degrees of freedom: the global x- and y-positional shift. Since the optical model is diffractive, the wavefronts are propagated from pupil to focal plane via a Fourier transform, and the PSF is the squared modulus of this focused wavefront – hence the phase information encoding the global position of the aperture is lost. Consequently, the model is invariant to x,y aperture shifts; however, noise in the gradients produced by autodiff can result in undesired drifts. By removing these degrees of freedom from the model, we can pin down both the global aperture mask position and the relative positions of each hole, avoiding this problem entirely. Figure 4 presents the residual between the recovered distorted aperture mask and its undistorted counterpart along with its effect on the PSF.

We identified an unexpected discrepancy in the reported size of the AMI mask holes across different online resources and software packages. The JDox documentation states that the face-to-face

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

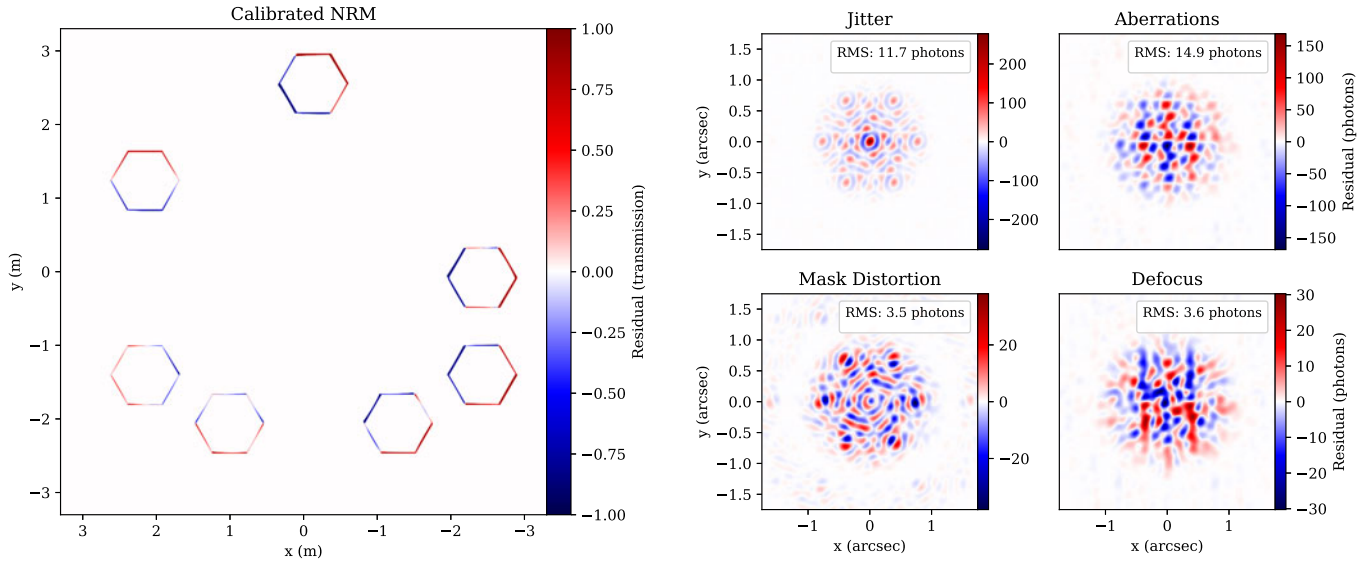


Figure 4. Left panel: Residual between the calibrated aperture mask and its idealised undistorted counterpart, discussed in Section 3.1.2. Right panel: PSF residuals of the four primary optical effects on the PSF. Top left: Instrumental jitter, applied through a convolution with a Gaussian kernel, discussed in Section 4.1 Top right: Primary mirror aberrations, modelled using Zernike polynomials on the primary, discussed in Section 3.1.3 Bottom left: Aperture mask distortions, discussed in Section 3.1.2, modelled by applying a distortion to the coordinates over which the aperture mask is calculated. Bottom right: Fresnel defocus modelled using a Fresnel propagation algorithm, discussed in Section 3.1.4 All effects are shown for the F430M filter, using a PSF with 10^6 total photons.

diameter of each hole, as projected onto the primary mirror, is 0.80 m. However, the mask definition file used by both WEBBPSF and IMPLANEIA adopts a slightly larger value of 0.82 m. Several other analysis pipelines in the AMI ecosystem were also found to prefer one of these two values. This discrepancy in the manufactured size of the mask holes was resolved in two ways: first, by allowing them to be inferred directly during the AMIGO calibration process, and second, by cross-checking the result against the original manufacturing specification file used. The AMIGO model confidently recovered a hole diameter of 0.80 m, consistent with both the JDox documentation and the manufacturing specification file, thereby validating the smaller of the two reported values.

3.1.3. *Wavefront model*

Accurate PSF modelling of any real system must also account for phase errors in the wavefront, accumulated by both the primary mirror and through the optical train. Despite its unprecedented optical stability, JWST is no exception. Observations made during its commission phase revealed two important time-dependent optical degradations: mirror tilt-events and micro-meteoroids Rigby et al. (2023). These effects necessitate the recovery of wavefront phases between, and sometimes within, any given observing program, a task made far simpler by both the non-redundant aperture mask and the gradients provided by the differentiable model. While most optical systems suffer from a sign degeneracy within certain phase modes, NRMs offer unambiguous recovery of all phase modes, making them an ideal calibration tool (Cheetham et al. 2012; Pope et al. 2014).

To recover the Optical Path Difference (OPD) state of the cumulative optical surfaces we employ Zernike polynomials, a set of orthogonal functions defined over the unit disk, commonly used to represent wavefront aberrations in optical systems. Their orthogonality and correspondence with classical aberration types (e.g. defocus, astigmatism, coma) make them particularly useful for decomposing and quantifying optical distortions. Due to these properties, Zernike expansions provide a compact and physically interpretable basis for modelling and correcting wavefront errors across a wide range of optical applications (Lakshminarayanan & Fleck 2011).

Given that JWST has a segmented primary mirror, we model a unique set of moderate order Zernike polynomials over each hole within the aperture which are then fit to any given observation. Choosing a total of 4 radial orders, we get 10 Zernike modes per hole for a total of 70 phase modes across the aperture. The inclusion of higher-order effects is trivial but was found to be unnecessary, as these terms remained statistically insignificant in the calibration data, up to $\sim 10^{10}$ photons. In data sets where these terms become important, either a larger number of Zernike modes can be solved for or their effects can be captured by the interferometric visibility model, that has greater resolution over the wavefront. By default, AMIGO will recover wavefront phases for all observed data, providing a potential way to do long-term wavefront sensing independent of dedicated observations. The effect these aberrations have on the resulting PSF as found in the calibration data is shown in Figure 4.

Somewhat surprisingly, slightly different wavefront phases are recovered across the three filters, with the most significant deviation found in the F430M filter. These differences, visualised in Figure 10, we believe to arise from imperfections in the optical surfaces of the filters. The differences found in F430M can also be found through a significant deviation in the recovered Fresnel defocus, discussed in the next section.

3.1.4. *Propagation model*

Under the Fraunhofer approximation, optical systems are typically modelled using two conjugate planes: the pupil and the focal plane. This approach assumes that light propagates as planar wavefronts, which is valid when the observation point is in the far-field

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

or at the focal point. However, many practical optical systems exhibit complexities such as a misalignment along the optical axis which necessitate modelling wavefront propagation to intermediate planes. In these scenarios, Fresnel diffraction theory (Morse & Feshbach 1953; Born et al. 1999; Hecht 2002; Goodman 2005) becomes essential, as it accounts for the coupling between wavefront phase and amplitude variations, providing a more accurate representation of PSF behaviour in these regimes.

The Fresnel diffraction integral, best expressed through Fourier Transforms, which describes the complex field E at point (x, y, z), is given by

$$E(x, y, z) = \mathcal{F}^{-1}\left[\mathcal{F}[E(x', y')](k_x, k_y) \cdot e^{ik_z z}\right] \tag{2}$$

where $E(x', y')$ is the field in the pupil plane, $k_z = \sqrt{k^2 - k_x^2 - k_y^2}$, and k_x, k_y correspond to the angular spectrum (spatial frequencies) of the wave. A notable application of Fresnel diffraction modelling is in the analysis of the HST optical performance. Thermal fluctuations in the HST structure cause 'breathing' modes, leading to temporal variations in the telescope's focus. By employing Fresnel-based models, researchers have been able to accurately characterise and correct these focus variations, thereby enhancing the quality of the scientific data obtained from HST (Krist, Hook, & Stoehr 2011).

Fresnel diffraction algorithms were found to be a necessary component to recover accurate PSF morphology with the AMI observing mode, with significant differences found across filters. The recovered defocus values were found to be 0.017, 0.050, and 0.010 μm in the F480M, F430M, and F380M filters, respectively. F430M was found to have the largest defocus, with more than double that found in the other filters. The unique effect induced to the PSF in the F430M filter is shown in Figure 4. It is worth noting that due to the need to recover both primary mirror aberrations and Fresnel defocus values, which have a high degree of covariance in the small defocus regime, that these values cannot be recovered fully independently. Considerable testing was done and accurate PSFs could only be recovered when considering both of these effects in tandem.

The coupling of amplitude and phase effects in wavefronts introduced by Fresnel effects is typically not important to downstream analysis, however careful consideration of its implication on interferometric observables is essential to recovering well-calibrated and high-precision visibilities. This is discussed more in Section 3.2.2.

We enhance the fidelity of the modelled PSF by applying a cubic spline interpolation to up-sample the image from a 3× to a 9× resolution. The result is then down-sampled back to the original 3× grid. While this may appear redundant, the process improves realism by more accurately accounting for the finite area over which detector pixels integrate light – an effect that is not captured by simple point sampled PSF models. This method was benchmarked against a reference PSF generated at 30× oversampling and was found to match the fidelity of a ∼12× oversample, while requiring only a 3× optical propagation. The choice of 9× for the up-sampling factor reflects a practical compromise between accuracy and computational efficiency. The resulting increase in spatial resolution reveals finer PSF structure, which is especially important given the BFE's strong coupling to PSF curvature.

3.2. Interferometric visibility model

Optical interferometry is typically conceived of as a purely inverse problem: each image has a set of complex visibilities that can be reduced down to the non-redundant baselines and analysed. This paradigm is insufficient for forward modelling approaches because in reality the visibility signal is defined across the entire support of the Optical Transfer Function (OTF), something that must be captured by our modelling approach. A visibility model without the flexibility to reproduce the full behaviour of the PSF as observed through the optical system will introduce biases and non-physical signals in the predicted PSF that further couples through the downstream non-linear detector model. This problem has mandated a re-think of the concept of a visibility as it applies to forward modelling, as well as the analysis methods used on the recovered observables.

In order to achieve the required behaviour our model must produce a continuous array of interferometric amplitudes and phases, appropriately conjugated about the origin, that can be multiplied by the PSF in the *uv*-plane. This is done by defining a set of knots across the OTF, whose values (one amplitude and one phase each) can be interpolated to the appropriate *uv*-coordinates as defined by the Fast Fourier Transform (FFT) of each monochromatic PSF. The resulting complex visibility map is then multiplied by the complex PSF splodges and transformed back to the image plane. The resulting PSF can now express complex instrumental and astrophysical effects while remaining firmly grounded in the diffractive physics that governs imaging systems. Figure 5 presents a diagram of this process from the wavefront to the final interferogram.

3.2.1. Latent visibility model

This simple visibility modelling method satisfies the requirements for use in a forward model, but comes with some draw-backs. The pixel basis used for the visibilities results in high-dimensional and covariant parameters, with many pixels falling out outside the OTF and therefore being unconstrained by the PSF. Highly dimensional problems present no inherent issues for differentiable models, but strongly covariant or unconstrained parameters can cause problems for optimisation algorithms. This mandated a deeper and more principled consideration of forward modelled visibilities, underpinned by Reduced Order Modelling (ROM; Benner et al. 2015).

ROM is a set of methods that can reduce parameter dimensionality in a way that preserves the essential dynamics or structure of the original system. We apply this concept to our visibility model, using autodiff to design an orthogonal low-dimensional latent set of parameters that fully capture the dynamics of the pixelised visibility model. However, ROM methods rely on linear model assumption, a property exhibited by visibility phases, but not amplitudes. This mandates a treatment of the *logarithm* of the complex visibilities

$$\ln\left(Ae^{i\phi}\right) = \ln\left(A\right) + i\phi \tag{3}$$

reformulating the mapping between the parameters in the *uv* and focal plane to be linear though the real and imaginary components of the logarithmic complex visibility (Pope 2016). Using this construction, we now seek a matrix $\mathbf{V}$ that can map between our full set of log amplitudes and phases

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

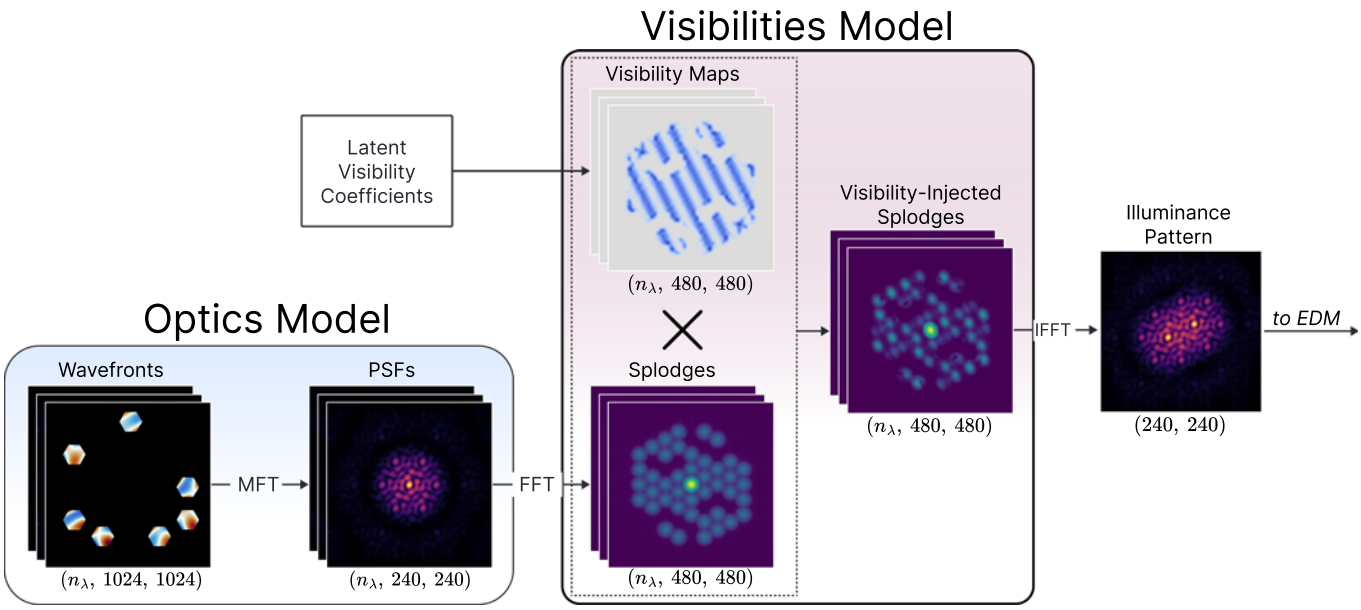


Figure 5. Flow chart of the injection of visibility signals to forwards-modelled PSFs. This demonstrates how high-resolution visibility signals can be directly injected into any PSF model provided the appropriate set of visibility basis vectors. An example binary-star signal is injected as a demonstrator.

$$\ln(\mathbf{A}), \boldsymbol{\phi} \in \mathbb{R}^N \tag{4}$$

and a latent set

$$\mathbf{A}_l, \boldsymbol{\phi}_l \in \mathbb{R}^M \tag{5}$$

where $M < N$.

To construct the matrix $\mathbf{V}$, best envisioned as a set of basis vectors, we compute the Jacobian matrix $\mathbf{J}_{\rm psf} \in \mathbb{R}^{N\times L}$ of the PSF, where L is the number of pixels in the PSF, with respect to the pixelised log visibilities $\ln(\mathbf{A}), \boldsymbol{\phi}$. This Jacobian captures the local sensitivity of each pixel in the PSF to each visibility log amplitude and phase parameter. Using the Gauss-Newton Hessian approximation

$$\mathbf{H} = \mathbf{J} \cdot \mathbf{\Sigma}^{-1} \cdot \mathbf{J}^{-1} \in \mathbb{R}^{N\times N} \tag{6}$$

we can estimate the Hessian of the pixelised log visibility parameters under a log-likelihood. An eigen-decomposition of this Hessian matrix

$$\mathbf{H} = \mathbf{Q}\Lambda\mathbf{Q}^T \tag{7}$$

returns eigenvectors $\mathbf{Q}$ that form an orthonormal basis for the *uv*-plane, ordered by their influence on the interferogram. By selecting out the top M eigenvectors corresponding to the largest eigenvalues Λ (i.e. ones with high Fisher information, that are constrained by data), we tailor a low dimensional latent set of orthogonal basis vectors $\mathbf{V} \in \mathbb{R}^{N\times M}$ that exhibit favourable properties for modelling.

This method is very powerful in many modelling regimes and enables customisation of the properties of the latent basis vectors through the choice of $\mathbf{\Sigma}$ in Equation (6). In the context of interferometric visibilities we consider two choices for $\mathbf{\Sigma}$. The first is the PSF itself, providing basis vectors that best constrain point sources – ideal for high-contrast companion recovery. This nevertheless could introduce difficulties when recovering extended, resolved, or medium-high contrast sources, where the noise distribution may be very different from the simulated point source used to build the basis. The second option is the identity matrix, providing an equal weighting to all pixels in the image plane. This provides higher expressiveness in the resulting basis vectors that remain capable of the high-contrast sources often targeted by interferometric observations.

In order to maximise generality, we chose an identity pixel covariance matrix to produce a set of latent visibility vectors, visualised in Figure 6. Interestingly but perhaps unsurprisingly, the low order basis vectors appear to select out the classic interferometric baselines. Examining the eigenvalues in Figure 6, we can see there is a knee around index 600 in both the amplitudes and phases. These reflect the visibility values *outside* of the OTF, where the PSF is unresponsive to any signal - as it is outside of the range of frequencies that the optical system is responsive to. We therefore model only those modes that are actually constrained by the optical configuration; an approach which implies a general method for describing the information content of diffraction-limited images, generalising the speckle statistics described by Mawet et al. (2014); a full exploration of this idea is beyond the scope of the present work.

3.2.2. *Visibility amplitudes and phases coupling*

An interesting consequence of the observed Fresnel effects discussed in Section 3.1 is the coupling of amplitude and phase effects in the wavefront. This paradigm further translates to the complex visibilities – meaning that the visibility amplitudes and phases do not cleanly separate and act independently on the PSF, instead having non-insignificant covariances. This was directly observed in the estimated Hessian of the pixelised visibilities as calculated by Equation 6. This implies the true orthonormal visibility basis vectors live in *complex* space, and is therefore composed of both visibility amplitudes *and* phases. However keeping the classically understood visibility amplitudes and phases separated eases both their interpretation and comparisons to existing methods, and is a convention adopted through the rest of this work.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

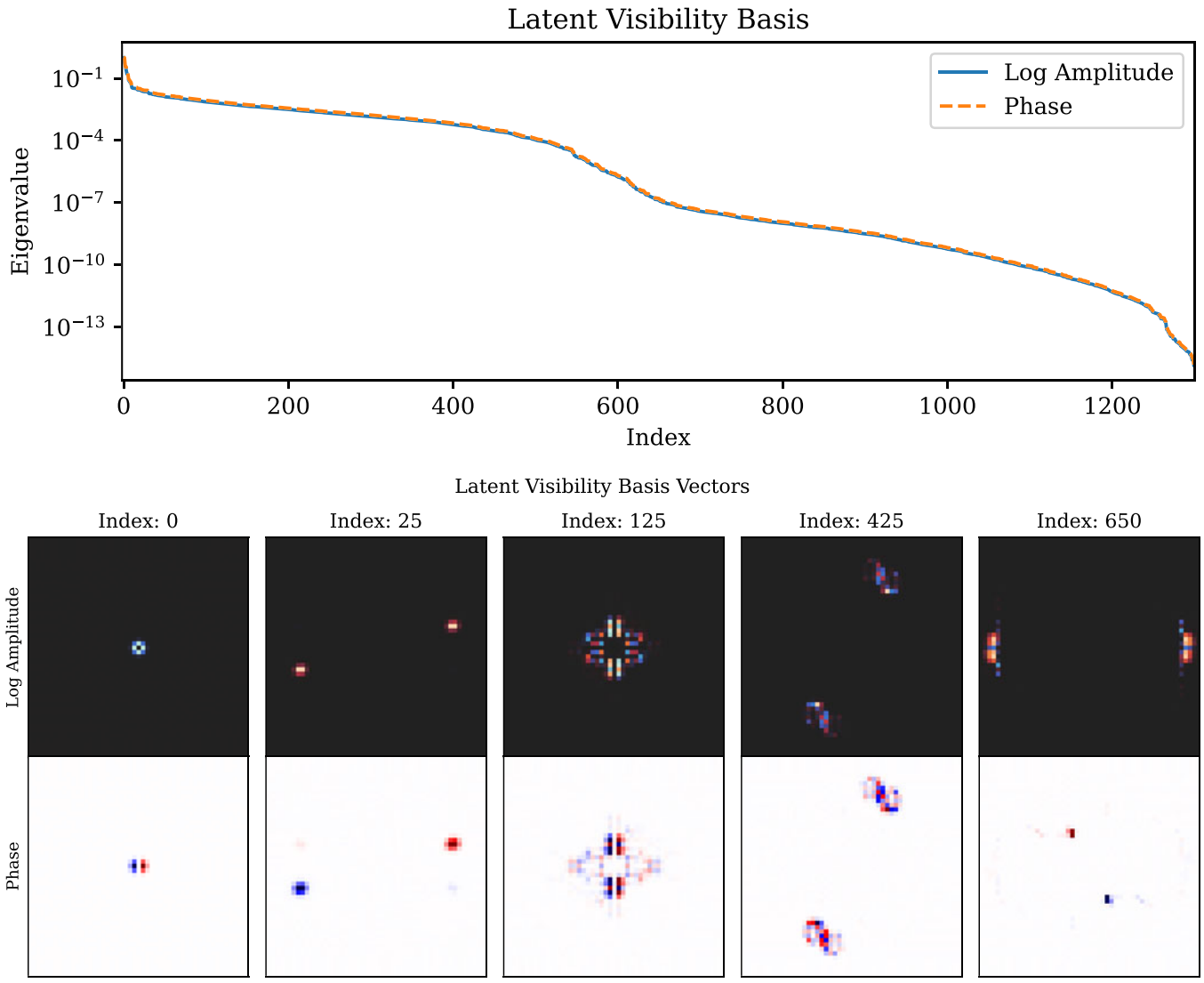


Figure 6. Demonstration of the produced latent visibility basis used for model-fitting. Top panel: The normalised eigenvalues for each visibility basis vector, ordered by their impact on the PSF in the image plane. Bottom: Representative log amplitude and phase basis vectors over a range of indexes. We can see that higher basis indices have increasing spatial resolution over the OTF, with low order ones picking out the classical interferometric baselines and their conjugates. Indices above ∼600 start to put power outside the OTF, are un-sensed by the optical system, and are excluded from the model, but shown here to demonstrate how the basis can be restricted to inside the OTF.

3.2.3. *Kernel amplitudes and phases*

Kernel phase analysis offers a powerful method to extract interferometric observables robust to residual wavefront error, even through clear aperture optical system with redundant *uv* baselines (Martinache 2010), and conversely to infer wavefront error from science data by the same approximation (Martinache 2013; Pope et al. 2014). It generalises the idea of closure phases (Jennison 1958), by approximating the effect of phase aberrations on pupil elements ϕ_j on the phases measured on baselines Φ_i, using a Jacobian matrix $\mathbf{J} \equiv \partial\Phi_i/\phi_j$

$$\Phi^{\text{meas}} = \Phi^{\text{sky}} + \mathbf{J}\cdot\phi. \tag{8}$$

The kernel phase idea is then to use singular value decomposition to find a kernel matrix $\mathbf{K}$ such that $\mathbf{K}\cdot\mathbf{J} = 0$, so that

$$\mathbf{K}\cdot\Phi^{\text{meas}} = \mathbf{K}\cdot\Phi^{\text{sky}} + \underbrace{\mathbf{K}\cdot\mathbf{J}\cdot\phi}_{=0} \tag{9}$$

and therefore recovering 'kernel phases' $\mathbf{K}\cdot\Phi^{\text{meas}}$ which are self-calibrating with respect to wavefront noise.

In its first implementations (Martinache 2010) this is calculated analytically by autocorrelation of a grid of points, and later by autodiff (Pope et al. 2021). This idea can be extended to kernel amplitudes (Pope 2016), by performing analysis on the log-visibilities, a practice we adopt here, which has the advantage of making gain errors additive rather than multiplicative and therefore putting phase and log-amplitude on an equal footing as the complex logarithm of a wavefront or visibility. The kernel phase and amplitude relations are approximate and rely on linearisation for redundant apertures, but are exact and recover closure phases as a special case for non-redundant apertures.

Even though we infer Zernike coefficients for each mirror fairly precisely in the workflow above, the kernel phase idea is still required here: instrumental degrees of freedom are linearly indistinguishable from a subspace of astrophysical degrees of freedom

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

(i.e. linear combinations of complex visibilities that are sensitive to Zernikes, and cannot be inferred separately). Furthermore, some observations may require fitting a unique wavefront or visibilities to different detector positions, over observations made at different epochs, or may need to account for tilt-events during or between different exposures, as occurred between imaging AB Dor and its calibrator.

Further building on these ideas and leveraging autodiff we expand the null-space from pure wavefront phases on a Zernike basis to also null over small defocus, flux and spectral miscalibrations. Trivial to calculate using autodiff (Pope et al. 2021), this formulation enables fine-grained control over the order of wavefront error, and other nuisance optical effect over which to null the observables. In this work we only null over the same Zernikes modelled in the optical system (i.e. the first 10 modes, as discussed in Section 3.1), preserving more astrophysical information in the interferometric outputs than a pixel-basis would allow. In principle, this same approach could be used to generate kernels invariant with respect to any instrumental or nuisance degrees of freedom, such as mask rotation or jitter.

Kernel amplitudes must be used for a different set of reasons. As discussed in Section 3.1, accurate PSF metrology can only be recovered by modelling wavefront behaviour outside of the Fraunhofer regime, i.e. using a Fresnel propagation, which projects phase aberrations into effects both in phase and amplitude in the final *uv*-plane. Furthermore, the finite field of view causes spatial correlation of complex visibilities, convolving the real and imaginary parts with a window function and therefore mixing amplitude and phase. In developing AMIGO, we notice miscalibration in amplitudes unless we account for this effect.

This projection of our extracted visibilities into a kernel space provides one final statistical hurdle: the output kernel visibilities do not preserve the original statistical independence of the basis vectors. To circumvent this final issue, an eigen-decomposition is performed on the measured visibility covariance matrix to restore statistical independence (following Ireland 2013). If we were to keep the full covariance matrix and use this in the likelihood, this would not not change the information content, but projecting once to a basis of statistically independent observables does allow for accelerated downstream analysis in evaluating the likelihood without matrix inversion each time.

Each operation in this process from pixelised visibilities through to statistically independent observables is linear, enabling each projection to be mapped into a single matrix, the Delay-Invariant Subspace of Calibrated Observables (DISCO) matrix. Figure 7 shows the DISCO basis vectors re-projected onto the pixel basis with their corresponding variances. Interestingly, many of the produced basis vectors look similar to combinations of the latent visibility basis vectors, implying that much of the information is preserved even through projection to the null space.

We believe that this representation of our data is close to optimal: using latent visibilities derived from the Fisher matrix eigenvectors ensures inclusion of *all and only* the information passed by the optical system, so that unconstrained modes are ignored but able to express any detectable visibility pattern across the *uv* plane. Then by projecting to a kernel space these are protected from miscalibrated instrumental degrees of freedom; and the full information from correlated noise is preserved in the final representation. By using an accurately trained pupil model and an exact Fourier sampling to construct this basis, we alleviate the issue of model misspecification which has rendered the original HST/NICMOS kernel phase results suspect (Pope, Martinache, & Tuthill 2013; Martinache et al. 2020), by instead directly solving for the instrument metrology first.

4. The effective detector model

AMIGO treats the detector, with its complicated pattern of sensitivity and cross talk, with a non-parametric EDM. The JWST H2RG detectors in NIRISS are subject to the BFE, driven by electrostatic interactions within the detector substrate, which induces flux-dependent charge redistribution across neighbouring pixels. These effects evolve over time and are entangled with other systematics, including variations in the PSF and inter-pixel sensitivity. Importantly, they non-linearly act on a patch of pixels in the core of the PSF, depending sensitively on the illumination pattern over these pixels, which will differ from star to star. As a result, pipelines depending on calibration in the Fourier domain may be inadequate for recovering signals at the precision levels required for high-contrast exoplanet imaging.

The EDM circumvents these issues by modelling the detector as a coherent physical system. Its architecture mirrors key components of the standard JWST pipeline but employs the pixel-to-planets philosophy of end-to-end differentiable *forward* models. This construction enables further innovations ideally situated for complex or poorly understood physical processes through the direct integration of a NN.

The defining feature of the EDM is its integration of machine-learned components into a physics-based forward model, forming what is known as a *hybrid model* (Akhare et al. 2023; Karniadakis et al. 2021; Willard et al. 2022). In particular, a NN is embedded *inside* the detector model as a differentiable transformation, trained to capture the charge migration behaviour of the BFE. This breaks from common NN usage in astronomy, where models are trained independently and used as surrogates for entire processes. Here, the NN is treated as one operator in a long Markovian chain, calibrated through gradients propagated from the raw detector data all the way back to the astrophysical parameters.

This design yields several critical advantages:

- The NN can leverage accurate upstream physical predictions, reducing its complexity and improving interpretability.
- Training is conducted *entirely* on real, on-sky data, without the need for lab-based calibrations or curated training sets.
- Physical constraints can be imposed on the NN outputs, such as flux conservation and locality of charge migration, drawing inspiration from the field of physics-informed machine learning (Raissi, Perdikaris, & Karniadakis 2019; Kidger 2022).

The hybrid approach of this model introduces new challenges. Since the NN is not trained in isolation, the quality of its predictions is dependent on both the accuracy of the upstream optics and visibility models as well as the downstream electronics model. If other components are poorly calibrated, the NN may erroneously learn to compensate for them, degrading generalisation. This coupling mandates joint optimisation of all model components, enforcing physical realism at each stage.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

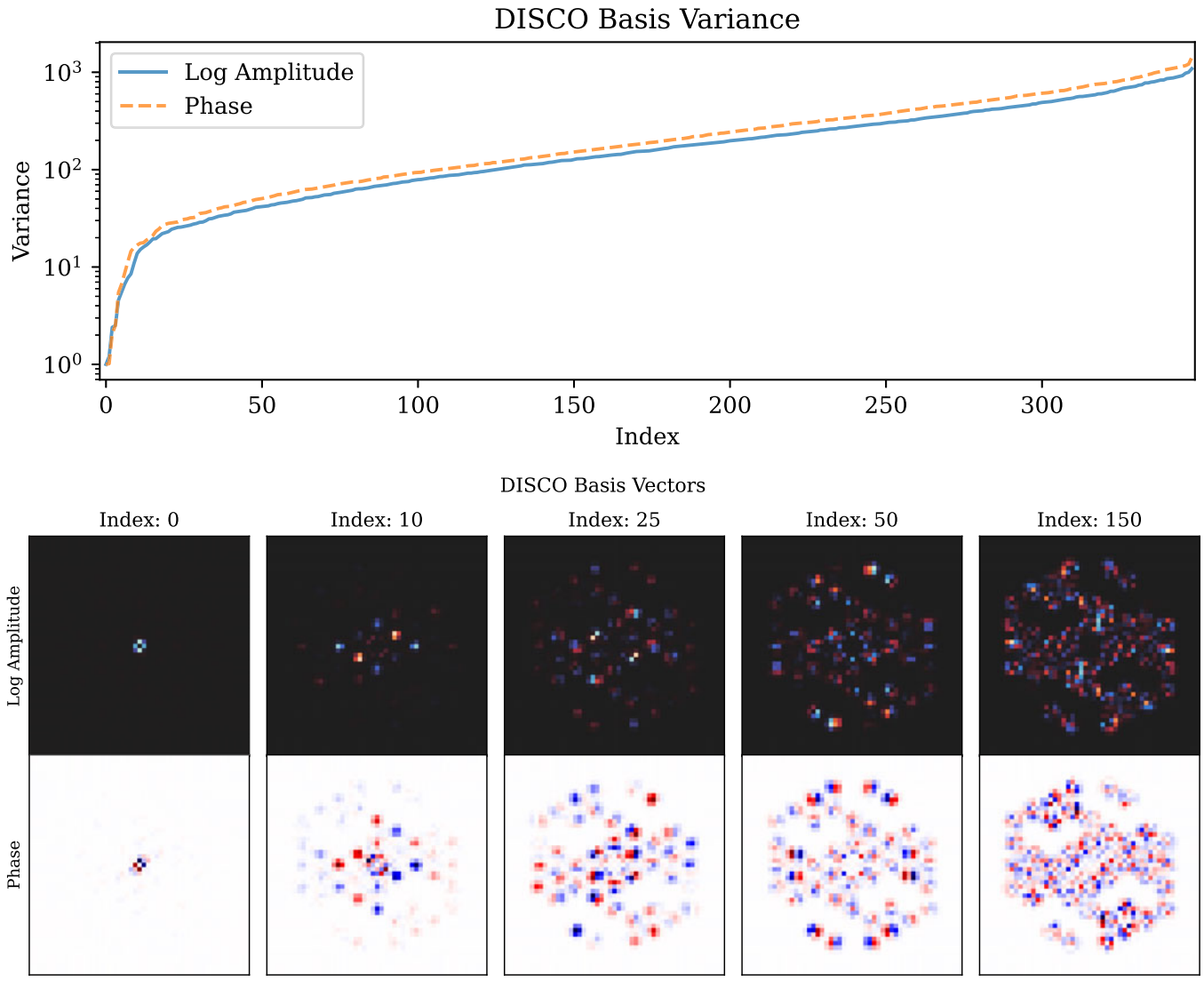


Figure 7. Example Delay-Insensitive Subspace of Calibrated Observables (DISCO) basis vectors found from the GO 1843 observation, discussed in Section 7.3. Top panel: Normalised log amplitude and phases basis vector variances. Bottom panel: Selection of representative DISCO basis vectors. Matching with the latent visibility basis vectors, low index vectors are better constrained and have lower spatial fidelity.

The EDM is comprised of three main stages:

1. Linear Detector Model
2. Non-linear Ramp Model
3. Electronics Model.

Together, these components form a structured pipeline that mirrors the physical data generation and readout process of H2RG detectors, as outlined in Figure 8. Because the BFE is spatially non-local, non-linear, and self-interacting, it poses the greatest challenge to conventional approaches out of the effects in this chain. Additionally, its interaction with other effects (such as flat field variations) renders modular correction ineffective. For example, an insensitive pixel will accumulate fewer electrons, which biases the electric field measurement and thus affects the migration behaviour of surrounding pixels.

To our knowledge, the EDM presented here is a novel tool in astronomy, analogous to the recent use of ML for sub-grid physics in otherwise physically-rigorous weather simulations (Kochkov et al. 2024). The result is a unified, physics-informed, and data-driven model capable of meeting the stringent requirements of exoplanet imaging.

4.1. Linear detector model

The first component of the EDM is the linear detector model which handles the transfer of the predicted PSF onto the pixels of the detector. This process includes two effects: instrumental jitter and PSF resampling.

Starting with the PSF predicted by the upstream optics and visibility model, instrumental jitter is applied at the calculated 3× PSF oversample by a Gaussian convolution. The instrumental jitter discussed here is distinct from the more commonly understood *pointing* jitter. Instrumental jitter is a blur arising from imperfect stability of all components in an optical train, as opposed to telescope pointing instability. JWST has an exquisite pointing accuracy of ∼1 mas, but the instrumental jitter recovered by AMIGO is a much larger at ∼21 mas. This blur can arise from

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

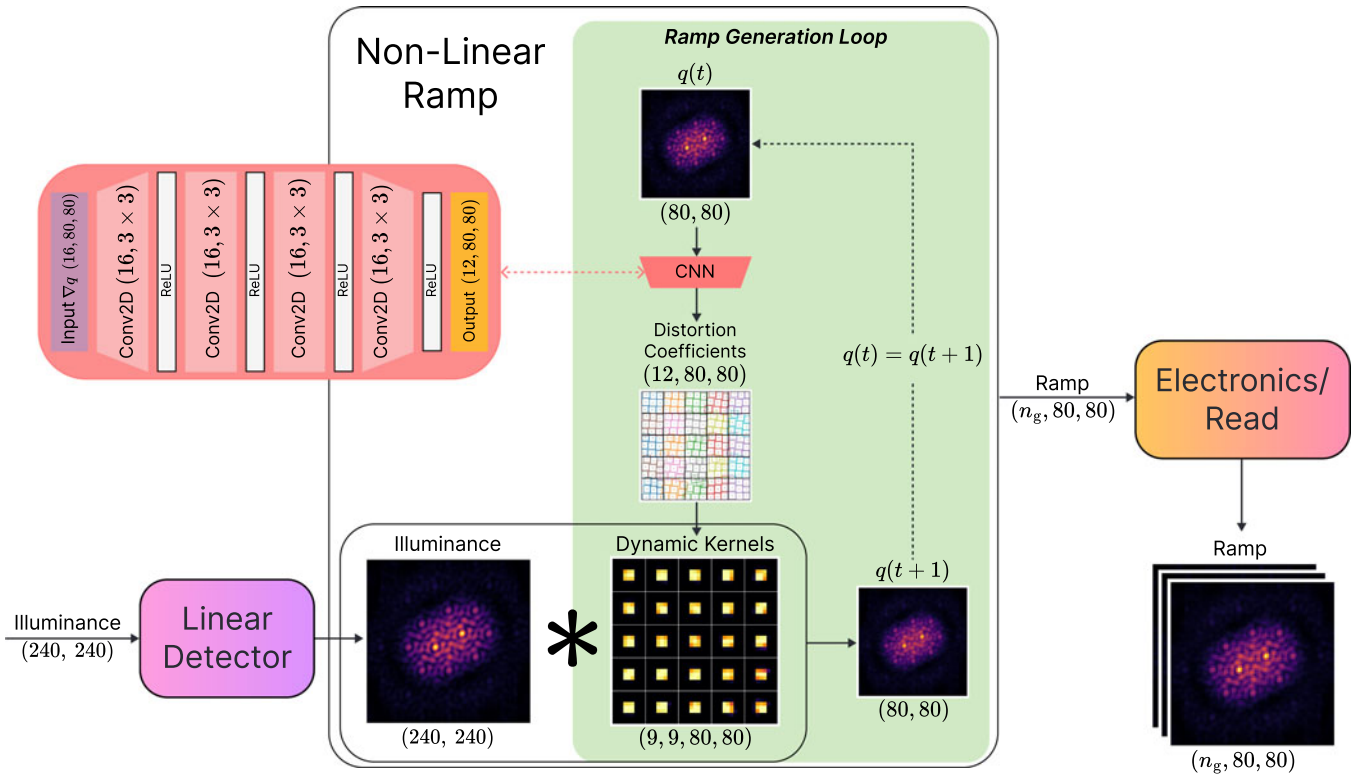


Figure 8. Schematic diagram of the EDM architecture, showing the three major components and the embedded neural network used to capture non-linear charge migration.

vibrational modes at frequencies much greater than a single integration in any of the optical components, such as the secondary mirror, or from electron diffusion in the detector substrate. It is not known which of these two potential sources is the origin of this blur. Instrumental jitter causes a substantial reduction in the sharpness of the PSF; this is visualised in Figure 4.

The next step requires resampling the predicted PSF from the optical coordinates to the detector pixels. The NIRISS detectors exhibit a documented anisotropy, as well as a small $\sim 0.5^\circ$ rotation with respect to the optical axis. Both of these operations are simultaneously applied via a cubic interpolation of the oversampled PSF, while preserving the $3\times$ oversample. Note that AMIGO does not use existing distortion solutions for NIRISS, since the PSF distortion effects in AMIGO are modelled via the pupil rather than focal plane. The final output is the oversampled illuminance pattern incident to the detector pixel.

4.2. Non-linear ramp model

The greatest challenge in the entire AMIGO pipeline is accurately capturing the dynamics of the BFE, handled by the non-linear ramp component of the EDM. The dynamics of the BFE should be expressible as a differential equation that can be evaluated and solved directly (as in MIRI, at longer wavelengths: Argyriou et al. 2023), but this is at present not tractable for the BFE in HgCdTe detectors. The approach used in that work, while accurate and crucial to understanding the true physics behind the BFE, falls short of the key requirements in this project: that it be differentiable and computationally efficient. Instead, we build a hybrid model, using an NN to capture the BFE dynamics without explicitly solving the complex differential equation.

The non-linear ramp model transforms the static 2D oversampled illuminance pattern into a 3D time-evolved charge accumulation ramp. Its construction is informed by the work on modelling the MIRI BFE (Argyriou et al. 2023), aiming to predict an evolving pixel area describing how charge transports from the excited photo-electrons in the photosensitive region down into the individual depletion layers of each pixel.

The time evolution of the charge accumulation is addressed by using a recurrent architecture, adding charge to each pixel in a fixed number of small time-steps. The measured charge at each group-read is found by interpolating the fixed number of time steps to the correct time stamp of each group read.

The evolving pixel areas are modelled with a dynamic filter (Brabandere et al. 2016), with each individual pixel in the 80×80 sub-array used for AMI observations using its own predicted unique convolutional kernel at each time step. Further complexities arise from the unique sensitivity of each pixel in this process (the flat-field). These variations result in either more or less charge accumulating in any one pixel, which then influences the resulting collecting area of both it and its neighbouring pixels. This mandates that both the inter- and intra-pixel sensitivity be modelled separately.

4.2.1. Neural network implementation and architecture

The non-linear ramp employs a recurrent neural architecture, behaving as a discretised neural ordinary differential equation (Kidger 2022) and mirroring the differential equation time evolution of the BFE as shown in Argyriou et al. (2023). Provided with the oversampled illuminance I and a charge distribution q_t, it seeks to predict the charge distribution at the time step q_{t+1} 10.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Using a Convolutional Neural Network (CNN) applied to the current charge distribution q_t, a set of polynomial distortion coefficients are predicted for each individual pixel. These distortion coefficients describe how each individual oversampled illuminance pixel is distorted as it travels through the electric fields of the detector substrate. Starting with a 3×3 rectilinear coordinate grid over each output pixel the distortions coefficients are applied, producing a local spatial transformation. These distorted coordinates are then used to calculate the overlap fraction from the new position with the neighbouring detector pixels. These overlap fractions then form the individual weights of the dynamic per-pixel kernels used to transport charge from the illuminance down to the detector pixels.

This construction is very important to ensuring physically constrained predictions by the CNN, since the overlap fractions from the illuminance pixels can only sum to one – directly enforcing flux conservation from the input illuminance. By wrapping the NN outputs in a kernel model that enforces these physical constraints the NN has greatly favourable properties to the system as a whole, easing the calibration process of both it and the wider forward model.

Inter- and intra-pixel sensitivity variations are also simple to include in this architecture. Since the model predicts where excited photo-electrons will be measured by the sensor, the quantum efficiency of this location can be baked directly into the kernels through a multiplication. A unique sensitivity value is used for each pixel in the detector, i.e. the flat-field. Intra-pixel sensitivity variations are parameterised by a simple quadratic function, using the same value for all pixels. Figure 9 shows the recovered intra- and inter-pixel sensitivity variations. These flat-field values are only calibrated using the publicly available in-flight flat-fielding data. This is a necessity in this process, as even though the AMI PSF covers many more pixels than clear aperture PSFs, many pixels in the AMI calibration data have little to no incident flux.

Using the dynamic per-pixel kernels that include pixel sensitivity variations, a dynamic convolution with the illuminance array I is performed to transport the charge from the detector surface down to the individual pixels. This output charge distribution can then be directly added to the present charge distribution q_t in order to predict the next charge state q_{t+1} and the processes repeated in order to predict the time evolving charge distribution in each pixel.

This recurrent processes is visualised in Figure 8, showing how the dynamical kernels are generated and applied. Figure 9 shows the various steps in the process. It presents the resulting pixel distortions and dynamic kernels as produced from an example input charge realisation. It also shows the final effect on the final measured PSF and how the charge bleeds between neighbours around the brightest pixels in the detector.

The embedded CNN in the non-linear ramp model is small with a simple architecture. Given that the charge bleeding between pixels is dominated by the electric field between each pixel, the input charge array is first convolved with 16 different spatial gradient kernels which are then fed to the CNN. These kernels are kept static and not trained. The CNN employs a total of 3 convolutional layers of width 16, followed by a final layer of width 12, and relu activations. All layers use a 3×3 kernel size and do not use bias terms, giving a total of 8 150 parameters – very small in the context of modern deep-learning. This structure is shown in Figure 8.

The recurrent and convolutional structures of the model were chosen to emulate the physical processes governing the BFE. Recurrent and convolutional networks behave as differential equations, encoding the temporal and spatial dynamics, respectively. The CNN depth was chosen to constrain the receptive field (the area of influence of a pixel a single step) to 7 pixels, and width of the layers were chosen ad-hoc, seeking to minimise the network complexity. Similarly, the number of recursions steps was chosen ad-hoc to reduce the model evaluation time while still able to accurately emulate the physical processes.

The choice to avoid biases in the layers ensures that the NN returns zero migration for an input array with zero charge, matching the inductive bias of our problem since zero charge results in zero electric field gradients and no charge bleeding. This further aids training as the CNN does not need to learn that no charge results in no bleeding. The 12 output channels of the final layer are the 2D polynomial distortion coefficients which are used to generate the normalised convolution kernels used to transport charge into each pixel. This hybrid model approach demonstrates the power that is gained by offloading the known physics to an encompassing forwards model that is able to provide both high-quality inputs and later transformations to outputs, while leaving the complex or unknown physics to a learned component.

This construction of a data-driven model of the BFE proves highly effective, but warrants careful examination in light of potentially spurious signals, particularly in the context of exoplanetary science. The hybrid approach of combining a forwards model with a NN enables outputs to be restricted to those consistent with the known physics, an outcome not commonly found within either physics-informed or traditional ML approaches. Standard practice in ML is to allow a NN to learn the physical laws that govern a system, or to penalise outputs that do not obey them. While often successful, such models remain capable of producing non-physical or invalid predictions, which is a potentially major failure mode in fields that rely on minute signals obscured by various noise sources (i.e. exoplanetary detection and characterisation).

The EDM designed in this work instead reformulates the NN as a latent data-driven module that predicts parameters of a larger *physically bounded* model. By producing the coefficients of a set of convolutional kernels that can only generate effects consistent with the known physics of the BFE — namely flux conservation and local influence – the EDM emulates the form of the differential equation governing the BFE while restricting outputs to physically valid behaviour. The failure mode of this model, arising from either poor upstream PSF prediction or poorly calibrated NN parameters, is a poor fit to the data, *not* invalid physical predictions or the injection of spurious signals. This formulation also simplifies NN architectures, reduces the number of parameters required, and enables training on very small unlabelled datasets.

4.3. Electronics and read model

The non-linear ramp model is able to produce highly-accurate charge evolution prediction in each pixel, but there remains one last set of transformations that are applied, arising from the electronic devices that measure the voltage in each pixel. This model handles four primary effects:

1. Dark current
2. Inter-Pixel Capacitance (IPC)
3. Non-linear gain
4. Amplifier noise.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

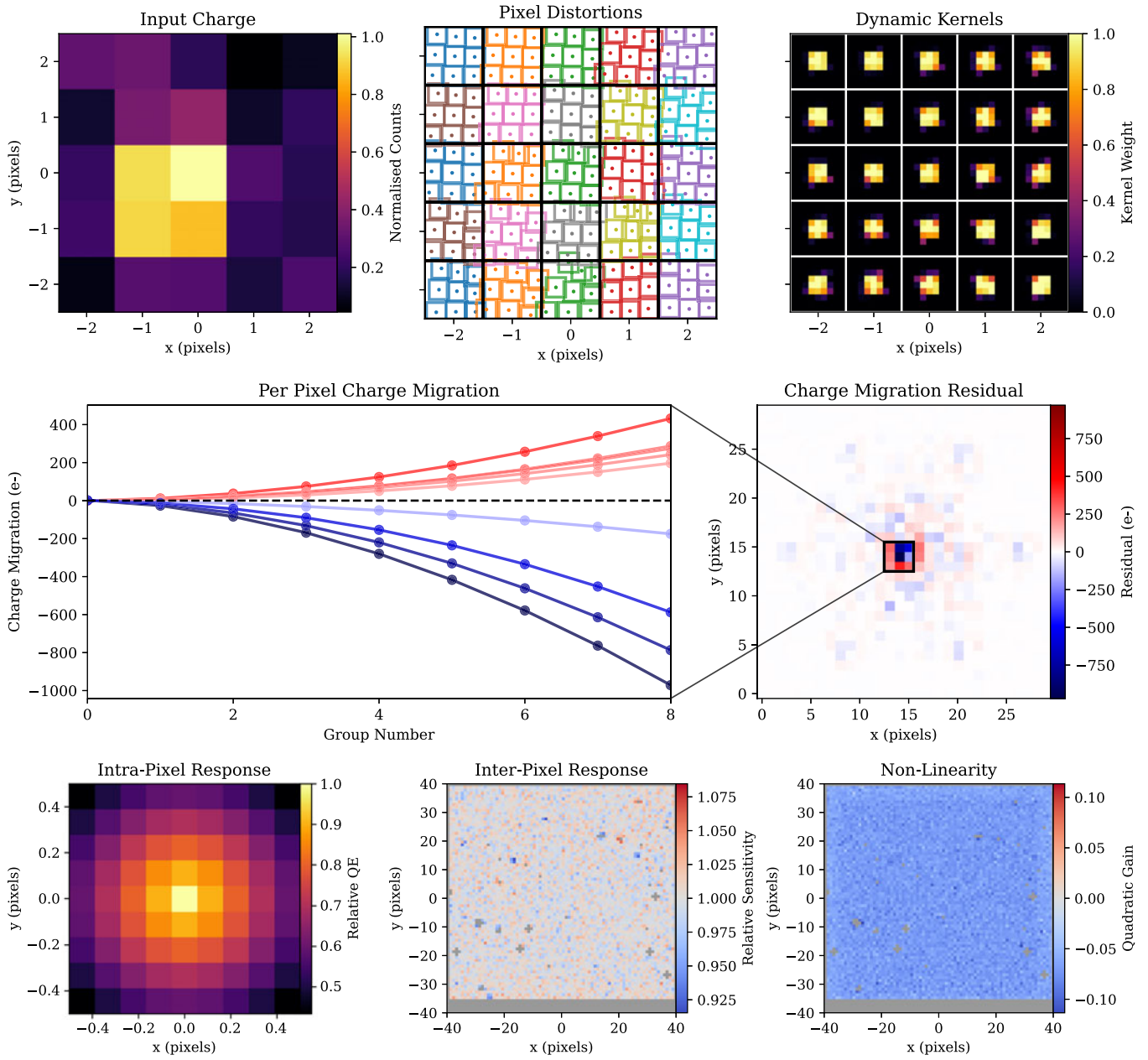


Figure 9. Demonstration of the EDM BFE model and recovered detector parameters. Starting with the normalised input charge distribution (top left), the CNN predicts a series of distortion coefficients that are applied to a 3× oversampled set of coordinates for each pixel. The distorted output pixel positions are visualised in the top middle panel. These output positions have their overlap fractions with each neighbouring pixel calculated in order to produce a set of flux conservative convolutional kernel for each pixel, shown in the top right panel. This dynamic convolution is applied to the predicted input illuminance pattern for a fixed number of time steps. The middle left panel shows the charge migration between a set of pixels around the brightest region of the PSF as charge is accumulated up the ramp. The middle right panel shows the cumulative charge migration between pixels for each individual pixel. The bottom row shows the various recovered pixel-level effects. The bottom left panel shows the recovered intra-pixel sensitivity variations, applied through a simple quadratic function. The AMIGO model only uses a 3× oversample, but is visualised here with 9× oversample. The bottom middle panel shows the inter-pixel sensitivity variations, ie the flat-field. The resulting distribution of sensitivities shows similar properties to the existing JWST pipeline calibrations. The bottom right panel shows the recovered quadratic non-linear gain term applied through the electronics model. Every pixel shows the expected negative gain as a function of pixel depth, in line with existing calibrations.

4.3.1. *Dark current*

Dark current effects are quite simple to implement. Individual contacts for each pixel heat up as they measure voltage and consequently can emit photons from the back of the detector. In practice the probability of this event varies from pixel to pixel, however AMIGO uses a constant value for all pixels both for simplicity and to avoid the including dark-current data in the calibration processes. This is desirable as dark-current calibration requires large files with many group reads in order to get sufficient signal on the singular photons emitted. AMIGO applies this dark current additively to each group read and a final recovered value of $\sim 0.45\ e^-$ is found.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

4.3.2. Inter-pixel capacitance

IPC is a generally well understood process in detectors – capacitive coupling between pixels results in the charge of a pixel influencing the measured value of its neighbours. IPC is conventionally modelled as a convolution; AMIGO employs this by convolving each group read with a static kernel. This is the only step where AMIGO uses ground-based calibration data products as the IPC is expected to be highly covariant with the charge migration effects of the BFE and therefore difficult to isolate. Using the ground-based calibration eases the training of the NN embedded in the EDM. The 5×5 IPC kernel, primarily composed of coupling to the directly adjacent neighbours, used is taken from the calibration-reference data system (CRDS).

4.3.3. Non-linear gain

NIRISS' H2RG detectors also exhibit a non-linear gain. This is modelled with a unique per-pixel quadratic polynomial response to the input charge. Typically, a much higher order polynomial is used (4th or 5th order), however higher order effects are small until pixel charges exceeds half their full well depth, a regime not recommended for AMI observations and that we have no calibration data for. As each pixel has a set of unique coefficients, it is only calibrated from the in-flight flat-fielding data, the same processes used for the intra-pixel sensitivity variations described in Section 4.2. This helps to avoid any biases introduced by uncalibrated optical or charge-migration systematics. The recovered per-pixel quadratic term is presented in Figure 9.

4.3.4. $1/f$ noise

$1/f$ noise, or red noise, is the thermal drift of the amplifiers as they read along the pixel columns (Rauscher et al. 2017). As the amplifier temperature drifts over time, the measured voltage drifts in kind, producing a visible vertical striping in recovered images. The magnitude of $1/f$ noise is larger for observations with fewer integrations and groups as the effect averages out to zero with an increasing number of reads. AMIGO has the ability to model $1/f$ noise with a low order polynomial added to each column for each group, however in practice this is unnecessary and usually skipped.

5. Base model training and calibration

The AMIGO model is calibrated using gradient-based optimisation, in which all model parameters (including both physical parameters and the embedded NN weights) are jointly fit. The validation datasets are simultaneously fitted (but not used for model calibration) to ensure model generalisation; these datasets are detailed in Section 2.1. The model is implemented in a fully differentiable framework, enabling end-to-end training of all components. This approach allows the full model – including optical system and EDM – to act as a single coherent system, representative of the true chain of physics leading to any observation.

Native gradient descent methods can be troublesome for data that has not been normalised, particularly when different datasets have different levels of total signal. Rather than normalising the data, AMIGO circumvents this issue by using a natural gradient descent approach (Martens 2020) which involves approximating the Fisher matrix of the parameters under the Laplace approximation (Kass, Tierney, & Kadane 1991; MacKay 2002). This enables the model *gradients* to be normalised, as opposed to the data. Fisher matrices are a natural tool for AMIGO as they can be efficiently calculated with autodiff for each dataset, added in order to reflect hierarchically constrained parameters, and approximated through Jacobians. The use of natural gradient descent speeds up convergence and transforms most parameter learning rates to order $\sim$ unity, all while preserving the Bayesian weighting between disparate datasets.

Each trained non-NN parameter uses a momentum-based optimiser and is assigned a unique learning rate. Complex visibilities are not fit during training. This is a deliberate choice: visibilities are highly covariant with optical parameters and 'soak up' residual systematics. All model calibration is done without ever transforming to the *uv*-plane.

The NN component is optimised with the Adam optimiser (Kingma & Ba 2014). Training is performed in batches to enable stochastic updates and efficient GPU usage. The calibration data was split into five batches, and their order randomised each epoch. Validation data is not batched because it does not influence the calibration of the model and stochastic behaviour has no effect. The NN is strongly covariant with physical model components, particularly during early stages of training where it may compensate for optical/detector effects. To mitigate this, the model is retrained multiple times from different initial weights while preserving the best-fit instrumental parameters. This staged optimisation process improves convergence and helps to separate the roles the NN and physical parameters play in capturing systematics.

Detector-specific parameters that are *unique* per-pixel are fit in parallel with, but independently from, the rest of the model. These components are trained exclusively on flat-field calibration data to avoids over-fitting arising from miscalibration elsewhere in the model and ensures that each pixel has the same total signal. The resulting flat-field maps as shown in Figure 9 exhibit smooth variations consistent with standard JWST pipeline calibrations, while the per-pixel non-linear gain terms match expected trends with signal depth.

All model training was performed on a NVIDIA RTX 4 080 GPU, with convergence reached after $\sim$20 000 steps across multiple re-trainings. Figure 10 presents the recovered wavefront state after the training process from the calibration data. There are only small differences between filters, which is expected given the different defocus values found in each filter (see Section 3.1.4). The resulting model shows a statistically good fit to the calibration data. Figure 11 shows a summary of the fits to the F430M filter for all five sub-pixel dither positions. Little to no PSF structure is present in the residuals. The residuals are close to being distributed as a unit-normal, indicating neither an under- nor over-fit. Identical summary plots for the other two filters are shown in Appendix A.

6. Inference from science data

The AMIGO framework is built around a fully differentiable forward model, which enables efficient gradient-based inference on both astrophysical and instrumental parameters as a digital twin of AMI. First we build this twin from high-quality calibration data; then, this same *base model* can be used in Bayesian inference by holding most parameters fixed and only fitting parameters of astrophysical interest like wavefronts, spectra, and source intensity distributions by gradient descent or HMC. A serialised version of this trained base model is included in the AMIGO repository and does not need to be retrained by users.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

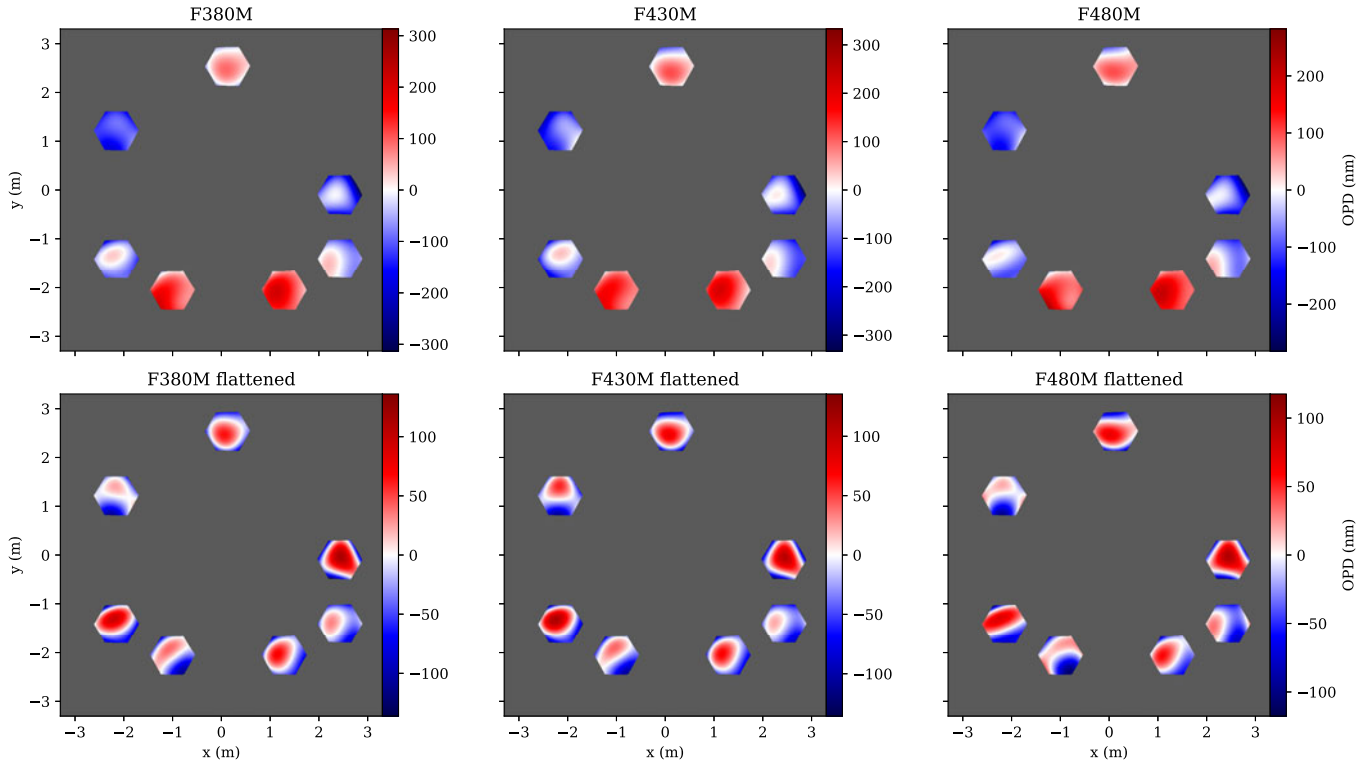


Figure 10. Recovered OPD maps from the calibration data for each filter. The top row shows the full OPD maps recovered for the F380M, F430M, and F480M filters, revealing large-scale piston, tip, and tilt terms across the segmented aperture. These low-order aberrations are expected given the off-axis sub-array placement used in AMI mode. The bottom row displays the same OPD maps with piston, tip, and tilt removed from each sub-aperture, enhancing the visibility of higher-order surface errors across individual segments. Residual aberrations at the ∼50–100 nm level are evident, and are consistent across wavelengths, suggesting a static contribution from optical surface figure errors or segment alignment offsets. Small differences are observed across filters, consistent with the observed Fresnel defocus discussed in Section 3.1.

In order to make this efficient for the end user as a black box, AMIGO caches large Jacobians and estimates per-dataset Fisher matrices which are used to normalise gradients, implementing natural gradient descent that improves convergence speed and robustness across diverse observations. For situations where further stability or precision is required, AMIGO also provides access to second-order optimisation routines and the ability to control learning rates over each recovered parameter. Users can choose to infer parameters either jointly or independently across multiple exposures – for instance, sharing optical aberrations across a sequence of observations while allowing astrophysical parameters to vary – and disentangling instrumental and scientific effects over multiple epochs.

By default, AMIGO only recovers a focused set of parameters: source positions, brightnesses, spectra, optical aberrations, and complex interferometric visibilities. Calibration parameters, such as pixel sensitivities, detector non-linearities, or the weights of the neural network within the EDM, fixed when fitting science data. However, thanks to its modular and generative structure, AMIGO allows users to selectively activate inference on any of these parameters when needed, or implement custom source parametrisations and modelling routines. Furthermore, while these are the default set of parameters, any parameter of the model can be solved for and constrained using any data. In practice this is necessary, but does enable the extension and re-calibration of the AMIGO model on data sets not well described by the default training data, such as non-sub-array image.

For example, as discussed in Section 4.3, AMIGO can model $1/f$ detector noise, but by default does not. Another example involves the documented physical alignment of the NIRISS aperture mask. Small differences in the mask rotation between calibration and science observations, especially when taken across different epochs, can introduce characteristic residuals. These are directly visible in the pixel-level residuals of the model fit. When present, AMIGO can infer a unique mask rotation angle per exposure, allowing this systematic effect to be characterised as part of the forward model.

In general, as a forwards model that infers parameter via Bayesian statistics, AMIGO is capable of placing priors over any set of parameters during its inference workflow. This can be achieved via a simple modifications of the likelihood functions used for optimisation through the inclusion of a prior over any model parameters. This also applies to model-fitting to the extracted visibilities – AMIGO is an end-to-end forwards model, enabling Bayesian methods to be applied widely to any component of the system or its outputs. This generalised formulation enables the AMIGO model to be with custom source parametrisation. For example rather than encoding source intensity distributions through interferometric observables, users can directly model a star-exoplanet system as two point sources modelled on the detector pixels, with particular spectral characteristic if desired.

6.1. Observable extraction workflow

The AMIGO pipeline is designed to extract the DISCO observables discussed in Section 3.2.3, starting from uncalibrated JWST data. Here we detail the basic workflow used for this process, comprised of three steps:

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

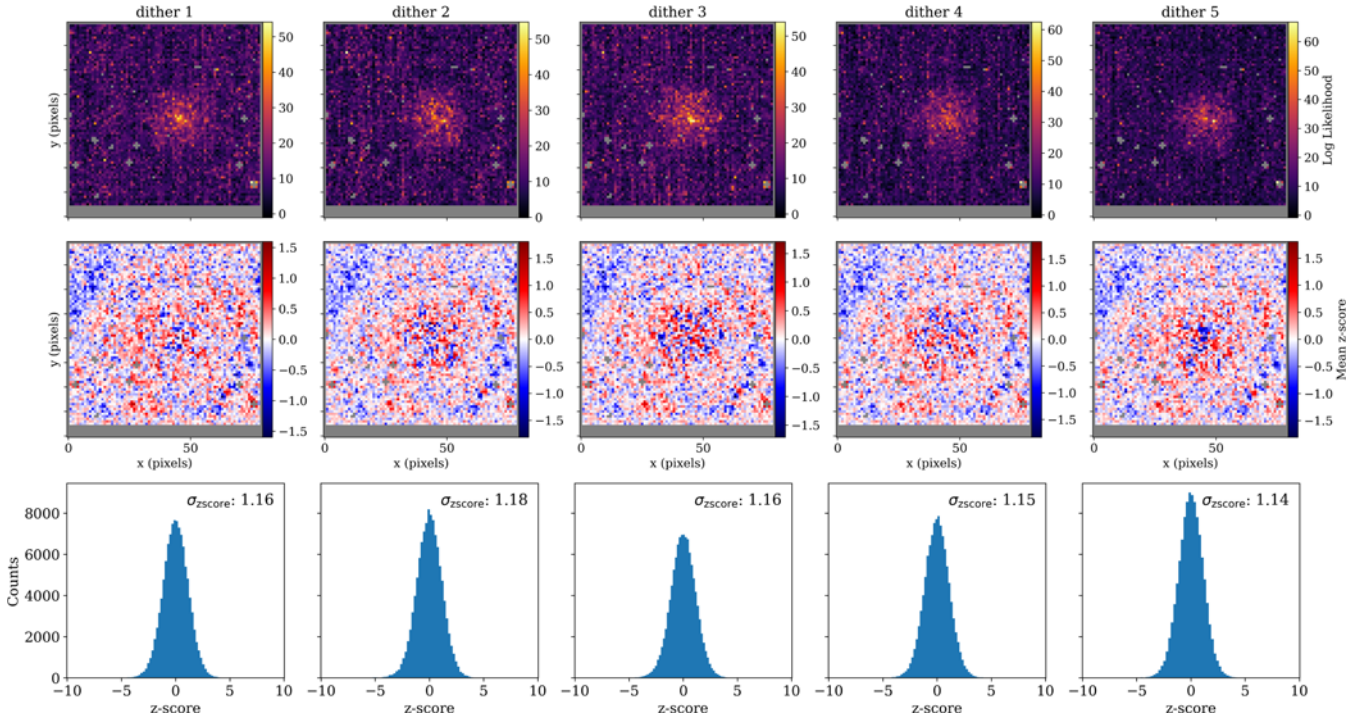


Figure 11. Summary of AMIGO model fit diagnostics across all five dithers for the F430M calibration data. Each column corresponds to a single dither position. The top row shows the per-pixel log-likelihoods from the final fit, highlighting the location of the target PSF. The middle row displays the average residual z-score per pixel, computed by averaging the uncertainty-normalised residuals across all groups in the ramp, revealing the spatial structure of any systematic model misfit. The bottom row shows histograms of all z-scores across pixels and groups for each dither, without any averaging over the groups. A perfect fit would have a standard deviation in the z-score be 1; we recover values between 1.1 and 1.2 in all three filters, indicating a good fit that has not learnt any noise characteristics. We note that the full likelihood is described with a covariance matrix that accounts for the anti-correlation between adjacent group-reads seen in slope data. Consequently, these summary statistics are only an helpful approximation and correct performance can only be described through the likelihood.

1. Data processing
2. Model fitting
3. Visibility reduction.

Each of these steps are detailed in the following sections, with Figure 12 detailing the overall workflow from data to DISCOs. As a forward-modelling pipeline, AMIGO operates differently to most existing tools. While default operation with little user input should suffice for the majority of use cases, the model provides great flexibility over how observables are extracted and is able to directly solve for multiple non-standard effects when they influence the data.

Despite modelling the full system, AMIGO still requires a PSF reference star in order to provide high-quality calibrated observables. As such, standard interferometric observing programs are recommended following the recommendation in JDox. Note that a further set of recommendations are provided in Section 8 designed to mitigate any potential model micalibrations.

6.1.1. Data processing

The initial data processing step is to reduce the 4D `uncal` JWST data to the `calslope` data produced used by AMIGO as described in Section 2.2. This is an automated procedure, and unlike the standard JWST data processing pipelines does not require any external calibration files. Users are able to chose the Nσ threshold used for outlier rejection, but defaults to 3σ. Default operation will also apply the ADC correction as described in Section 2.2, which should always be performed. Both the PSF reference and target stars should use the same data processing configuration. A notebook detailing this process is provided.

6.2. Model fitting

The model fitting step is the most important to the recovery of high-quality observables. As a forwards model of the instrument, the AMIGO pipeline can fit a number of different source (i.e. astrophysical) models to any provided `calslope` data, such as a point source or visibility model. Standard operation for interferometric observable recovery should use the `SplineVisFit` class on both the PSF reference and target stars, which take the `calslope` file as it input and produces an `Exposure` object. The `Exposure` class operates on the AMIGO model in order to produce a predicted observation, and we maximise the log-likelihood of this model with respect to data. Flags in this class control which parameters in the base model are fixed and which are recovered from the data, for example the recovery of the $1/f$ stripe signal.

As discussed in the previous section, under default settings we recover source position, flux, spectral slope, visibilities, and wavefront phase. The AMIGO pipeline will not recover the wavefront phases from non-PSF reference stars, as the wavefront phase and interferometric signal are strongly covariant. The wavefront phase is recovered only from the PSF reference stars, and fixed in inferring visibilities from the science target and in generating the kernel observable operator. Users have full control over what is recovered from data, and for example can infer more complex spectral models; or different wavefront states can be used for observations taken in different epochs.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

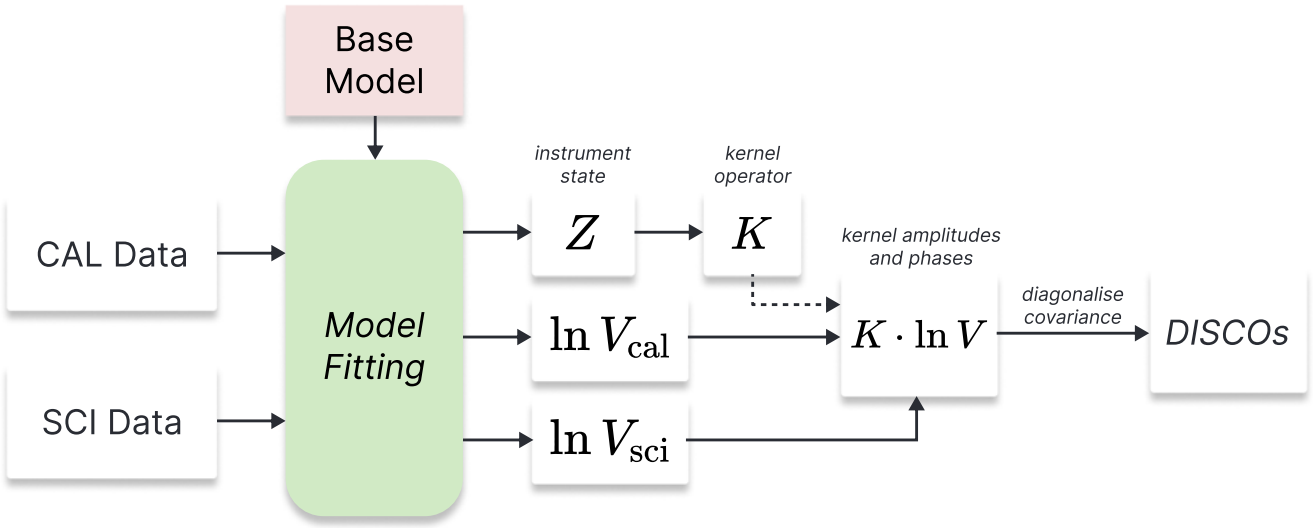


Figure 12. Schematic diagram of the basic workflow of extracting the DISCOs from data.

With these model fit objects, users now need to construct the AMIGO base model used by the fit to predict observations by loading a pre-calculated calibration file (produced as described in Section 5), which is released for each version of AMIGO.

With the calibrated base model and the model fit objects the observables can be recovered from the data. This is done via natural gradient descent, minimising the log-likelihood, using the Fisher matrix (which can be constructed approximately from pre-calculated Jacobians) to normalise each parameter of the model so that we require only one gradient descent hyperparameter. Model fits to both the target and PSF reference stars must be done at the same time to ensure accurate wavefront phases are used on the target stars. These can be done in batches to reduce both computational and memory overheads. The model should be optimised for ∼200 epochs for good convergence, however this can vary depending on the specifics of the input data. There is no fixed termination criterion, and convergence must be determined by the user. In some cases tweaks to the parameter learning rates is required. While it is currently experimental, AMIGO also provides an interface to the BFGS optimiser (Broyden 1970; Fletcher 1970; Goldfarb 1970; Shanno 1970) as implemented in OPTIMISTIX (Rader, Lyons, & Kidger 2024) which provides improved convergence guarantees.

With the fitted model, the parameter uncertainty covariance matrices are estimated via the Hessian of the log-likelihood evaluated at the best-fit parameters. For computational efficiency this is only done over the recovered visibility parameters. These parameters and their uncertainties are then saved to an intermediate data product which is next used for interferometric observable reduction and calibration. A notebook detailing this process is provided.

AMIGO also supports the fitting of data with or without dithers. The initial observing strategy for AMI recommended not using dithers due to the pointing precision of JWST, however, as is recommended in Section 8 employing a 5-point sub-pixel dither is recommended to mitigate pixel-level effects and improve the precision of AMIGO observables. AMIGO is able to either jointly solve for interferometric observables jointly across all dithers, or solve for them independently to later be combined into a single set of observables per source as is seen in traditional inverse pipelines.

6.3. Visibility reduction

The next step projects the recovered observables to the kernel space and calibrates them. The recovered wavefront state is loaded into the AMIGO model and we calculate the Jacobian of the visibilities with respect to not just pupil Zernike phase modes but also flux, spectral slope, and Fresnel defocus. A singular value decomposition is performed on the Jacobian and used to construct a kernel basis and used to project the complex visibilities to kernel observables. This projection is done on both the PSF reference and target stars, where the target star visibilities are calibrated by a subtraction of the reference star visibilities. This operation is subtraction, not a division, as AMIGO recovers the log complex visibilities (see Section 3.2.1). The same projection is performed on the parameter covariance matrices to propagate uncertainties correctly.

The final step in this procedure is the projection to the statistically-independent DISCO space, as detailed in Section 3.2.3. This is found via an eigendecomposition of the calibrated kernel visibility uncertainty covariance matrix, where the eigenvectors are used to project both the visibilities and covariance matrices into a statistically-independent space (Ireland 2013), yielding the DISCO observables. Finally, a `calvis` data product is produced that reduces all of the remaining relevant information needed for downstream interferometric analysis. It provides the *uv* coordinates of the observation, mean wavelength, a single DISCO matrix that maps from the pixelised *uv* coordinates back to the DISCO space, as well as all the intermediate data products. A notebook detailing this process is provided.

This is considered the end of the basic AMIGO pipeline, which is designed to extract calibrated interferometric observables. However, given the unique data product of AMIGO, a number of classes and interfaces useful for fitting interferometric models to the DISCOs is also provided. These are simply example classes to help users with their final scientific inference,

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

as in practice the particulars of any given data set can require the construction of parametric models, or more complex analysis routines.

7. Results

This section offers a quantitative assessment of the AMIGO model's performance on both medium and high contrast interferometric imaging data from JWST NIRISS AMI mode. Rather than a reduction to calibrated visibilities and closure phases, AMIGO infers DISCOs (described in Section 3.2). These latent observables are robust to both low- and moderate-order wavefront miscalibration and capture the complex visibilities over the full OTF and sources across the FOV. As such, direct comparisons using conventional metrics like closure phase scatter are not applicable. Instead, performance is evaluated empirically via the recovery of known companions in representative and publicly available archival data. Two observing programs are used for this comparison: COM 1093 (PI: Thatte), a preliminary exploration of performance on the AB Dor system during the commissioning phase, and GO 1843 (PI: Kammerer), a deep observation pushing the limits of the instrument seeking to characterise the red substellar companions around HD 206893.

The theoretical performance of an AMI system is commonly benchmarked using the expected contrast floor for detecting a point source near a host star. This is determined by the closure phase uncertainty $\sigma_{\rm CP}$ (Ireland 2013)

$$\sigma_{\rm CP} = \frac{N_h}{N_p}\sqrt{1.5N_p} \tag{10}$$

where N_p is the number of photons collected and N_h is the number of holes in the NRM. In the high-contrast regime $\sigma_{\rm CP}$ is used as an approximation to the 1-σ per-observable detection limit in *contrast* of a companion (Sivaramakrishnan et al. 2023). For a 7-hole mask like AMI this provides an approximate contrast limit cited in JDox as a recommendation for exposure time calculations

$$\text{contrast limit} \approx \sqrt{100/N_p}. \tag{11}$$

This value does not directly translate to a principled 'Nσ detection' limit as it is not cognisant of mask or signal geometry, nor analysis methodology; it also does not appear in this form in the Ireland (2013) paper. Nevertheless as shown below, this rule of thumb for the contrast limit proves to be accurate.

7.1. Interferometric model fitting

Model fits to the calibrated observables are performed with the software DrPangloss, an under-development Jax accelerated interferometric analysis package (Blakely et al. 2024). While the final output product of the official AMIGO pipeline are the DISCOs, AMIGO also provides basic interfaces for integration with DrPangloss given its natural synergy as Jax + Zodiax-based software and to facilitate seamless analysis of its novel output products.

To fit interferometric models to the data, first the calibrated DISCOs are loaded into the `AmigoOIData` class which holds the recovered observables, the uncertainties, the *uv* coordinates, and the DISCO matrix that maps between the *uv* plane and the DISCO space. There are two main procedures used to fit high-contrast models to the DISCOs, which are used in both of the observing programs analysed in this manuscript. To identify the existence of any companions, a grid-search is performed over a 1 arcsec radius of the primary star. This grid search is done by solving for the best-fitting contrast at each position in the grid by minimising the log-likelihood of the predicted binary via a BFGS algorithm. This fit is performed via a vectorised operation in Jax, providing orders of magnitude speed up over comparable software like CANDID (Gallenne et al. 2015), taking ∼20 s in each filter on an NVIDIA RTX 4080 Graphics Processing Unit (GPU). This produces a log-likelihood detection map that can be used to identify companions through the maximum likelihood estimate. Both the best fit companion position and contrast at the maximum likelihood estimate can then be used to infer its properties via an Monte Carlo Markov Chain (MCMC; Metropolis et al. 1953) sampler: we use HMC as implemented in the probabilistic programming language NumPyro. HMC achieves much better performance on high-dimensional problems than other MCMC samplers by taking advantage of derivative information.

For companion astrometry, we adopt a uniform prior over position in RA and Dec with width 150 mas, and a uniform prior over log-contrast with width 2 orders of magnitude, in both cases where the means of the priors are determined by the best-fit values found from the grid search. The HMC chain also fits a multiplicative error term $\sigma_{\rm scale}$ to both the DISCO amplitude and phase uncertainties. This has the effect of pushing the χ^2_ν to ∼1 during the MCMC sampling. This operates under the assumption that the model being fit (a point source in this case) is accurate, allows for the sampler to operate more efficiently, and provides an understanding over the quality of the resulting fit, the uncertainties of the recovered parameters, and the quality of the estimated errors on the data. As our uncertainties are derived from propagating photon noise uncertainty through the pipeline, $\sigma_{\rm scale}$ is approximately the ratio of our data scatter to the ideal photon noise limit. This process is performed both jointly across all filters, and uniquely per filer.

In the case of multiple companions, as with the GO 1843 program, the best-fit primary companion found via MCMC can be subtracted from the data. Then the grid-search can be repeated in order to find the position and contrast of any other companions. Next the MCMC fit is performed simultaneously using a joint model of both companions in order to recover precise parameters for both companions.

7.2. COM 1093: Commissioning data – AB Doradus AC

The AB Dor system was the first observed by AMI as part of the COM 1093 program during JWSTs commissioning phase. This program aimed to demonstrate the detection of point source companions at moderate contrasts and test target placement precision through primary and sub-pixel dithering operations. Preliminary analysis of this data produced results inconsistent with pre-flight estimates (Sivaramakrishnan et al. 2023). The recovered contrast limits fell short of expectations by ∼0.5–1 mag, making this program ideal to benchmark AMIGO.

To test AMIGO's performance on point-source companion recovery, we use a subset of the data in line with AMIs original recommended observing strategy of single exposures with no dithers. Table 3 details the subset of exposures used for analysis. In practice better parameter fits could be extracted by employing all of the data and dithers across the program. However, here we only use a subset of the available data in order to match the analysis done in Sivaramakrishnan et al. (2023) and provide a fair benchmark.

Default operation of AMIGO takes likelihood statistics over *slopes* rather than ramps. This aids problems that arise from predicting the pixel-to-pixel bias level that can result in model

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 3. Summary of COM 1093 program observations used to test the AMIGO model. The number of photons is an estimation from the raw data, and details the number of usable photons after accounting for the 1/groups fractional loss from discarding the first group of an integrations. The percentage loss of photons is also shown.

Observation	Type	Star	Filter	Groups	Integrations	Photons	Loss (%)
12	SCI	AB Dor	F480M	5	69	92×10^6	20.0
12	SCI	AB Dor	F430M	4	82	94×10^6	25.0
12	SCI	AB Dor	F380M	2	160	97×10^6	50.0
15	CAL	HD 37093	F480M	12	61	85×10^6	8.3
15	CAL	HD 37093	F430M	9	78	104×10^6	11.1
15	CAL	HD 37093	F380M	4	118	100×10^6	25.0

over-fitting as discussed in Section 2.2. While this is not a fundamental problem and future releases of AMIGO plan to directly address this, presently AMIGO discards the first group read from an exposure resulting in a loss of data. Exposures with few groups are more affected as the first group forms a larger fraction of the overall data. As such, AMIGO as currently working operates on *less* overall signal than conventional pipelines. Improving use of low group count data in an important step in AMIGO development. Table 3 also presents the approximate photon counts in the COM 1093 dataset used for analysis, along with the fraction that AMIGO uses. Examination reveals that each filter has a total of $\lesssim 10^8$ photons each, which yields an estimated contrast floor of $\lesssim 7.5$ mag from Equation (11).

7.2.1. COM 1093: Analysis

The raw `uncal` exposures were downloaded from Mikulski Archive for Space Telescopes (MAST) and processed into the `calslope` format using the default operation of the AMIGO data processing pipeline, described in Section 2.2. The calibrated AMIGO model, described in Section 2, was then fit to the processed data to recover the latent visibilities described in Section 3.2.1. The recovered wavefront phases were used to calculate the DISCO basis used for the output observables of the AMIGO model.

Figure 13 shows the recovered log-likelihood detection maps in each filter, calculated with a binary fit to contrast at each RA, Dec with the models provided by DrPangloss. These show a confident and independent detection of the known companion in all three filters.

To recover posterior samples of the detected companion we run an HMC sampler using `numpyro` (Phan, Pradhan, & Jankowiak 2019; Bingham et al. 2019) and DrPangloss. While the companion is detected independently in all filters, we perform a broadband joint-fit to the astrometry of the companion to ensure the best constraints on both the astrometry and photometry. To assess the quality of the fit and the recovered parameters we perform a predictive posterior check – visualising the correlation and residual between the recovered and predicted DISCOs found with the MCMC samples and shown in Figure 13. Multiplicative error inflation factor $\sigma_{\rm scale}$ is fit to account for and quantify uncertainty in the AMIGO estimated errors. These terms remain close to one, indicating both a good fit and approximately accurate uncertainty quantification. The full posterior samples are shown in Figure 14, and the resulting fit in Table 4. The apparent parameter degeneracy seen between the position angle and separation in Figure 14 is due to high signal-to-noise of the data and the projection geometry from Cartesian to polar coordinate systems. A well-constrained companion will produce circular likelihood contours in Cartesian coordinates, but when projected into polar coordinates results in banana-like contours. However, given how well the companion is localised, only a very small range of position angles form the posterior and only a small arc of this banana is realised, resulting in a Gaussian-like apparently degeneracy.

To quantify the overall detection limits the best-fit companion is subtracted from the recovered DISCOs, while applying the error jitter term $\sigma_{\rm scale}$. Using these cleaned DISCO values we calculate the azimuthal Bayesian upper limits following the methods introduced by Ruffio et al. (2018), shown in Figure 15. This demonstrates that the AMIGO model performs at the expected contrast limits of the instrument, indicating it is well calibrated for high-contrast studies.

7.3. GO 1843: HD 206893 at high contrast

To demonstrate the performance of AMIGO on a deep AMI dataset, we analyse archival images of the HD 206893 exoplanetary system at three wavelengths (F380M, F430M, F480M). These images are part of GO 1843 program with the ultimate goal of detecting the brown dwarf HD 206893 B and planet HD 206893 c at angular separations of ∼200 and ∼100 mas, respectively. The three NIRISS AMI filters chosen for this program probe the presence of CO and CO_2 at 3–5 μm wavelengths. Together with the data at shorter wavelength measured from the ground, this constrains the carbon and oxygen chemistry in the atmospheres of the two substellar companions and provides insights into their formation history. Additionally, photometric constraints at the AMI wavelengths can constrain the origin of the extremely red near-infrared colour of HD 206893 B, which is also the reddest known substellar companion. This system is ideal for testing performance because prior observations with VLTI/GRAVITY provide highly-precise constraints on the companion orbits (Hinkley et al. 2023). This paper presents the detection of the HD 206893 B/c companions with AMIGO as a demonstration of its performance close to the theoretical contrast detection limits based on photon noise. A future publication will present the scientific analysis of the JWST photometry in the context of the atmospheric carbon chemistry and the dust cloud properties (Kammerer et al. in preparation).

This dataset represents a far greater challenge to AMI, seeking the recovery of companions close to the theoretical limit of recoverable contrasts and pushing beyond the diffraction limit for the inner companion. A classical analysis of the HD 206893 images with Amical (Soulain et al. 2020) provides a tentative detection of the brighter B companion in all three bands (without c), but

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 4. Best-fit relative joint astrometry and per-filter photometry for the AB Dor C companion, showing the $+/-1\sigma$ bounds. Reported quantities include the on-sky separation (Sep), position angle (PA), and Δ-magnitude in each of the F380M, F430M, and F480M filters. Results from two different fit types are compared: A joint fit in all three filters and a fit to each filter uniquely.

Fit type	Sep (mas)	PA (°)	Δmag (F380M)	Δmag (F430M)	Δmag (F480M)
Joint	$328.85^{+0.48}_{-0.85}$	$169.482^{+0.020}_{-0.022}$	$4.313^{+0.013}_{-0.023}$	$4.354^{+0.016}_{-0.012}$	$4.381^{+0.017}_{-0.016}$
F380M	$329.9^{+1.3}_{-1.2}$	$169.521^{+0.045}_{-0.033}$	$4.301^{+0.021}_{-0.016}$	–	–
F430M	$328.58^{+0.98}_{-1.18}$	$169.468^{+0.033}_{-0.036}$	–	$4.357^{+0.014}_{-0.015}$	–
F480M	$327.5^{+1.3}_{-1.1}$	$169.447^{+0.041}_{-0.036}$	–	–	$4.377^{+0.018}_{-0.015}$

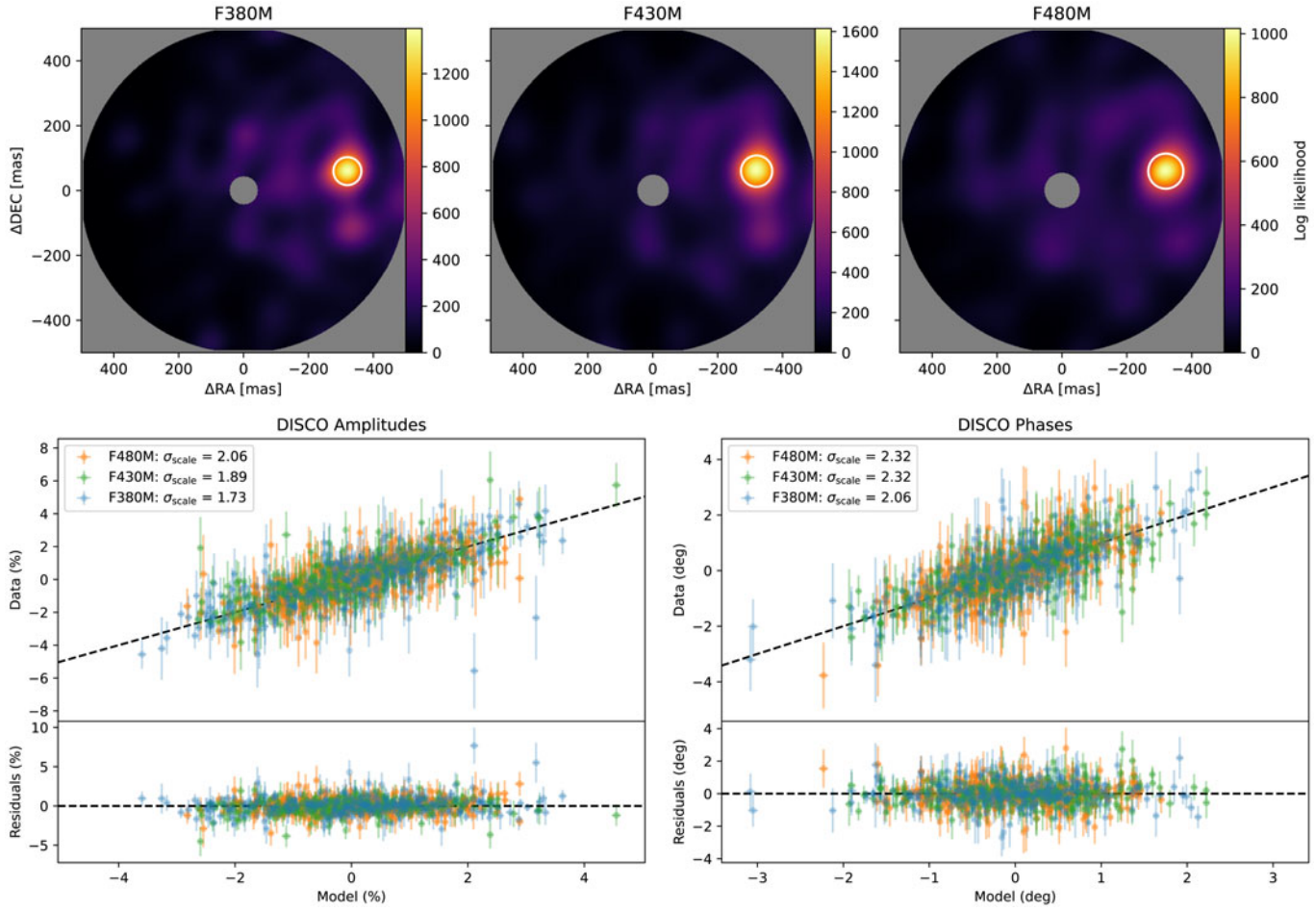


Figure 13. Summary of the fits to the AB Dor C companion. The top panels shows the log-likelihood detection maps over companion position for each filter (F380M, F430M, F480M) from a grid search over the recovered DISCOs. Each panel shows the marginalised log-likelihood surface as a function of companion offset in ΔRA and ΔDEC, revealing a strong and consistent peak in all filters. The greyed central region denotes the IWA masked by the interferometric null. The recovered peak likelihood location is consistent across filters, indicating the reliability of the model prediction. The bottom panels show the predictive posterior check comparing AMIGO recovered DISCO amplitudes (left) and phases (right) against the predicted values from the MCMC samples, for the three filters: F380M, F430M, and F480M. Each panel shows individual measurements with 1σ error bars, with the top panels plotting model predictions against data and the lower panels showing residuals. The black dashed line represents the 1:1 line for a perfect model prediction. Scale factors $\sigma_{\rm scale}$ are applied to the observational uncertainties with the effect of ensuring $\chi^2_\nu \approx 1$ during MCMC fitting and are quoted in the legend for each filter. The agreement across all filters and observables confirms both the validity of the forward model and appropriate uncertainty quantification in the recovered posterior.

at a position that is not compatible with the orbit of the object – known to $\sim$1 mas from GRAVITY observations – by several tens of mas. This is likely due to uncorrected charge migration and distortion effects. When fixing the position of the companion in the fit, reasonable photometric constraints in all three bands can be obtained, but at much lower signal to noise than what would be expected without detector systematics. The photometry is also biased by these systematic effects. The fainter c companion remains undetected below the systematic noise when using the existing analysis pipelines.

Analysis with AMIGO yields significantly better results than classic approaches. We find a confident detection of both companions in agreement with GRAVITY data, with well constrained photometry even when analysed per filter.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

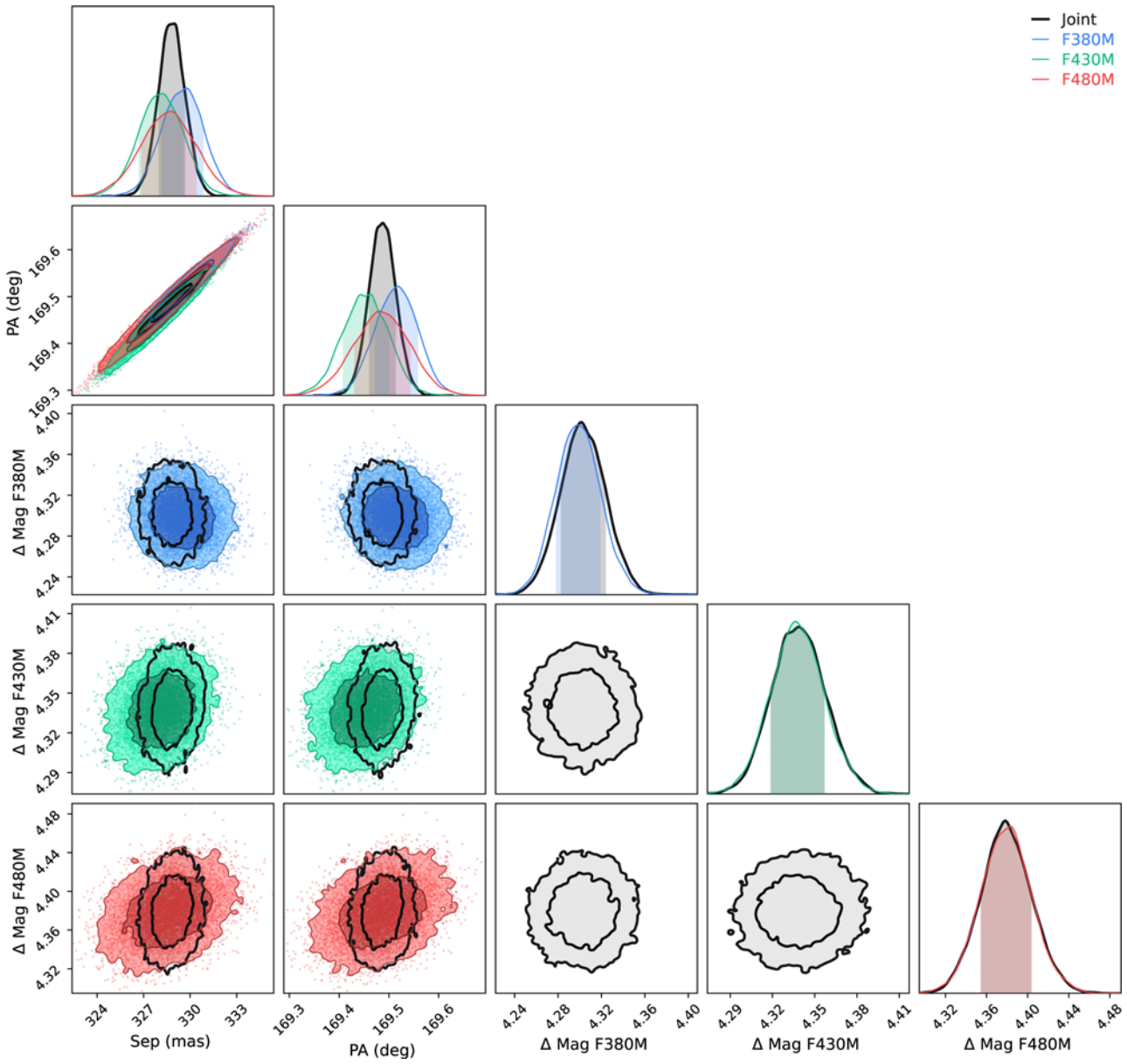


Figure 14. Joint posterior distributions from an MCMC fit to the AB Dor C companion. Two fit types are shown: Joint – simultaneously modelling astrometry and photometry across all three filters – and per-filter fits. The parameters shown include the separation (mas), position angle (degrees), and contrasts (Δmag) in the F380M, F430M, and F480M bands. The joint-fit samples are shown in black. One- and two-σ credible regions are shown in dark and light shades, respectively. Strong correlation is observed between separation and position angle due to the projection geometry, while photometric parameters are weakly correlated and independently constrained in each filter. The tight constraints in both astrometry and photometry reflect the high signal-to-noise of the companion detection.

The targeted observational approach of the GO 1843 program provides a standard template for high-contrast companion recovery: one target and one calibrator in the three primary filters, F380M, F430M, F480M, without any dithers. Here we use all of the exposures, summarised in Table 5. The GO 1843 data targets are dimmer than those in the COM 1093 program (with more photons at shorter wavelengths), providing more groups per integration and preserving more photons. Table 5 also provides an overview of the approximate photon counts in each exposure as well as the fraction that is lost through the present AMIGO pipeline. Shorter wavelengths lose more photons due to having fewer groups but most of the signal is preserved. From these values and Equation (11) this gives an expected contrast limit of ∼9.66, 9.24, and 8.78 Δ mags in the F380M, F430M and F480M filters respectively.

7.3.1. GO 1843: Analysis

The same process used in Section 7.2.1 was used for the GO program, taking `uncal` exposures from MAST and processing into `calslopes` with the default AMIGO data processing pipeline, then fiting and recovering the instrument state, latent visibilities, and DISCOs.

As a high-contrast companion observation we use the same methods for analysis of the DISCOs, expanded to account for the second companion. Figure 16 shows the log-likelihood detection

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 5. Summary of GO 1843 program observations used to test the AMIGO model. The number of photons is an estimation from the raw data, and details the number of usable photons after accounting for the 1/groups fractional loss from discarding the first group of an integrations. The percentage loss of photons is also shown.

Obs	Type	Star	Filter	Groups	Integrations	Photons	Loss (%)
1	SCI	HD 206893	F480M	11	720	1.05×10^9	9.09
1	SCI	HD 206893	F430M	8	2 177	2.47×10^9	12.5
1	SCI	HD 206893	F380M	4	7 046	5.31×10^9	25.0
2	CAL	HD 205827	F480M	10	641	0.86×10^9	10.0
2	CAL	HD 205827	F430M	7	1 885	2.15×10^9	14.3
2	CAL	HD 205827	F380M	3	7 800	4.74×10^9	33.3

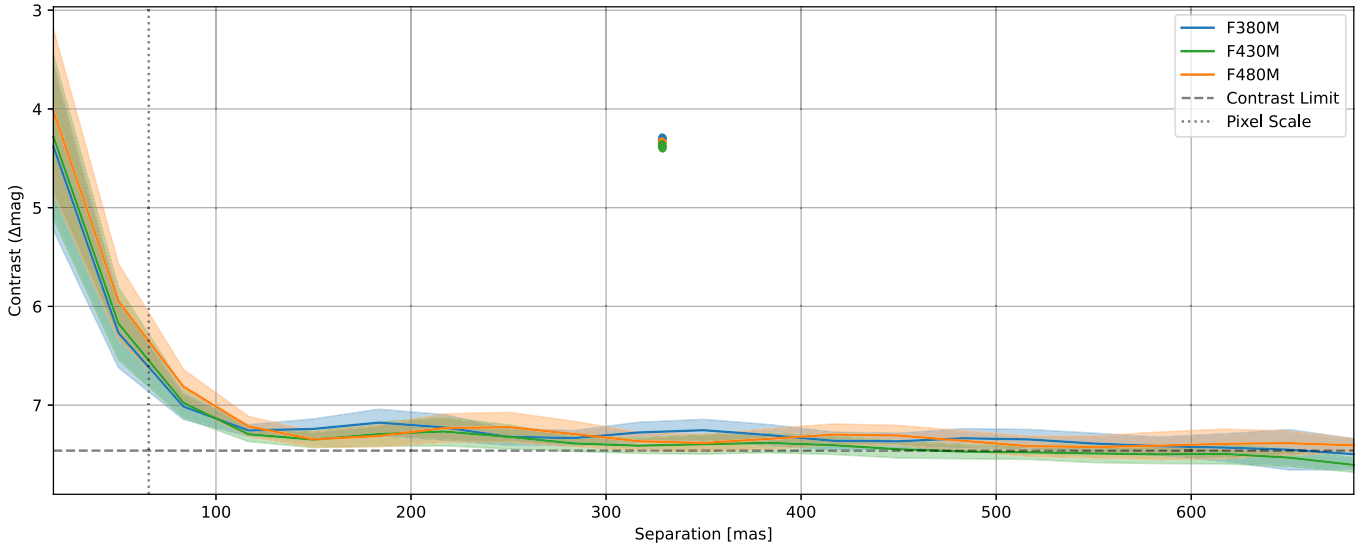


Figure 15. Sensitivity limits derived using the Ruffio et al. (2018) method for calculating 3σ contrast upper limits as a function of angular separation, applied to the F380M, F430M, and F480M filters. The shaded regions denote the $\pm 1\sigma$ variation across azimuthal annuli. The dashed black line indicates the approximate contrast limit as calculated from Equation (11), confirming that model performance is consistent with the expected limits. The grey dotted line shows the size of a single pixel for reference. The recovered companion 3σ contour plots are shown in their respective colours but are so well constrained they only appear as single dots on the figure. These curves quantify the AMI contrast performance and establish detection limits for additional companions in the field.

maps for the full data, along with the data after the best-fit B companion has been subtracted off. Subtracting the B companion helps to reveal the dimmer c companion that can be obscured by the signal of the B companion. This shows a strong detection of both companions with statistically consistent astrometry (as compared against VLTI/GRAVITY constraints provided in Hinkley et al. 2023), independently in all filters.

Full inference is performed with a joint MCMC fit to both the B and c companions with astrometry constrained across all filters to ensure the most accurate recovery of photometry. The quality of this fit is assessed in Figure 16 via a correlation plot between the measured and predicted DISCOs. The recovered error scaling terms $\sigma_{\rm scale}$ are all close to unity, and reasonably consistent with the values found in from the fit to the AB Dor companion, indicating consistent performance in the high contrast regime and a good fit to the data. While both the B and c companions are simultaneously fit in the HMC, we present the samples for their parameters independently to prevent clutter. Examination of the full samples reveal insignificant correlation between the parameters of the two companions. Figure 17 shows the parameter posteriors for the B companion and Figure 18 shows the same for the c companion. A summary of the recovered parameters is presented in Tables 6 and 7.

Quantification of the detection limits in this dataset is found by subtracting out the best-fit values from the MCMC while also applying the error scaling term to the data uncertainty. The azimuthal upper-limits are calculated on the cleaned DISCOs via the Ruffio et al. (2018) method and shown in Figure 19. This reveals that the performance of the AMIGO model remains in strong agreement with the expected contrast limits even down to ∼9.5 mag, inside the average diffraction limit of of all three filters, ∼120 mas.

8. Discussion

The results presented in this work demonstrate that a robust analysis of JWST AMI data through differentiable forward model provides a new way to calibrate and extract observables. These methods provide insight into a new set of observing strategies better suited to these approaches. Prior investigations by Sallum et al. (2024) have led to a set of three recommendations for future observations to mitigate the effect of charge migration:

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

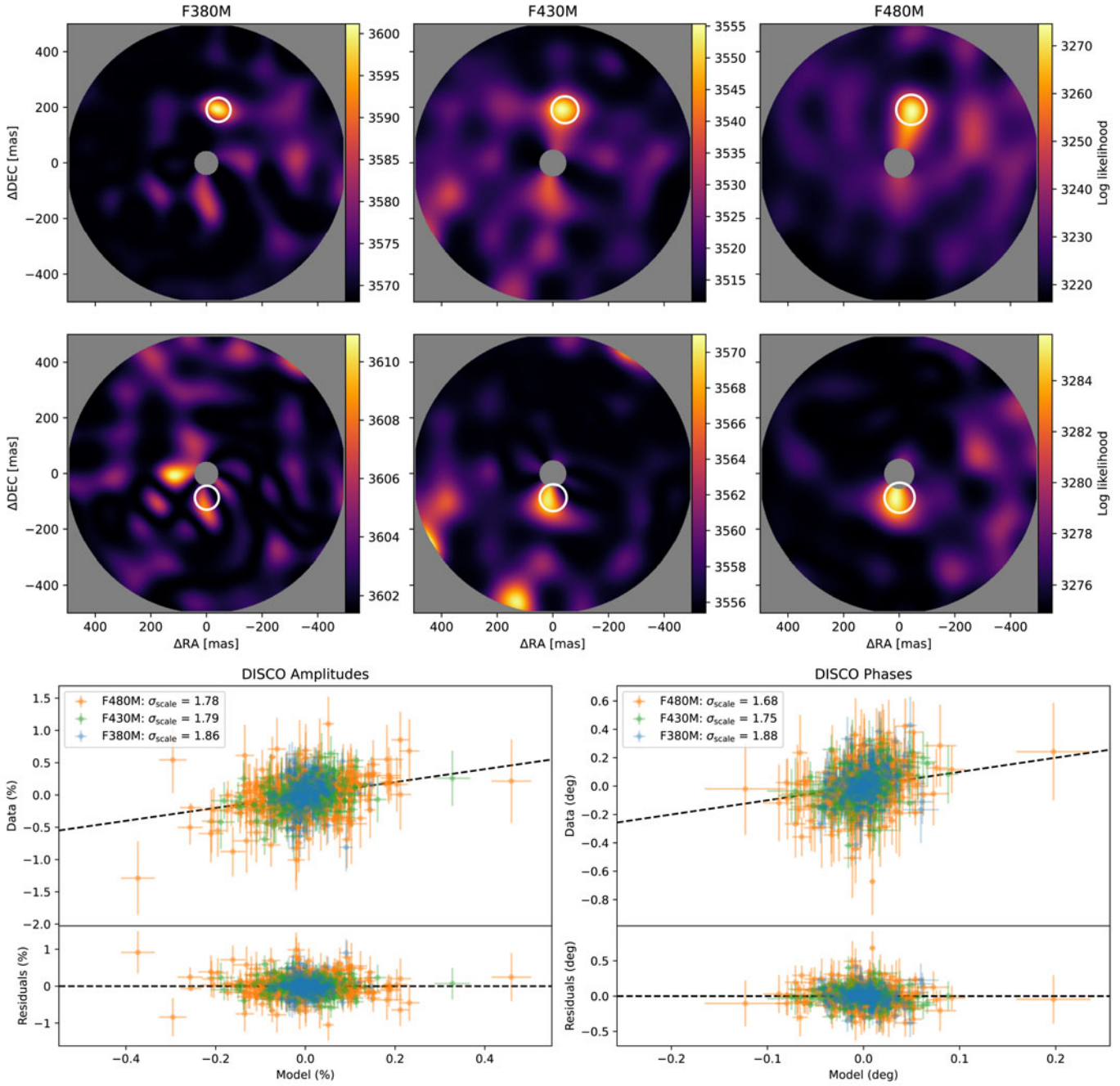


Figure 16. Summary of the fits to the HD 206893 companions. The top two panels shows the marginalised log-likelihood surface as a function of companion offset in ΔRA and ΔDEC, revealing a strong and consistent peak in all filters. The top row shows the detection maps for the full data, with the GRAVITY prediction for the B companion shown as a white circle. The middle row shows the detection map after the best-fit B companions has been subtracted from the data, revealing the inner c companion being consistently detected in all three filters, with the GRAVITY prediction overlaid with a white circle. The peaks in each filter for both companions can be seen matching the expected positions. The greyed central region denotes the IWA masked by the interferometric null. The bottom panels show the predictive posterior check comparing AMIGO recovered DISCO amplitudes (left) and phases (right) against the predicted values from the MCMC samples, for the three filters: F380M, F430M, and F480M. Each panel shows individual measurements with 1σ error bars, with the top panels plotting model predictions against data and the lower panels showing residuals. The black dashed line represents the 1:1 line for a perfect model prediction. Scale factors σ_{scale} are applied to the observational uncertainties in MCMC with the effect of ensuring $\chi^2_\nu \approx 1$ and are quoted in the legend for each filter. The agreement across filters and observables confirms the validity of the forward model and appropriate uncertainty quantification in the recovered posterior.

1. Maximise the number of groups per integration.
2. Observing calibrator and science targets to a similar well depth.
3. Selecting calibrators with a similar brightness to science targets.

We agree with these recommendations. We also provide two further recommendations for using AMIGO, while mitigating its weaknesses:

1. Remaining comfortably below half pixel well depth (∼50 k e^-, ∼30 k counts): this is already a recommended strategy

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 6. Best-fit relative joint astrometry and per-filter photometry for the HD 206893 B companion, showing the $+/-$ 1σ bounds. Reported quantities include the on-sky separation (Sep), position angle (PA), and Δ-magnitude in each of the F380M, F430M, and F480M filters. Results from two different fit types are compared: A joint fit in all three filters and a fit to each filter uniquely.

Fit type	Sep (mas)	PA (°)	Δmag (F380M)	Δmag (F430M)	Δmag (F480M)
Joint	$197.1^{+4.3}_{-4.1}$	$102.3^{+1.2}_{-1.6}$	$8.33^{+0.18}_{-0.16}$	7.98 ± 0.15	$7.49^{+0.11}_{-0.12}$
F380M	199.3 ± 6.2	$101.9^{+2.7}_{-2.6}$	$8.33^{+0.21}_{-0.14}$	–	–
F430M	$192.1^{+9.4}_{-8.4}$	$99.4^{+2.4}_{-3.5}$	–	$7.96^{+0.16}_{-0.14}$	–
F480M	$199.4^{+9.0}_{-8.4}$	$104.0^{+2.0}_{-2.1}$	–	–	$7.47^{+0.142}_{-0.096}$

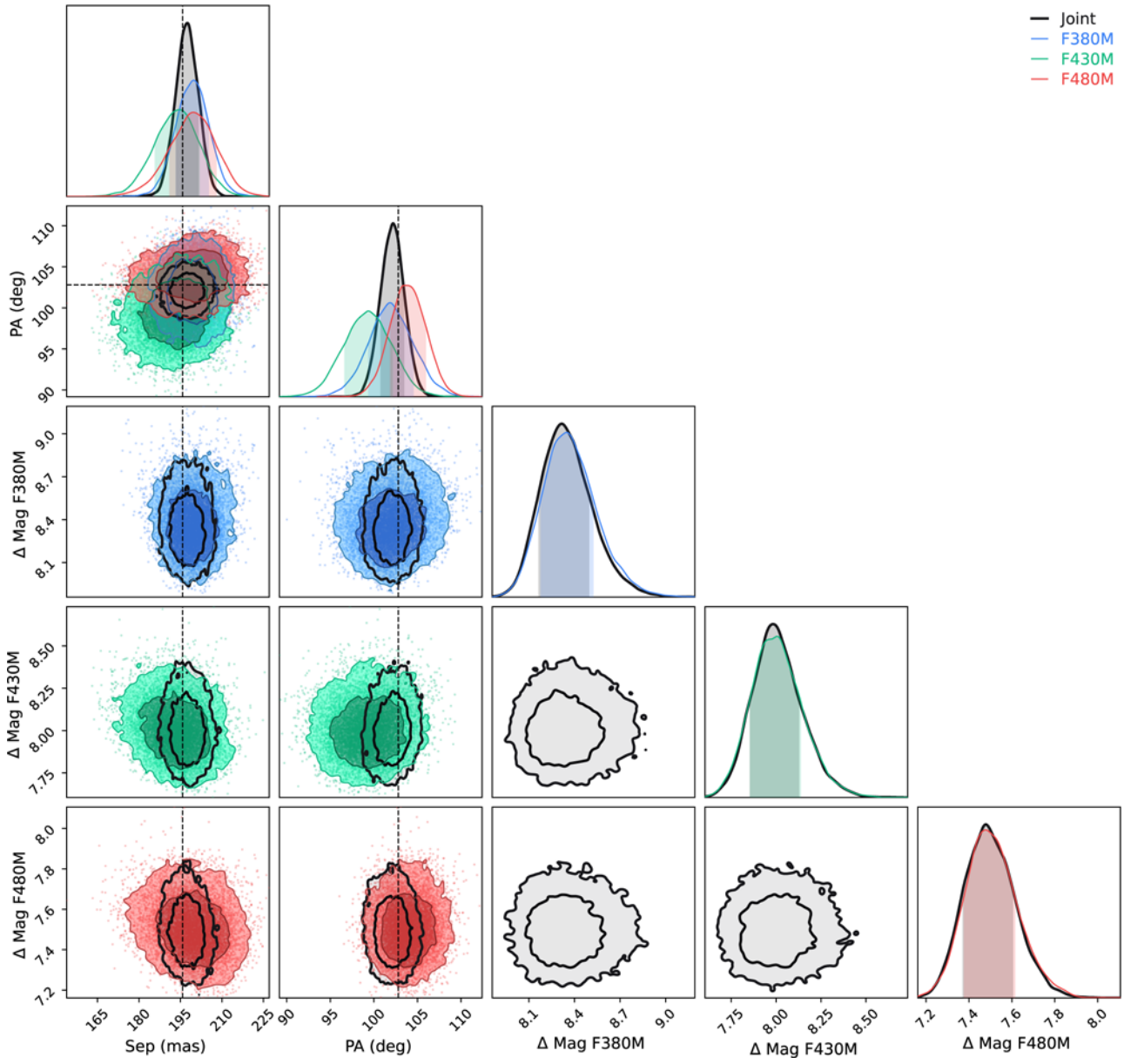


Figure 17. Joint posterior distributions from an MCMC fit to the HD 206893 B companion. Two fit types are shown: Joint – simultaneously modelling astrometry and photometry across all three filters – and per-filter fits. While both companions are fit simultaneously, only the B companions samples are shown here for clarity. The parameters shown include the separation (mas), position angle (degrees), and contrasts (Δmag) in the F380M, F430M, and F480M bands. The joint-fit samples are shown in black. One- and two-σ credible regions are shown in dark and light shades, respectively. The tight constraints in both astrometry and photometry reflect the high signal-to-noise of the companion detection. The expected position predicted by GRAVITY orbit fits (with a precision of $\sim$1 mas) are overlaid in a black dashed line, showing strong agreement with AMI, both in each filter and in the joint-fit.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Table 7. Best-fit relative joint astrometry and per-filter photometry for the HD 206893 c companion, showing the +/- 1σ bounds. Reported quantities include the on-sky separation (Sep), position angle (PA), and Δ-magnitude in each of the F380M, F430M, and F480M filters. Results from two different fit types are compared: A joint fit in all three filters and a fit to each filter uniquely.

Fit Type	Sep (mas)	PA (°)	Δmag (F380M)	Δmag (F430M)	Δmag (F480M)
Joint	103^{+16}_{-14}	$-80.7^{+4.1}_{-5.1}$	$9.14^{+0.43}_{-0.30}$	$8.32^{+0.23}_{-0.29}$	$8.07^{+0.34}_{-0.28}$
F380M	125^{+30}_{-20}	$-88.0^{+10.4}_{-8.9}$	$9.17^{+0.48}_{-0.30}$	—	—
F430M	102^{+24}_{-19}	$-79.6^{+7.4}_{-7.3}$	—	$8.36^{+0.28}_{-0.34}$	—
F480M	91^{+26}_{-36}	$-83.8^{+9.6}_{-8.4}$	—	—	$8.03^{+0.51}_{-0.70}$

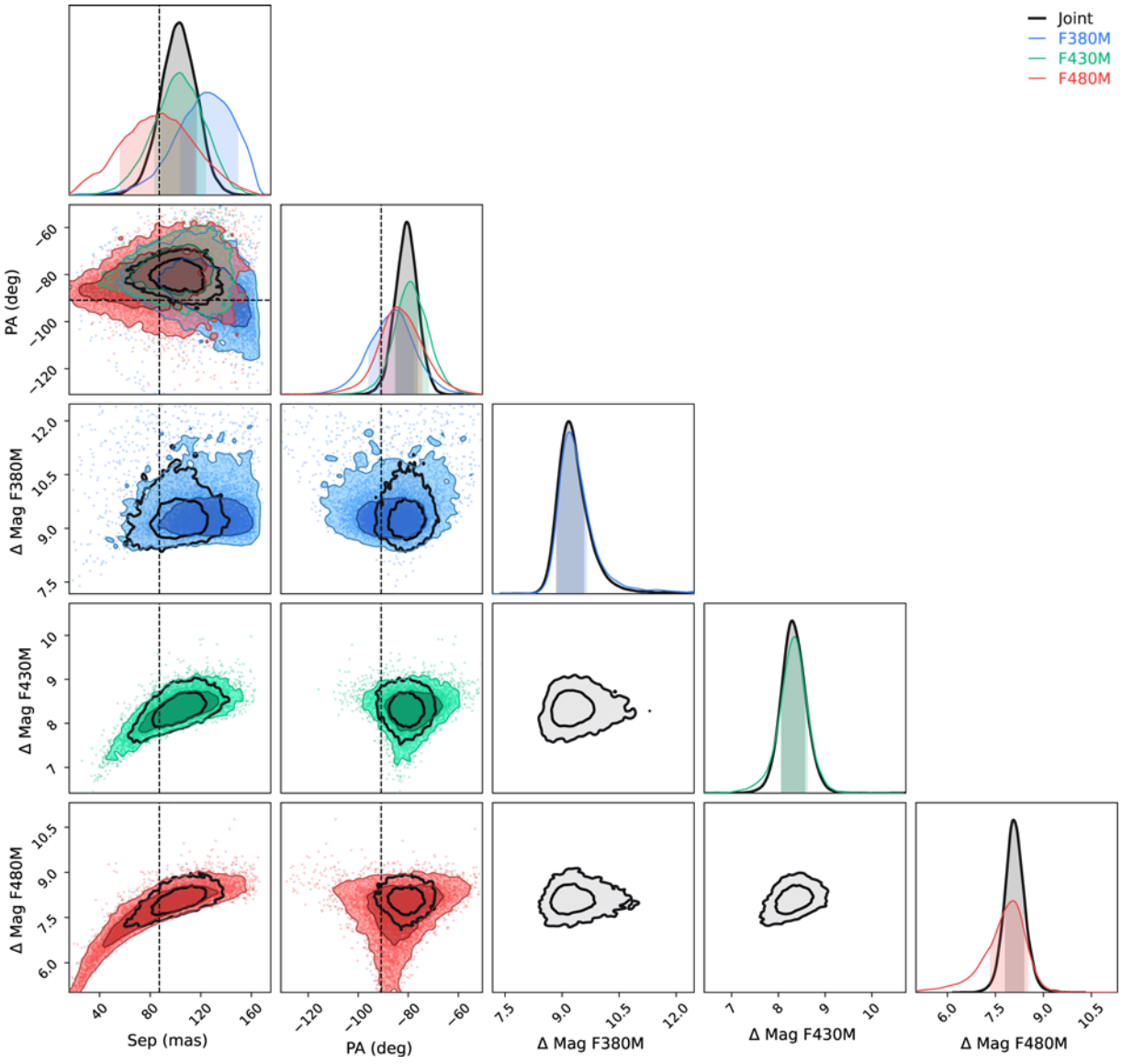


Figure 18. Joint posterior distributions from an MCMC fit to the HD 206893 c companion. Two fit types are shown: Joint – simultaneously modelling astrometry and photometry across all three filters – and per-filter fits. While both companions are fit simultaneously, only the B companions samples are shown here for clarity. The parameters shown include the separation (mas), position angle (degrees), and contrasts (Δmag) in the F380M, F430M, and F480M bands. The joint-fit samples are shown in black. One- and two-σ credible regions are shown in dark and light shades, respectively. The degeneracy between separation and contrast inside the diffraction limit is clearly shown in the F480M filter, demonstrating how photometry can be improved with a multi-band fit. The fit to the F380M filter shows that it falls right on the border of detection, with chains becoming poorly constrained above ~10 mag. The expected position predicted by GRAVITY orbit fits are overlaid in a black dashed line, showing decent agreement with AMI. Deviations between the expected and recovered positions could arise from either statistical noise or coupling to non-linear distortion arising from an imperfect calibration of the BFE, however given the relatively large astrometric uncertainty, we expect this to arise largely from low signal from the dim companion.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

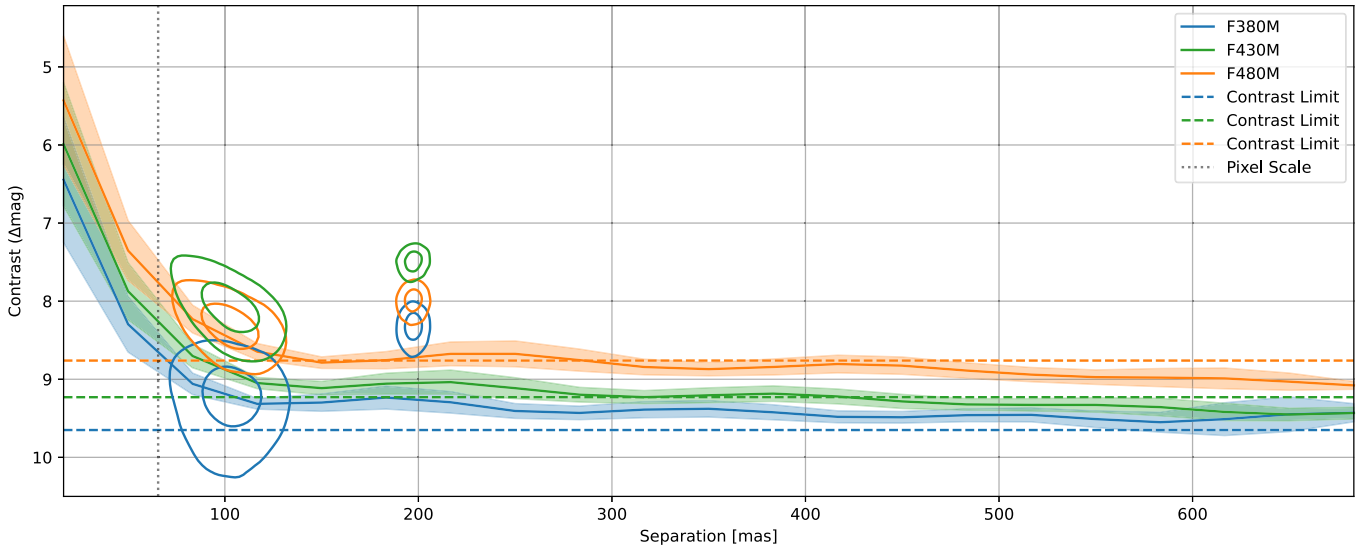


Figure 19. Sensitivity limits for HD 206893 derived using the Ruffio et al. (2018) method for calculating 3σ contrast upper limits as a function of angular separation, applied to the F380M, F430M, and F480M filters. The shaded regions denote the $\pm 1\sigma$ variation across azimuthal annuli. The dashed black line indicates the approximate contrast limit as calculated from Equation (11), confirming that model performance is consistent with the expected limits. The grey dotted line shows the size of a single pixel for reference. The recovered companion 1 and 2σ contour plots are shown in their respective colours. Recovery of the dim inner companion near the IWA and inside of the diffraction limit quantifies the AMI contrast limits as well calibrated and close to the theoretical limits of performance for AMI.

for AMI observations given the worsening charge migration above these values. Adding to this the AMIGO model is only trained on exposures below these values, making data with deeper pixel wells fall into a regime where model behaviour is unexplored and unlikely to perform well. Future calibration data with deeper exposures should provide a path towards precise observables with forward modelling methods, but until calibrations are performed in this regime, observations should remain below these levels.

2. A 5-point sub-pixel dither pattern: provided precise telescope pointing, conventional interferometric calibration should remove the effect of pixel-level miscalibrations making dithering an unnecessary complication to analysis. In practice non-linear and pixel to pixel detector effects can couple strongly to measured PSF shapes through the BFE. While AMIGO provides a direct model of these effects, nonlinearity makes precise calibration challenging without a more thorough treatment and better calibration data. A 5-point sub-pixel dither increases computational costs but provides a direct way to decouple pixel-level miscalibration from science observables and is expected to improve precision through the AMIGO model and pipeline.

These items provide a generally achievable set of guidelines to aide overall precision in output observables. Faster read-out modes for AMI are also being developed that will make it easier to select acceptable targets that fit within these guidelines.

The case study of HD 206893 highlights the strengths of AMIGO and also provides insight into potential limitations. The B companion is recovered with high confidence in both the astrometric and photometric precision. There is very strong agreement between the AMI astrometry and the predictions made by GRAVITY (Hinkley et al. 2023) indicating that AMI is now capable of confident results in the high-contrast regime up to Δ mag ~ 9. There is also a confident detection of the c companion, however it shows a $\sim 2\sigma$ deviation in the astrometry. This companion presents the greatest challenge to the AMIGO model, with its harsh contrast and separation well beyond the capabilities of any other JWST observing modes, and placement inside the diffraction limit of all three filters. These deviations could be consistent with an astrometric bias induced from this minute separation, given its position well inside of the pixels most affected by charge migration from the BFE. This could also explain the significantly larger uncertainty in its recovered position. Given these considerations the possibility of biases in the recovered photometry will remain until further study can be performed on other well constrained systems. A promising sign is that both of these companions remain detectable independently in each filter indicating that these results can be still be considered credible.

AMIGO has shown it can recover high-contrast companions at the diffraction limit. We recover a 3σ contrast limit through the Ruffio et al. (2018) method in approximate agreement with the scaling from Equation (11) (Ireland 2013) and recommended in JWST proposal materials. We adopt this as a good rule of thumb, but would also remind readers of the proposed observing strategies for AMIGO, emphasising that the photon loss arising from the current implementation is not directly accounted for in this equation and should be treated with care. This remains as a practical estimate of AMI performance though the methods explored in the work. We expect future calibrations and refinement of AMIGO to reclaim the photons lost by the first group, to improve performance across the board, and push beyond these estimates.

9. Conclusions and future work

The AMIGO model and pipeline provide a comprehensive recommissioning of JWST/AMI by harnessing end-to-end differentiable models and the novel techniques that they enable. It has enabled a robust recovery of the unrealised potential of high-contrast interferometric imaging at the highest angular resolutions

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

probed by JWST. This is the first pipeline in astronomy to fully integrate a differentiable physical forward model, including a full treatment of non-linear detector systematics and an embedded NN, and trained entirely from on-sky data. Through case studies of AB Dor AC and HD 206893, we demonstrate that AMIGO can recover faint companions in the high-contrast and high-angular resolution regime while preserving precision in both astrometry and photometry. Its performance approaches the theoretical photon-noise limits of the instrument, and surpasses the performance of available pipelines and approaches. This first version of the model is also only the start of its journey; many significant improvements can be made and are planned, placing it as a platform for both future high quality calibrations of JWST and improvements to existing and future AMI data and instruments.

Quantitative detection limits inferred from the AMIGO model provide a new characterisation of its true performance limits to date. We find that AMI can achieve contrasts of $\gtrsim$9 mag with astrometry up to the diffraction limit $\sim \lambda/D$. Detail is also recovered very close to an inner working angle of $\sim$100 mas, but with unknown potential biases introduced through the non-linear detector effects. The contrast limits found match closely with the Equation (11), giving a detectable contrast of $\sim 10/\sqrt{N_{\text{photons}}}$. These values are suitable for inclusion in future proposals, offering guidance on expected sensitivity, inner working angle and spatial resolution.

Beyond its immediate utility for science, AMIGO showcases the broader impact that differentiable forward models are poised to make on modelling in astronomical instrumentation and imaging as a whole. By treating systems in a physically principled end-to-end manner spanning astrophysics, optics, and detectors, AMIGO enables efficient and interpretable inference, rigorous uncertainty quantification, and coherent system calibration. This 'pixels to planets' philosophy – avoiding data reduction, and modelling data directly in their rawest possible form – represents a promising new direction for data analysis for other instruments affected by complex or non-linear systematics.

Concrete calibrations provided by this approach, namely the integral non-linearity of the ADC, are common mode to all observations made by NIRISS. We have already found identical signals in other non-AMI modes with the potential to induce biases in downstream scientific analysis. Cursory examinations of NIRCam and NIRSpec data do not show the same signal. Nonetheless, its presence in NIRISS data warrants further consideration of its effects in broader JWST calibration efforts, including SOSS transmission spectroscopy observations.

Looking forwards, several avenues of improvement and wider development are clear. The architecture of the embedded NN is both compact and physically constrained, but remains limited by the availability of extensive calibration data. While the PSFs produced by AMI present highly structured and well known illuminance pattern, most flux remains concentrated in the core. The calibration data used for training fills peak pixels to approximately half their maximum depth, and only provides a handful of pixels that fill beyond much smaller fractions of their maximum, and are suitable for calibration of the strongest BFE regime. This makes generalisation to the dynamics of charge migration for resolved sources an unexplored space, where model performance could degrade. Our team's Cycle 4 program GO 8330 plans to observe a bright, wide binary source at multiple primary and sub-pixel dither positions. These exposures will fill many pixels to half well depths while allowing the core to saturate. This should provide a dataset far richer in the complex non-linear dynamics of the BFE, while partially exploring the full pixel well depth in the saturated core. Further training of AMIGO on this data should provide a substantial improvement to its overall ability to directly model and understand the physics of this problem.

Provided a high fidelity EDM trained on this more informative dataset, it may become possible to develop a self-contained detector model able to *directly* invert the effects of the detector entirely by sampling input photon distributions. Such a 'de-BFE' approach has the potential to plug into a classic detector calibration pipeline that inverts all physical degradations, returning high quality super-resolved PSFs that accurately capture the statistical correlations induced by its measurements. Models trained on lab data during instrument pre-commissioning could be released in tandem with detector hardware itself as part of the supplier's service.

The calibrated EDM model can further serve as a platform to probe the solid-state physics of the H2RG family of detectors and their counterparts. Deeper and more targeted calibration programs should enable a fast, accurate, and differentiable model that can be used to hypothesis test solid-state architectures against and better understand the true form on the underlying physics.

The AMIGO model and this manuscript fully exclude all calibration and analysis in the F277W filter. This comes down to a lack of available high quality training data, along with a more complex set of physics associated with a larger spectral range. Despite its potentially unique benefits, the calibration dataset used for AMIGO did not observe in this filter as very few programs use it. Furthermore, the relative short spectral range and small band passes across the three primary filters F380M, F430M, and F480M allows for chromatic effects in the BFE to be ignored. Furthermore, the implementation of the interferometric visibility model considers all sources to have negligible spectral variation across any filter. While these limitations are not fundamental, the present AMIGO model is neither designed nor calibrated to operate at the F277W filter and is expected to perform poorly. Future version of AMIGO may be expanded to operate at these wavelengths, although this would require dedicated calibrated data and change to the structure of various sub-modules within AMIGO.

The DISCO observables introduced in this work represent a new approach to information-preserving projections of interferometric data. Building on Fisher information and kernel phase and amplitudes, and accelerated through autodiff, it enables the exploration of the information theoretic limits of imaging systems. By restricting the basis vectors to the OTF through their eigenvalues and varying the number of resolution elements and eigenvectors used in the *uv*-plane, the subspace of recoverable images can be quantified and concrete informational limits on instrumental performance can be explored. These ideas parallel limits on point-source detection set by small-sample statistics near the diffraction limit (Mawet et al. 2014), recast in the Fourier domain. In this paper we have not delved into depth to establish precise limits for DISCO linearity, scaling laws, or calibration strategies; future work will explore these ideas concretely for AMI (Charles et al. 2025). Given the excellent long-term stability and frequent monitoring of the wavefront using NIRCam (Lajoie et al. 2023), it may even be possible to do calibrator-free AMI imaging by trusting the NIRCam-measured OPD between the observation epoch and the wavefront map in the AMIGO base model and relying on the DISCOs to deal with residual wavefront error.

The current AMIGO model is only calibrated for the AMI specific sub-array, however, it is also possible to observe with AMI

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

using the full frame of the detector. The AMIGO model in its current state is not expected to be performant if directly translated to these observations for a few reasons. Over the field of view of the full frame, PSF variations seen as changes in both the wavefront error and mask distortion both become important in precisely describing the PSF at any point. While the wavefront variations can be solved for directly for each PSF, the flexibility of distorted aperture models is likely to result in over-fitting from data if treated in the same way. Calibrating a spatially varying mask model is possible and just requires a high-quality calibration data set. Furthermore, AMIGO does not have calibrated pixel-level effects (such as sensitivity and linearity variations) for the full-frame as this was not included in the training data. It is also expected that the read characteristics of the pixels (such as bias-voltage resets and $1/f$ noise) are different for the full-frame. Extending the AMIGO model to this regime faces no fundamental challenges, but was excluded from this work due to the lack of full-frame existing calibration and science data.

This more general approach may also be able to improve the performance of spectrally-dispersed wavefront sensing and kernel phase on IFU data (Martinache 2016; N'Diaye et al. 2022; Chaushev et al. 2023). The NIRSpec Integral Field Unit (IFU) has been used to achieve extremely high contrast ($\sim 3 \times 10^{-6}$), for which Ruffio et al. (2024) depend on a forward model of the IFU PSF using WEBBPSF; a trained ∂LUX model as in AMIGO will be more flexible and, we expect, enable us to push to deeper systematics floors.

An advantage of modelling visibilities in a latent basis which very finely samples the *uv* plane is that the interferometric field of view (FOV) (which corresponds to the Nyquist limit of this *uv* sampling) is large, potentially as large as the native 80×80 sub-array FOV if the *uv* sampling is the FFT coordinate grid conjugate to the raw image. This is ideal for image reconstruction/deconvolution. An implementation in JAX of regularised maximum likelihood image reconstruction, similar to MPOL (Czekala et al. 2025), is likely to be straightforward, and better able to recover complex scenes than reconstructions limited to one visibility per baseline and the basic closure phases. It remains an open question whether this is better suited to a DISCO likelihood or direct to pixels. As an extension, it may be advantageous to apply score-based methods to very expressively build priors over images (Feng et al. 2023; Dia et al. 2025), and likelihoods over detector effects (Adam et al. 2025), applications which will only be possible with differentiable instrument models.

The methods, ideas, and philosophy laid out in this work are not limited to either JWST nor interferometric systems. Natural extensions to NIRCam and archival HST NIR camera and multi-object spectrometers are straightforwardly possible. The tighter PSFs and redundant apertures provide significant but not insurmountable challenges. Various methods harnessing kernel phase or data driven empirical PSF subtraction remain effective and have been used on both JWST and HST data to reveal faint companions at high resolution (De Furio et al. 2023; Calissendorff et al. 2023). These observations provide an ideal testing ground for deeper exploration of the ideas and models presented in this work to other imaging modes and instruments. One set of low-hanging fruit may be the HST/NICMOS LT dwarfs, which were previously studied with kernel phase but without accurate aperture calibration (Pope et al. 2013; Martinache et al. 2020).

Ground-based observatories also serve to benefit from the methodology explored in this work, as many remain limited by instrumental miscalibrations and non-linear effects. Furthermore, the lack of well-calibrated differentiable forwards models necessitates the use of linear and empirical methods. As the next generation of Extremely Large Telescopes start to come online and the field turns to telemetry-based post-processing (Males et al. 2021; Guyon et al. 2022), high-fidelity and well calibrated end-to-end instrumental models will become essential to achieving the contrast goals of these instruments. This work lays the foundation for the methods, models, and calibrations strategies essential to their success.

Finally, there is growing potential to apply the pixels-to-planets philosophy to coronagraphic observations. PSF subtraction using the conventional optical modelling software WEBBPSF can improve recovered image deconvolution, accounting for the spatially-varying PSF (Balmer et al. 2025). Replacing these nominal PSF models with on-sky calibrated predictions from a model like AMIGO with data-driven estimates of the coronagraph metrology, and directly accounting for and calibrating non-linear detector effects, could unlock deeper contrasts and higher precision observables. This is likely to be helpful for the Roman coronagraph's wavefront ground-in-the-loop control (Bailey et al. 2023), and future work with HabWorlds (National Academies 2023).

Acknowledgments. We would like to thank Steph Sallum, Laurent Pueyo, Marshall Perrin, Eddie Bergeron, Joel Sanchez-Bermudez, Thomas Vandal, Loïc Albert, Laurence Perreault-Levasseur, Alexandre Adam, Noé Dia, Adam Taras, Lucinda Lilley, Hayden Greer, Shashank Dholakia, and Shishir Dholakia for their helpful discussions, advice, and contributions.

We acknowledge and pay respect to the traditional owners of the land on which the University of Sydney, Macquarie University, and the University of Queensland are situated, upon whose unceded, sovereign, ancestral lands we work. We pay respects to their Ancestors and descendants, who continue cultural and spiritual connections to Country. We acknowledge and respect the Lək̓ʷəŋən (Songhees and Esquimalt) Peoples on whose territory the University of Victoria stands, and the Lək̓ʷəŋən and W̱SÁNEĆ Peoples whose historical relationships with the land continue to this day.

We are grateful in particular for the support of the Space Telescope Science Institute for their extensive work in understanding, commissioning, and supporting JWST and its instruments, including significant correspondence in this paper and our related works.

BP, PT, and SR have been supported by the Australian Research Council grant DP230101439 and BP by DE210101639; and LD and MC have been supported by the Australian Government Research Training Program (RTP) award. We are grateful to the Australian public for enabling this science. BP and SR would like to thank the Big Questions Institute for their philanthropic support. DB and DJ acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). DJ also acknowledges support from NRC Canada. AS and KV acknowledge support from the NSF, NASA, and the STScI Director's Discretionary Fund. Development of ∂LUX has been supported by the Breakthrough Foundation through their Toliman project as a part of the Breakthrough Watch initiative.

This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The NIRISS instrument was funded by the Canadian Space Agency and built in Canada by Honeywell. The AMIGO base model is constructed from observations associated with JWST program CAL 4481 (PI: A. Sivaramakrishnan). The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST.

We would also like to thank the reviewer for their insightful and highly constructive comments and feedback. These have helped strengthen this paper and its interpretability to the wider community.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

This research made use of NumPy (Harris et al. 2020); Matplotlib (Hunter 2007); Jax (Bradbury et al. 2018); `numpyro` (Phan et al. 2019; Bingham et al. 2019); `equinox` (Kidger & Garcia 2021); `optax` (Hessel et al. 2020); and ChainConsumer (Hinton 2016).

Code and data availability. All codes and data used to produce this work are publicly available and open source. The AMIGO model and pipeline are hosted on GitHub, along with a series of example notebooks used in this paper's results. All data are available from MAST under the appropriate proposal numbers; in this pipeline paper, we have relied only on publicly available data with no proprietary data included. We encourage researchers to adapt and apply AMIGO to their own interferometric datasets, and to contact the development team with questions and contributions.

References

Adam, A., et al. 2025, AJ, 169, 254
Akhare, D., Luo, T., & Wang, J.-X. 2023, CMAME, 406, 115902
Antilogus, P., Astier, P., Doherty, P., Guyonnet, A., & Regnault, N. 2014, JI, 9, C03048
Argyriou, I., et al. 2023, AAP, 680, A96
Bailey, V. P., et al. 2023, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 12680, 126800T
Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Natur, 320, 595
Balmer, W. O., et al. 2025, AJ, 169, 209
Benner, P., Gugercin, S., & Willcox, K. 2015, SIAMRv, 57, 483
Betancourt, M. 2017, arXiv e-prints, arXiv:1701.02434
Beuzit, J.-L., et al. 2019, A&A, 631, A155
Bingham, E., et al. 2019, JMLR, 20, 28:1
Blakely, D., et al. 2022, APJ, 931, 3
Blakely, D., et al. 2025, AJ, 169, 152
Blakely, D., et al. 2024, The James Webb Interferometer: Space-based interferometric detections of PDS 70 b and c at 4.8 μm, arXiv:2404.13032
Bonneau, D., Delfosse, X., Clausse, J. M., & Mourard, D. 2012, SearchCal: The JMMC Evolutive Search Calibrator Tool, ASCL, record ascl:1210.012
Born, M., et al. 1999, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light (7th edn.; Cambridge: Cambridge University Press)
Brabandere, B. D., Jia, X., Tuytelaars, T., & Gool, L. V. 2016, Dynamic Filter Networks, arXiv:1605.09673
Bradbury, J., et al. 2018, JAX: Composable transformations of Python+NumPy programs
Broyden, C. G. 1970, IMAJAM, 6, 76
Calissendorff, P., et al. 2023, APJl, 947, L30
Campagne, J.-E., et al. 2023, OJA, 6, 15
Charles, M., et al. 2025, Image Reconstruction with the JWST Interferometer, arXiv:2510.10924
Chaushev, A., et al. 2023, JATIS, 9, 028004
Cheetham, A., et al. 2016, in SPIE Conference Series, Optical and Infrared Interferometry and Imaging Vol, 9907, ed., M. Fabien, M. J. Creech-Eakman & Peter G. Tuthill, 99072T
Cheetham, A. C., Tuthill, P. G., Sivaramakrishnan, A., & Lloyd, J. P. 2012, OExpr, 20, 29457
Czekala, I., et al. 2025, arXiv e-prints, arXiv:2502.00100
De Furio, M., et al. 2023, 948, 92
Desdoigts, L., Pope, B., Gully-Santiago, M., & Tuthill, P. 2024, Differentiable Optics with dLux II: Optical Design Maximising Fisher Information, arXiv:2406.08704
Desdoigts, L., Pope, B. J. S., Dennis, J., & Tuthill, P. G. 2023, JATIS, 9, doi:10.1117/1.jatis.9.2.028007
Dholakia, S., Dholakia, S., & Pope, B. J. S. 2024, arXiv e-prints, arXiv:2410.03449
Dia, N., Yantovski-Barth, M. J., Adam, A., et al. 2025, arXiv e-prints, arXiv:2501.02473
Doyon, R., et al. 2012, in Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, Vol. 8442, International Society for Optics and Photonics (SPIE), ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. M. O. Jr., 84422R
Doyon, R., et al. 2023, arXiv:2306.03277
Dube, B. 2019, JOSS, 4, 1352
Feng, B. T., et al. 2023, arXiv e-prints, arXiv:2304.11751
Fletcher, R. 1970, CJ, 13, 317
Follette, K. B. 2023, An Introduction to High Contrast Differential Imaging of Exoplanets and Disks, arXiv:2308.01354
Gallenne, A., et al. 2015, AAP, 579, A68
Gardner, J. P., et al. 2006, SSRv, 123, 485
Gardner, J. P., et al. 2023, PASP, 135, 068001
Goldfarb, D. 1970, MC, 24, 23
Goodman, J. W. 2005, Introduction to Fourier Optics (3rd edn.; Greenwood Village, CO: Roberts and Company Publishers)
Goudfrooij, P., Grumm, D., Volk, K., & Bushouse, H. 2024, PASP, 136, 014503
Greenbaum, A. Z., Pueyo, L., Sivaramakrishnan, A., & Lacour, S. 2015, APJ, 798, 68
Greenbaum, A. Z., et al. 2019, AJ, 157, 249
Griewank, A., & Walther, A. 2008, Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation, Other Titles in Applied Mathematics (2nd edn.; Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM))
Gully-Santiago, M., & Morley, C. V. 2022, APJ, 941, 200
Guyon, O., Pluzhnik, E. A., Kuchner, M. J., Collins, B., & Ridgway, S. T. 2006, APJs, 167, 81
Guyon, O., et al. 2022, High Contrast Imaging at the Photon Noise Limit with WFS-based PSF Calibration, arXiv:2208.01806
Guyonnet, A., Astier, P., Antilogus, P., Regnault, N., & Doherty, P. 2015, AAP, 575, A41
Haguenauer, P., et al. 2010, in Optical and Infrared Interferometry II, Vol. 7734, International Society for Optics and Photonics (SPIE), ed. W. C. Danchi, F. Delplancke, & J. K. Rajagopal, 773404
Harris, C. R., et al. 2020, Natur, 585, 357
Hattori, S., et al. 2024, exoplanet-dev/jaxoplanet: Astronomical Time Series Analysis with JAX, doi:10.5281/zenodo.10736936
Hecht, E. 2002, Optics (4th edn.; San Francisco: Addison-Wesley)
Hessel, M., et al. 2020, Optax: Composable Gradient Transformation and Optimisation, in JAX!
Hinkley, S., et al. 2023, AAP, 671, L5
Hinton, S. R. 2016, JOSS, 1, 00045
Hinz, P. M., et al. 2016, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9907, Optical and Infrared Interferometry and Imaging V, ed. F. Malbet, M. J. Creech-Eakman, & P. G. Tuthill, 990704
Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv e-prints, arXiv:1008.4686
Hogg, D. W., & Villar, S. 2024, Is Machine Learning Good or Bad for the Natural Sciences?, arXiv:2405.18095
Horta, D., Price-Whelan, A. M., Hogg, D. W., Ness, M. K., & Casey, A. R. 2025, AJ, 169, 314
Hunter, J. D. 2007, CSE, 9, 90
Ireland, M. J. 2013, MNRAS, 433, 1718
Jefferys, W. H., Benedict, G. F., Hemenway, P. D., Shelus, P. J., & Duncombe, R. L. 1985, CM, 37, 299
Jennison, R. C. 1958, MNRAS, 118, 276
Kammerer, J., et al. 2023, PASP, 135, 014502
Karniadakis, G. E., et al. 2021, NatRvP, 3, 422
Kass, R., Tierney, L., & Kadane, J. 1991, Laplace's Method in Bayesian Analysis, doi:10.1090/conm/115/07
Kawahara, H., et al. 2022, ApJS, 258, 31
Kawahara, H., et al. 2025, ApJ, 985, 263
Kidger, P. 2022, On Neural Differential Equations, arXiv:2202.02435
Kidger, P., & Garcia, C. 2021, Equinox: Neural Networks in JAX via Callable PyTrees and Filtered Transformations, arXiv:2111.00254
Kingma, D. P., & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980
Kochkov, D., et al. 2024, Natur, 632, 1060

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

Krist, J. E., Hook, R. N., & Stoehr, F. 2011, in Optical Modeling and Performance Predictions V, Vol. 8127, International Society for Optics and Photonics (SPIE), ed. M. A. Kahan, 81270J
Lacour, S., et al. 2011, A&A, 532, A72
Lajoie, C.-P., et al. 2023, OTE Science Performance Memo 2 - A Year of Wavefront Sensing with JWST in Flight: Cycle 1 Telescope Monitoring & Maintenance Summary, Technical Report JWST-STScI-008497, arXiv:2307.11179
Lakshminarayanan, V., & Fleck, A. 2011, JModOptic, 58, 1678
Lau, R. M., et al. 2023, arXiv:2311.15948
LeCun, Y., Bengio, Y., & Hinton, G. 2015, Natur, 521, 436
Liaudat, T., Starck, J.-L., Kilbinger, M., & Frugier, P.-A. 2023, IP, 39, 035008
Lucas, M., et al. 2024, PASP, 136, 114504
MacKay, D. J. C. 2002, Information Theory, Inference & Learning Algorithms (USA: Cambridge University Press)
Males, J. R., Fitzgerald, M. P., Belikov, R., & Guyon, O. 2021, PASP, 133, 104504
Margossian, C. C. 2019, WIREs Data Mining and Knowledge Discovery, 9, e1305. doi:10.1002/widm.1305
Martens, J. 2020, New Insights and Perspectives on the Natural Gradient Method, arXiv:1412.1193
Martinache, F. 2010, APJ, 724, 464
Martinache, F. 2013, PASP, 125, 422
Martinache, F. 2016, in SPIE Conference Series, Vol. 9907, Optical and Infrared Interferometry and Imaging V, ed. F. Malbet, M. J. Creech-Eakman, & P. G. Tuthill, 990712
Martinache, F., et al. 2020, AAP, 636, A72
Mawet, D., et al. 2014, APJ, 792, 97
McDougall, H. G., Davis, T. M., & Pope, B. J. S. 2025, arXiv e-prints, arXiv:2505.09832
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, JCP, 21, 1087
Monnier, J. D. 2003, RPPh, 66, 789
Monnier, J. D., Tuthill, P. G., & Danchi, W. C. 1999, APJ, 525, L97–L100
Morse, P. M., & Feshbach, H. 1953, Methods of Theoretical Physics, Vol. 1 (New York: McGraw-Hill)
National Academies. 2023, Pathways to Discovery in Astronomy and Astrophysics for the 2020s (National Academies Press), doi:10.17226/26141
N'Diaye, M., et al. 2022, in SPIE Conference Series, Vol. 12183, Optical and Infrared Interferometry and Imaging VIII, ed. A. Mérand, S. Sallum, & J. Sanchez-Bermudez, 121831M
O'Brien, M. J., et al. 2026, ACS, 82, doi:10.1107/S2059798326000550
Page, J., & Favaros, P. 2020, in 2020 IEEE International Conference on Image Processing (ICIP), 2995
Paszke, A., et al. 2019, in Advances in Neural Information Processing Systems 32 (Red Hook, NY: Curran Associates, Inc.), 8024
Perrin, M. D., et al. 2014, in SPIE Conference Series, Vol. 9143, Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave, ed. J. M. Oschmann, Jr., M. Clampin, G. G. Fazio, & H. A. MacEwen, 91433X
Perrin, M. D., Soummer, R., Elliott, E. M., Lallo, M. D., & Sivaramakrishnan, A. 2012, in SPIE Conference Series, Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. M. Oschmann, Jr., 84423D
Phan, D., Pradhan, N., & Jankowiak, M. 2019, arXiv preprint arXiv:1912.11554
Pope, B., et al. 2014, MNRAS, 440, 125
Pope, B., Martinache, F., & Tuthill, P. 2013, APJ, 767, 110
Pope, B. J. S. 2016, MNRAS, 463, 3573–3581
Pope, B. J. S., Pueyo, L., Xin, Y., & Tuthill, P. G. 2021, APJ, 907, 40
Rader, J., Lyons, T., & Kidger, P. 2024, arXiv:2402.09983
Raissi, M., Perdikaris, P., & Karniadakis, G. 2019, JCPh, 378, 686
Rauscher, B. J., et al. 2007, PASP, 119, 768
Rauscher, B. J., et al. 2017, PASP, 129, 105003
Ray, S., et al. 2023, MNRAS, 519, 2718
Ray, S., et al. 2025, arXiv:2310.11508
Readhead, A. C. S., Walker, R. C., Pearson, T. J., & Cohen, M. H. 1980, Natur, 285, 137
Rigby, J., et al. 2023, PASP, 135, 048001
Robertson, J., et al. 1991, PASA, 9, 162
Rowlands, N., Midwinter, C., & Warner, G. 2018, in High Energy, Optical, and Infrared Detectors for Astronomy VIII, Vol. 10709, International Society for Optics and Photonics (SPIE), ed. A. D. Holland, & J. Beletic, 107091K
Ruder, S. 2016, arXiv e-prints, arXiv:1609.04747
Ruffio, J.-B., et al. 2018, AJ, 156, 196
Ruffio, J.-B., et al. 2024, AJ, 168, 73
Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986, Natur, 323, 533
Sallum, S., Eisner, J., Skemer, A., & Murray-Clay, R. 2023, ApJ, 953, 55
Sallum, S., Ray, S., & Hinkley, S. 2022, in Optical and Infrared Interferometry and Imaging VIII, Vol. 12183, International Society for Optics and Photonics (SPIE), ed. A. Mérand, S. Sallum, & J. Sanchez-Bermudez, 121832M
Sallum, S., et al. 2024, arXiv:2310.11499
Shanno, D. F. 1970, MC, 24, 647
Siskind, J. M., & Pearlmutter, B. A. 2008, HSC, 21, 361
Sitzmann, V., et al. 2018, ACMTOG, 37, 114
Sivaramakrishnan, A., et al. 2012, in Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave, Vol. 8442, International Society for Optics and Photonics (SPIE), ed. M. C. Clampin, G. G. Fazio, H. A. MacEwen, & J. M. O. Jr., 84422S
Sivaramakrishnan, A., et al. 2023, PASP, 135, 015003
Soulain, A., et al. 2020, in SPIE Conference Series, Vol. 11446, Optical and Infrared Interferometry and Imaging VII, International Society for Optics and Photonics (SPIE), ed. P. G. Tuthill, A. Mérand, & S. Sallum, 1144611
Stolker, T., et al. 2024, A&A, 682, A101
ten Brummelaar, T. A., et al. 2005, ApJ, 628, 453
Tuthill, P., Monnier, J., Danchi, W., Wishnow, E., & Haniff, C. 2000, PASP, 112, 555
Twiss, R. Q., Carter, A. W. L., & Little, A. G. 1960, Obs, 80, 153
Vides, C. L., Sallum, S., Eisner, J., Skemer, A., & Murray-Clay, R. 2023, APJ, 958, 123
Wagner, K., Apai, D., & Kratter, K. M. 2019, APJ, 877, 46
Willard, J., Jia, X., Xu, S., Steinbach, M., & Kumar, V. 2022, ACMCS, 55, doi:10.1145/3514228
Willson, M., et al. 2016, A&A, 595, A9
Wizinowich, P., et al. 2000, PASP, 112, 315
Wong, A., et al. 2021, JOSAB, 38, 2465

Appendix A. Calibration data fit

The full AMIGO model calibration was completed on the dataset described in Section 2.1, with the process detailed in Section 5. Given the similarity of the fits across all three filters, the F380M and F480M fits to the calibration have been placed here. Figures A1 and A2 show the calibrated AMIGO model fit to the remaining calibration data.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194

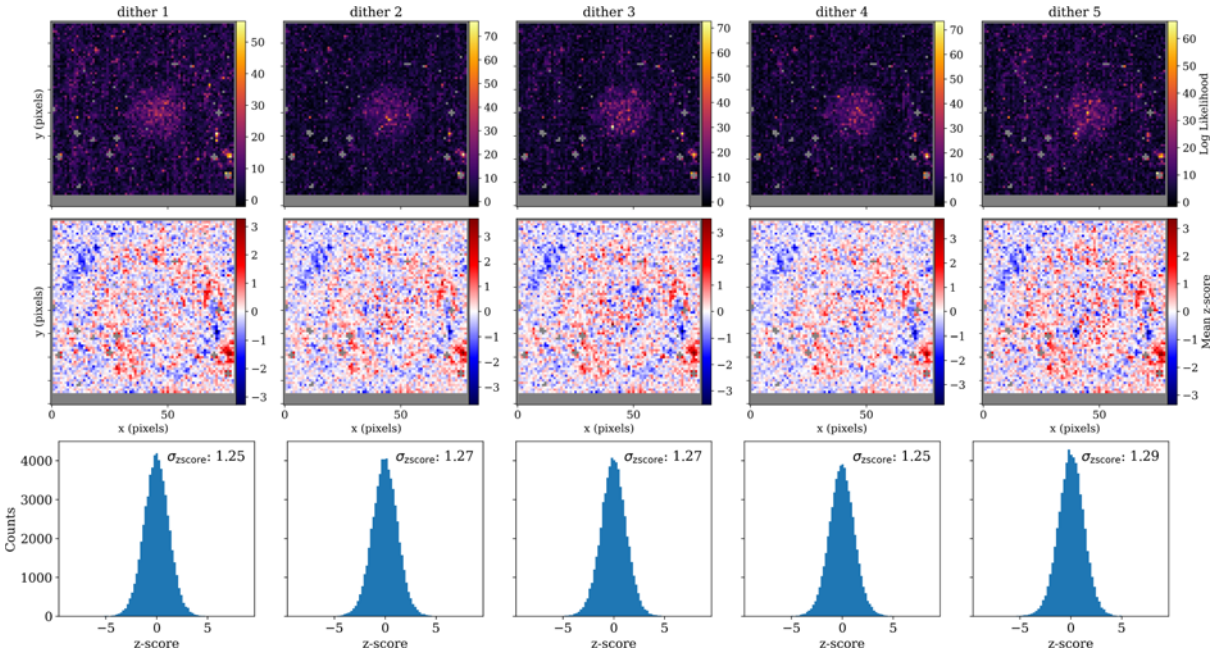


Figure A1. Summary of AMIGO model fit diagnostics across all five dithers for the F380M calibration data. Each column corresponds to a single dither position. The top row shows the per-pixel log-likelihoods from the final fit, highlighting the location of the target PSF. The middle row displays the average residual z-score per pixel, computed by averaging the uncertainty-normalised residuals across all groups in the ramp, revealing the spatial structure of any systematic model misfit. The bottom row shows histograms of all z-scores across pixels and groups for each dither, without any averaging over the groups. A perfect fit would have a standard deviation in the z-score be 1; we recover values between 1.1 and 1.2 in all three filters, indicating a good fit that has not learnt any noise characteristics. We note that the full likelihood is described with a covariance matrix that accounts for the anti-correlation between adjacent group-reads seen in slope data. Consequently, these summary statistics are only an helpful approximation and correct performance can only be described through the likelihood.

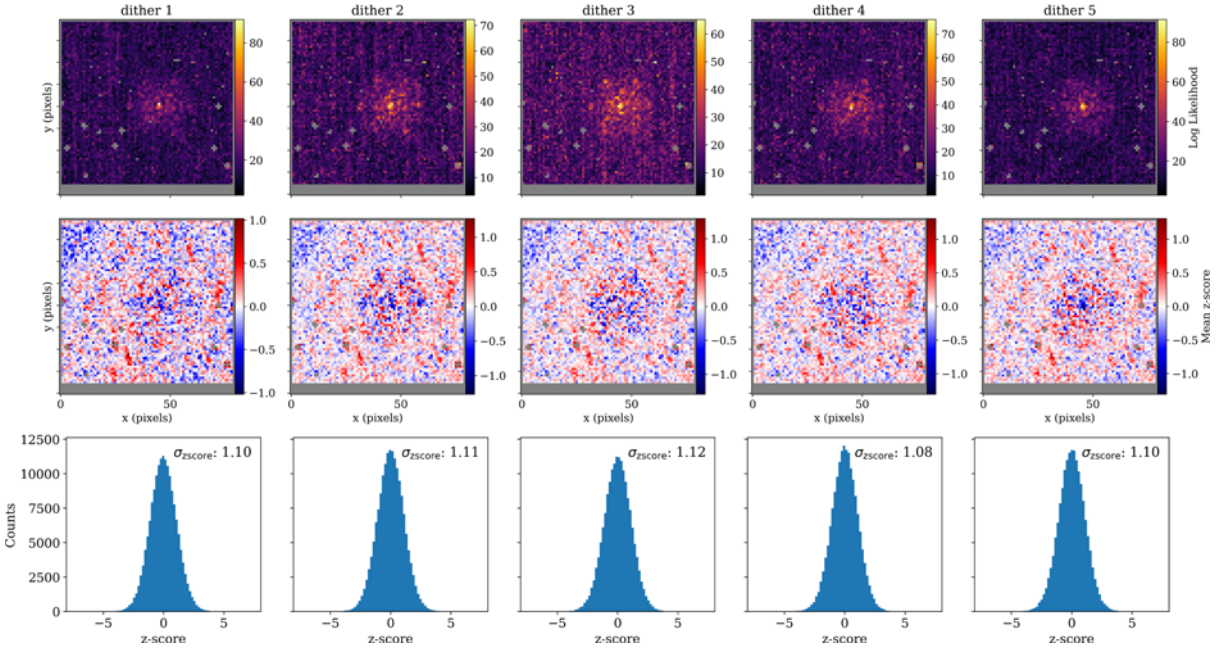


Figure A2. Summary of AMIGO model fit diagnostics across all five dithers for the F480M calibration data. Each column corresponds to a single dither position. The top row shows the per-pixel log-likelihoods from the final fit, highlighting the location of the target PSF. The middle row displays the average residual z-score per pixel, computed by averaging the uncertainty-normalised residuals across all groups in the ramp, revealing the spatial structure of any systematic model misfit. The bottom row shows histograms of all z-scores across pixels and groups for each dither, without any averaging over the groups. A perfect fit would have a standard deviation in the z-score be 1; we recover values between 1.1 and 1.2 in all three filters, indicating a good fit that has not learnt any noise characteristics. We note that the full likelihood is described with a covariance matrix that accounts for the anti-correlation between adjacent group-reads seen in slope data. Consequently, these summary statistics are only an helpful approximation and correct performance can only be described through the likelihood.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 21:50:57, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/pasa.2026.10194