The JWST Early Release Science Program for Direct Observations of Exoplanetary Systems. IV. NIRISS Aperture Masking Interferometry Performance and Lessons Learned

Steph Sallum[1], Shrishmoy Ray[2], Jens Kammerer[3], Anand Sivaramakrishnan[4], Rachel Cooper[4], Alexandra Z. Greebaum[5], Deepashri Thatte[4], Matthew De Furio[6], Samuel M. Factor[6], Michael R. Meyer[7], Jordan M. Stone[8], Aarynn Carter[9], Beth Biller[10,11], Sasha Hinkley[2], Andrew Skemer[9], Genaro Suárez[12], Jarron M. Leisenring[13], Marshall D. Perrin[4], Adam L. Kraus[6], Olivier Absil[14], William O. Balmer[4,15], Sarah K. Betti[4], Anthony Boccaletti[16], Mariangela Bonavita[17], Mickael Bonnefoy[18], Mark Booth[19], Brendan P. Bowler[6], Zackery W. Briesemeister[20], Marta L. Bryan[21], Per Calissendorff[7], Faustine Cantalloube[22], Gael Chauvin[23], Christine H. Chen[4,15], Elodie Choquet[22], Valentin Christiaens[14], Gabriele Cugno[7], Thayne Currie[24,25], Camilla Danielski[26], Trent J. Dupuy[27], Jacqueline K. Faherty[12], Michael P. Fitzgerald[28], Jonathan J. Fortney[9], Kyle Franson[6,74], Julien H. Girard[4], Carol A. Grady[29], Eileen C. Gonzales[30], Thomas Henning[31], Dean C. Hines[4], Kielan K. W. Hoch[32], Callie E. Hood[9], Alex R. Howe[20], Markus Janson[33], Paul Kalas[21,34,35], Grant M. Kennedy[36], Matthew A. Kenworthy[37], Pierre Kervella[16], Daniel Kitzmann[38], Masayuki Kuzuhara[39], Anne-Marie Lagrange[16], Pierre-Olivier Lagage[40], Kellen Lawson[20], Cecilia Lazzoni[2], Ben W. P. Lew[41], Michael C. Liu[42], Pengyu Liu[10,11], Jorge Llop-Sayson[43], James P. Lloyd[44], Anna Lueber[45], Bruce Macintosh[46], Elena Manjavacas[47], Sebastian Marino[2], Mark S. Marley[48], Christian Marois[49], Raquel A. Martinez[1], Brenda C. Matthews[50], Elisabeth C. Matthews[31], Dimitri Mawet[43,51], Johan Mazoyer[16], Michael W. McElwain[20], Stanimir Metchev[52], Brittany E. Miles[13], Maxwell A. Millar-Blanchaer[53], Paul Molliere[31], Sarah E. Moran[48], Caroline V. Morley[6], Sagnick Mukherjee[9], Paulina Palma-Bifani[23], Eric Pantin[54], Polychronis Patapis[55], Simon Petrus[56,57], Laurent Pueyo[4], Sascha P. Quanz[55], Andreas Quirrenbach[58], Isabel Rebollido[4], Jea Adams Redai[59], Bin B. Ren[18], Emily Rickman[60], Matthias Samland[31], B. A. Sargent[4,61], Joshua E. Schlieder[20], Glenn Schneider[13], Karl R. Stapelfeldt[51], Ben J. Sutlieff[10,11], Motohide Tamura[62], Xianyu Tan[63], Christopher A. Theissen[64], Taichi Uyama[5], Arthur Vigan[22], Malavika Vasist[14], Johanna M. Vos[66], Kevin Wagner[13,75], Jason J. Wang[43,67], Kimberly Ward-Duong[68], Niall Whiteford[12], Schuyler G. Wolff[13], Kadin Worthen[15], Mark C. Wyatt[65], Marie Ygouf[51], Xi Zhang[69], Keming Zhang[21], Zhoujian Zhang[9], Yifan Zhou[70], and Alice Zurlo[71,72,73]

[1] Department of Physics and Astronomy, University of California, Irvine, 4129 Frederick Reines Hall, Irvine, CA 92697, USA; ssallum@uci.edu
[2] University of Exeter, Astrophysics Group, Physics Building, Stocker Road, Exeter, EX4 4QL, UK
[3] European Southern Observatory, Karl-Schwarzschild-Straße 2, 85748 Garching, Germany
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[5] IPAC, California Institute of Technology, 1200 E. California Blvd., Pasadena, CA 91125, USA
[6] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Stop C1400, Austin, TX 78712, USA
[7] Department of Astronomy, University of Michigan, 1085 S. University, Ann Arbor, MI 48109, USA
[8] Naval Research Laboratory, Remote Sensing Division, 4555 Overlook Ave. SW, Washington, DC 20375, USA
[9] Department of Astronomy & Astrophysics, University of California, Santa Cruz, 1156 High St., Santa Cruz, CA 95064, USA
[10] Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK
[11] Centre for Exoplanet Science, University of Edinburgh, Edinburgh EH9 3HJ, UK
[12] Department of Astrophysics, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024, USA
[13] Steward Observatory and the Department of Astronomy, The University of Arizona, 933 N. Cherry Ave., Tucson, AZ 85721, USA
[14] Space Sciences, Technologies & Astrophysics Research (STAR) Institute, Université de Liège, Allée du Six Août 19c, B-4000 Sart Tilman, Belgium
[15] Department of Physics & Astronomy, Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA
[16] LESIA, Observatoire de Paris, Université PSL, CNRS, Université Paris Cité, Sorbonne Université, 5 place Jules Janssen, 92195 Meudon, France
[17] School of Physical Sciences, Faculty of Science, Technology, Engineering and Mathematics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK
[18] Université Grenoble Alpes, Institut de Planétologie et d'Astrophysique (IPAG), F-38000 Grenoble, France
[19] UK Astronomy Technology Centre, Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK
[20] NASA-Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD 20771, USA
[21] Department of Astronomy, 501 Campbell Hall, University of California Berkeley, Berkeley, CA 94720-3411, USA
[22] Aix Marseille Univ, CNRS, CNES, LAM, Marseille, France
[23] Laboratoire J.-L. Lagrange, Université Cote d'Azur, CNRS, Observatoire de la Cote d'Azur, 06304 Nice, France
[24] Department of Physics and Astronomy, University of Texas-San Antonio, 1 UTSA Circle, San Antonio, TX, USA
[25] Subaru Telescope, National Astronomical Observatory of Japan, 650 North A'ohōkū Place, Hilo, HI 96720, USA
[26] Instituto de Astrofísica de Andalucía, CSIC, Glorieta de la Astronomía s/n, 18008, Granada, Spain
[27] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ, UK
[28] University of California, Los Angeles, 430 Portola Plaza, Box 951547, Los Angeles, CA 90095-1547, USA
[29] Eureka Scientific, 2452 Delmer. St., Suite 1, Oakland CA, 96402, USA
[30] Department of Physics and Astronomy, San Francisco State University, 1600 Holloway Ave., San Francisco, CA 94132, USA
[31] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[32] Center for Astrophysics and Space Sciences, University of California, San Diego, La Jolla, CA 92093, USA
[33] Department of Astronomy, Stockholm University, AlbaNova University Center, SE-10691 Stockholm, Sweden
[34] SETI Institute, Carl Sagan Center, 189 Bernardo Ave., Mountain View, CA 94043, USA
[35] Institute of Astrophysics, FORTH, GR-71110 Heraklion, Greece
[36] Department of Physics, University of Warwick, Gibbet Hill Road, Coventry, CV4 7AL, UK
[37] Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands

[38] Center for Space and Habitability, University of Bern, Gesellschaftsstrasse. 6, 3012 Bern, Switzerland
[39] Astrobiology Center of NINS, 2-21-1, Osawa, Mitaka, Tokyo, 181-8588, Japan
[40] Université Paris-Saclay, Université Paris Cité, CEA, CNRS, AIM, 91191, Gif-sur-Yvette, France
[41] Bay Area Environmental Research Institute and NASA Ames Research Center, Moffett Field, CA 94035, USA
[42] Institute for Astronomy, University of Hawai'i, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[43] Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA
[44] Department of Astronomy and Carl Sagan Institute, Cornell University, 122 Sciences Drive, Ithaca, NY 14853, USA
[45] Ludwig Maximilian University, University Observatory Munich, Scheinerstrasse 1, Munich D-81679, Germany
[46] University of California Observatories, 1156 High St., Santa Cruz, CA, 95064, USA
[47] AURA for the European Space Agency (ESA), ESA Office, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD, 21218 USA
[48] Dept. of Planetary Sciences; Lunar & Planetary Laboratory; Univ. of Arizona; Tucson, AZ 85721, USA
[49] National Research Council of Canada, Canada
[50] Herzberg Astronomy & Astrophysics Research Centre, National Research Council of Canada, 5071 West Saanich Road, Victoria, BC V9E 2E7, Canada
[51] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[52] Western University, Department of Physics & Astronomy and Institute for Earth and Space Exploration, 1151 Richmond Street, London, ON N6A 3K7, Canada
[53] Department of Physics, University of California, Santa Barbara, Santa Barbara, CA, 93106, USA
[54] IRFU/DAp Département D'Astrophysique CE Saclay, Gif-sur-Yvette, France
[55] Institute of Particle Physics and Astrophysics, ETH Zurich, Wolfgang-Pauli-Str. 27, 8093 Zurich, Switzerland
[56] Instituto de Física y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Av. Gran Bretaña 1111, Valparaíso, Chile
[57] Núcleo Milenio Formacíon Planetaria—NPF, Universidad de Valparaíso, Av. Gran Bretaña 1111, Valparaíso, Chile
[58] Landessternwarte, Zentrum für Astronomie der Universität Heidelberg, Königstuhl 12, D-69117 Heidelberg, Germany
[59] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[60] European Space Agency (ESA), ESA Office, Space Telescope Science Institute, 3700 San Martin Drive, Baltimore 21218, MD, USA
[61] Center for Astrophysical Sciences, The William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[62] The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[63] Tsung-Dao Lee Institute, Shanghai Jiao Tong University, 520 Shengrong Road, Shanghai, People's Republic of China
[64] Department of Astronomy \& Astrophysics, University of California, San Diego, La Jolla, CA 92093, USA
[65] Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK
[66] School of Physics, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland
[67] Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208, USA
[68] Department of Astronomy, Smith College, Northampton, MA 01063, USA
[69] Department of Earth and Planetary Sciences, University of California, Santa Cruz, Santa Cruz, CA, 95064 USA
[70] University of Virginia, Department of Astronomy, 530 McCormick Rd., Charlottesville, VA 22904, USA
[71] Instituto de Estudios Astrofísicos, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile
[72] Escuela de Ingeniería Industrial, Facultad de Ingeniería y Ciencias, Universidad Diego Portales, Av. Ejército Libertador 441, Santiago, Chile
[73] Millennium Nucleus on Young Exoplanets and their Moons (YEMS), Chile

Abstract

We present a performance analysis for the aperture masking interferometry (AMI) mode on board the James Webb Space Telescope Near Infrared Imager and Slitless Spectrograph (JWST/NIRISS). Thanks to self-calibrating observables, AMI accesses inner working angles down to and even within the classical diffraction limit. The scientific potential of this mode has recently been demonstrated by the Early Release Science (ERS) 1386 program with a deep search for close-in companions in the HIP 65426 exoplanetary system. As part of ERS 1386, we use the same data set to explore the random, static, and calibration errors of NIRISS AMI observables. We compare the observed noise properties and achievable contrast to theoretical predictions. We explore possible sources of calibration errors and show that differences in charge migration between the observations of HIP 65426 and point-spread function calibration stars can account for the achieved contrast curves. Lastly, we use self-calibration tests to demonstrate that with adequate calibration NIRISS F380M AMI can reach contrast levels of ∼9–10 mag at $\gtrsim \lambda/D$. These tests lead us to observation planning recommendations and strongly motivate future studies aimed at producing sophisticated calibration strategies taking these systematic effects into account. This will unlock the unprecedented capabilities of JWST/NIRISS AMI, with sensitivity to significantly colder, lower-mass exoplanets than lower-contrast ground-based AMI setups, at orbital separations inaccessible to JWST coronagraphy.

Unified Astronomy Thesaurus concepts: James Webb Space Telescope (2291); Interferometry (808); Direct imaging (387); High contrast techniques (2369); Observational astronomy (1145)

1. Introduction

The James Webb Space Telescope (JWST; Gardner et al. 2006, 2023) is now the first space-based infrared interferometer thanks to the aperture masking interferometry (AMI) mode on the Near Infrared Imager and Slitless Spectrograph (NIRISS) instrument (e.g., Sivaramakrishnan et al. 2012; Kammerer et al. 2022; Sivaramakrishnan et al. 2023). AMI transforms a conventional telescope into an interferometric array via a pupil-plane mask (e.g., Tuthill et al. 2000). By blocking the majority of the light, AMI enables the calculation of self-

calibrating observables that provide sensitivity to asymmetries close to and even within the classical diffraction limit. This technique has been thoroughly demonstrated from the ground via observations of stellar and substellar companions (e.g., Ireland & Kraus 2008; Biller et al. 2012; Hinkley et al. 2015), circumstellar disks (e.g., Tuthill et al. 2001; Sallum et al. 2019, 2023) and more.

Theoretical predictions of JWST/NIRISS AMI contrast have highlighted its potential for direct exoplanet studies (e.g., Soulain et al. 2020), including observations of actively forming planets (e.g., Sallum & Skemer 2019) and young planets near the water-ice line that are colder than those accessible with ground-based AMI (e.g., Ray et al. 2023a). JWST/NIRISS AMI has now been explored as part of the Early Release Science (ERS) 1386 program (described in its entirety in Hinkley et al. 2022). In Ray et al. (2023b), we demonstrate the unique science capabilities of NIRISS AMI, with a deep nondetection that places limits on close-in planets in the HIP 65426 exoplanetary system. Here we present an accompanying study to assess in detail the achievable contrast and noise characteristics of the NIRISS AMI mode.

1.1. NIRISS Aperture Masking Interferometry

On NIRISS, the AMI mode passes light from 7 of the 18 JWST primary mirror segments. The images on the detector are then the sums of interference fringes from the resulting 21 hole pairs (baselines), each of which probes a unique spatial frequency. Analyzing these data via Fourier techniques or fringe fitting enables the calculation of complex visibilities, which are the amplitudes and phases of the fringes associated with the mask baselines. From the complex visibilities we can calculate closure phases (e.g., Baldwin et al. 1986), sums of phases around baselines that form triangles, and squared visibilities, the powers associated with the various baselines (e.g., Jennison 1958). These and related quantities can be used to infer the source brightness distribution via model fitting and image reconstruction (often applied together to understand the limitations of each approach; Sallum & Eisner 2017).

1.2. Outline of This Paper

Here we discuss in detail the data reduction steps taken in the ERS 1386 AMI observations, their achieved contrast, and possible limiting noise sources for NIRISS AMI based on the analysis of the ERS 1386 data. The outline of the paper is as follows: In Section 2 we provide an overview of AMI noise sources and introduce the theoretical framework that we use to quantify AMI noise and performance. In Section 3 we briefly describe the observations of HIP 65426 and its two point-spread function (PSF) references. In Section 4 we describe and justify the AMI data reduction steps that were taken. There we also describe how we generate the contrast curves presented here and in Ray et al. (2023b). In Section 5 we discuss the observed performance of the ERS 1386 AMI observations and use a variety of tests to explore the limiting noise sources for NIRISS AMI. We discuss the results of these tests more generally in Section 6, before concluding with recommendations for future NIRISS AMI programs in Section 7.

2. AMI Noise Sources

AMI's sensitivity to close-in asymmetries is largely enabled by self-calibrating phase observables (e.g., closure phase; Baldwin et al. 1986). Due to the importance of phase and closure phase for achieving AMI resolution, in this work we focus primarily on phase and closure phase noise properties, leaving amplitude characterization to future studies. Here we briefly describe different sources of phase errors in AMI data and establish the theoretical framework that we use to assess and predict NIRISS AMI performance. Following Ireland (2013), below we describe three types of AMI phase noise sources: random, static, and calibration errors.

2.1. Random Phase Errors

Random errors are noise sources that contribute to variations around the mean. These include shot noise from the star and background and detector noise sources such as read noise and dark current. For an N_h-hole mask, we express the complex visibility phase scatter induced by these random noise terms in a manner consistent with Ireland (2013):

$$\sigma(\phi) = \frac{N_h}{N_p V} \sqrt{0.5(N_p + n_p \sigma_{\mathrm{pix}})}, \tag{1}$$

where V is the fringe visibility of the observation compared to a perfect point source and N_p is the total number of photons collected from the science target.

The σ_{pix} and n_p terms capture detector-level noise contributions, with σ_{pix} giving the pixel-level noise in units of e⁻. We use a general term for σ_{pix} since measurements of NIRISS's combined read noise, $1/f$ noise (correlated noise between pixels caused by the readout electronics configuration; e.g., Schlawin et al. 2020), and dark current noise exist in units of e⁻.[76] We note that the magnitude of these combined detector noise terms is low enough that, for the observations conducted here, the N_p term will dominate the expression for $\sigma(\phi)$. We include the σ_{pix} term for completeness and since this noise source has been well characterized for NIRISS.

Lastly, the n_p term is the total number of pixels in the image (s). For example, if two 80×80 NIRISS AMI subframes were used to collect a total of N_p photons, n_p would be set to $2 \times 80 \times 80$. Furthermore, any windowing (e.g., with a super-Gaussian; see Section 4) effectively decreases the number of pixels n_p, since windows suppress detector-level noise where their throughput is low. When we apply a window function, we thus calculate n_p by summing the pixels with the window throughput as a weight for each pixel. Extending Equation (1) to closure phase and allowing for each baseline to have its own visibility (V_i), the random closure phase noise for a single triangle is

$$\sigma(\phi_{\mathrm{CP}}) = \frac{N_h}{N_p} \sqrt{[0.5(N_p + n_p \sigma_{\mathrm{pix}})] \sum_{i=1}^{3} \frac{1}{V_i^2}}. \tag{2}$$

Here we are neglecting photon-noise contributions from the thermal background, since this noise contribution is small and difficult to accurately measure given the small size of the NIRISS AMI subframes. We also note that random temporal phase variations and flat-field errors (if images are not acquired

---

Table 1
Summary of Observations

Star	Readout	$N_{\rm groups}$	$N_{\rm ints}$	$t_{\rm exp}$ (s)
HD 115842	NISRAPID	2	10,000	2468.00
	NISRAPID	2	5500	1357.40
HIP 65426	NISRAPID	13	10,000	10,766.40
	NISRAPID	13	950	1022.81
HD 116084	NISRAPID	3	10,000	3222.4
	NISRAPID	3	6000	1933.44

on identical pixels) can contribute to statistical closure phase errors. We neglect these in our theoretical statistical error expression since JWST's stability should mean that these effects do not dominate the error budget. Furthermore, by following the recommended NIRISS AMI observing strategy of placing science targets and calibrators on identical pixels, flat-field errors can be mitigated. We discuss this simplification in the context of the observed phase scatter in Section 5.2.

2.2. Static Phase Errors

Static errors cause the mean closure phase signal to deviate from zero. Following the treatment in Ireland (2013), we describe these errors as being third-order and higher in phase (after Taylor expanding the expression for the complex visibility phase). These can arise from complex spatial variations in the wavefront phase across each mask subaperture. Time-invariable correlated and asymmetric detector systematics such as $1/f$ noise and alternating column pattern noise (e.g., Rauscher et al. 2017) can also contribute to static closure phase errors. These higher-order static terms must be adequately calibrated to maximize AMI achievable contrast, which is why traditional AMI campaigns (including ERS 1386) observe dedicated PSF reference targets.

2.3. Calibration Errors

If the static noise sources described above are not adequately removed during the calibration process, then the residual "calibration error" will cause the achievable contrast to differ from the photon-noise expectations defined in Equations (1) and (2). There are a variety of sources of calibration errors, including quasi-static speckles, which are slowly varying spatial aberrations. For JWST such quasi-static errors could be caused by thermal drift of the telescope between observations of the science target and the PSF stars. Calibration errors can also be caused by wavelength-dependent interactions between the light and optics that cause PSF shapes (and thus asymmetries) to depend on the target spectral type. Any detector-level systematics that change from science target to calibrator would also induce calibration errors (with variability from frame to frame for a single object causing inflated random errors; see, e.g., Section 2.1).

3. Observations

The ERS 1386 AMI observations were executed in the F380M filter on 2022 July 30 UTC (Table 1). The NISRAPID readout mode was used, and the images were acquired using the 80×80 pixel SUB80 subarray. Observations of HIP 65426 were preceded by observations of the PSF reference HD 115842 and followed by observations of a second PSF reference HD 116084. The HIP 65426, HD 115842, and HD

116084 integrations consisted of 13, 2, and 3 groups up the ramp, respectively. The numbers of total integrations recorded for the three objects were 10,950, 15,500, and 16,000, respectively. As described in Ray et al. (2023b), the PSF references were chosen to be close on-sky ($<5°$ separation) and with similar spectral types (B0.5-2I) to HIP 65426 (A2V). They were executed in a noninterruptible sequence and placed on identical detector pixels to the science target to minimize calibration errors.

4. NIRISS AMI Data Reduction and Analysis

4.1. Detector-level Corrections

We use the `jwst` pipeline (version 1.7.1; Bushouse et al. 2022) to perform standard detector-level corrections for each group. These include superbias, reference pixel, linearity, persistence, and dark current corrections. For any given reduction, we can apply some subset of the above corrections to the group-level data products. We then use stage 1 of `jwst` (i.e., by running `calwebb_detector1`) to generate rate images, which serve as the starting point for custom AMI reduction steps using `SAMpy` (e.g., Sallum et al. 2022). We tested various combinations of the `jwst` detector-level corrections and did not find significant differences between the AMI observables. The final reduction presented here and in Ray et al. (2023b) thus includes all of the `jwst` reduction steps.

We use the `jwst`-generated data quality, saturation, and jump detection arrays as the starting point for custom bad-pixel corrections. We also flag additional bad pixels applying our own criteria:

1. Pixels in individual integrations that deviate by more than 5σ from their median value (measured across all integrations). This selects high and low value outliers compared to an individual pixel's median behavior.
2. Pixels in individual integrations that differ by $>100\times$ from their location in a median-filtered image with a 3×3 kernel size. This effectively selects high and low value outliers compared to pixel neighbors in individual images.
3. Pixels with standard deviations (measured across all integrations) that differ by 5σ compared to their neighboring pixels. This selects "flaky" pixels by comparing a pixel's variability across all images to that of its immediate neighbors.

Together these criteria flag 1.5%–2.2% of pixels depending on the object. We apply a bad-pixel correction in Fourier space following the methodology described in Kammerer et al. (2019) (i.e., minimizing the bad pixels' contributions to the Fourier-plane power outside the spatial frequencies sampled by the pupil). Figure 1 shows example frames before and after bad-pixel correction for HIP 65426 and the PSF reference HD 116084.

Figure 1 also illustrates another low-level detector effect that can contribute to noise in the AMI observables. These are the sub-percent-level vertical striping features seen in the PSF reference stars, which were observed with fewer groups than the science target. These may be caused by $1/f$ noise and/or alternating column pattern noise (see, e.g., Section 2; Rauscher et al. 2017). Such effects can be mitigated by the use of a window function that tapers the image away from the center of
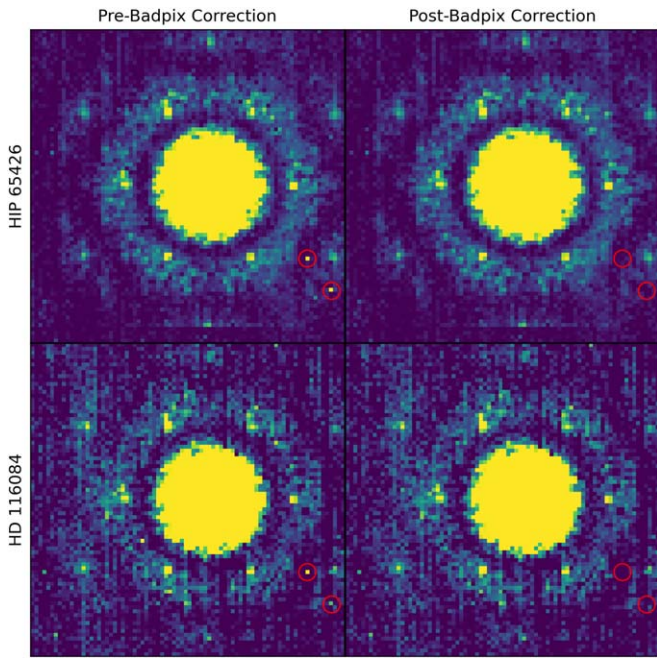
Figure 1. Example images of HIP 65426 (top panels) and HD 116084 (bottom panels) before and after bad-pixel correction (left and right panels, respectively). Red circles indicate two bad pixels in the left panels that have been corrected in the right panels. The color scale has been stretched so that it ranges from 0% to 1% of the peak counts in each image. Residual striping features are seen in the HD 116084 images, which we discuss in Section 4.1. The pixel size in all images is 65 mas, for a total field of view of $5{.}''2 \times 5{.}''2$.

the PSF. We thus explore the use of a super-Gaussian window function, which has the following form: $\exp(-(r/\sigma)^n)$, where r is the radius from the center of the image and σ determines the FWHM. For all of the tests that follow and for the final reduction we use a radial exponent of $n = 4$ for the super-Gaussian.

An aggressive super-Gaussian (e.g., small FWHM) would mean masking some portion of the PSF. For reference, the extent of the central portion of the AMI interferogram (determined by the first null in the PSF of a single mask hole) has a size of $\sim$36 pixels in diameter. The outer, ring-like portion of the interferogram, which corresponds to the first Airy ring in a single mask hole PSF (where signal is visible in each image), exists between $\sim$36 and $\sim$66 pixels in diameter. Thus, super-Gaussian functions with FWHMs significantly smaller than 36 pixels will attenuate the central portion of the interferogram, and those with FWHMs smaller than 66 pixels will attenuate the outer ring.

Given the dimensions described above, the choice of super-Gaussian FWHM must be made to balance including more photons (large FWHM) against eliminating pixels that contribute excess noise (small FWHM). We evaluate the choice of FWHM by quantifying the level of random error in the closure phases. To do this, we measure the scatter around the mean of each closure phase by taking the standard deviation across all of the integrations for each object. We explore super-Gaussian FWHMs between 28 and 48 pixels and normalize the standard deviations by that for the smallest FWHM of 28 pixels. The upper FWHM limit of 48 pixels is chosen so that the attenuation at large distances is high enough to avoid systematics due to sharpness at the edge of the 64×64 pixel centered and cropped images. A change in FWHM that

increases the signal-to-noise ratio would lead to a decrease in the scatter around the mean, and vice versa.

Figure 2 shows examples of the results of these tests. For the majority of the closing triangles, the scatter about the mean decreases as the size of the super-Gaussian window increases. However, this is not the case for closure phases whose triangles include horizontal baselines, in the two PSF reference data sets. This is because the spatial frequencies probed by horizontal baselines are particularly sensitive to asymmetries caused by vertical striping. The effect is not seen in the science target closure phases with horizontal baselines, because this striping is mitigated when more groups are used up the ramp.

The results shown in Figure 2 seem to motivate applying different windowing settings to different closure phase calculations. However, as we discuss in Section 5.1, the noise in the calibrated closure phases was minimized in the reduction where all objects and closing triangles have the 48-pixel FWHM window applied. Furthermore, the possibility of inflated systematic errors due to this striping in the PSF references, as well as the resulting calibration errors, motivates detector-level removal of the stripes. We thus tested a simple destriping approach similar to the channel bias correction described in Sallum & Eisner (2017). We used the outskirts of the images to estimate individual column biases, which we then subtracted, but this did not lower the closure phase scatter for these triangles. This is perhaps due to the fact that the SUB80 subarray is relatively small compared to the NIRISS AMI PSF. In future work more sophisticated destriping approaches will be explored.

4.2. Extraction of Fourier Observables

SAMpy can extract Fourier observables by sampling either single pixels at the center of each baseline's location in Fourier space or multiple pixels for each baseline, with an extent defined by the user. The reductions presented here and in Ray et al. (2023b) take the second approach, sampling all pixels within 25% of the peak signal for each baseline. This flux cut is arbitrary and changeable; a 25% value ensures that most of the (u, v) space for each baseline is sampled, but without beginning to sample nearby baselines. Sampling the Fourier transform within a 50% flux cut did not change the results significantly.

The visibility amplitudes, visibility phases, and squared visibilities are calculated by averaging the observables measured from these regions of the Fourier transform. When using multiple pixels, the closure phases can be calculated in two ways: (1) by finding all individual triangles of pixels that close, calculating their bispectra, and averaging the bispectra for each closing triangle before taking the phase as the closure phase; or (2) by calculating an average complex visibility for each baseline in the closing triangle, calculating one bispectrum value using the three average visibilities, and then taking its phase as the closure phase.

While closure phase approach 1 is strictly correct, it can be computationally expensive since many closing triangles of pixels may connect the three baselines. We thus test whether there is a significant difference between the observables calculated between approaches 1 and 2, and we find that they are nearly identical. This may be because we first apply a window function that creates correlations between pixels in Fourier space. We thus apply approach 2 for the closure phase tests shown here and for the final reduction presented in Ray et al. (2023b). However, we note that an additional advantage
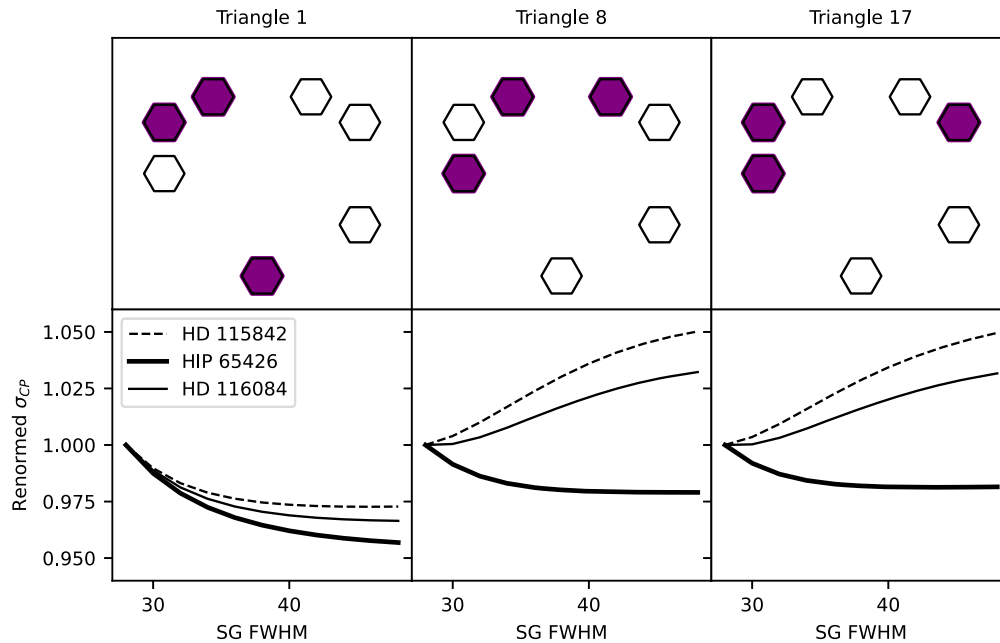
Figure 2. Top panels depict three example closing triangles from the NIRISS AMI mask. The black outlines show the positions of all seven holes, and the filled purple hexagons show the holes used to calculate the closure phases in the bottom panels. Bottom panels show the closure phase standard deviation around the mean (σ_{CP}) as a function of super-Gaussian FWHM (SG FWHM), for the closing triangles depicted in the top panels. The σ_{CP} values are normalized by the value for the smallest SG FWHM of 28 pixels. The thick solid line shows HIP 65426. The thin solid and dashed lines show the two PSF references HD 116084 and HD 115842, respectively. Increasing the super-Gaussian FWHM size for triangles with horizontal baselines leads to an increase in the random error of the PSF reference closure phases. This is because the larger super-Gaussian FWHMs allow for increased contributions by vertical striping on the detector (which has the largest effect on the horizontal baselines).

of approach 1 is that it can be applied without averaging all the pixel-triangle bispectra, to preserve phase information that changes within the mask subapertures. This may be useful for trying to recover wide-separation signals that would be aliased by the minimum mask baselines, such as the signal from the wide-separation companion HIP 65426 b (e.g., Ray et al. 2023b).

For all observables, we assign statistical error bars capturing the random phase and amplitude errors. We estimate these by measuring the standard deviation of each observable across all integrations in the data set and then dividing by the square root of the number of integrations. This gives the standard error of the mean, or the statistical uncertainty in the measurement of the mean observable. As discussed in Section 2, these errors may not dominate if calibration errors exist, which is often the case with AMI observables. For fits to calibrated data, following, e.g., Ireland (2013), we inflate error bars to match the scatter measured across the closing triangles (such that the reduced χ^2 of the best-fit model is roughly equal to 1).

4.3. Calibration

To calibrate the HIP 65426 observables, we subtract the (closure) phases of a reference from the HIP 65426 phases and divide the (squared) visibilities of a reference into the HIP 65426 visibilities. Since these data consist of a single science pointing sandwiched between two reference star pointings, we test four different calculations of that reference. The first two are simply using each individual PSF star as the calibration reference. We calculate a third reference by fitting a linear trend in time to the two PSF star observables for each baseline/closing triangle and then sampling that trend at the time of the HIP 65426 observable. Lastly, we calculate a calibration

reference by averaging (with equal weighting) the observables of the two reference stars. As discussed in Section 5.1, we find that the calibration quality is maximized when only the HD 116084 PSF star is used as the calibration reference.

4.4. Modeling and Contrast Curve Generation

Ray et al. (2023b) describe the procedure we apply for model fitting and the results of model fits to the calibrated HIP 65426 data set. For this paper we focus on contrast curves as one method of quantifying NIRISS AMI performance. To measure achievable contrast, following Sallum & Skemer (2019), we fit a grid of single-companion models to the data, varying their separation, position angle, and contrast. We average over position angle, resulting in a grid in just separation and contrast. For each separation, we then calculate the 5σ detectable contrast by finding the companion model with a χ^2 that is greater than the null (no companion) model χ^2 by $\Delta\chi^2 = 25$. A $\Delta\chi^2$ interval of 25 (corresponding to 5σ for a distribution with one degree of freedom) is used to calculate the achievable contrast since, for each separation, only one parameter (contrast) determines the difference between each model and the null model.

This procedure assumes that the reduced χ^2 of the best-fit companion model is equal to 1 and that the observables' error bars are Gaussian. The first assumption is valid since we use the inflated error bars described in Section 4.2, which are constructed such that the best-fit reduced χ^2 is roughly equal to 1. Additionally, to calculate the contrast curves, we explicitly rescale the χ^2 values so that the best-fit has a reduced χ^2 of exactly 1 (effectively assuming that any poor fit quality is caused by imperfect error bar estimation). For the second assumption, as discussed in detail in Section 5, the AMI
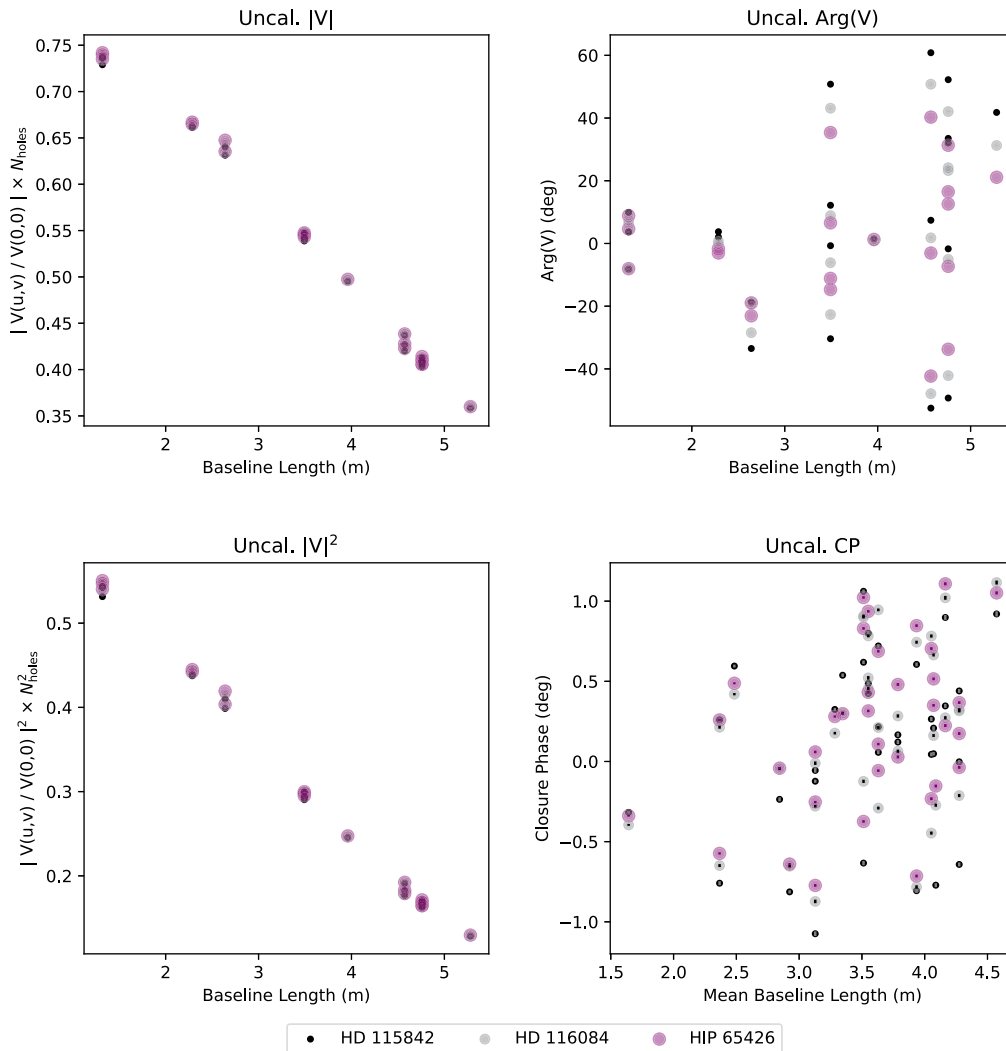
Figure 3. Uncalibrated Fourier observables. The small black circles, medium gray circles, and large purple circles show observables for HD 115842, HD 116084, and HIP 65426, respectively. The top left panel shows complex visibility amplitudes vs. baseline length, normalized by the total signal in the images and multiplied by the number of mask holes. These values should all be equal to 1 for a perfect-Strehl point-source observation. The bottom left panel shows squared visibilities vs. baseline length with the same normalization applied. The top right panel shows complex visibility phases vs. baseline length, and the bottom right panel shows closure phases vs. the average length of the three baselines in each triangle. In all panels statistical error bars (estimated from the standard error measured across all integrations) are smaller than the size of the symbols.

observables' uncertainties are dominated by calibration errors, which may not necessarily follow a Gaussian distribution.

While the possibility of non-Gaussian errors means that these contrast curves may be optimistic estimates for practical AMI data sets, we use them here for a few reasons. The first is that the distributions of calibrated closure phases and squared visibilities can be shown to be roughly Gaussian (albeit with a higher standard deviation than the statistical errors, which we capture in the error bar estimation described in Section 4.2). Furthermore, self-calibrated data sets that can be used to explore potential performance will have calibration errors equal to 0, making their error bars better modeled by Gaussian distributions. Lastly, the computational expediency of this method compared to, e.g., companion injections and recoveries also means that a large number of contrast curves can be generated quickly to enable the wide variety of calibration tests performed here. However, we note that for an AMI data set dominated by calibration errors, injection and recovery tests are useful both for exploring the validity of the $\Delta\chi^2$ approach and for determining robust scientific detection limits.

5. Results: NIRISS AMI Performance and Noise Characteristics

5.1. Fully Reduced Fourier Observables

Figure 3 shows the uncalibrated Fourier observables for HIP 65426 and the two reference stars HD 115842 and HD 116084. In all four panels, the sizes of the statistical error bars are significantly smaller than the scatter in the observables. This shows that significant static errors exist, which require calibration for removal. Figure 4 shows the best calibration of the squared visibilities and the closure phases, which is achieved when the HD 116084 observables are used as the calibration reference and when the 48-pixel FWHM super-Gaussian window is applied to all images. We show just these two observables since they are used to generate the detection maps in Ray et al. (2023b). Comparing the calibrated data in Figure 4 to the uncalibrated data in Figure 3 shows that the calibration significantly reduces the static errors in the data. In the following subsections we discuss the random, static, and calibration errors in more detail.
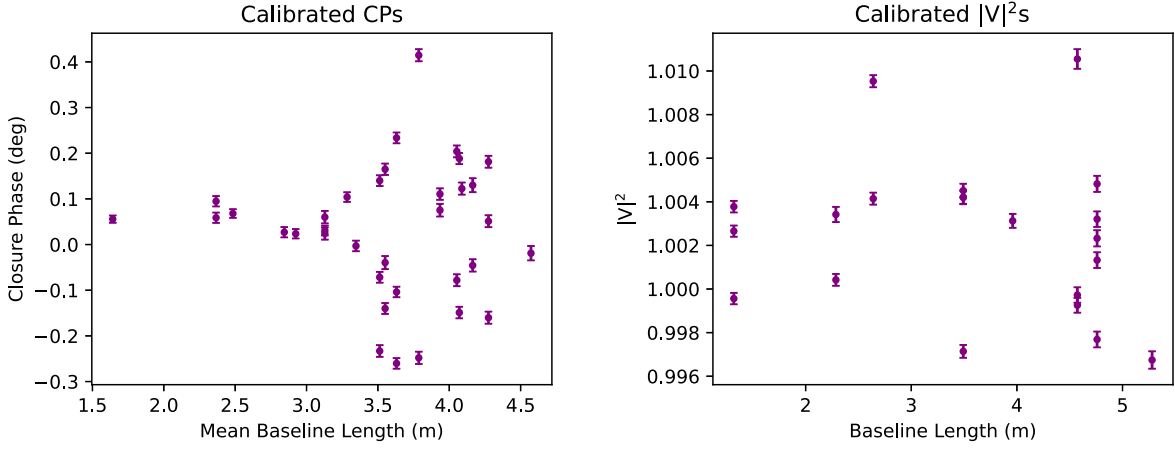
Figure 4. Calibrated HIP 65426 Fourier observables. The left panel shows the HIP 65426 closure phases vs. the mean baseline length of each triangle, after calibrating with the HD 116084 reference. The right panel shows the HIP 65426 squared visibilities after calibrating with the HD 116084 reference. The error bars show the statistical errors that have been propagated taking the calibration procedure into account.
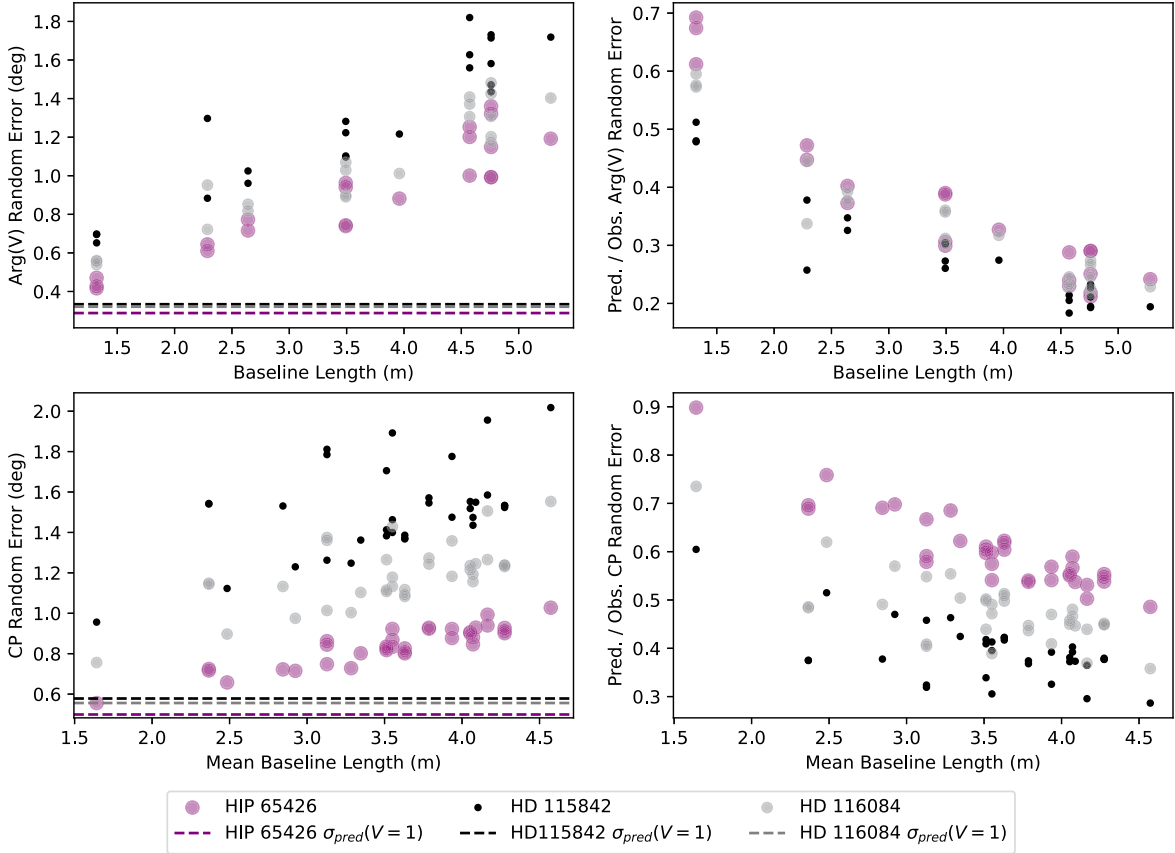

Figure 5. Left panels show the observed standard deviation of each visibility phase (top) and closure phase (bottom) vs. baseline length. The dashed lines show the predicted scatter for phase observables with a fringe visibility equal to 1. Right panels show the ratio between the predicted and observed scatter for the two phase quantities. The decrease in this ratio as a function of baseline length is consistent with a systematic decrease in fringe visibility. In all panels, the small black circles, medium gray circles, and large purple circles show observables for HD 115842, HD 116084, and HIP 65426, respectively.

5.2. Random Errors, Static Errors, and Fringe Visibility

The magnitude of the static phase errors can be roughly estimated by taking the standard deviation of the final, time-averaged phases. For complex visibility phases this standard deviation is measured across the baseline index, and for the closure phases it is measured across the triangle index. These values are $\sim 21°$–$31°$ for the complex visibility phases, with the scatter increasing from HD 115842 to HD 116084 to HIP

65426. The closure phase static errors range from $0°\!.53$ to $0°\!.58$ for the three objects. The fact that the errors are inflated for the phases compared to the closure phases is due to the self-calibrating properties of the closure phases.

The statistical error bars between the phases and closure phases are similar, with values of $\sim 0°\!.008$–$0°\!.012$ for the three objects. If these errors came only from the terms in Equations (1) and (2), the closure phase errors should be
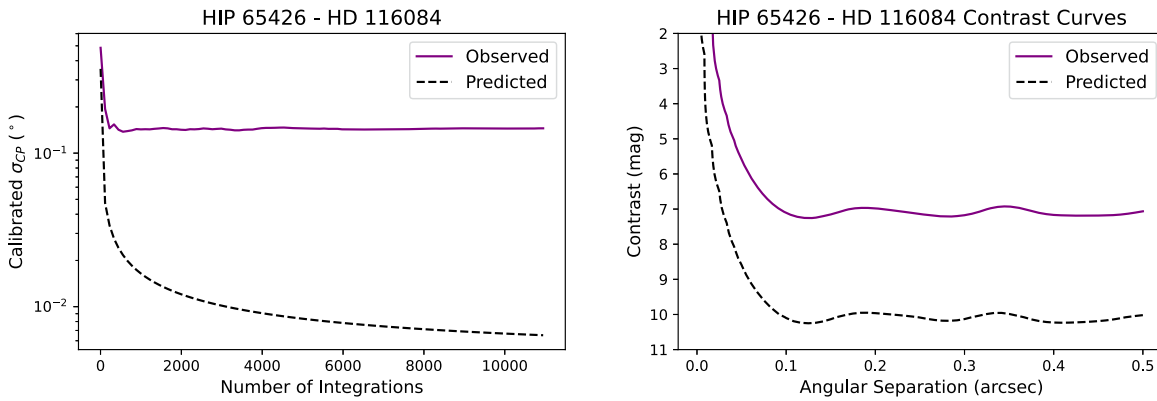
Figure 6. Left: comparison between HIP 65426 calibration quality and expectations for photon noise. The purple solid line shows the scatter in the time-averaged calibrated HIP 65426 closure phases, as a function of the number of integrations that were included in the average. The dashed black line shows the expected closure phase scatter for photon noise. The noise in the HIP 65426 closure phases calculated using all 10,950 integrations is inflated by a factor of 22.25 compared to photon-noise expectations. Right: contrast curves calculated from the observed HIP 65426 observations (using all 10,950 integrations) calibrated with the entire HD 116084 data set, compared to theoretical predictions. The purple solid line shows the observed contrast, and the black dashed line shows the predicted contrast. The observations underperform compared to predictions by a factor of ∼3 mag.

approximately a factor of $\sqrt{3}$ higher than the phase errors. This suggests that there is an additional random noise term affecting the visibility phases and not the closure phases. One possible explanation for this is pointing jitter, which would calibrate down in the closure phases (since it introduces a ramp in Fourier phase for each integration) but not in the phases. Jitter combined with flat-field variation would also contribute first-order phase errors to the visibility phases, but it would only contribute at third order and higher for closure phases. A quantitative comparison of pointing jitter and flat-field errors' effects on statistical phase and closure phase errors will be the subject of future work.

The left panels of Figure 3 show that the static errors in the (squared) visibility amplitudes vary as a function of baseline length. The normalization shown in Figure 3 (dividing the amplitudes by the zero-spacing amplitude, and multiplying by the number of holes) is constructed so that a perfect-Strehl point source would have all visibilities equal to 1. The dropoff as a function of baseline length suggests that the coherence decreases systematically for the longer baselines, which is a well-known phenomenon in ground-based AMI data sets (e.g., Monnier 1999; Sallum et al. 2021).

The drop in raw visibility amplitudes is relevant for the random noise predictions given by Equations (1) and (2), since the visibilities map directly to the V terms in their denominators. This implies that the statistical errors of the (closure) phases should increase as a function of the (mean) baseline length. Figure 5 shows that this is the case. The random variations of each complex visibility phase and each closure phase increase as a function of baseline length.

If all count-rate quantities in Equation (1) were properly estimated and all random noise sources were taken into account, the ratio between the random complex visibility phase errors and the Equation (1) predictions for 100% Strehl should equal the fringe visibility for each baseline. These quantities are shown in the top right panel of Figure 5, and they show a similar characteristic decrease to the raw visibilities shown in Figure 3. However, two important differences exist between the two. The first is that the raw visibility amplitudes are always higher than the visibilities that would be inferred from the ratios. This indicates that there are additional random noise

sources for all three objects that are not captured by Equation (1). The second is that there are variations in the ratios from object to object that are not seen in the raw visibility amplitudes. This indicates that the additional random noise sources are more severe for the two calibrator stars than for HIP 65426, and most severe for the HD 115842 calibrator. We explore whether this could stem from the number of groups used for these two objects in Section 5.4.

The tests in Figure 5 and the raw visibilities in Figure 3 suggest that we should not be adopting $V = 1$ in Equations (1) and (2). However, adopting the raw visibility values in Figure 3 causes the photon-noise predictions from Equation (2) to be larger than the observed scatter about the mean closure phases for all three objects. This may be due to the fact that both windowing and averaging over multiple pixels in the Fourier transform can smooth out noise variations in a way that is not captured by Equations (1) and (2). Since this is the case, and since AMI observers often assume $V = 1$ during expected signal-to-noise ratio calculations, in the closure phase noise comparisons that follow we assume $V = 1$ for theoretical noise calculations. This can be thought of as a conservative choice for comparing measured performance to theoretical predictions, since assuming $V = 1$ results in the lowest predicted closure phase scatter and thus the worst underperformance relative to theoretical predictions. We note that an intermediate assumption that takes into account the slope of the visibilities versus baseline (but not the overall normalization) may be a more accurate predictor of the expected closure phase noise levels.

5.3. Reference PSF Calibration Quality

Here we quantify the closure phase calibration quality by making comparisons to the theoretical predictions of Equation (2) assuming $V = 1$. With a large enough number of collected photons, static errors will dominate over random errors. The observables can thus be expected to have a $\sigma \propto \sqrt{n_{int}}$ trend, where n_{int} is the number of integrations, until the systematic noise floor is reached. With perfect calibration, all underlying systematics would be removed, and the calibrated data would follow the $\sigma \propto \sqrt{n_{int}}$ trend.

To assess the calibration quality, we compare the scatter in the calibrated HIP 65426 data to expectations for photon noise.
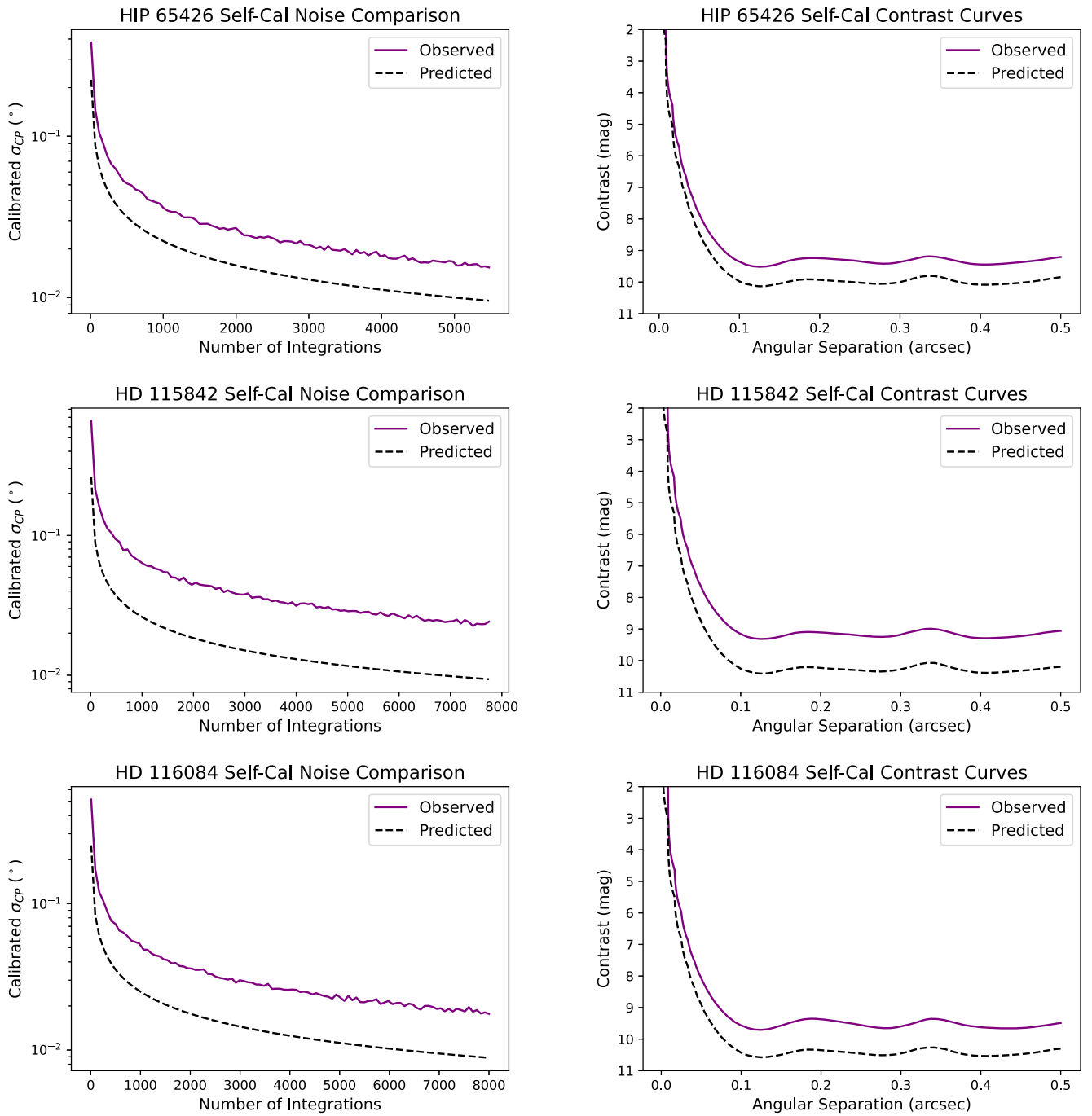
Figure 7. Left panels: comparisons between self-calibration quality and expectations for photon noise, for HIP 65426, HD 115842, and HD 116084 from top to bottom. In each panel, the purple solid line shows the scatter in the time-averaged calibrated closure phases for a jackknifed data set, as a function of the number of integrations that were included in the average. The dashed black line shows the expected closure phase scatter for photon noise. The ratio of these two quantities (observed/predicted) is steady for all three objects, at $\sim$1.6 for HIP 65426, $\sim$2.5 for HD 115842, and $\sim$2.1 for HD 116084. Right panels: contrast curves calculated from the self-calibrated observations using the maximum number of integrations in each jackknifed data set, compared to theoretical predictions. Top to bottom panels show HIP 65426, HD 115842, and HD 116084. In each panel, the purple solid line shows the observed contrast, and the black dashed line shows the predicted contrast.

For this test and in some of the analysis that follows, we use jackknife resampling (e.g., Quenouille 1949) to create shallower data sets by systematically removing individual integrations. For example, to calculate the calibrated closure phase scatter in a data set with n_{int} integrations, we create a large number of "jackknifed" data sets by selecting different subsets of n_{int} integrations pulled from the entire data set for a given object. For each subset of n_{int} integrations, we calculate the mean set of 35 closure phases. We then perform calibration

as described in Section 4.3 and calculate the standard deviation of the 35 calibrated closure phases. We take the average of that standard deviation across all jackknifed n_{int}-integration data sets as representative of the calibrated closure phase scatter.

We apply this jackknife resampling procedure to the HIP 65426 integrations and calibrate using all of the integrations from the HD 116084 PSF calibrator. We use HD 116084 since it provided better calibration quality (final closure phase scatter of 0°.14). We split the HIP 65426 data up into shallower data
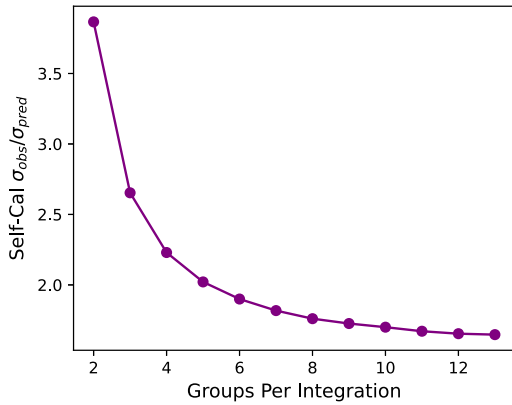
Figure 8. The ratio between the observed closure phase scatter and the theoretical noise floor for self-calibrated HIP 65426 data sets, as a function of the number of groups included per integration. In this test, the jackknifed data sets have the maximum number of integrations (5475, half of the HIP 65426 total integrations). The calculation of the statistical noise floor includes both the higher read noise and lower photon count when fewer groups are used.

sets with depths ranging from $n_{int} = 2$ integrations to the total number of $n_{int} = 10,950$. For each data set we then subtract the average of all of the HD 116084 integrations to calibrate. This creates a large number of HIP 65426 data sets with varying depths that have all been calibrated with the highest signal-to-noise ratio HD 116084 measurements. We calculate the scatter for each set of calibrated closure phases by taking the standard deviation across the different triangles.

To calculate expectations for the noise floor, we calculate the total counts in the subframed, super-Gaussian windowed images. We convert to the total number of collected photons accounting for the exposure parameters ($t_{int} = 0.22632$ for HD 116084, $t_{int} = 0.98072$ for HIP 65426) and the NIRISS gain (1.61 e$^-$/ADU). This results in 9.87e5 photons per integration for HIP 65426 (for a total of 1.08e10 photons for all 10,950 integrations) and 7.99e5 photons per integration for HD 116084 (for a total of 1.28e10 photons for all 16,000 integrations). We then use Equation (2) to predict the closure phase scatter in the photon-noise limit. For each HIP 65426 data set consisting of n_{int} averaged integrations, we also calculate the ratio between the resulting scatter and the predicted scatter given the total number of photons for those n_{int} integrations.

Figure 6 shows the results. The scatter in the calibrated HIP 65426 closure phases decreases to a depth of about 200 integrations and then flattens out at $\sim 0°.14$. This shows that there are residual systematic errors present in the calibrated data, at a level consistent with expectations for $\sim$2e8 photons. The ratio between the observed scatter and the predicted scatter for the full data set depth ($n_{int} = 10,950$) is 22.25. This amounts to a degradation in achievable contrast of $\sim$3 mag.

As a consistency check, we performed the same tests as above using HD 115842 as a reference for HIP 65426 and using HD 116084 as a reference for HD 115842. When the full HD 115842 data set is used as a PSF reference for HIP 65426, the calibrated closure phase scatter reaches a level of $\sim 0°.26$ by the same depth of $\sim$200 integrations. When HD 116084 is used as a PSF reference for HD 115842, the calibrated scatter reaches $\sim 0°.25$ by roughly the same number of integrations. These tests, especially the $\sim 2\times$ higher scatter in the HIP 65426 data calibrated with HD 115842, show that varying static errors

are present across the three data sets. This is also evidenced by the differing single-companion fit results for the HIP 65426 data set calibrated with each reference PSF (Ray et al. 2023b).

5.4. Self-calibration Tests: Achievable Contrast with Perfect Calibration

We use "self-calibration" tests to explore the source of these inflated errors and resulting lower contrast. To check that the underlying problem is indeed an uncalibrated static error, we jackknife each data set into pairs of shallower data sets and use those pairs to self-calibrate. We then perform the same tests as described in Section 5.3. For example, for HIP 65426, we create pairs of data sets with total numbers of integrations between 2 and 5475 (half the number of integrations of the complete HIP 65426 data set). We average each of the two data sets to calculate two sets of mean closure phases, which we then calibrate against each other. We note that this test is only useful for characterizing the random noise level of the data set, since the self-calibration process removes both the static errors and any science signal present in the data.

Following self-calibration, we compare the results to photon noise following the procedure in Section 5.3. Figure 7 shows the results. For all three objects, the scatter in the self-calibrated data set decreases proportionally to $\sqrt{n_{int}}$, where n_{int} is the number of integrations included in each jackknifed data set. The calibration quality compared to photon noise varies from object to object, with HIP 65426 reaching a noise level that is $\sim 1.6\times$ the noise floor, HD 115842 at $\sim 2.5\times$, and HD 116084 at $\sim 2.1\times$.

We note that, given the visibility caveats discussed at the end of Section 5.2, the overall normalization of these values is likely overly conservative, but their relative values indicate increased random errors for the two reference PSFs. In order to explore the source of this additional noise in a more controlled way, we calculate the scatter in the self-calibrated HIP 65426 data after varying the number of groups used per integration. As shown in Figure 8, the ratio between the observed scatter and the noise floor is worse when fewer groups are used. This suggests that the additional noise may be a detector-level effect that is somewhat mitigated by increasing the number of reads up the ramp.

Figure 8 also shows that the ratio between the observed scatter and the noise floor for the self-calibrated HIP 65426 data is higher than that for the calibrators when identical numbers of groups are considered. The ratio for HIP 65426 is ~ 3.9 for a two-group reduction and ~ 2.7 for a three-group reduction, compared to 2.5 for the two-group HD 115842 data set and 2.1 for the three-group HD 116084 data set. This is also consistent with the additional noise being a changing detector-level effect such as time-variable pattern noise (visible as the vertical stripes in Figure 1, which change from image to image). Images with fewer collected photons will suffer more from these time-variable detector systematics, which are independent of the target brightness. This is likely the cause of the inflated noise for the two- and three-group reductions of HIP 65426, which is significantly fainter than the two calibrators.

These self-calibration tests show that, with adequate removal of static errors, NIRISS can reach unprecedented broadband AMI contrasts despite the slightly inflated random errors for the few-group integrations. This is illustrated in the right panels of Figure 7. These show contrast curves for the deepest data sets
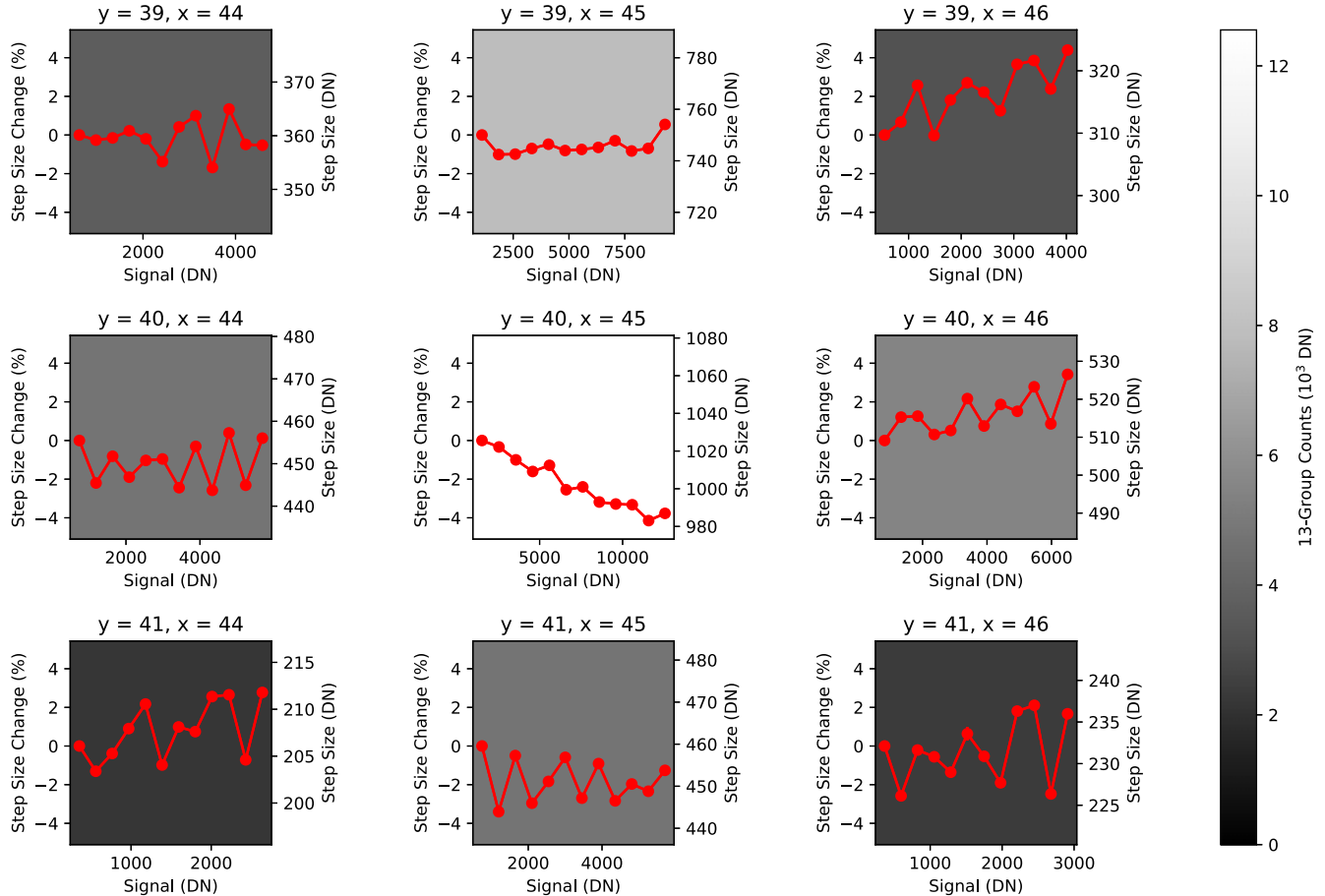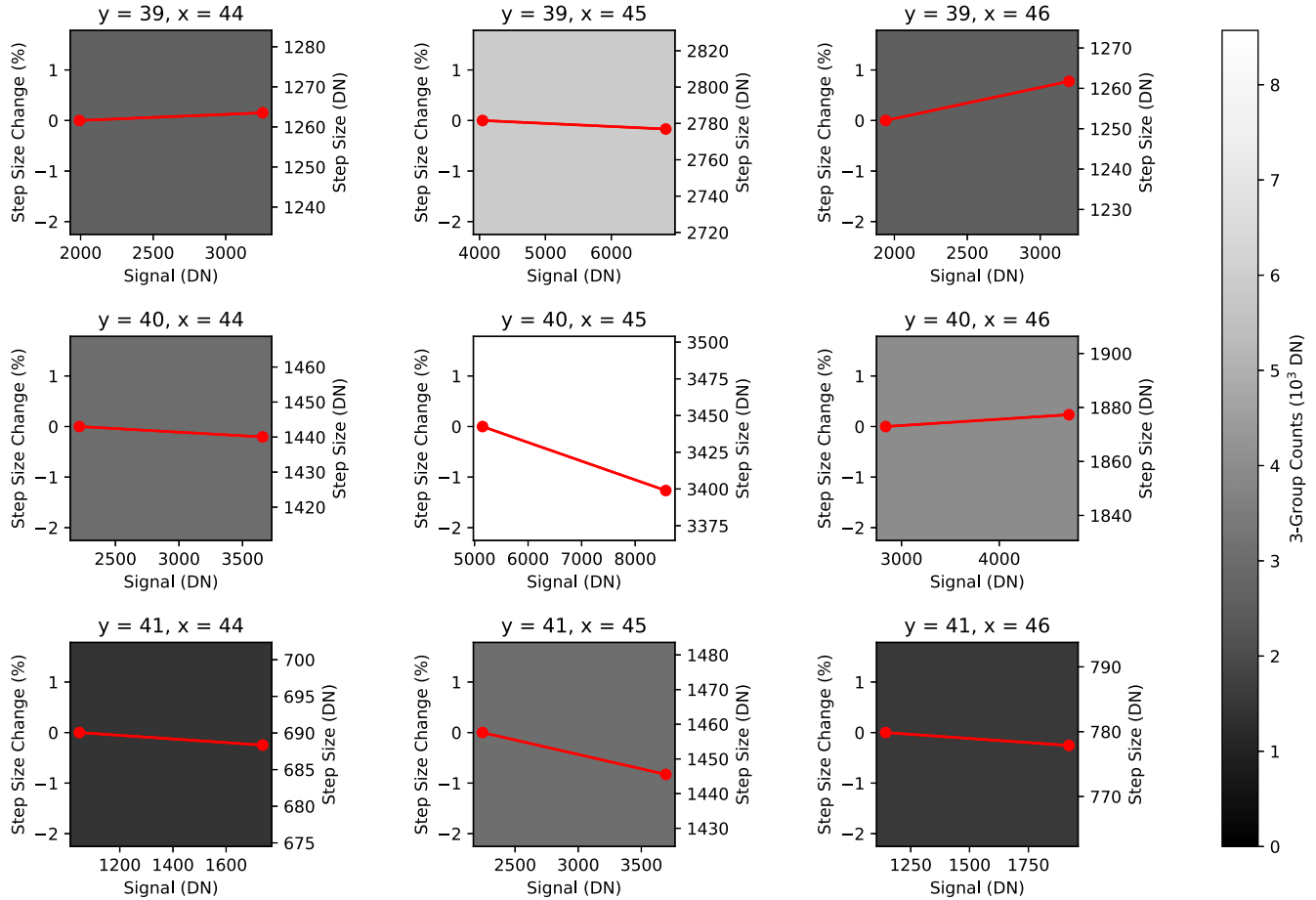
Figure 9. Charge migration analysis for HIP 65426. Each panel corresponds to a single pixel, with the nine panels centered around the brightest pixel in the AMI PSF. In each panel, the red line shows the change in accumulated charge relative to the first group pair (left axis), as well as the absolute amount of charge accumulated by each group pair (right axis). The gray scale indicates the maximum accumulated charge in each pixel (i.e., measured by the final two group pairs), with the color scale set by the peak counts in the central, brightest pixel in the PSF. Comparing the lines in the various panels shows that the charge accumulation in the central pixel decreases by ~4% over the course of an integration, while the charge accumulation in the pixels right and diagonally right of center increases by ~4%. This charge migration causes signal-dependent PSF changes.

in the left panels, where the two jackknifed data sets consist of half the total number of integrations for each object. The achievable contrasts at separations $\gtrsim \lambda/D$ range from 9 to 10 mag. The self-calibrated HIP 65426 contrast, at ~9.5 mag, outperforms the final HIP 65426 reduction (calibrated with HD 116084) by ~2.5 mag. This motivates the need to better understand and calibrate the residual errors shown in Figure 6.

5.5. Charge Migration Levels and PSF Effects

One possible source of systematic differences in PSF shape is charge migration. This is a physical phenomenon where the electric field generated by accumulated charges in one location on the detector leads to the deflection of newly incident charges onto a slightly different location on the detector (e.g., Plazas et al. 2018; Hirata & Choi 2020). These charge deflections cause the PSF to smear out spatially, with charges from the brightest pixels "spilling over" onto adjacent ones. The physical mechanism behind charge migration means that it becomes more dramatic at higher levels of total accumulated charge, causing the PSFs for brighter objects to smear out

more. These brightness-dependent PSF changes are often referred to as the "brighter-fatter effect."

H2RG detectors have been demonstrated in the laboratory to exhibit charge migration for high levels of accumulated charge (~20,000 e⁻; Plazas et al. 2018). This ~20,000 e⁻ signal limit is generally used during JWST/NIRISS observation planning, and the ERS 1386 observations were designed to be below this count level for individual integrations. While severe charge migration is thus unlikely, achieving exquisite contrast of ~9–10 mag could be negatively affected by very low levels of charge migration. We thus analyze the `jwst` 4D data products to assess the degree of charge migration in the data for each object and to quantify its effect on their relative PSFs, following a procedure identical to that in Plazas et al. (2018).

The 4D `jwst` data products have shapes [n_{int}, n_g, n_y, n_x], where n_{int} is the number of integrations, n_g the number of groups, and n_y and n_x the number of pixels in the y- and x-directions. We start with the 4D data cubes that have had all standard reduction steps applied, including superbias, reference pixel, linearity, persistence, and dark current corrections. These would normally next undergo ramp fitting to produce rate images. For each integration, we take the differences between

Figure 10. Charge migration analysis for HD 116084. Each panel corresponds to a single pixel, with the nine panels centered around the brightest pixel in the AMI PSF. In each panel, the red line shows the change in accumulated charge relative to the first group pair (left axis), as well as the absolute amount of charge accumulated by each group pair (right axis). The gray scale indicates the maximum accumulated charge in each pixel (i.e., measured by the final two group pairs), with the color scale set by the peak counts in the central, brightest pixel in the PSF. Comparing the lines in the various panels shows that the charge accumulation in the central pixel decreases by ∼1% over the course of an integration, while the charge accumulation in the pixels right and diagonally right of center increases by ∼1%. This charge migration causes signal-dependent PSF changes.

adjacent groups to calculate the total charge accumulated (in DN) at each step up the ramp. For any given pixel, we can then plot the accumulated charge at each step, as a function of the total accumulated signal at that step (calculated as the average of the two adjacent groups).

Figure 9 shows the results for the central 9 pixels of the 13-group observations of HIP 65426. This analysis shows that the charge accumulation in the central, brightest pixel ($x = 45$, $y = 40$) decreases slightly as the total signal increases from ∼1500 to ∼13,000 DN. The absolute step size decreases by ∼40 DN from ∼1025 to ∼986 DN. These changes correspond to a fractional decrease relative to the first group pair of ∼4%. These decreases in the central pixel are accompanied by ∼15–20 DN (few percent level) increases in the pixels directly right ($x = 46$, $y = 40$) and up and right ($x = 46$, $y = 39$) of center.

Figure 10 shows the results for the same 9 pixels in the three-group observations of the HD 116084 PSF reference. In this case, a similar absolute decrease occurs for the central pixel, with step sizes decreasing from ∼3442 to ∼3398 DN as the total charge increases from ∼5150 to ∼8570 DN. Since the individual steps up the ramp are larger, this decrease corresponds to a fractionally lower change in step size of

∼1%. For these data, the greatest changes in adjacent pixels are a ∼15 DN (∼1%) decrease in the pixel directly below center ($x = 45$, $y = 41$) and a ∼10 DN (∼1%) increase in the pixel up and right of center ($x = 46$, $y = 39$). Since the HD 115842 observations only utilized two groups, we cannot investigate charge migration levels for that object.

These tests show that small amounts of charge move between pixels as integrations move up the ramp. To investigate the degree to which this charge migration affects the PSF shape, we compare the central 9 pixels of the PSF for different reductions of the HIP 65426 data set that include varying numbers of groups. Since AMI observations are sensitive to fractional changes in the PSF (i.e., overall normalization of the image does not change the closure phase signal), we compare the fluxes in the central 9 pixels after normalizing by the central, brightest pixel. In this case, we use `jwst` to produce 3D `rateints` data cubes from subsets of the HIP 65426 4D cubes, varying the maximum number of groups ($n_{g,\mathrm{max}}$) used for ramp fitting. We then average the central 9 pixels over all integrations and normalize them by the signal in the central brightest pixel. We take differences between reductions with different $n_{g,\mathrm{max}}$ values to quantify the fractional changes in the PSF cores.
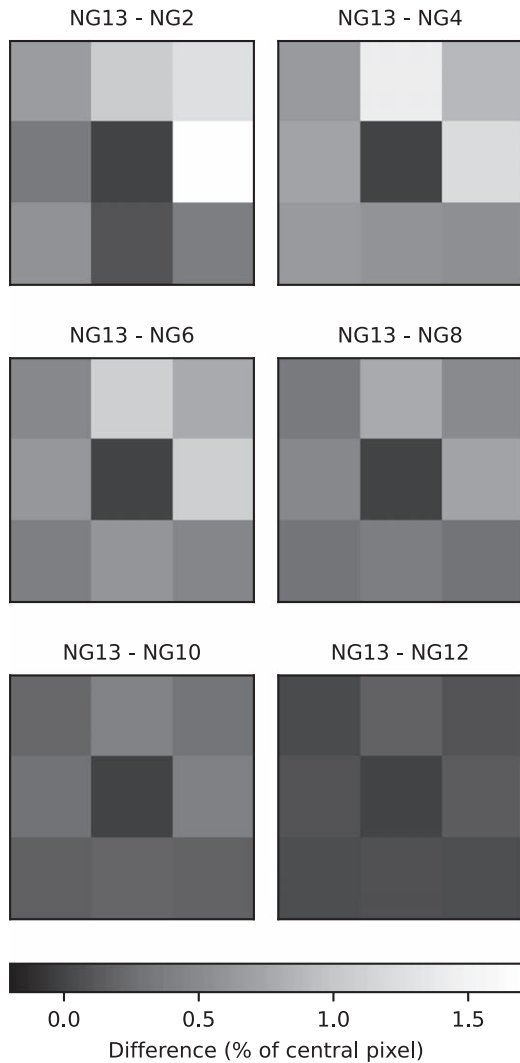
Figure 11. Changes in HIP 65426 PSF shape as a function of the maximum number of reads up the ramp. Individual panels show the difference between the central 9 pixels in the average images for the full 13-group HIP 65426 reduction and reductions with 2, 4, 6, 8, 10, and 12 maximum groups. All images are normalized to the central brightest pixel before differencing, in order to capture changes in the shape of the PSF. Comparison of the different panels shows that charge migration causes ~1% level changes in the central brightest region of the PSF.

Figure 11 shows the differences between the $n_{g,\mathrm{max}} =$ 2, 4, 6, 8, 10, 12 reductions and the full $n_{g,\mathrm{max}} = 13$ HIP 65426 reduction. The results show that the pixels up and to the right of center are ~0.7%–1.7% brighter in the $n_{g,\mathrm{max}} = 13$ reduction compared to the reductions with fewer groups. This is consistent with the charge migration results in Figure 9, which shows charge accumulation increasing in those pixels and decreasing in the central pixel as the reads up the ramp increase. The differences become smaller as the $n_{g,\mathrm{max}}$ values approach 13, which is also consistent with this scenario since adjacent group pairs will have the most similar charge accumulation properties.

5.6. Charge Migration Systematic Errors and Self-calibration Tests

We perform further self-calibration tests to explore whether the signal-dependent PSF changes in Figure 11 could cause the calibration errors and contrast degradation in Section 5.3. We

perform 11 different calibrations of the 13-group HIP 65426 data set, each one using an HIP 65426 reduction with a smaller number of groups (2–12) to create a calibration reference. We then compare the resulting closure phase scatter to theoretical noise predictions, taking into account the smaller number of collected photons and higher read noise for the fewer-group HIP 65426 reductions when evaluating Equation (2). We also calculate the achievable contrast of each calibrated HIP 65426 data set and the predicted achievable contrast. We note that comparing the self-calibration quality as a function of number of groups is valuable for isolating the effects of charge migration, since factors that normally change between a science target and PSF calibrator observation (such as target acquisition) are held constant.

Figure 12 shows the results. As the number of groups in the calibration reference decreases, the ratio between the observed scatter and the theoretical prediction increases. The difference between the achievable contrast and the predicted contrast also becomes more drastic as the number of calibration groups decreases. The nine-group HIP 65426 reduction has the most similar peak counts to the HD 116084 reference star, at ~8597 DN compared to ~8570 DN. Calibrating the 13-group HIP 65426 data set with its own nine-group reduction is thus a representative case for understanding the charge migration contributions when calibrating with HD 116084, since the severity of charge migration depends on the number of accumulated counts. The contrast curve when the nine-group data set is used as a reference is nearly identical to that observed for HIP 65426 calibrated with HD 116084, with a maximum achievable contrast of ~7.2 mag compared to ~7 mag.

The 12-group calibration is also significantly worse than the 13-group self-calibration in Figure 7, with an achievable contrast of ~7.6 mag compared to ~9.5 mag. This makes sense given that a ~7.6 mag companion at λ/D would correspond to a ~0.1% flux ratio at a distance of <2 pixels. The differences in pixel fractional fluxes between the 12- and 13-group reductions are comparable to this flux ratio at these separations (Figure 11). The achievable contrast results when the fewer-group reductions are used, as calibrators behave similarly to this case. For example, the achievable contrast when the two-group HIP 65426 reduction is used as a reference for the 13-group reduction, of ~4.5 mag, is comparable to the fractional flux differences between those two reductions ($\gtrsim 1\%$) shown in Figure 11.

One caveat for interpreting Figure 12 is that, as shown in Figure 8, the few-group reductions have inflated random errors compared to the predictions of Equation (2). This implies that any degradation in calibration quality when few-group reductions of HIP 65426 are used as references could be partially caused by those inflated random errors. However, even when this inflation is taken into account, the results in Figures 8 and 12 show that charge migration still dominates. If there were no charge migration, then the ratio of the observed to predicted scatter for an n_g-group reference in Figure 12 should be equal to the 13-group ratio from Figure 8 added in quadrature to the n_g-group ratio from Figure 8. This is because Figure 8 isolates the increased random errors as a function of groups up the ramp, while Figure 12 includes contributions from both inflated random noise levels and calibration errors due to charge migration. The fact that the noise ratios shown in Figure 12 are significantly higher than those in Figure 8 shows
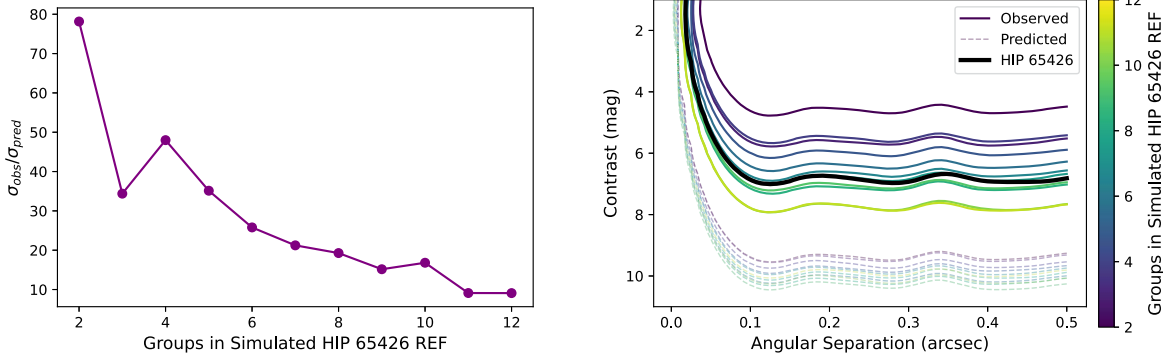
Figure 12. Self-calibration charge migration tests. The left panel shows the ratio between the observed closure phase scatter and the prediction for photon noise of the HIP 65426 13-group data set calibrated using HIP 65426 data sets with varying numbers of groups. The right panel shows the resulting achievable contrast for each calibration (solid lines with color scale), the predicted achievable contrast (dashed lines with color scale), and the HIP 65426 contrast when calibrated with HD 116084 (thick black line).

that charge migration is the dominant source of scatter in the calibrated closure phases for this test.

6. Discussion

The statistical tests carried out here demonstrate that the underperformance of the HIP 65426 NIRISS AMI data set (relative to photon noise with $V = 1$) results from a combination of inflated random and calibration errors. The random visibility phase errors are larger than theoretical predictions and increase systematically with baseline length, consistent with a drop in raw visibility amplitude. This suggests that the perfect-Strehl ($V = 1$) assumption in predicting photon-noise-limited phase scatter is overly optimistic. However, this is not the only source of inflated random errors, since the relative inflation varies from object to object in this study (Figure 5). This may be due to increased levels of changing detector systematics for the objects where fewer numbers of groups were used (e.g., Figure 1). Indeed, the behavior of the complex visibility phase scatter best matches the predictions given the raw visibility amplitudes for HIP 65426, where 13 groups were used per integration.

The self-calibration tests demonstrate that systematic differences between HIP 65426 and the PSF stars contribute significantly to the calibration errors in Figure 6. While one obvious and exciting source of systematic PSF differences is intrinsic phase signals (such as companions), several lines of evidence show that this is not the case here. As described in Ray et al. (2023b), no global χ^2 minimum exists for a single-companion fit to HIP 65426 calibrated with HD 116084. This is also the case when HIP 65426 is calibrated with HD 115842, and the best-fit single-companion model has different parameters from the other calibration.

Furthermore, neither HD 115842 nor HD 116084 had significant single-companion signals in SPHERE AMI vetting data (Ray et al. 2023b). The NIRISS observations of the two objects calibrated against one another result in high closure phase scatter (Section 5.3) with no global single-companion χ^2 minimum. The fact that the reference stars do not calibrate well against each other also argues against spectral type mismatches as the source of the HIP 65426 calibration errors. HD 115842 and HD 116084 are spectral type B0.5 and B2, respectively, while HIP 65426 is spectral type A2. If spectral type mismatches were to blame, the scatter in the HIP 65426 closure phases calibrated with HD 116084 (0°.14) should be higher than the scatter of the two reference stars calibrated against each other (0°.25).

The jackknife tests in Section 5.6 show that differing amounts of charge migration can account for the HIP 65426 calibration errors (e.g., Figure 12). Differences of ∼0.5%–1% in the central 9 pixels of the PSF (compared to the peak pixel) lead to a factor of ∼10–20 inflation compared to photon-noise expectations. This produces achievable contrast nearly identical to that for HIP 65426 calibrated using HD 116084. Furthermore, the fact that the nine-group HIP 65426 data set is just as bad a calibration reference as the three-group HD 116084 data set suggests that charge migration dominates over differing detector systematics associated with groups up the ramp. Indeed, the severity of the calibration errors when the 12-group HIP 65426 data set is used as a reference may indicate that charge migration is the dominant noise source when PSF references are approximately within a spectral type and are observed close in time to the science target.

7. Conclusions and Recommendations for Future NIRISS AMI Programs

We conducted a systematic investigation of the noise properties and performance of the JWST/NIRISS aperture masking interferometry mode in the F380M filter. The results showed that both the random and calibration errors are inflated compared to photon-noise predictions. The random error inflation as a function of baseline length also showed that theoretical noise predictions assuming $V = 1$ are overly optimistic. Furthermore, the self-calibration tests demonstrated that differences in charge migration between the science target and PSF references can account for the observed calibration errors and achievable contrast in the HIP 65426 data set.

Based on this study, and in the absence of sophisticated charge migration calibration strategies, we make the following observation planning recommendations for future NIRISS AMI programs:

1. To mitigate inflated random phase errors that may be caused in part by detector systematics, we recommend maximizing the groups per integration.
2. To mitigate the effects of charge migration, we recommend observing PSF calibrators to a similar well depth to that of the science target.
3. When possible, we recommend selecting PSF calibrators with similar brightnesses to the science target. This would allow for both the peak counts and number of groups to be matched as well as possible between the two data sets.

15

The above recommendations may produce overly stringent requirements on PSF calibrator stars for some science targets. This motivates future studies to quantitatively explore the relative importance of factors such as differing detector readout settings, charge migration levels, calibrator spectral types, and thermal drifts between science targets and calibrators. It also demonstrates the importance of developing sophisticated calibration strategies that can account for such differences between science targets and calibrators. These calibration strategies could be applied in preprocessing steps, such as including a PSF correction taking into account charge migration effects. With enough archival calibrator data sets, they could also be applied directly to the Fourier observables after building a calibration reference from a library of observations (e.g., Kammerer et al. 2019). Such future calibration programs and analyses will work toward the exquisite contrast levels expected from NIRISS AMI, enabling unprecedented thermal-infrared exoplanet science at small angular separations.

Acknowledgments

ORCID iDs

Steph Sallum ⓘ https://orcid.org/0000-0001-6871-6775
Shrishmoy Ray ⓘ https://orcid.org/0000-0003-2259-3911
Jens Kammerer ⓘ https://orcid.org/0000-0003-2769-0438
Alexandra Z. Greebaum ⓘ https://orcid.org/0000-0002-7162-8036
Matthew De Furio ⓘ https://orcid.org/0000-0003-1863-4960
Samuel M. Factor ⓘ https://orcid.org/0000-0002-8332-8516
Michael R. Meyer ⓘ https://orcid.org/0000-0003-1227-3084
Jordan M. Stone ⓘ https://orcid.org/0000-0003-0454-3718
Aarynn Carter ⓘ https://orcid.org/0000-0001-5365-4815
Beth Biller ⓘ https://orcid.org/0000-0003-4614-7035
Sasha Hinkley ⓘ https://orcid.org/0000-0001-8074-2562
Andrew Skemer ⓘ https://orcid.org/0000-0001-6098-3924

Genaro Suárez ⓘ https://orcid.org/0000-0002-2011-4924
Jarron M. Leisenring ⓘ https://orcid.org/0000-0002-0834-6140
Marshall D. Perrin ⓘ https://orcid.org/0000-0002-3191-8151
Adam L. Kraus ⓘ https://orcid.org/0000-0001-9811-568X
Olivier Absil ⓘ https://orcid.org/0000-0002-4006-6237
William O. Balmer ⓘ https://orcid.org/0000-0001-6396-8439
Sarah K. Betti ⓘ https://orcid.org/0000-0002-8667-6428
Anthony Boccaletti ⓘ https://orcid.org/0000-0001-9353-2724
Mariangela Bonavita ⓘ https://orcid.org/0000-0002-7520-8389
Mickael Bonnefoy ⓘ https://orcid.org/0000-0001-5579-5339
Mark Booth ⓘ https://orcid.org/0000-0001-8568-6336
Brendan P. Bowler ⓘ https://orcid.org/0000-0003-2649-2288
Zackery W. Briesemeister ⓘ https://orcid.org/0000-0002-1764-2494
Marta L. Bryan ⓘ https://orcid.org/0000-0002-6076-5967
Per Calissendorff ⓘ https://orcid.org/0000-0002-5335-0616
Faustine Cantalloube ⓘ https://orcid.org/0000-0002-3968-3780
Gael Chauvin ⓘ https://orcid.org/0000-0003-4022-8598
Christine H. Chen ⓘ https://orcid.org/0000-0002-8382-0447
Elodie Choquet ⓘ https://orcid.org/0000-0002-9173-0740
Valentin Christiaens ⓘ https://orcid.org/0000-0002-0101-8814
Gabriele Cugno ⓘ https://orcid.org/0000-0001-7255-3251
Thayne Currie ⓘ https://orcid.org/0000-0002-7405-3119
Camilla Danielski ⓘ https://orcid.org/0000-0002-3729-2663
Trent J. Dupuy ⓘ https://orcid.org/0000-0001-9823-1445
Jacqueline K. Faherty ⓘ https://orcid.org/0000-0001-6251-0573
Michael P. Fitzgerald ⓘ https://orcid.org/0000-0002-0176-8973
Jonathan J. Fortney ⓘ https://orcid.org/0000-0002-9843-4354
Kyle Franson ⓘ https://orcid.org/0000-0003-4557-414X
Julien H. Girard ⓘ https://orcid.org/0000-0001-8627-0404
Eileen C. Gonzales ⓘ https://orcid.org/0000-0003-4636-6676
Thomas Henning ⓘ https://orcid.org/0000-0002-1493-300X
Dean C. Hines ⓘ https://orcid.org/0000-0003-4653-6161
Kielan K. W. Hoch ⓘ https://orcid.org/0000-0002-9803-8255
Callie E. Hood ⓘ https://orcid.org/0000-0003-1150-7889
Alex R. Howe ⓘ https://orcid.org/0000-0002-4884-7150
Markus Janson ⓘ https://orcid.org/0000-0001-8345-593X
Paul Kalas ⓘ https://orcid.org/0000-0002-6221-5360
Grant M. Kennedy ⓘ https://orcid.org/0000-0001-6831-7547
Matthew A. Kenworthy ⓘ https://orcid.org/0000-0002-7064-8270
Pierre Kervella ⓘ https://orcid.org/0000-0003-0626-1749
Daniel Kitzmann ⓘ https://orcid.org/0000-0003-4269-3311
Masayuki Kuzuhara ⓘ https://orcid.org/0000-0002-4677-9182
Kellen Lawson ⓘ https://orcid.org/0000-0002-6964-8732
Cecilia Lazzoni ⓘ https://orcid.org/0000-0001-7819-9003
Ben W. P. Lew ⓘ https://orcid.org/0000-0003-1487-6452
Michael C. Liu ⓘ https://orcid.org/0000-0003-2232-7664
Pengyu Liu ⓘ https://orcid.org/0000-0001-7047-0874
Jorge Llop-Sayson ⓘ https://orcid.org/0000-0002-3414-784X
Anna Lueber ⓘ https://orcid.org/0000-0001-6960-0256
Bruce Macintosh ⓘ https://orcid.org/0000-0003-1212-7538
Elena Manjavacas ⓘ https://orcid.org/0000-0003-0192-6887
Sebastian Marino ⓘ https://orcid.org/0000-0002-5352-2924
Mark S. Marley ⓘ https://orcid.org/0000-0002-5251-2943

Christian Marois https://orcid.org/0000-0002-4164-4182
Raquel A. Martinez https://orcid.org/0000-0001-6301-896X
Brenda C. Matthews https://orcid.org/0000-0003-3017-9577
Elisabeth C. Matthews https://orcid.org/0000-0003-0593-1560
Dimitri Mawet https://orcid.org/0000-0002-8895-4735
Johan Mazoyer https://orcid.org/0000-0002-9133-3091
Michael W. McElwain https://orcid.org/0000-0003-0241-8956
Stanimir Metchev https://orcid.org/0000-0003-3050-8203
Brittany E. Miles https://orcid.org/0000-0002-5500-4602
Maxwell A. Millar-Blanchaer https://orcid.org/0000-0001-6205-9233
Paul Molliere https://orcid.org/0000-0003-4096-7067
Sarah E. Moran https://orcid.org/0000-0002-6721-3284
Caroline V. Morley https://orcid.org/0000-0002-4404-0456
Sagnick Mukherjee https://orcid.org/0000-0003-1622-1302
Paulina Palma-Bifani https://orcid.org/0000-0002-6217-6867
Polychronis Patapis https://orcid.org/0000-0001-8718-3732
Simon Petrus https://orcid.org/0000-0003-0331-3654
Sascha P. Quanz https://orcid.org/0000-0003-3829-7412
Isabel Rebollido https://orcid.org/0000-0002-4388-6417
Jea Adams Redai https://orcid.org/0000-0002-4489-3168
Bin B. Ren https://orcid.org/0000-0003-1698-9696
Emily Rickman https://orcid.org/0000-0003-4203-9715
Matthias Samland https://orcid.org/0000-0001-9992-4067
B. A. Sargent https://orcid.org/0000-0001-9855-8261
Joshua E. Schlieder https://orcid.org/0000-0001-5347-7062
Karl R. Stapelfeldt https://orcid.org/0000-0002-2805-7338
Ben J. Sutlieff https://orcid.org/0000-0002-9962-132X
Motohide Tamura https://orcid.org/0000-0002-6510-0681
Xianyu Tan https://orcid.org/0000-0003-2278-6932
Christopher A. Theissen https://orcid.org/0000-0002-9807-5435
Taichi Uyama https://orcid.org/0000-0002-6879-3030
Arthur Vigan https://orcid.org/0000-0002-5902-7828
Malavika Vasist https://orcid.org/0000-0002-4511-3602
Johanna M. Vos https://orcid.org/0000-0003-0489-1528
Kevin Wagner https://orcid.org/0000-0002-4309-6343
Jason J. Wang https://orcid.org/0000-0003-0774-6502
Kimberly Ward-Duong https://orcid.org/0000-0002-4479-8291

Niall Whiteford https://orcid.org/0000-0001-8818-1544
Schuyler G. Wolff https://orcid.org/0000-0002-9977-8255
Kadin Worthen https://orcid.org/0000-0002-5885-5779
Mark C. Wyatt https://orcid.org/0000-0001-9064-5598
Marie Ygouf https://orcid.org/0000-0001-7591-2731
Xi Zhang https://orcid.org/0000-0002-8706-6963
Keming Zhang https://orcid.org/0000-0002-9870-5695
Zhoujian Zhang https://orcid.org/0000-0002-3726-4881
Yifan Zhou https://orcid.org/0000-0003-2969-6040
Alice Zurlo https://orcid.org/0000-0002-5903-8316

References

Baldwin, J. E., Haniff, C. A., Mackay, C. D., & Warner, P. J. 1986, Natur, 320, 595
Biller, B., Lacour, S., Juhász, A., et al. 2012, ApJL, 753, L38
Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2022, JWST Calibration Pipeline, v1.7.0, Zenodo, doi:10.5281/zenodo.7059013
Gardner, J. P., Mather, J. C., Abbott, R., et al. 2023, PASP, 135, 068001
Gardner, J. P., Mather, J. C., Clampin, M., et al. 2006, SSRv, 123, 485
Hinkley, S., Carter, A. L., Ray, S., et al. 2022, PASP, 134, 095003
Hinkley, S., Kraus, A. L., Ireland, M. J., et al. 2015, ApJL, 806, L9
Hirata, C. M., & Choi, A. 2020, PASP, 132, 014501
Ireland, M. J. 2013, MNRAS, 433, 1718
Ireland, M. J., & Kraus, A. L. 2008, ApJL, 678, L59
Jennison, R. C. 1958, MNRAS, 118, 276
Kammerer, J., Girard, J., Carter, A. L., et al. 2022, Proc. SPIE, 12180, 121803N
Kammerer, J., Ireland, M. J., Martinache, F., & Girard, J. H. 2019, MNRAS, 486, 639
Monnier, J. D. 1999, PhD thesis, Univ. of California, Berkeley
Plazas, A. A., Shapiro, C., Smith, R., Huff, E., & Rhodes, J. 2018, PASP, 130, 065004
Quenouille, M. H. 1949, Ann. Math. Stat., 20, 355
Rauscher, B. J., Arendt, R. G., Fixsen, D. J., et al. 2017, PASP, 129, 105003
Ray, S., Hinkley, S., Sallum, S., et al. 2023a, MNRAS, 519, 2718
Ray, S., Sallum, S., Hinkley, S., et al. 2023b, arXiv:2310.11508
Sallum, S., & Eisner, J. 2017, ApJS, 233, 9
Sallum, S., Eisner, J., Skemer, A., & Murray-Clay, R. 2023, ApJ, 953, 55
Sallum, S., Eisner, J. A., Stone, J. M., et al. 2021, AJ, 161, 28
Sallum, S., Ray, S., & Hinkley, S. 2022, Proc. SPIE, 12183, 121832M
Sallum, S., & Skemer, A. 2019, JATIS, 5, 018001
Sallum, S., Skemer, A. J., Eisner, J. A., et al. 2019, ApJ, 883, 100
Schlawin, E., Leisenring, J., Misselt, K., et al. 2020, AJ, 160, 231
Sivaramakrishnan, A., Lafrenière, D., Ford, K. E. S., et al. 2012, Proc. SPIE, 8442, 84422S
Sivaramakrishnan, A., Tuthill, P., Lloyd, J. P., et al. 2023, PASP, 135, 015003
Soulain, A., Sivaramakrishnan, A., Tuthill, P., et al. 2020, Proc. SPIE, 11446, 1144611
Tuthill, P. G., Monnier, J. D., & Danchi, W. C. 2001, Natur, 409, 1012
Tuthill, P. G., Monnier, J. D., Danchi, W. C., Wishnow, E. H., & Haniff, C. A. 2000, PASP, 112, 555