Robust Deep Contrast Imaging with Self-Calibrating Coronagraph Systems

Milestones White Paper

Milestone #1

High contrast self-calibration: 10x contrast gain to 1e-9 raw contrast

Team:
Olivier Guyon (PI)
Niyati Desai, Julien Lozi, Jared Males, Phillip K. Poon, Susan Redmond, Garreth Ruane, Kent Wallace, Alex Walter

May 20, 2025

Approvals

Released by:

E-SIGNED by Olivier Guyon
on 2025-05-21 23:02:17 GMT

, 2025

Olivier Guyon | Date
Principal Investigator, University of Arizona & Subaru Telescope

Approved by:

E-SIGNED by Nicholas Siegler
on 2025-05-21 17:13:15 GMT

Nicholas Siegler | Date
Exoplanet Exploration Program Chief Technologist, NASA JPL

E-SIGNED by Brendan Crill
on 2025-05-21 15:40:58 GMT

Brendan Crill | Date
Exoplanet Exploration Program Deputy Chief Technologist, NASA JPL

E-SIGNED by Lucas Paganini
on 2025-05-21 19:25:06 GMT

Lucas Paganini | Date
Exoplanet Exploration Program Executive, NASA HQ

1. Introduction, Background

This milestone is aimed at advancing the self-calibrating high contrast imaging (SC-HCI) concept. In a SC-HCI system, unwanted starlight is accurately and precisely measured, so that it can then be reliably identified and subtracted from science data (images or spectra) to unambiguously reveal the faint planet light. The technique aims at achieving detection limits set by photon noise instead of speckle noise.
To realize this goal, the HCI system must be designed to support wavefront sensing using starlight, operating simultaneously with science data acquisition.
The self-calibrating property will benefit exoplanet-imaging missions in four interconnected ways:

1. Provides improved detection limits, approaching or reaching the photon-noise limit.
2. Offers immunity from speckle noise which often sets the detection limit of HCI systems.
3. Improves observing efficiency by relaxing the need for off-target PSF reference acquisitions. Does not require science data acquisition interruptions for wavefront control.
4. Can relax telescope stability and instrument raw contrast requirements.

The photon noise limit may not necessarily be due to residual stellar light, as other terms (exozodi light especially) may dominate in some cases. Ensuring that residual stellar light only contributes as photon noise would enable optimizations associated with increased stellar light, such as broader spectral band, smaller inner-working angle and higher coronagraph throughput.

In the first milestone of this project, we seek to demonstrate post-processing calibration of high contrast images with a >10x gain over conventional PSF subtraction approaches.

2. Milestone Description and Rationale

The milestone definition is as follows:

Demonstrate a >10x gain in contrast, over standard PSF subtraction, by post-processing self-calibration using an out-of-band wavefront sensor, reaching a post-processed contrast level below 1e-9.

Rationale: The 10x ratio will bring a significant gain in exoplanet imaging, and can also relax some of the most challenging optics stability requirements imposed on high contrast imaging systems. Demonstrating that residual speckles can be calibrated and subtracted will relax the mission raw contrast requirement, allowing for the tradeoff between the starlight suppression performance characteristics (raw contrast, inner-working angle, sensitivity to low-order aberration, throughput, spectral bandwidth) to be revisited. For example, maintaining deep raw contrast becomes challenging in broader spectral bands, so relaxing the raw contrast requirement allows for a wider spectral band to be used for science acquisition.

2.1. Relevance to Exoplanet Imaging and Spectroscopy

In a Sun-Earth system analog, the planet-to-star contrast in reflected light flux is ~1.5e-10. Achieving this deep contrast remains very challenging, placing tight requirements on wavefront stability. Two noise terms can set the contrast detection limit: photon noise and speckle noise. Reliable high signal-to-noise ratio requires pushing these noise terms ~10x below the planet flux level. Photon-noise, being temporally uncorrelated, is reduced by increasing the total exposure time. Speckle noise, originating from wavefront error (static and dynamic) can be mitigated by either wavefront control, or by calibration. This milestone will focus on the latter approach, also referred to as post-processing.

The most commonly used post-processing techniques are to compare the science frame(s) with PSF reference frame(s); we refer to this approach as standard PSF subtraction in this document (see Reference subtraction section for algorithm description). The reference PSF can be either acquired on a target other than the science target (reference differential imaging - RDI), or can be the science target observed at a different rotation angle (angular differential imaging - ADI). In both approaches, the post-processing residual is limited by wavefront variations between the science and reference frames, placing tight constraints on wavefront stability.

In this milestone, we will demonstrate a 10x gain in post-processing residual compared to standard PSF subtraction, in the regime where wavefront changes are setting the contrast limit in post-processing residuals. The comparison between the two approaches will be done by using the same dataset for both techniques.

In this milestone, we will reach a 1e-9 post-processed contrast residual and focus on demonstrating the algorithm. The 1e-9 residual is approximately two orders of magnitude short of the ~1e-11 residual noise level required for observing Earth-like planets around Sun-like stars. We will push the contrast residual toward 1e-11 in subsequent milestones.

2.2. Milestone requirements

Scoring region
Contrast shall be evaluated in the ***scoring region***, an area of at least 10 sq-λ/D. This will keep experimental runs short, as the corresponding number of modes (~20) can be fully probed with a moderate number of wavefront realizations.

Wavelength range
The wavelength shall be within the 400 nm to 1700 nm range, in accordance with testbed capabilities and future envisioned coronagraph instruments on space telescope(s). The primary experimental setup for this milestone will operate in visible light (<780nm) in vacuum, with supporting in-air testbed extending to the nearIR (up to 1700 nm).

Wavefront aberrations will be injected
Each frame corresponds to an independent, static injected wavefront error (WFE), and consists of a high contrast (HC) frame and a wavefront sensor (WFS) frame (both with the same injected WFE). A frame may be a coadd of multiple camera reads, as needed to reach the required measurement SNR.
Each injected WFE realization will be a random linear superposition of WF modes. The modes shall provide full coverage of spatial frequencies within the scoring region, in both phase and amplitude. With a 10 sq-λ/D scoring region, there will be 20 independent modes (one cosine and one sine per spatial frequency).

Number of frames = 100 or more
The observation set shall consist of at least 100 consecutive frames. The WFS and HC data acquisition cameras/sensors may be running asynchronously, with the WFS camera acquiring data at higher frequency than the HC camera. In this case, a frame corresponds to a time interval defined by one or a few HC camera exposures; the HC frame is the average of these exposures, and the corresponding WFS frame will be constructed by averaging all WFS camera exposures acquired over this time interval.

Number of successful demonstrations = 3 or more
The milestone requirements shall be met on at least 3 separate datasets.

Photometric efficiency
Data calibration by reference subtraction or other techniques incurs a cost in **photometric efficiency** (defined later in this whitepaper). There is a tradeoff between optimizing calibration accuracy and preserving photometric efficiency. We require >25% photometric efficiency, equivalent to a < 2x increase in noise level.

2.3. Definitions and Framework

The milestone's goal is to demonstrate that **self-calibration** brings a >10x contrast improvement over a standard **reference subtraction** approach.
Notations, definitions and data processing framework are summarized in Fig. 1 and used throughout this document.

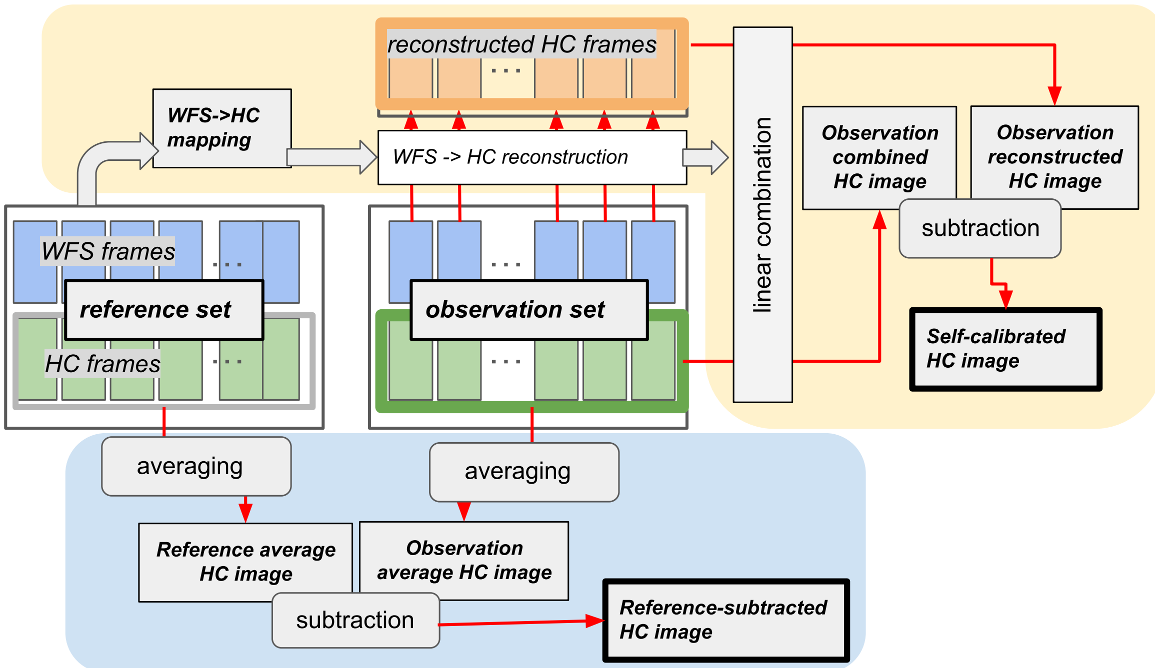


Figure 1: Data collection and analysis workflow. The input data collected for the experiment consists of synchronized wavefront sensing (WFS) and high contrast (HC) frames, shown near the center of the figure respectively as blue and green rectangles. Frames are subdivided in a reference set for calibration and an observation set emulating the astrophysical observation. For clarity, the figure considers that reference set frames are acquired before the observation set, but the reference set may consist of frames acquired both before and after the observation set. Two separate processing tasks will be performed and compared for the purpose of this milestone: reference subtraction (bottom part of the figure, light blue) and self-calibration (top & right part, light orange). See text for details.

2.3.1. Input data

- A ***frame*** refers to all data taken during a time interval over which WF aberrations are considered static. Frame numbers start at 0 and are incremented over time.
- A high contrast ***HC frame*** is a measured focal plane high contrast image corresponding to a single WF state (no WF change introduced during frame).
- A wavefront sensing ***WFS frame*** is the corresponding wavefront sensor image, ideally acquired simultaneously with the science frame.
- The ***observation set*** consists of at least 100 consecutive pairs of HC/WFS frames.
- The ***reference set*** consists of frames that are distinct from the observation set, and consist of an arbitrary number of frames.

Example input data measurement sequence:
300 consecutive frames are acquired, each consisting of a science frame and a WFS frame. Frames 0-99 and 200-299 are the reference set, and frames 100-199 are the observation set. Splitting the reference set in two continuous sequences, one before and one after the observation set, helps ensure that long-term drifts are adequately captured.

2.3.2. Reference subtraction

The reference subtraction approach (lower part of figure 1) does not use WFS information and serves as a baseline against which the contrast improvement is measured. It is similar to reference differential imaging (RDI) approaches commonly used in high contrast imaging.
We define reference subtraction as follows:

- All HC frames within the reference set are averaged. The result is the ***reference HC image***
- All HC frames within the observation set are averaged. The result is the ***observation HC image***
- The ***reference HC image*** is subtracted from the ***observation HC image***. The result is the ***reference-subtracted HC image***

2.3.3. Self-calibration

Self-calibration, developed here, uses WFS information to enhance calibration of high contrast images.
The steps to self-calibration (upper/right part of figure 1) are as follows:

- A ***mapping function*** from WFS frames to HC frames is derived (learned) from the reference set frames. This mapping allows for a WFS frame to be transformed to a corresponding HC frame. The mapping function is analogous to a lookup table, but is in practice encoded in modal space for efficiency (see Fig. 5).
- For each frame in the observation set, the ***reconstructed HC frame*** is computed solely from the WFS frame, by running the mapping function on the WFC frame. The mapping function also returns uncertainties (modal error bars).
- The ***observation reconstructed HC image*** is computed as a linear combination of reconstructed HC frames. The coefficients to the linear combination are derived from reconstruction uncertainties, so that poorly constrained information is weighted down.
- The ***observation combined HC image*** is computed as the same linear combination applied to the observation set HC frames
- The observation reconstructed HC image is subtracted from the observation HC combined image. The result is the **self-calibrated HC image**.

2.3.4. Contrast measurements

- The ***frame raw contrast value*** (scalar for each frame) is the average surface brightness in the scoring area of the HC frame, normalized to the unocculted PSF core surface brightness.

The following two contrast quantities are measured on the post-processed images, which are meant to be free of starlight. These are root-mean square (RMS) quantities as they measure the a residual starlight component that could be zero-mean:

- The ***self-calibrated HC image contrast residual value*** (scalar) is the root-mean square (RMS) of the self-calibrated HC image, computed over the scoring area.

- The ***reference-subtracted HC image contrast residual value*** (scalar) is the root-mean square (RMS) of the reference-subtracted HC image, computed over the scoring area.

The following gain quantity is derived from the contrast residual values:

- The ***self-calibration post-processing gain*** (scalar) is the ratio of the two quantities above. It is >1 if self-calibration outperforms reference subtraction.

<u>For the milestone to be achieved, the self-calibrated HC image contrast residual value must be <1e-9, and the self-calibration post-processing gain must be >10x.</u>

2.3.5. Photometric efficiency

Calibration techniques involving image subtraction can amplify the photon and readout noise terms beyond the noise level expected by simply averaging all HC frames in the observation set.

To illustrate this effect, we consider a weighted average of a set of N frames I_0, I_1, … I_{N-1}:

$I_s = a_0\, I_0 + a_1\, I_1 + \ldots\, a_{N-1}\, I_{N-1}$

With:

$a_0 + a_1 + \ldots\, a_{N-1} = 1$

We assume that each frame I_i has the same noise level $\boldsymbol{\sigma}$ (standard deviation).

The noise level in the average image is:

$\boldsymbol{\sigma}_s = C\, \boldsymbol{\sigma}$

with: $C = (a_0^2 + a_1^2 + \ldots\, a_{N-1}^2)^{1/2}$

C is minimal and equal to $C_{min}=1/N^{1/2}$ for:

$a_0 = a_1 = \ldots = a_{N-1} = 1/N$

corresponding to a straight average of all input frames (equal weights). Non-equal weights yield $C>C_{min}$ (noise amplification).

This effect is quantified here as a ***photometric efficiency*** $\boldsymbol{(C_{min}/C)^2}$, which measures the equivalent fractional efficiency in exposure time. A 10% photometric efficiency means that the final (after processing) noise level is equivalent to averaging only 10% of the observation time.

The photometric efficiency is optimal with the reference subtraction approach, for which it is 50% if the reference set has the same number of frames as the observation set (noise is amplified by sqrt(2) by the subtraction, equivalent to halfing the observation time).
Photon and readout noise will be propagated through the self-calibration algorithm to compute its photometric efficiency.

2.3.6. HC/WFS frame synchronization

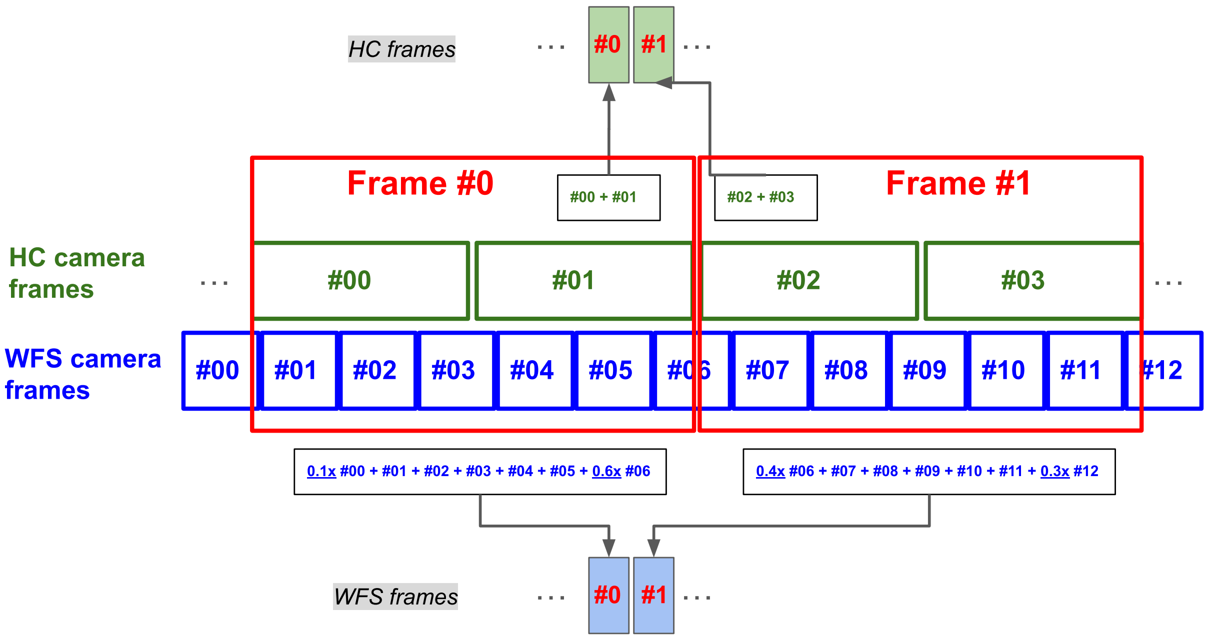


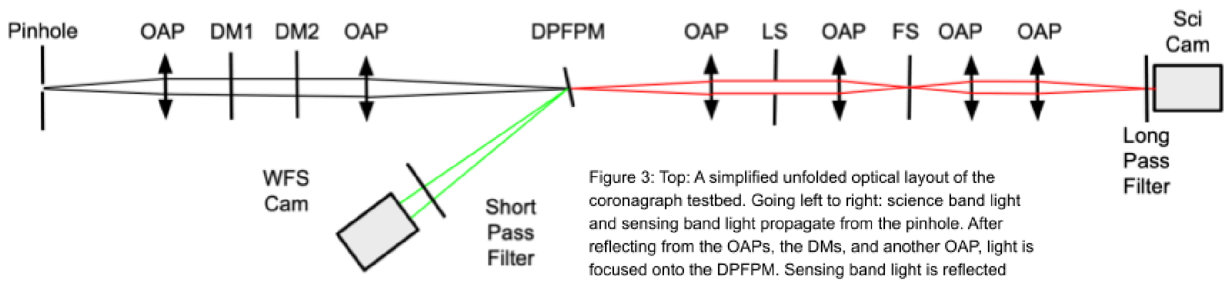
Figure 2. Synchronization between high contrast (HC) and wavefront sensing (WFS) data is performed at the frame level. Frames are defined as consecutive time intervals within which raw data (camera frames) are averaged. In this example, the HC camera framerate is slower than the WFS camera framerate, and each frame consists of two HC camera frames. WFS camera frames are averaged to the corresponding time intervals. If a WFS camera frame spans across two frames, its signal is split between the frames according to the fractional time spent in each frame.

Our proposed approach requires pairs of HC and WFS frames corresponding to the same wavefront error. This is ensured by synchronization between the HC and WFS frames. The raw camera frames are not assumed to be synchronized, and the HC camera operates at a slower frame rate than the WFS camera due to significantly lower flux level. The first step to our algorithm is to create synchronized HC/WFS frames from non-synchronized HC/WFS raw camera images, as described in Fig. 2. This is done by defining each frame as a time interval consisting of a number of consecutive HC camera frames (2 for the example shown in Figure 2), and binning WFS camera frames according to this timing. WFS camera frames falling in between time intervals may be split between the two frames (as shown in Figure 2), or allocated to the nearest frame (this solution is acceptable if the WFS camera exposure is significantly shorter than the frame duration).

3. Experimental Setup

This milestone will be demonstrated at the High Contrast Imaging Testbed Facility (HCIT) at the NASA Jet Propulsion Laboratory (JPL), on a coronagraph testbed which will use a Dual-Purpose Focal Plane Mask (DPFPM). The DPFPM combines the ability to perform high contrast imaging using a Lyot-style focal plane mask (FPM) and wavefront sensing using a Zernike wavefront sensor (Ruane et al, 2020, Wallace et al, 2023). The DPFPM enables both without the need to physically insert or remove masks into or out-of the optical path when performing either task, allowing each task to be performed asynchronously.

A simplified and unfolded optical path of a coronagraph with a DPFPM is shown in Figure 3. Coherent light is focused onto a pinhole, which acts as a simulated star, then propagates to a series of off-axis parabolas (OAPs). The first OAP forms a pupil plane where the first deformable mirror (DM) is located. A second DM is located shortly after the first DM at a non-pupil plane. The fourth OAP then focuses the light on the DPFPM.



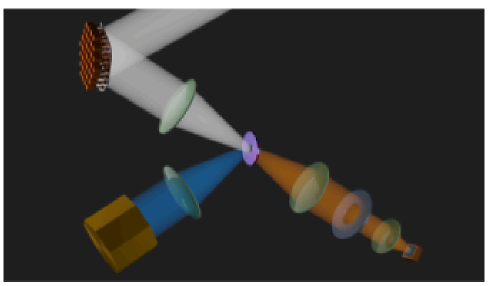


Figure 3: Top: A simplified unfolded optical layout of the coronagraph testbed. Going left to right: science band light and sensing band light propagate from the pinhole. After reflecting from the OAPs, the DMs, and another OAP, light is focused onto the DPFPM. Sensing band light is reflected from the dichroic of the DPFPM onto a WFS camera. Any residual science band that is not occulted by the Lyot-style FPM travels passes through the dichroic of the DPFPM. Residual science band light that passes through the FPM substrate is then collimated and blocked by a circular Lyot Stop (LS) at a pupil plane. Any remaining light passes through a field stop (FS) located at the focus following the Lyot stop. This field stop (FS) blocks stray and rejected light that exists outside of the contrast region-of-interest called the "dark hole". This is then imaged onto the science camera. Spectral filters are used to prevent cross-talk between the two bands.
Left: 3D rendering of the area around the DPFPM (center).

The DPFPM design, shown in Figure 4, uses a dichroic coating on the focal plane mask substrate to separate the science band (620-700 nm) and the sensing band light (500-550 nm). Science band light is occulted by the Lyot-style focal plane mask. Any remaining science band light is transmitted by the DPFPM and propagates to a series of off-axis parabolas, a Lyot stop, a field stop, and finally is imaged onto the science camera.

Light from the sensing band is reflected from the dichroic coating of the DPFPM. The phase dimple is used to convert wavefront variations in the sensing band into intensity variations in the pupil plane where the wavefront sensing camera is then located (Steves et al, 2020). Thanks to this approach, no additional focal plane wavefront sensing optics is needed.

The Zernike wavefront acts in the focal plane and senses in the pupil plane. In the focal plane are two regions that are phase shifted with respect to one another. The central region has a diameter of roughly 2 lambda/D, and has an optical path - or phase shift - of lambda/4. This is called ‘the dimple’. It is worth noting that this dimple is physically quite small and pancake shaped. Typical dimensions are a diameter of ~ 20 to 30 um, and a depth of ~ 300 to 400 nm. The on-axis light that is focused in this plane (the point spread function or PSF for short) is centered on this dimple. The part of the light that passes through this dimple is phase-shifted and it diffracts as if it is passing through a small pinhole of the same diameter. This light is called the reference beam. The light passing around this dimple is otherwise unaffected. As the light propagates away from the focal plane, the reference beam light expands and starts to overlap the light that passes by the dimple. In the subsequent pupil plane, the beam overlap is nearly complete (some reference light falls outside geometric pupil aperture). Because these beams overlap, and are coherent, they interfere with one another thereby creating an intensity modulation, or fringe pattern. The phase shift on the reference beam makes the response of the fringe intensity straightforward - pupil phase errors that are greater than the mean result in a brighter signal. Likewise, phase errors less than the mean result in a dimmer fainter signal.The spatial frequencies in the pupil plane are only limited by the number of detector pixels across the pupil. For the dual purpose mask, the wavefront sensing light - between 500 and 550 nm - is much shorter than the science band (625 to 700 nm). The dichroic coating on the substrate acts as a mirror in the wavefront sensing band, and it acts like a window for the science light in transmission. Thus, we capture all spatial frequencies that are reflected from the focal plane..

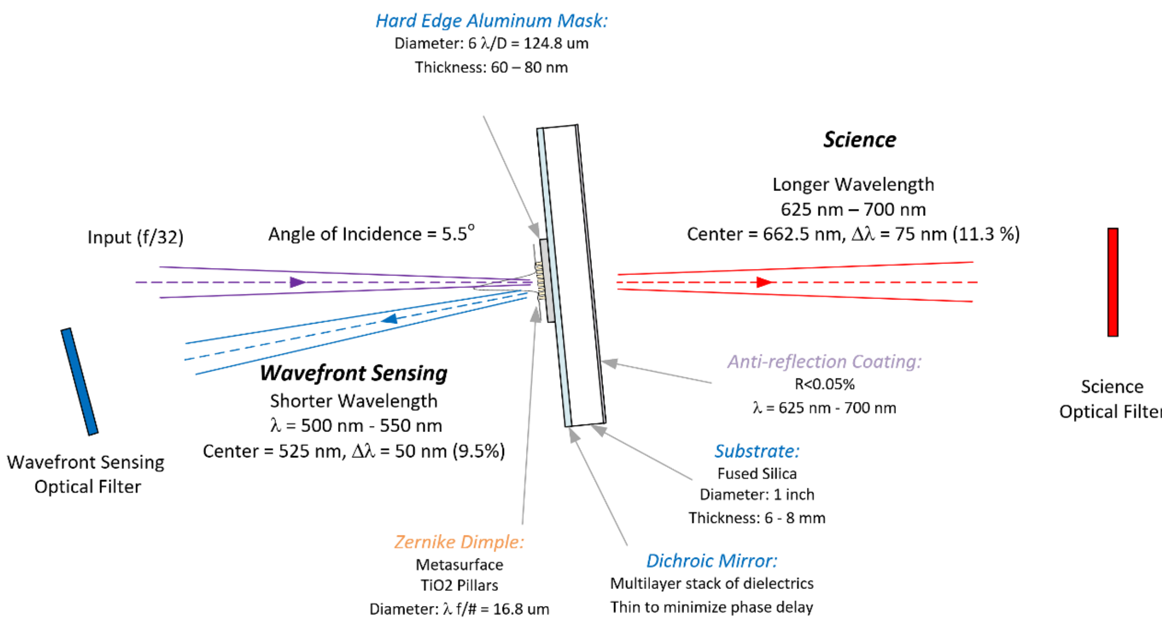


Figure 4: Dual-purpose focal plane mask (DPFPM) design (SAT PI Kent Wallace). Incoming light from the left is split into science (transmitted) and wavefront sensing (reflected) beams, according to wavelength. In transmission at the science wavelength, the DPFPM acts as a r=3 lambda/D coronagraph focal plane mask. In reflection at the wavefront sensing wavelength, it acts as a Zernike-type wavefront sensing focal plane mask and is specifically designed to match the reflectivity and phase shift on reflection between the metal (r < 3 lambda/D) and dichroic (r > 3 lambda/D) regions in the wavelength range used for wavefront sensing; thus the transition has little impact on the wavefront sensing capability of the Zernike sensor. An occultor radius of 3 lambda/D in the high-contrast spectral channel is typical for a Lyot coronagraph; this choice builds on historical success of the previous Lyot coronagraph demonstrations with similar design parameters. The Zernike dimple diameter is lambda/D in the wavefront sensing spectral channel, which is the preferred size for the Zernike sensor.

If we detect any cross-talk between the two bands, we can mitigate it by inserting a longpass filter, which blocks wavelengths less than 600 nm, before the science camera and insert a shortpass filter, which blocks wavelengths greater than 600 nm, before the sensing camera.

We propose a setup to take advantage of DPFPM while eliminating the need for any physical insertion or removal of optics during the science and wavefront sensing exposures. Figure 5 shows the setup which allows both the science and sensing bands to be measured simultaneously and asynchronously. The broadband tunable laser would generate the science band and a dedicated narrowband laser generates light within the sensing band. The science and sensing bands are combined in a wavelength division multiplexer (WDM). Unlike a traditional fiber splitter, a WDM can support both bands with over 50% throughput.

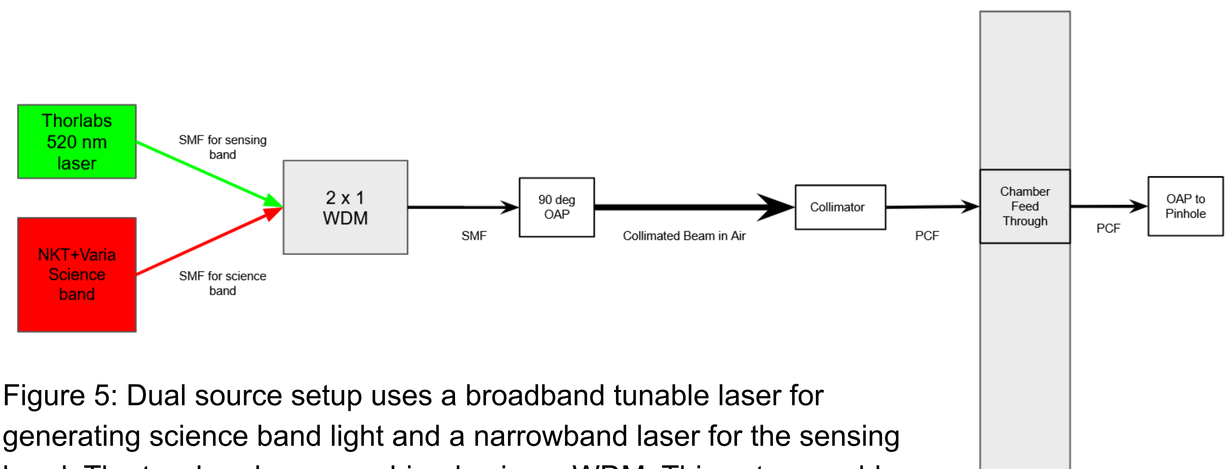


Figure 5: Dual source setup uses a broadband tunable laser for generating science band light and a narrowband laser for the sensing band. The two bands are combined using a WDM. This setup would allow for simultaneous wavefront sensing and science imaging.

By measuring the wavefront and performing high contrast science imaging (or spectroscopy), we ensure no WF evolution between WF measurements and PSF acquisitions. This is the configuration envisioned for on-orbit science data collection, when exposure time is precious, especially for exoplanet spectroscopy. Our demonstration will help define the optimal architecture to spread light between science and WF sensing.

To achieve deep raw contrast, the coronagraph testbed is housed inside a vacuum chamber and placed on vibration isolation devices, called Minus-Ks, to reduce uncontrollable dynamic wavefront error during the process of establishing the region of high contrast. The coronagraph is installed on a carbon fiber optical table and covered in multi-layer insulation to improve thermal stability. Using the DPFPM, the coronagraph testbed has demonstrated (Wallace et al. 2024) raw contrast levels of 10^-9 Normalized Intensity (NI) by using Pairwise Probing (PWP) and Electric Field Conjugation (EFC). While the formal report on these results is still under preparation by the Wallace SAT team, we provide a preview in Fig. 6. The first generation of the Dual Purpose Lyot Coronagraph focal phase mask achieved a mean normalized intensity of 4×10^{-10} in a 10% bandwidth centered at 650nm.

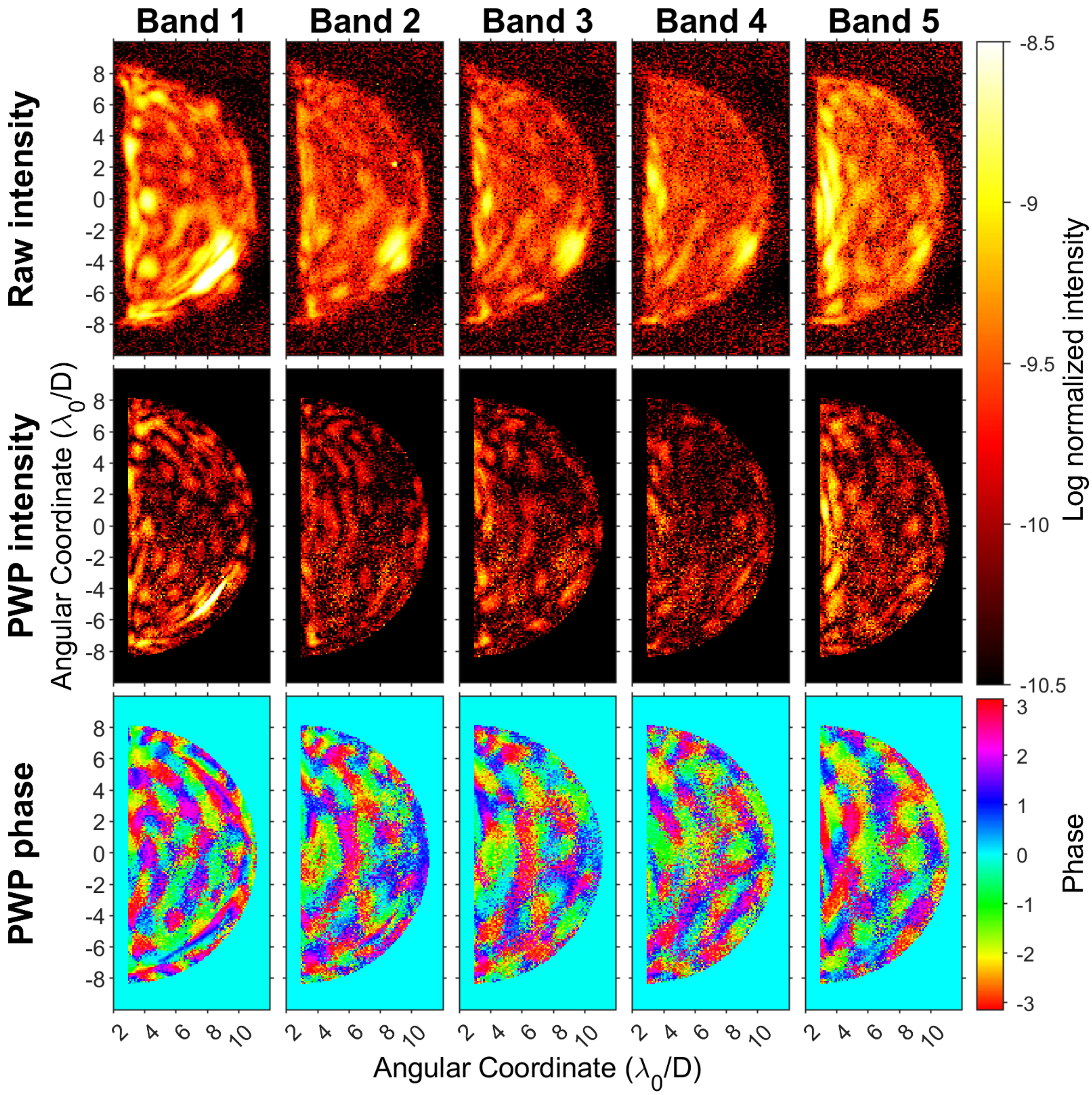


Figure 6. (Top row) The raw normalized intensity in five sub-bands with 2% spectral bandwidth demonstrating a 4e-10 mean raw contrast in a 10% bandwidth. (Middle and bottom rows). The intensity and phase as determined from the pair-wise probing (PWP) focal plane E-field estimation process. The pseudo-star is located at the origin and the optical bench is parallel to the horizontal axis

4. Algorithm Description

The main steps to the algorithm are:

1. Building a mapping function between WFS frames and HC frames from the reference set
2. Computing, for each WFS frame, a reconstructed HC frame
3. Computing the ***reconstructed HC image*** by linear combination of the reconstructed HC frames
4. Computing the ***observation HC image*** by running the exact same linear combination on the input observation HC frames
5. Subtracting the reconstructed HC image from the observation HC image

The two most delicate steps to the algorithm are building the mapping function (step #1) and choosing the linear coefficients for steps #3 and #4. These involve tradeoffs that affect the accuracy, reliability and photometric efficiency.

4.1. Mapping function

The mapping function takes as an input a WFS frame and computes the corresponding estimated HC frame. There are two main concepts behind this step: performing linear-quadratic (linQ) modal extrapolation, and enforcing locality to ensure the model reliability.

4.1.1. Efficient extrapolation with the linQ model

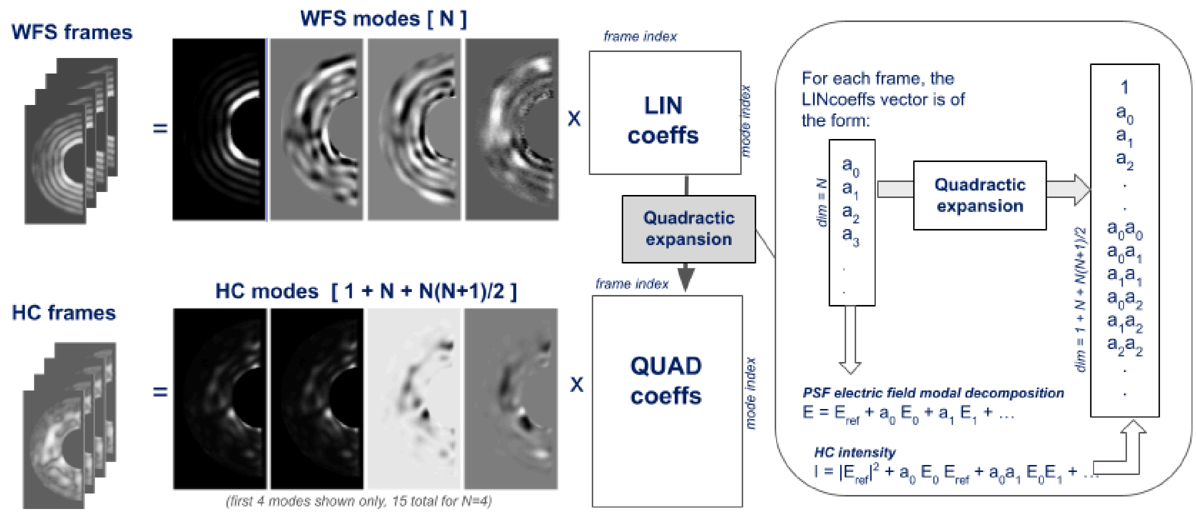


Fig 7: Overview of the linQ modal mapping model. WFS frames (top) are decomposed in WFS modes and their corresponding coefficients (LIN coeffs) for each frame. The linQ model assumes that HC frames are quadratically related to WFS frames, so there exists a linear decomposition of HC frames for which the corresponding coefficients (QUAD coeffs) are a quadratic expansion (right) of the LIN coeffs.

The linQ modal extrapolation (Fig. 7) assumes that WFS frames are responding linearly to small wavefront perturbations, while HC frames respond quadratically to the same wavefront perturbations. The linQ approach was validated in the Linear Wavefront Control PSF calibration milestone demonstration (Guyon et al. 2024).
Its main steps are:

1. Identify dominant modes of variation (principal components) in the set of *WFS frames*. These are the *WFS modes*. Each WFS frame is then represented as a vector of modal coefficients. The full set of reference frames is stored in the *LIN coeffs* matrix.
2. Quadratically expand the WFS frame vectors: append the vector with squares and products of its coefficients. The full set of reference vectors is stored in the *QUAD coeffs* matrix.
3. Solve for the HC modes by multiplication of the *HC frames* with the pseudo-inverse of the *QUAD coeffs* matrix.

.
The mapping function is encoded by the set of input (WFS) modes and the set of the output (HC) modes. Reconstructing a HC frame is done by:

1. Projection of the input WFS frame onto the WFS modal basis. The input WFS frame is then represented as a vector
2. Performing the quadratic expansion of the above WFS vector
3. Expanding the output vector to a reconstructed HC frame by multiplying by the HC modes

The linQ modal extrapolation is efficient: the mapping from WFS to HC frames can be constrained with a relatively small number of frames thanks to the strong linear-quadratic model constraint.

4.1.2. Enforcing Locality

The linear-quadratic relationship underlying the linQ approach is a local approximation. With sufficiently large WF changes, the WFS frame response is no longer linear. Locality constraints must be checked and enforced to prevent unreliable extrapolations beyond the WFS linearity domain.

When processing an observation WFS frame, locality must be enforced by first checking and possibly selecting the reference WFS frames that are most similar (smallest euclidean distance) to the observation WFS frame. The linQ algorithm may then operate on this smaller, but more local, subset of frames. A distance-based clustering algorithm will be run on the set of WFS frames to define the subset of frames to be included in the reconstruction. The subset will be the largest cluster of frames within the WFS linearity domain. It may include all input frames if the input disturbances are small, or may reject outliers. If frames are rejected, the photometric efficiency (included in the milestone definition) will be reduced.

We will explore running independent linQ reconstructions on separate clusters of WFS frames as a way to mitigate loss of WFS linearity. This partitioning step may be required to meet the milestone photometric efficiency requirement.

4.2. Linear Combination & Photometric Efficiency

The reconstructed HC frames are linearly combined to produce the final reconstructed HC image. This linear combination is not an average, as the reconstructed HC frames vary in noise level and reconstruction robustness: the most reliable and lowest noise reconstructed frames are given higher weight. Favoring the most reliable reconstructed frames however comes at a cost in photometric efficiency.

To illustrate this tradeoff, as an example, we can consider a search for a close match between a set of M=1000 reference frames and M=1000 observation frames. An exhaustive search through the 1e6 possible reference/observation pairs would reveal the –most likely excellent– closest match between reference and observation WFS frames. All other frames would be discarded, yielding excellent calibration accuracy, but with poor photometric efficiency (0.05%). We note that the occurrence of lucky pairs of matching WFS frames between the observation and reference sets is significantly more likely than the occurrence of a lucky imaging instance, as the number of pairs is M^2, providing ~M x more opportunities for matches than for lucky imaging (see Appendix for details). Yet, in high-dimension space, relying on lucky matches alone is not sufficient to ensure high photometric efficiency, and the lin-Q model is required for efficiency. Our goal is to maintain high photometric efficiency while providing an accurate HC image reconstruction.

5. Supporting Activities and Future Work

Algorithms will be validated and exercised on two air testbeds: the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) and Magellan MagAO-X extreme adaptive optics systems. Both systems are configured to operate as in-air laboratory testbeds as well as on-sky instruments, and share the same software infrastructure. While they operate at more moderate contrast than required for this milestone, they provide greater flexibility, faster turn-around time, and a wider range of optical sensors/cameras options. Specifically, the lin-Q mapping algorithm and optimal linear combination techniques at the core of our proposed approach are being tested on datasets we acquire on SCExAO and MagAO-X systems. The fast turn-around time for experiments (~minutes) is especially valuable, as experimental parameters (for example: amplitude and dimensionality of input disturbances, wavelength bandwidth choices) can be tested and scanned.

A number of issues that could adversely affect the reconstruction accuracy are not addressed this in milestone:

- Temporal variations within each frame
- Low flux level in both HC and WFS frames
- Dissimilarity between reference and observation sets. The two sets may sample non-overlapping distributions of wavefront errors.

These will be addressed in future work and milestones.

A limitation of MS#1 is that the set of perturbations injected by DM actuation may not be representative of the full set of disturbances in a complete optical system including the instrument's upstream optics and telescope. In parallel with in-vacuum testing at JPL, we will deploy and validate our algorithms on datasets acquired in-air and on-sky with HCI systems operating at a much shallower contrast level. This will help validate that our approach can handle a wide range of perturbations in a realistic environment.

6. References

Steves, John, et al. “**Picometer wavefront sensing using the phase-contrast technique.**” Optica, vol. 7, no. 10, 2020, pp. 1267-1274.

Ruane, Garreth, et al. “**Wavefront sensing and control in space- based coronagraph instruments using Zernike’s phase-contrast method**.” Journal of Astronomical Telescopes, Instruments, and Systems, vol. 6, no. 4, 2020, 045005-1 to 045005-15. SPIE Digital Library, hyperlink.

Ruane, Garreth, and Jorge Llop-Sayson. Conversations. 2024.

Wallace, James K., et al. “**Technology Milestone White Paper Dual-Purpose Focal Plane Mask.**” Strategic Astrophysics Technology Awards, NASA Exoplanet Program, 9 November 2023, https://exoplanets.nasa.gov/internal_resources/2885/ . Accessed Dec 17, 2024.

Guyon et al., “**Linear Wavefront Control PSF Calibration**” Strategic Astrophysics Technology Awards, NASA Exoplanet Program, Dec 18, 2024, https://exoplanets.nasa.gov/internal_resources/3157/ . Accessed March 21, 2025.

James Kent Wallace, Garreth Ruane, Nasrat Raouf, Tobias S. Wenger, A.J. Riggs, Jeffrey B. Jewell, Daniel N. Shanks “**Dual purpose focal plane masks for contemporaneous wavefront sensing and high-contrast imaging: hardware status**”, Poster presentation, SPIE Aug 2024. hyperlink

Appendix: Probabilistic Considerations in PSF Subtraction

Our approach to PSF subtraction relies on matching observation frames to reference frames, so that for most of the observation frames, a reasonably good subtraction can be done. The quality of the subtraction relies on a combination of luck (is/are there reference frame(s) that is/are similar to the observation frame?) and extrapolation (how well can we extrapolate from reference frames an observation frame that has not been encountered in the reference set). We discuss here how the probability of lucky matches scales with the number of dimensions and the number of samples. This probability is related to the PSF subtraction’s photometric efficiency (= fraction of the dataset that contributes to the final PSF-subtracted science image). We represent each frame as a multidimensional variable, with each dimension encoding a wavefront mode.

We compare here the “***lucky matching***” probability (having a match between reference and observation frames) to the more familiar “***lucky imaging***” case (probability that a frame is good).

Both concepts can be expressed as a tradeoff between quality and photometric efficiency, where a fraction of the data is discarded to ensure the remaining frames meet a threshold.

For simplicity, we assume that the multidimensional points encoding frames sample the same zero-mean distribution (uniform distribution from -1 to +1 along each dimension), for both the reference and observation sets, with no correlation between the dimensions.

Lucky Imaging
The probability of obtaining a diffraction-limited short exposure image through atmospheric turbulence was quantified by Fried ("Probability of getting a lucky short-exposure image through turbulence", Fried, JOSA, 1978). A similar approach can be applied here.
Assuming N modes (number of dimension), each with uniform distribution (from -1 to +1), the probability that a point lands within a radius r of origin follows the N-ball volume equation.
$P(r) \ = \ \boldsymbol{\pi}^{N/2} \ / \ \boldsymbol{\Gamma}(N/2 \ + \ 1) \ \ (r/2)^{N}$ [eq A1]
With M samples, the number of points within r of the origin is thus M P(r)

Lucky Matching
For PSF subtraction, in the context of our approach, a closely related concept is the probability that two frames (one drawn from the reference set, one from the observation set) are within a distance r of each other. With M points in N dimensions for the reference and observation sets, there are M^2 pairs of (reference,observation) points, with the difference between the two points uniformly distributed in a r=2 hypercube. The probability that this difference is less than r is $P(r)/2^{N}$, with P(r) from equation A1.
With M^{2} pairs, the number of such matching pairs is $M^{2}P(r)/2^{N}$ so any point in the observing set has a probability of having a match in the reference set:
$PM(r) \ = \ M/2^{N} \, P(r)$ [eq A2]

Comparing equations A1 and A2, we draw the following conclusions:

- The photometric efficiency of both lucky imaging and lucky matching decreases steeply with the number of dimensions N
- While the photometric efficiency of lucky imaging is independent of the number of samples M, the photometric efficiency of lucky matching increases with the number of samples

For example, with M=1e6 samples in N=8 dimensions, and r=0.2, we have P(r)~4.1e-8, so the expected number of lucky samples are:

Lucky imaging: M P(r) = 0.04

Lucky matching: M (M/2^N P(r)) = 158

In this example, there is a 4% probability of finding a single lucky imaging frame in the sample of one million points, while the expected number of observation samples having a match in the reference set is 158. While the lucky matching efficiency is much higher than the lucky imaging efficiency, it is still too small (photometric efficiency ~ 1.6e-4) to be useful, highlighting the need for local fitting (lin-Q model) which is analogous to extending the matching radius.