applied optics

Extending capture range for piston retrieval in segmented systems

Scott W. Paine* and James R. Fienup

The Institute of Optics, University of Rochester, Rochester, New York 14627, USA
**Corresponding author: scott.paine1@gmail.com*

Received 2 August 2017; revised 18 October 2017; accepted 18 October 2017; posted 19 October 2017 (Doc. ID 303869); published 14 November 2017

Image-based wavefront sensing is a powerful technique for measuring the aberrations of optical systems and surfaces. It often fails for segmented systems with large piston errors per segment. We propose a method for finding these errors using broadband light and a specialized grid search as part of a more global search. We show that this method has a high rate of success for a case where nonlinear optimization gets stuck in local minima. We also explore points of failure.

OCIS codes: (010.7350) Wave-front sensing; (100.5070) Phase retrieval; (110.6770) Telescopes; (120.5050) Phase measurement; (350.1260) Astronomical optics.

https://doi.org/10.1364/AO.56.009186

1. INTRODUCTION

Image-based wavefront sensing methods often use nonlinear optimization to attempt to match measured point-spread function (PSF) data with a physical model of the PSF computed from wavefront estimates. These algorithms are advantageous for space-based systems as they require no additional wavefront-sensing hardware, and are currently used during the test and commissioning of the James Webb Space Telescope (JWST) [1,2]. Gradient-based search methods such as the ones outlined in Ref. [3] are used to perform phase retrieval, and by using modular gradient methods with algorithmic differentiation they can be implemented rapidly without the need to derive a full analytic gradient [4].

Gradient-based, as well as other phase-retrieval methods, often encounter a "capture range" problem. These methods use gradient information to establish a search path through the given high-dimensional parameter space of an error metric. When an error metric has multiple points where the gradient is zero, performing optimization can get stuck in a local minima that is not the true solution. At these local minima, the error metric value is larger than that at the global minimum, which is the true solution. Getting stuck in a local minimum commonly happens if the starting point (the initial estimate of the unknown phase parameters) for optimization is "too far" in parameter space from the true solution: the solution is outside of the capture range. The capture range for a given system can be extended through taking a diversity of additional measurements and adapting the model to be consistent with these measurements [5–7]. Random starting points or auxiliary system information can also be used to perform a global search without changing the capture range of the optimization algorithm itself [8,9].

Segmented systems such as the JWST require phase retrieval algorithms that can capture surface errors in individual primary segments (piston, tip, tilt, etc.) as well as aberrations that are not limited to individual segments, like global defocus, astigmatism, and coma. An important aberration of the first type is per-segment piston aberrations. This kind of retrieval can be done through the use of additional experimental components like pseudo non-redundant masks (PRNMs) or grisms with closure phase [10–12]. Löfdahl and Eriksson have shown that it is also possible to take multiple images at predetermined wavelengths to determine per-segment piston aberrations [13]. Currently, the JWST uses dispersed fringe sensing (DFS) with a dispersed Hartmann sensor (DHS) to reduce segment piston aberration anywhere from a maximum of 350 μm down to approximately 1 μm, which is considered coarse phasing [1,2,14]. From this point, the RMS wavefront error (WFE) is small enough that image-based wavefront sensing can be performed to lower the overall RMS WFE to 50 nm [2,15]. Without the coarse phasing step, the WFE falls outside of the capture range of image-based wavefront sensing due to 2π piston phase ambiguities, which cause local minima [2]. We seek to overcome this ambiguity with polychromatic light and use a grid search to search the segment piston parameter space for a global minimum.

A. Grid Searches Using Polychromatic Models

In monochromatic systems, a 2π phase ambiguity arises due to the periodic nature of the phases. If a simple monochromatic two-segment system is considered with a phase difference Δp rad

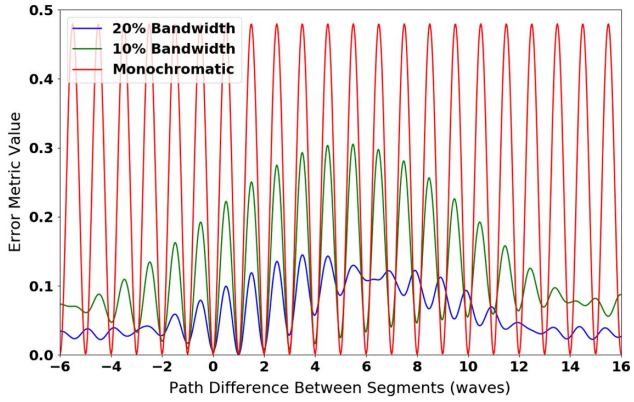

Fig. 1. Data consistency metric values for a simple two-segment system with a square spectrum at different bandwidths and a piston path difference of 1 wave. The path difference values are in units of the central wavelength λ_0.

between the two segments, then $\Delta p + 2\pi n$, for all integers n, will produce the same PSFs, since $\exp(i\Delta p) = \exp[i(2\pi n + \Delta p)]$. This means that a measured PSF for a monochromatic system will have several simulated PSFs that will cause a data consistency metric, such as a sum of square differences between the computed and measured PSFs, to be zero for any segment piston values of p_1, p_2 such that $p_1 - p_2 = \Delta p + 2\pi n$. Such an effect is shown by the red (upper) line in Fig. 1, for which the true value of Δp is 2π rad, or 1.0 waves.

It is possible to remove this ambiguity with bandwidth. The phase in wavelengths (cycles) for a given wavelength λ is equal to OPD/λ, where the OPD is the physical optical path difference. Introducing different values of λ leads to different phases, making Δp vary with wavelength. This removes the ambiguity and leads to only one correct value of Δp that reduces the data consistency metric to near zero. Figure 1 shows the effects of 10% (middle green line) and 20% (lower blue line) bandwidth having a flat-top spectra used in simulating PSFs and computing estimated PSFs. Here, we define bandwidth of the flat-top spectra as the ratio of the total difference in wavelength divided by a central wavelength, which we call λ_0. Our simulated PSFs are built using discretely sampled versions of a continuous spectrum, which allows us to make individual PSFs corresponding to each wavelength component and perform a sum weighted by the spectrum to create a broadband PSF, which we compare to broadband data [16]. The true solution in Fig. 1 lies where the OPD between segments is λ_0. We see that the error metric for broadband light is zero only for an OPD of λ_0 and is nonzero for all other phase difference values.

In this example, removing the phase ambiguity by using broadband light left a number of local minima in the data consistency metric values. A gradient-based nonlinear optimization procedure will typically move to the local minimum nearest to a starting point. Therefore, if we have a starting guess for the path difference of our two-segment system that is not within the range $0.5\lambda_0$–$1.5\lambda_0$, we will end up in a local minimum that is not the global solution. To overcome this, we note that, within a certain range, all local minima are roughly separated by a path difference that is an integer multiple of λ_0. Therefore, in our optical model in the computer, we can adjust the optical path values for a single segment by amounts $n\lambda_0$ (where n is an integer) and check our data consistency metric for an improvement. In a system with more segments, each path difference between segments will cause different interference fringes in the PSF in different directions, so we can attempt to use this technique on each segment to recover a solution for the per-segment phase difference for all segments.

Figure 1 also demonstrates that the grid search itself has a capture range. The 20% bandwidth case has local minima that do not align with the monochromatic case in regions where the difference between segments is greater about $\pm 5\lambda_0$ from the true value; similarly, the 10% bandwidth has different local minima where the path difference is greater than about $\pm 10\lambda_0$ from the true value. This capture range of the grid search is determined by the coherence length of the light allowed through the optical system. When the path difference between two segments is greater than the coherence length, the fields from the two segments become incoherent with respect to one another and lose the coherent effects such as fringe patterns in the PSF that help our optimizer estimate the phase differences.

2. ALGORITHM

A brute force algorithm, trying $-5\lambda_0, -4\lambda_0, \ldots, +4\lambda_0, +5\lambda_0$ piston values for all combinations of each of the 17 non-phased segments of the JWST would require $\sim 10^{17}$ PSF evaluations, which is impractical. Our efficient algorithm can be split into three basic parts: an initial nonlinear optimization, a grid search, and a final nonlinear optimization. For our nonlinear optimization we use the limited-memory BFGS algorithm with bounds, which is a gradient-based optimization algorithm [17]. For a data consistency metric, we utilize a gain-and-bias-invariant normalized mean squared error metric, given by

$$\Phi = \frac{\sum_{x,y} w(x,y)[\alpha M(x,y) + \beta - D(x,y)]^2}{\sum_{x,y} w(x,y)[D(x,y)]^2}, \quad (1)$$

where $D(x,y)$ is our broadband data, $M(x,y)$ is our modeled broadband intensity, $w(x,y)$ is a weighting mask (used, for example, to ignore the effects of bad pixels), and α and β are gain and bias terms determined to minimize Φ as in Ref. [18]. For our work here, we have no weighting on our metric, so $w(x,y) = 1$. For our initial guesses of the WFE in all the simulations in this paper, except for those at the end of Section 3, we included known terms in our physical model (such as defocus) and set all other wavefront parameters to zero, in order to have a consistent starting point for all simulations. It is also feasible to use a random starting guess for the WFEs with appropriate RMS WFE. To place ourselves in a local minimum close to the starting guess, we enforced bounds on the per-segment piston phase values, forcing them to lie between $-2\lambda_0$ and $2\lambda_0$. This prevented the optimizer from giving large per-segment piston differences. We then use the parameters from this local minimum and the original optical model in our grid search algorithm. We first establish a baseline data consistency metric value using the PSF generated by our forward model with the stagnated parameters and the measured data. We then alter the path difference value for a single segment by an integer

amount of λ_0, and use our physical model to generate a PSF, and compute its data consistency metric value. If the new value is less than the baseline, then we accept the new value as the baseline and update our parameters to keep the new piston value. We repeat this for several integer values of λ_0 for a single segment, and then repeat this over all segments. This grid search is run until no improvement is made to the baseline error metric value, which means that it must be run at least twice. Once this is complete, we use the new values for our parameters as a starting point for a final gradient-based nonlinear optimization without enforcing any piston bounds to "fine-tune" the result.

3. MONTE CARLO ANALYSIS

To improve and analyze this algorithm, we performed several Monte Carlo analyses. First, we examined how the algorithm with the grid search included compared with a simple nonlinear optimization approach. We examined 1000 cases that included Poisson noise with 40,000 peak photo-electrons and zero-mean Gaussian read noise with a standard deviation of 12 photo-electrons. Our simulated model had an exit pupil like that of the JWST, with 18 hexagonal segments, and included a 20% bandwidth flat-top spectrum discretely sampled with 15 equally spaced points centered around a central wavelength λ_0. We simulated PSFs using this system where the only aberrations on the telescope were segment piston errors, with amounts that were randomly chosen from a uniform distribution between 0 and $3.5\lambda_0$. This allows the PSFs to statistically be outside of the capture range of nonlinear optimization but inside the capture range of our grid search. Our optimizer was given three images with defocus amounts of $4\lambda_0$, 0, and $-4\lambda_0$, which is consistent with the imaging capabilities of the JWST and improves the probability of convergence for image-based wavefront sensing [1,2,5,19].

In these preliminary trials, none of the cases converged when only nonlinear optimization was used. The grid search implementation found the correct phases in 874 of the 1000 test cases. By including two more images corresponding to $8\lambda_0$ and $-8\lambda_0$ defocus frames, the grid search implementation found the correct phases in 903 of the 1000 test cases. Finally, by allowing for an additional grid search after our second nonlinear optimization step, the correct phases were obtained in 925 of the 1000 test cases.

We considered the effect of the shape of the spectrum on the probability of convergence by implementing physical models with discretely sampled rectangular, triangular, and Gaussian spectra, shown in Fig. 2. The spectra used for these models were shaped so that they all had identical coherence lengths, using the coherence time $\tau_c = \int_0^\infty |S(\nu)|^2 d\nu / (\int_0^\infty S(\nu) d\nu)^2$, where $S(\nu)$ is our spectrum. This definition was adapted from Eq. (5) in Ref. [20], using the Fourier relationship between the complex degree of self-coherence $\Gamma(\tau)$ and the spectrum $S(\nu)$. We used the same noise parameters and defocus planes as in the previous trial, but we chose random segment piston values uniformly distributed between $-5\lambda_0$ and $5\lambda_0$, ensuring that there would be cases that fall outside of the capture range. We also allowed for two grid searches in our optimization. Figure 3 shows the results of these Monte Carlo simulations. In these simulations, both triangular and Gaussian spectra performed significantly better than the square spectrum when the path differences fell outside of the coherence length of the spectrum. When the path differences were within the coherence length of the spectrum, there was no appreciable difference in convergence probability with respect to the type of spectrum. However, there was notably lower convergence for the square spectrum relative to the other two when the coherence length was shorter.

For the longest coherence length (narrowest spectrum), we believe the decrease in convergence rates results from data consistency metric values at or near the noise floor for minima near the solution, as can be seen on the green line in Fig. 1. It became difficult to differentiate between solutions, and therefore the algorithm became stuck in a local minima. There are values around the global minimum consisting of errors of $\pm\lambda_0$ difference for segments where the local minima are very close in magnitude. Noise will make these values even closer, leading to local minima that are difficult to discern from the true global minima. At the shortest coherence length, we believe differences in convergence come from the coherence effects of the spectra. Recall that the complex degree of self-coherence $\Gamma(\tau)$ is proportional to the Fourier transform of the spectrum $S(\nu)$. Of the three spectra in these simulations, only the square spectrum has a Fourier transform with negative sidelobes. We believe this causes the local minima of our error metric to change to

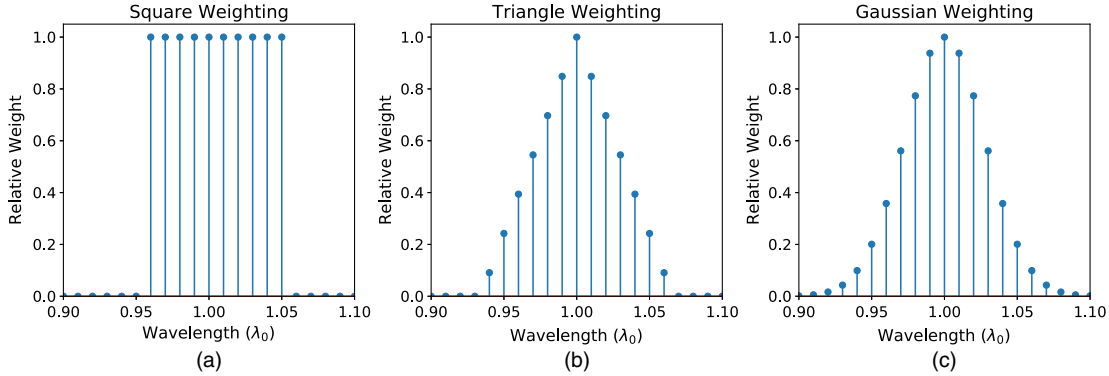

Fig. 2. Discretely sampled (a) square, (b) triangular, and (c) Gaussian spectra used for polychromatic models for convergence comparison. The wavelength samples are in terms of the central wavelength λ_0. Each sample has the same coherence length, and has 15 equally spaced samples.

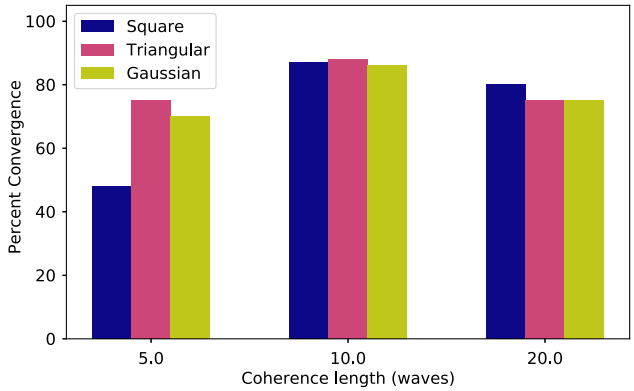

Fig. 3. Convergence comparisons for different spectra with path differences up to $10\lambda_0$ at different coherence lengths. All results are from simulations that use two grid searches.

non-integer amounts past the coherence length, as seen by the red line in Fig. 4 at approximately 5 waves of path difference. In these simulations, we used discrete approximations to the triangular and truncated Gaussian spectra, for which $\Gamma(\tau)$ will be only approximately non-negative. As a result, for the Gaussian spectrum, the spacings between the local minima increased slowly, making the spacing between the global minimum and the local minima a non-integer amount. As seen by the blue line in Fig. 4 in the region of 6.0–8.0 waves of path difference, the error metric value has minima that have distances between them greater than λ_0, and the "depth" of these minima becomes shallow outside of the coherence length. Despite these deviations of the local minima from integer values, our grid search could still recover the true solution since the perturbations of the local minima were not large. Indicated by the black dashed curve, the triangular spectrum has local minima that, in the region from 6.0 to 8.0 waves, also deviate from integer wave amounts from the true solution; however, at 9.0 waves and 14.0 waves, the local minima return to being nearly integer spaced, indicating that this triangular spectrum may have a wider capture range than both Gaussian and rectangular spectra for the same coherence length.

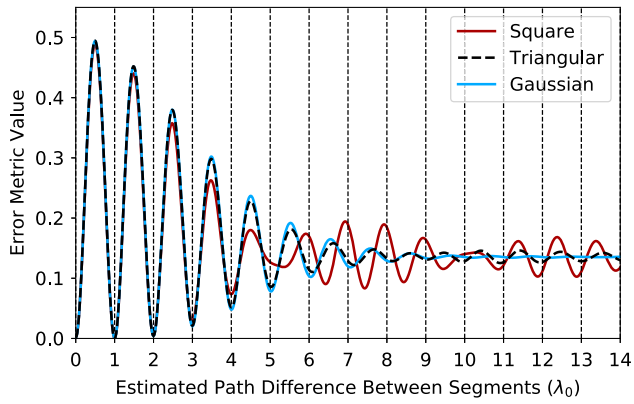

Fig. 4. Data consistency metric values for different spectra for a simple two-segment system that has a piston path difference of 1 wave. The path difference values are in units of the central wavelength λ_0.

We performed Monte Carlo analysis to determine the robustness of our algorithm in a system with both large piston and additional aberrations. We used our forward model of the JWST with a Gaussian spectrum of width $0.05\lambda_0$. For this analysis, we allowed all of the following to vary: the maximum difference in piston, global RMS WFE, and per-segment RMS WFE other than piston. Figure 5 shows examples of data PSFs. All simulations included the same noise statistics and defocus planes as described previously, and used two grid searches. We implemented bootstrapping techniques during our nonlinear optimization algorithm to combat stagnation of the given algorithm in a local minima. We first optimized only the per-segment WFE and then both the global and per-segment WFE contributions. The resulting wavefront was then used as a starting guess for our grid search algorithm. Our criteria for successful reconstruction was a residual RMS WFE less than $0.02\lambda_0$, which was determined using error metric versus residual RMS WFE plots like those seen in Fig. 6. The results of these simulations are summarized in Table 1. Simulations were done with values of segment RMS WFE out to $0.5\lambda_0$, but all simulations with greater than $0.25\lambda_0$ waves RMS of per-segment aberration failed to converge. We note that there is lower probability of convergence when the maximum path difference is $3.0\lambda_0$ as compared to $6.0\lambda_0$. The results also show lower likelihood of convergence for small per-segment WFE, with the simulations converging the most when the per-segment RMS WFE is $0.15\lambda_0$. This did not change when the order of bootstrapping was changed. We hypothesize that this is due to the nonlinear optimization algorithm's using low-order aberrations to attempt to fit piston errors. For small amounts of per-segment WFE, this would cause the algorithm to move to local

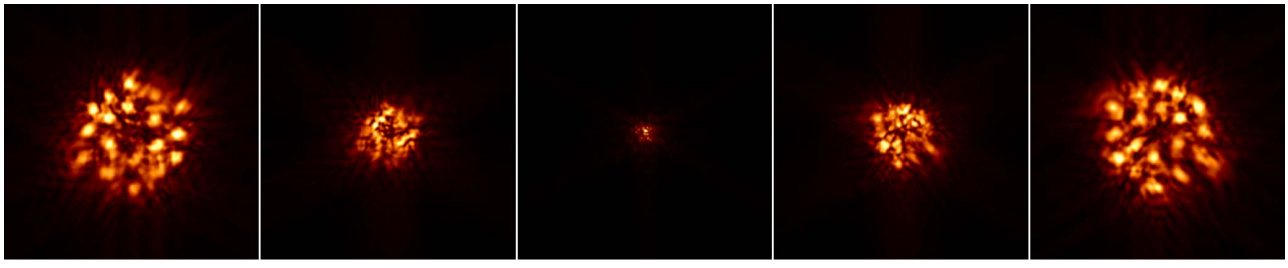

Fig. 5. Sample data PSFs prior to applying noise, with known defocuses of $-8\lambda_0$, $-4\lambda_0$, 0, $4\lambda_0$, and $8\lambda_0$ from left to right. For these PSFs, the maximum path difference is $3.0\lambda_0$, the global RMS WFE is $0.15\lambda_0$, and the (non-piston) per-segment RMS WFE is $0.05\lambda_0$. Each PSF array has been square-rooted element-wise to show dimmer features.

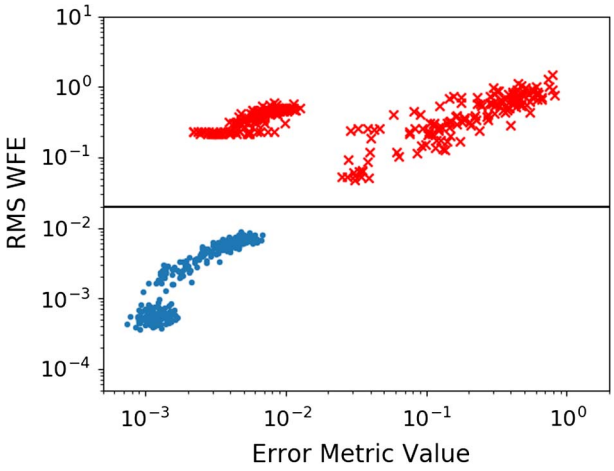

Fig. 6. Plot of RMS WFE in λ_0 versus error metric value for solutions, which was used to determine cutoff criterion for successful wavefront recovery for 582 simulations with a global RMS WFE of $0.15\lambda_0$ and maximum path difference of $3.0\lambda_0$. The cutoff line can be seen at $0.02\lambda_0$ RMS WFE. Blue dots (300) indicate successful retrieval cases and red Xs (282) indicate unsuccessful retrievals.

minima with much higher per-segment RMS WFE than in the true solution, which would be used to try to fit features that were not attributed to per-segment wavefront. It was common in the failed retrievals that a single segment would have a large amount of tilt, which would cause the energy to overlap with another segment's energy in some frames. This has been observed before, and can be countered by using a geometrical optics approach that ensures that light from neighboring segments does not overlap in the modeled PSFs if it does not overlap in the data [8,21]. The results in Table 1 show overall that the algorithm is robust to additional per-segment RMS WFE of up to $0.2\lambda_0$, and additional global RMS WFE of up to $0.25\lambda_0$; that is, the probability of success for any one retrieval is high enough that with a few different random starting guesses, we can expect a successful retrieval with high probability. The starting point for the nonlinear optimizations in Table 1 was an aberration-free system, which means the optimizer can get stuck in local minima near this guess in parameter space. We chose a particular start, underlined in Table 1 with $0.05\lambda_0$ of per-segment WFE, $0.15\lambda_0$ of global WFE, and $3.0\lambda_0$ maximum path difference. Table 1 shows a 43.3% chance of success for an ensemble of cases having the same statistics. This particular case had stagnated at a wavefront that had $1.8\lambda_0$ residual RMS WFE from the true solution, and produced an estimate with residual global WFE and per-segment WFE. We then tried optimization with

Table 1. Monte Carlo Results for Determining Robustness of Grid Search Algorithm for 20% Bandwidth[a]

Maximum Segment Path Difference: $3.0\lambda_0$

Global RMS WFE (λ_0) \ Segment RMS WFE (λ_0)	0.00	0.05	0.10	0.15	0.20	0.25
0.00	26	46	81	98	79	10
0.05	34	52	72	96	80	10
0.10	43	52	73	94	71	7
0.15	45	43	69	93	53	6
0.20	38	51	72	79	37	3
0.25	36	50	66	57	19	2

Maximum Segment Path Difference: $6.0\lambda_0$

Global RMS WFE (λ_0) \ Segment RMS WFE (λ_0)	0.00	0.05	0.10	0.15	0.20	0.25
0.00	64	69	91	96	80	6
0.05	68	76	92	89	70	4
0.10	63	77	87	90	62	3
0.15	69	72	86	85	51	1
0.20	73	75	88	76	23	2
0.25	59	71	70	57	11	0

Maximum Segment Path Difference: $9.0\lambda_0$

Global RMS WFE (λ_0) \ Segment RMS WFE (λ_0)	0.00	0.05	0.10	0.15	0.20	0.25
0.00	40	50	75	81	51	3
0.05	38	45	74	84	59	3
0.10	53	53	68	81	45	5
0.15	47	57	74	78	30	3
0.20	43	39	63	61	19	3
0.25	39	59	53	44	15	1

Maximum Segment Path Difference: $12.0\lambda_0$

Global RMS WFE (λ_0) \ Segment RMS WFE (λ_0)	0.00	0.05	0.10	0.15	0.20	0.25
0.00	11	19	24	36	24	0
0.05	14	14	28	35	11	0
0.10	12	16	30	31	19	1
0.15	12	19	22	30	11	0
0.20	13	9	19	22	8	2
0.25	7	10	19	21	5	0

[a]Each cell contains the percentage of successful retrievals, with a total of 97 trials per cell.

this case again, but with 30 different random starting guesses that each had the same amount of per-segment WFE, global WFE, and maximum path difference as the true wavefront, but having different values of the particular coefficients. Of the 30 random starting guesses, 27 converged to a solution with less than $0.02\lambda_0$ residual RMS WFE, which was our criterion for convergence. This result demonstrates that utilizing a small number of random starting guesses can greatly increase the probability of a successful retrieval for this algorithm.

4. FAILURE MODES

This algorithm has two types of local minima that most commonly account for convergence failure. Figure 7(a) shows the difference between the true wavefront and an estimated wavefront for which a particular segment path difference falls outside of the coherence length. This minimum occurs due to a shift in local minima in the data consistency metric for large path separations. The blue line in Fig. 1 and the red line in Fig. 4 show how the local minima change for 20% bandwidth when the path difference between segments is longer than 5 waves for a flat-top spectrum: the local minima go from near-integer values to near-half-integer values. For the Gaussian spectrum, the distance between adjacent local minima slowly increases until they reach a difference that is far enough from integer λ_0 amounts that the grid search could not recover, as seen by the blue line in Fig. 4. The triangular spectrum has the same behavior near the coherence length, but then has the minima shift once more to be closer to integer λ_0 separations, as seen by the black dashed line in Fig. 4. For all spectra, this problem was accounted for algorithmically for our Monte Carlo simulations by setting constraints on the maximum path difference between any two segments in the system, and is detectable by the larger data consistency error value that is calculated from the simulated PSFs. Despite the larger error metric, however, one is still stuck in the local minimum, unable to find the global minimum.

Figure 7(b) shows the more difficult failure mode. Here, one or more contiguous group(s) of segments differ from the true wavefront solution by λ_0. The failure seen in Fig. 7(b) shows one of the more obvious groupings, but the differing groups can be as small as two segments. This type of failure is particularly difficult to detect, as the data consistency metric value does not differ greatly from the data consistency metric value at the true solution when enough noise is present. This is illustrated in Fig. 8, which shows that there are cases in which we could

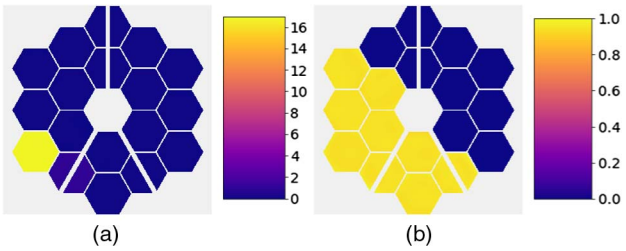

Fig. 7. Modes of failure for the grid search algorithm: (a) a segment outside the coherence width of the model, and (b) the more common failure mode, the "split telescope." Color bar values are in units of λ_0.

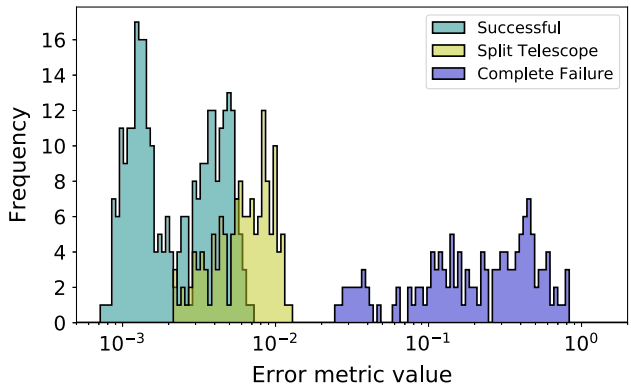

Fig. 8. Histogram of error metric values versus failure type for the same simulation group in Fig. 6. Note the overlap between the "split telescope" minima and the successful retrievals.

not differentiate the "split telescope" and true solutions by simply looking at the error metric. There are a set of cases in Fig. 8 where the error metric is below 2×10^{-3} and all of the cases are successful, indicating a possible threshold that guarantees success. With higher than $6\lambda_0$ path differences, the group of local minima corresponding to other types of failure have error metric values that overlap the "split telescope group" more, and it becomes even more difficult to determine if a retrieval is correct, a "split telescope," or a general failure. When we examined the number of these types of failures versus successful retrievals in our Monte Carlo simulations, we found that there were 2–5 times more successful retrievals than "split telescopes," depending on path difference.

We hypothesize that the "split telescope" occurs due to the use of defocus diversity. In planes with defocus, the fields for individual segments overlap less if the segments are farther away. Therefore, fringe patterns in the PSFs will be stronger when due to nearby pairs of segments, and a preference for correct fitting relative to neighbors occurs. The only frame where all segment fields overlap each other fully is at focus, which is the case for only one frame in our model, and this zero-defocus plane tends to be less informative for phase retrieval than the out-of-focus planes. Therefore, the preference of fitting to neighbors causes different groups to form and the solution falls into this local minimum. This can be overcome, once again, by using multiple starting guesses and placing more stringent requirements on acceptance for a solution: for example, having an error metric value below 2×10^{-3} for the cases in Fig. 8.

5. CONCLUSION

We presented a computational method for extending the capture range of image-based wavefront sensing for a segmented system with per-segment piston errors of multiple waves of the central wavelength λ_0. We accomplish this by taking advantage of the quasi-periodic shape of the error metric as a function of path difference between segments with polychromatic light and performing a grid search by adjusting segment piston values in integer units of λ_0. Our method works for piston errors within the coherence length of a given light source, and is robust to additional aberrations and noise. Defocus diversity is

used for improving convergence probability. The algorithm fails when segment path differences are outside of the coherence length of the source or when strong fitting preference to neighboring segments causes groups of segments to dephase from the remainder of the telescope, referred to as a "split telescope." We find that coupling this method with multiple random starting guesses for parameters other than per-segment piston error increases the probability of convergence, even in the presence of additional aberrations.

We are currently examining methods of preventing the "split telescope" failure by adjusting contiguous segment groups. It may also be possible to utilize mixed-integer nonlinear problem (MINLP) optimization to perform a tree search for the global minimum [22]. This method requires constant pruning of branches, or large amounts of memory to hold or cache possible solutions. We could also attempt to prevent the "split telescope" by adjusting the segments neighbor-first. In this way, we would not allow any distant segments to move until all of their neighbors were appropriately fit, preventing dephasing for groups. Additionally, we will examine the effects of random starting guesses across all of the cases in Table 1 to see if convergence can be improved across all trials.

Funding. NASA Goddard Space Flight Center (GSFC) (NNX15AE94A).

REFERENCES

1. D. S. Acton, P. D. Atcheson, M. Cermak, L. K. Kingsbury, F. Shi, and D. C. Redding, "James Webb Space Telescope wavefront sensing and control algorithms," Proc. SPIE **5487**, 887–896 (2004).
2. D. S. Acton, J. S. Knight, A. Contos, S. Grimaldi, J. Terry, P. Lightsey, A. Barto, B. League, B. Dean, J. S. Smith, C. Bowers, D. Aronstein, L. Feinberg, W. Hayden, T. Comeau, R. Soummer, E. Elliott, M. Perrin, and C. W. Starr, "Wavefront sensing and controls for the James Webb Space Telescope," Proc. SPIE **8442**, 84422H (2012).
3. J. R. Fienup, "Phase-retrieval algorithms for a complicated optical system," Appl. Opt. **32**, 1737–1746 (1993).
4. A. S. Jurling and J. R. Fienup, "Applications of algorithmic differentiation to phase retrieval algorithms," J. Opt. Soc. Am. A **31**, 1348–1359 (2014).
5. R. A. Gonsalves, "Phase retrieval and diversity in adaptive optics," Opt. Eng. **21**, 829–832 (1982).
6. R. G. Paxman, T. J. Schulz, and J. R. Fienup, "Joint estimation of object and aberrations by using phase diversity," J. Opt. Soc. Am. A **9**, 1072–1085 (1992).
7. G. R. Brady, M. Guizar-Sicairos, and J. R. Fienup, "Optical wavefront measurement using phase retrieval with transverse translation diversity," Opt. Express **17**, 624–639 (2009).
8. A. S. Jurling and J. R. Fienup, "Extended capture range for focus-diverse phase retrieval in segmented aperture systems using geometrical optics," J. Opt. Soc. Am. A **31**, 661–666 (2014).
9. D. Moore and J. R. Fienup, "Extending the capture range of phase retrieval through random starting parameters," in *Frontiers in Optics* (Optical Society of America, 2014), paper FTu2C.2.
10. E. H. Smith, G. Vasudevan, R. D. Reardon, R. Bernier, and K. J. Triebes, "Coarse phasing of a segmented mirror using a dispersed fringe sensor," Proc. SPIE **4850**, 469–477 (2003).
11. J. F. Simar, Y. Stockman, and J. Surdej, "Single-wavelength coarse phasing in segmented telescopes," Appl. Opt. **54**, 1118–1123 (2015).
12. F. Shi, G. Chanan, C. Ohara, M. Troy, and D. C. Redding, "Experimental verification of dispersed fringe sensing as a segment phasing technique using the Keck telescope," Appl. Opt. **43**, 4474–4481 (2004).
13. M. G. Löfdahl and H. Eriksson, "Resolving piston ambiguities when phasing a segmented mirror," Proc. SPIE **4013**, 774–782 (2000).
14. F. Shi, S. A. Basinger, and D. C. Redding, "Performance of dispersed fringe sensor in the presence of segmented mirror aberrations: modeling and simulation," Proc. SPIE **6265**, 62650Y (2006).
15. L. D. Feinberg, B. H. Dean, D. L. Aronstein, C. W. Bowers, W. Hayden, R. G. Lyon, R. Shiri, J. S. Smith, D. S. Acton, L. Carey, A. Contos, E. Sabatke, J. Schwenker, D. Shields, T. Towell, F. Shi, and L. Meza, "TRL-6 for JWST wavefront sensing and control," Proc. SPIE **6687**, 668708 (2007).
16. J. R. Fienup, "Phase retrieval for undersampled broadband images," J. Opt. Soc. Am. A **16**, 1831–1837 (1999).
17. R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," SIAM J. Sci. Comput. **16**, 1190–1208 (1995).
18. S. T. Thurman and J. R. Fienup, "Phase retrieval with signal bias," J. Opt. Soc. Am. A **26**, 1008–1014 (2009).
19. B. H. Dean, D. L. Aronstein, J. S. Smith, R. Shiri, and D. S. Acton, "Phase retrieval algorithm for JWST flight and testbed telescope," Proc. SPIE **6265**, 626511 (2006).
20. L. Mandel and E. Wolf, "The measures of bandwidth and coherence time in optics," Proc. Phys. Soc. London **80**, 894–897 (1962).
21. S. T. Thurman, "Method of obtaining wavefront slope data from through-focus point spread function measurements," J. Opt. Soc. Am. A **28**, 1–7 (2011).
22. P. Belotti, C. Kirches, S. Leyffer, J. Linderoth, J. Luedtke, and A. Mahajan, "Mixed-integer nonlinear optimization," Acta Numer. **22**, 1–131 (2013).