



HAL
open science

Wavefront control with algorithmic differentiation on the HiCAT testbed

Scott D. Will, Marshall D Perrin, Emiel H Por, James Noss, Ananya Sahoo, Peter Petrone, Iva Laginja, Raphael Pourcelot, Susan M. Redmond, Laurent Pueyo, et al.

► **To cite this version:**

Scott D. Will, Marshall D Perrin, Emiel H Por, James Noss, Ananya Sahoo, et al.. Wavefront control with algorithmic differentiation on the HiCAT testbed. SPIE Optical Engineering + Applications 2021, Aug 2021, San Diego, United States. pp.118230V, 10.1117/12.2594283 . hal-03461815

HAL Id: hal-03461815

<https://hal.science/hal-03461815>

Submitted on 1 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wavefront control with algorithmic differentiation on the HiCAT testbed

Scott D. Will^{a,b,c}, Marshall D. Perrin^b, Emiel H. Por^b, James Noss^b, Ananya Sahoo^b, Peter Petrone^{b,d}, Iva Laginja^{e,f}, Raphaël Pourcelot^g, Susan M. Redmond^h, Laurent Pueyo^b, Tyler D. Groff^c, James R. Fienup^a, and Remi Soummer^b

^aThe Institute of Optics, University of Rochester, Rochester, NY 14627, USA

^bSpace Telescope Science Institute, 3700 San Martin Drive, Baltimore MD 21218, USA

^cNASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

^dHexagon Federal, Chantilly, VA 20151, USA

^eAix Marseille Univ., CNRS, CNES, LAM, Marseille, France

^fDOTA, ONERA, Université Paris Saclay, F-92322 Châtillon, France

^gUniversité Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, France

^hDepartment of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08540, USA

ABSTRACT

Future space-based coronagraphs will rely critically on focal-plane wavefront sensing and control with deformable mirrors to reach deep contrast by mitigating optical aberrations in the primary beam path. Until now, most focal-plane wavefront control algorithms have been formulated in terms of Jacobian matrices, which encode the predicted effect of each deformable mirror actuator on the focal-plane electric field. A disadvantage of these methods is that Jacobian matrices can be cumbersome to compute and manipulate, particularly when the number of deformable mirror actuators is large. Recently, we proposed a new class of focal-plane wavefront control algorithms that utilize gradient-based optimization with algorithmic differentiation to compute wavefront control solutions while avoiding the explicit computation and manipulation of Jacobian matrices entirely. In simulations using a coronagraph design for the proposed Large UV/Optical/Infrared Surveyor (LUVOIR), we showed that our approach reduces overall CPU time and memory consumption compared to a Jacobian-based algorithm. Here, we expand on these results by implementing the proposed algorithm on the High Contrast Imager for Complex Aperture Telescopes (HiCAT) testbed at the Space Telescope Science Institute (STScI) and present initial experimental results, demonstrating contrast suppression capabilities equivalent to Jacobian-based methods.

1. INTRODUCTION

Future space coronagraphs attempting to image and characterize Earth-like planets around nearby solar-type stars will rely critically on closed-loop wavefront sensing and control (WFS&C) using deformable mirrors (DMs) to mitigate optical aberrations in the primary beam path. These aberrations, primarily mid-spatial frequency wavefront errors and optical misalignments in the telescope and coronagraph optics, give rise to a speckle floor that is coherent with the star and evolves slowly over time in response to minute drifts in the thermal and mechanical state of the observatory. If uncorrected, the speckle floor overwhelms the faint image of the orbiting planet, which is expected to be 10^{10} times fainter than the host star at 0.1 arcseconds of separation or less.¹

Space-based coronagraphy has a unique set of attributes that distinguish it from conventional ground-based adaptive optics, namely, extreme sensitivity to non-common path aberrations and a comparatively high degree of temporal stability due to the absence of atmospheric turbulence. The goal is not to maximize the Strehl ratio of the point-spread function (PSF), but rather to minimize starlight within a specified region of the image plane,

Further author information: (Send correspondence to S.D.W.)

S.D.W.: E-mail: scott.will@rochester.edu

called the dark zone, in which the signal from the orbiting exoplanet may be isolated and interrogated. These constraints have given rise to a unique set of focal-plane WFS&C algorithms that close the loop around the image plane directly rather than relying on an external wavefront sensor, utilizing phase diversity provided by the DMs² or inline holography³ to estimate the complex stellar electric field (E-field) from intensity measurements. The WFS&C loop drives the starlight inside the dark zone toward zero over a series of iterations, each of which consists of estimation of the E-field from the host star followed by an update to the DM correction. This is illustrated in Figure 1.

Current state-of-the-art wavefront control algorithms adopt an optimal control approach: in each WFS&C iteration, a DM command update is found by minimizing a scalar cost function that is parameterized by the most recent estimate of the E-field as well as a computer model of the coronagraph. The most common algorithms, stroke minimization (SM)⁴ and electric field conjugation (EFC),² leverage a first-order approximation to the stellar E-field to formulate cost functions that are quadratic functions of both the DM command updates and the DM-corrected E-field within the dark zone. In both cases, the optimal DM update is written down in closed form as the solution of a linear system of equations constructed from a Jacobian matrix that describes the impact of a change in DM actuator commands on the focal-plane E-field.

The complexity of SM and EFC is dominated by the cost of computing and manipulating the Jacobian matrix, which grows in size proportionally to the product of the DM actuator count and the dark-zone pixel count; simultaneous increases in both compound multiplicatively, and a separate Jacobian is required for each controllable wavelength for broadband imaging. The Jacobian is most often model-based, in which case an optical diffraction model of the coronagraph is evaluated repeatedly to predict the impact of each actuator. This requires the coronagraph model to be evaluated at least as many times as there are actuators, which is undesirable if the number of actuators is large, computational resources are limited, or both. Moreover, the Jacobian is a linearization of the true, nonlinear behavior of the DMs and must be recalculated periodically as the state of the DMs evolves over time.

Until recently, little attention has been paid to the computational demands of SM and EFC. However, over time as direct imaging missions demand DMs with ever-higher actuator density to enable wider and wider search areas, computational aspects will become an inevitable point of concern from a systems engineering standpoint. In on-orbit WFS&C, all sensing and control computations are processed by the flight computer; conversely, in ground-in-the-loop (GITL) scenarios, raw data is communicated to a ground-based computing node that calculates the DM correction and relays it back to the observatory. Though each approach has tradeoffs, a major advantage of on-orbit WFS&C is the ability to update DM commands more frequently without relying on the continuous availability of communication links with the ground stations. In turn, this can help to relax

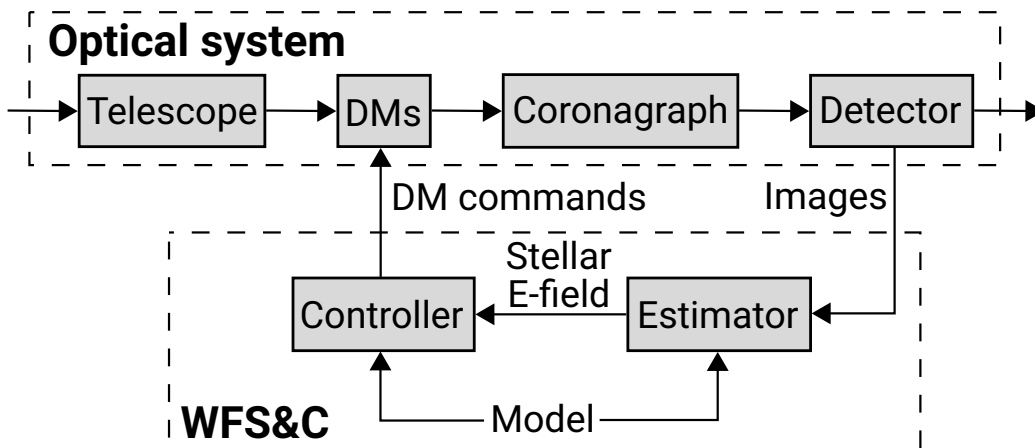


Figure 1. Closed-loop coronagraphic WFS&C uses images from the science camera, rather than an external wavefront sensing instrument, to estimate the E-field from the host star within the dark zone and drive it toward zero. Model-based WFS&C algorithms use a numerical model of the coronagraph to solve an inverse problem for the unknown E-field and corresponding DM correction, respectively.

observatory-level wavefront stability requirements by enabling the high-contrast dark zone to be maintained over shorter time intervals. Successful deployment of an on-orbit architecture is predicated on the availability of sufficient computational resources. However, radiation-hard, space-qualified computing hardware lags behind conventional hardware by decades and is extremely resource-limited, posing a substantial capability gap if the current algorithms are expected to be deployed on-orbit on future missions like the Large UV/Optical/Infrared Surveyor (LUVOIR) and Habitable Exoplanet Observatory (HabEx) flagship mission concepts formulated for the Astro2020 Decadal Survey.^{5,6}

As a case in point, the current state-of-the-art testbed for coronagraph laboratory demonstrations, the Decadal Survey Testbed (DST) at NASA’s Jet Propulsion Laboratory, has successfully demonstrated 3.82×10^{-10} instrumental contrast over an annular dark zone extending from $3 \lambda_0/D$ to $8 \lambda_0/D$ using two DMs each with 48×48 actuators.⁷ However, the baseline requirement for the HabEx mission is two 64×64 DMs, and a search area with a maximal outer radius of $32 \lambda_0/D$ - nearly double the total actuator count and a factor of 4 increase in search radius. Because the number of detector pixels in the dark zone scales with the area of the dark zone rather than its radius, this corresponds to an increase by nearly a factor of 32 in the worst-case Jacobian dimensionality. For LUVOIR Architecture “A”, with a pair of 128×128 DMs and a dark zone with a $64\lambda_0/D$ maximal outer radius, the corresponding worst-case increase is nearly a factor of 512.

An alternative to the analytical minimization approach of SM and EFC is to minimize the wavefront control cost function iteratively using gradient-based optimization. This eliminates the need to construct the Jacobian altogether, but instead requires a method for accurately and efficiently evaluating the gradient of the cost function with respect to the DM update. Algorithmic differentiation (AD), and in particular reverse-mode algorithmic differentiation (RMAD), is one such method.⁸ Recently, we formulated an algorithm based on this approach, which we refer to here as algorithmic differentiation penalty stroke minimization (AD-PSM).⁹ Our proof-of-concept simulations with AD-PSM using a small-angle design for the LUVOIR coronagraph indicated superior computational efficiency and comparable starlight suppression performance in comparison with SM for DM formats as small as 50×50 actuators, when taking into account the cost of calculating the Jacobian once prior to the WFS&C loop. However, while the CPU time and memory consumption of SM grew superlinearly with actuator count, the increase in both for AD-PSM was negligible. For 128×128 actuators, the largest format that we considered, AD-PSM utilized 95% less memory and overall CPU time than SM. This suggests that iterative methods are a promising alternative to analytical techniques that overcome the technical hurdle posed by the computational demands of wavefront control with high actuator counts and large dark zones in an on-orbit WFS&C scenario.

In this paper, we report on the first experimental demonstration of two algorithmic differentiation-based wavefront control algorithms, AD-PSM and algorithmic differentiation electric field conjugation (AD-EFC), using the High Contrast Imager for Complex Aperture Telescopes (HiCAT) at the Space Telescope Science Institute (STScI) in Baltimore, MD. We benchmark the contrast performance of AD-PSM and AD-EFC as a function several key parameters including regularization and the termination tolerance of the nonlinear optimizer, and compare it to SM and EFC.

This paper is structured as follows. In Section 2, we review concepts from our earlier work, including algorithmic differentiation and the mathematical formulation of AD-PSM and AD-EFC. In Section 3, we provide an overview of HiCAT and discuss our experimental setup, including algorithm implementation details that are pertinent to our demonstration. In Section 4, we present and discuss our experimental results. Finally, in Section 5, we present our conclusions and discuss our future work.

1.1 Notation

In this paper, our principal concern is with algorithms that operate on discrete vector-valued quantities, which are represented in boldface. Many of these quantities vary as a function of control iteration, and are denoted with the subscript k . These may be truly discrete, such as the vector \mathbf{a}_k of DM actuator command updates, or may represent discretizations of functions of continuous spatial variables, such as E-fields. We denote \mathbf{x} as a column vector, its transpose \mathbf{x}^T as a row vector, and $\|\mathbf{x}\|$ as its Euclidean length. For complex-valued quantities, \dagger denotes the Hermitian transpose. Scalar quantities are denoted in ordinary (i.e., non-boldface) typographic weight.

By convention, the derivative of a scalar with respect to a column vector is a row vector, i.e.,

$$\frac{\partial J}{\partial \mathbf{x}} = \left[\frac{\partial J}{\partial \mathbf{x}[1]} \quad \cdots \quad \frac{\partial J}{\partial \mathbf{x}[N]} \right], \quad (1)$$

where $\mathbf{x}[n]$ is the n th element of \mathbf{x} . Consequently, the derivative of a column vector with respect to another column vector is a matrix of row vectors:

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial \mathbf{y}[1]}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial \mathbf{y}[M]}{\partial \mathbf{x}} \end{bmatrix}. \quad (2)$$

2. WAVEFRONT CONTROL USING ALGORITHMIC DIFFERENTIATION

The goal of the WFS&C loop in coronagraphy is to drive starlight within the dark zone toward zero over time so that a faint orbiting companion becomes detectable against the reduced background. Each iteration of the WFS&C loop, indexed by the integer k , consists of two steps: an estimation step, in which the aberrated E-field $\widehat{\mathbf{E}}_{\text{ab},k}$ within the dark zone is estimated from focal-plane intensity measurements, and a control step, in which the DM correction is updated to reduce the energy in $\widehat{\mathbf{E}}_{\text{ab},k}$. In this paper, we focus principally on the control step.

Modern model-based wavefront control algorithms find the DM correction update \mathbf{a}_k by minimizing some cost function $J_k(\mathbf{a}_k; \widehat{\mathbf{E}}_{\text{ab},k})$ with respect to \mathbf{a}_k . Usually, J_k is constructed to trade off between minimizing starlight and minimizing the size of the correction, which helps to regularize the problem and stabilize the solution. In general, the true relationship between the starlight in the dark zone and the DM correction is highly nonlinear and nonconvex, owing to the fact that the DMs impart phase-only corrections of the form $\exp\{i\phi_{\text{DM}}\}$ at or near the coronagraph entrance pupil. However, when the optical aberrations are small, we can approximate the true E-field in the coronagraph entrance pupil with a first-order Taylor series expansion about the desired DM update. In this case, the corrected E-field in the dark zone has the form

$$\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k) \approx \widehat{\mathbf{E}}_{\text{ab},k} + \mathbf{E}_{\text{DM},k}(\mathbf{a}_k), \quad (3)$$

where $\widehat{\mathbf{E}}_{\text{ab},k}$ is the estimate of the aberrated dark-zone E-field produced by the estimation step, and $\mathbf{E}_{\text{DM},k}$ is the E-field resulting from the unknown update to the DM correction. We can also write $\mathbf{E}_{\text{DZ},k}$ in the form

$$\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k) = \widehat{\mathbf{E}}_{\text{ab},k} + \mathbf{G}_k \mathbf{a}_k, \quad (4)$$

where $\mathbf{G}_k \triangleq \partial \mathbf{E}_{\text{DZ},k} / \partial \mathbf{a}_k$ is the Jacobian matrix with dimensions $N_{\text{pix}} \times N_{\text{act}}$, N_{pix} is the number of dark-zone pixels, and N_{act} is the total number of controllable DM actuators. The intensity from the corrected E-field, integrated over the dark zone, can be written in terms of the Jacobian as

$$\|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k)\|^2 = \|\mathbf{G}_k \mathbf{a}_k\|^2 + \|\widehat{\mathbf{E}}_{\text{ab},k}\|^2 + 2 \text{Re}\left\{ \widehat{\mathbf{E}}_{\text{ab},k}^\dagger \mathbf{G}_k \right\} \mathbf{a}_k. \quad (5)$$

This is a purely quadratic function of \mathbf{a}_k , meaning that under this approximation, there exists a unique, optimal DM correction that minimizes the dark-zone starlight.

SM and EFC utilize the relationship in Eq. (5) to derive closed-form expressions for this optimal correction in terms of \mathbf{G}_k that can be evaluated by solving a linear system of dimension $N_{\text{act}} \times N_{\text{act}}$, as illustrated in Figure 2. The only remaining task is to construct \mathbf{G}_k . Typically, a computer model of the coronagraph is used to predict the individual impact of each DM actuator on $\mathbf{E}_{\text{DZ},k}$, which corresponds to building \mathbf{G}_k column-by-column. Accordingly, this involves evaluating the coronagraph model at least N_{act} times in total, which is on the order of thousands (with currently-available DM technology) to tens of thousands (with DM formats baselined for LUVOIR). As we show in Appendix A, this is equivalent to minimizing the cost function using Newton's method with an exact Hessian matrix. We discuss further in Section 2.3.

The Fast Linear Least-Squares Coronagraph Optimization (FALCO) package¹⁰ mitigates this problem by implementing a model specially optimized for Jacobian evaluation that propagates only a small region of the entrance pupil locally surrounding each DM actuator. The Jacobian can also be estimated empirically using the self-coherent camera (SCC),³ Zernike phase contrast,¹¹ or the expectation-maximization (EM) algorithm.¹²

Alternatively, we can find the DM correction by minimizing the cost function J_k with respect to \mathbf{a}_k iteratively, rather than analytically, using gradient-based nonlinear optimization as shown in Figure 3. To do so eliminates the need to evaluate the Jacobian matrix entirely, but requires a way of calculating the gradient vector $\partial J_k / \partial \mathbf{a}_k$. RMAD provides a way of doing so that is both computationally efficient and accurate, in the sense that the derivatives computed by RMAD are accurate to machine precision and do not utilize finite-difference approximations.⁸

The basic principle of RMAD is that any function that can be written down as a sequence of differentiable operations, called the forward model, can be transformed mechanically to construct a related function, called the adjoint model, that evaluates the derivative of the forward model with respect to any of the intermediate variables encountered during its evaluation, including its inputs. The derivative evaluated around any given value of the inputs is found by evaluating the forward model, passing the values of the intermediate values as parameters to the adjoint model, and then evaluating the adjoint model. If the output of the forward model is a scalar, then the derivatives with respect to all intermediate variables and inputs are computed simultaneously using only a single forward model and single adjoint model evaluation, regardless of the dimensionality of the derivative variables. Figure 4 illustrates this procedure for the wavefront control cost function J_k .

There are several advantages to adopting an iterative approach. First, it reduces up-front computation by eliminating the cost of precomputing the Jacobian matrix. Second, it scales much more gently as a function of problem size because it only involves operations on vectors rather than matrices, and increases in the dimensions of separate variables scale additively, rather than multiplicatively.⁹ In our earlier work, we showed that the number of floating-point operations and storage required by SM and EFC scale as N_{act}^3 and N_{act}^2 , respectively. For AD-PSM and AD-EFC, both scale as N_{act} . Finally, because control solutions are found by differentiating through the coronagraph model itself rather than a matrix representation, the model is always linearized around the current DM state, and it becomes trivial to update the model between WFS&C iterations as new information, such as data from a low-order wavefront sensor (LOWFS), becomes available. Both of these ultimately improve the convergence rate of the WFS&C loop to deep contrast.

The fundamental tradeoff is that by adopting an iterative method, one sacrifices a closed-form solution in order to gain better asymptotic computational efficiency. For small problem sizes or when the Jacobian can be evaluated once and reused for many WFS&C iterations or over a series of many WFS&C experiments, the analytic solutions from SM and EFC can be faster to evaluate, neglecting the up-front cost of computing the Jacobian. Because of their superior asymptotic efficiency, however, iterative methods become more attractive for

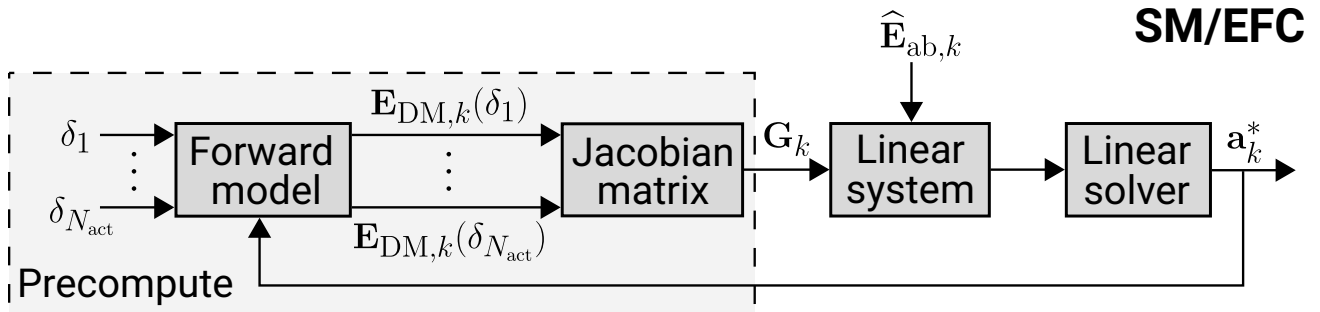


Figure 2. In EFC and SM, the Jacobian matrix is precomputed by using a computer model the coronagraph to predict the effect of an update to each of the N_{act} DM actuators individually on the dark-zone E-field, here represented by the Kronecker δ functions δ_n , where $\delta_n[i] = 1$ if $i = n$ and zero otherwise. The Jacobian \mathbf{G}_k and the estimate of the aberrated dark-zone E-field $\hat{\mathbf{E}}_{\text{ab},k}$ together are used to construct a linear system of equations whose solution is the desired DM update \mathbf{a}_k^* . We show in Appendix A that this is equivalent to minimizing the wavefront control cost function using Newton’s method, and discuss further in Section 2.3.

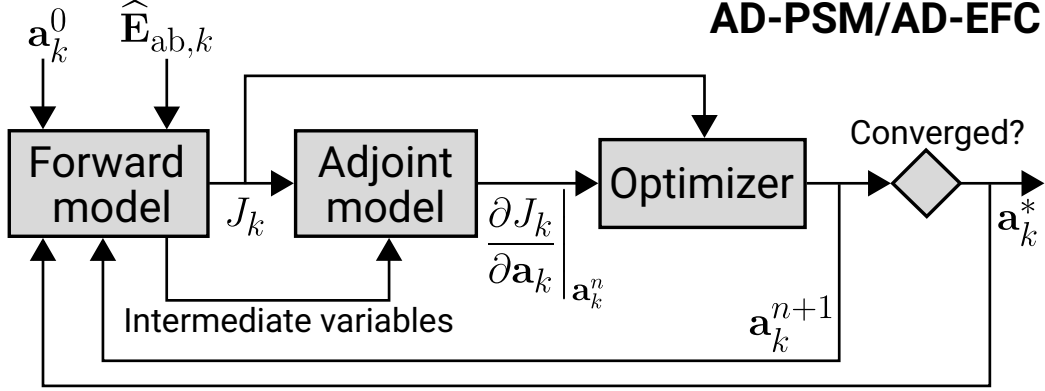


Figure 3. Our algorithmic differentiation-based wavefront controllers AD-PSM and AD-EFC use RMAD to differentiate the wavefront control cost function J_k with respect to the DM correction update \mathbf{a}_k , yielding the gradient vector $\partial J_k / \partial \mathbf{a}_k$ evaluated at the current iterate \mathbf{a}_k^n . A nonlinear optimization algorithm calculates a new iterate \mathbf{a}_k^{n+1} that reduces the value of the cost function, i.e., $J_k(\mathbf{a}_k^{n+1}) \leq J_k(\mathbf{a}_k^n)$. This procedure is repeated until the gradient becomes sufficiently small, indicating that the solution \mathbf{a}_k^* is at or near a local minimum of the cost function. A starting guess for the solution \mathbf{a}_k^0 as well as the aberrated E-field $\widehat{\mathbf{E}}_{ab,k}$ are the input parameters.

cases where the DM actuator count, the dark zone pixel count, or both, are very large, as will be the case for both LUVOIR and HabEx.

We conclude this section with a brief discussion of the merits of the small-aberration approximation. There has been recent interest in fully nonlinear control algorithms that do not utilize a first-order Taylor series approximation of the complex field in the coronagraph entrance pupil, and instead minimize starlight using the true, nonlinear phase effects of the DMs.^{13,14} While these methods in principle have the potential to enable convergence to deep contrast over very few WFS&C iterations, in real systems the convergence rate of the WFS&C loop is limited by noise, estimation error, and model mismatch rather than the accuracy of the first-order approximation to the pupil-plane E-field. There are also practical drawbacks that make them less attractive than

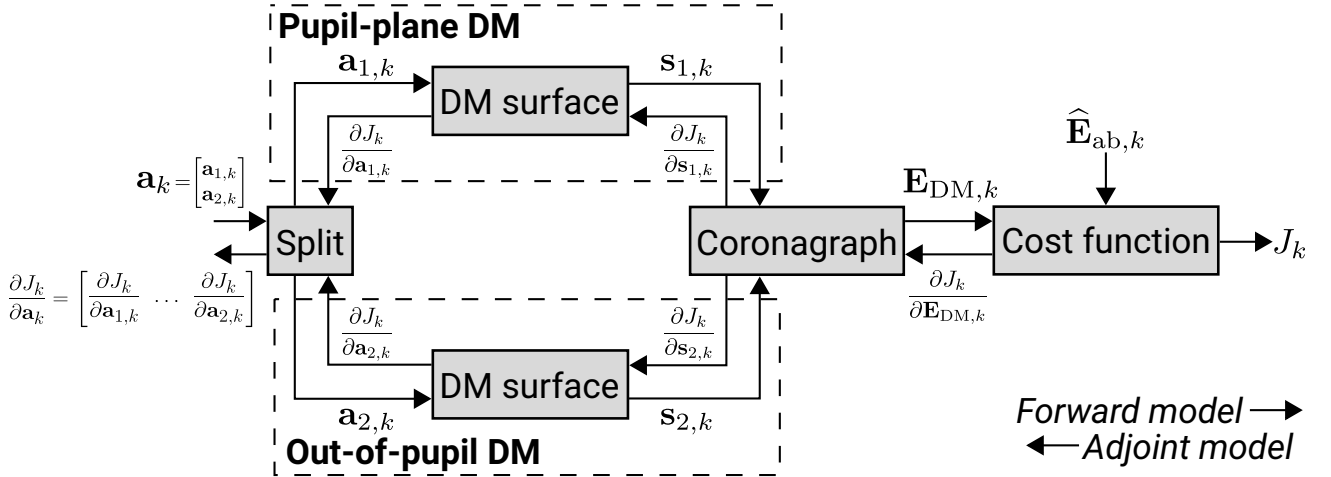


Figure 4. The forward model for the wavefront control problem maps DM command updates \mathbf{a}_k to a scalar cost function J_k . The DM command vector is split into independent command vectors $\mathbf{a}_{1,k}$ and $\mathbf{a}_{2,k}$ for the pupil-plane and out-of-pupil DM, respectively. These are mapped onto DM surfaces $\mathbf{s}_{1,k}$ and $\mathbf{s}_{2,k}$ using the model in Appendix B, and propagated through an end-to-end coronagraph model to predict the resulting dark-zone E-field $\mathbf{E}_{DM,k}$. Reverse-mode algorithmic differentiation transforms the forward model into an adjoint model that backpropagates the partial derivatives of J_k with respect to each intermediate variable $\mathbf{a}_{1,k}$, $\mathbf{a}_{2,k}$, $\mathbf{s}_{1,k}$, $\mathbf{s}_{2,k}$, and $\mathbf{E}_{DM,k}$ in reverse order, starting from the output on the right. The derivatives with respect to the individual DM command vectors $\mathbf{a}_{1,k}$ and $\mathbf{a}_{2,k}$ are concatenated to form the full gradient vector for optimization.

focal-plane algorithms such as SM, EFC, and our proposed controllers AD-PSM and AD-EFC. First, as described earlier, stellar dark-zone intensity is a nonconvex function of DM actuation, making these methods susceptible to convergence to suboptimal local minima in any given WFS&C iteration. Second, because the corrected focal-plane E-field can no longer be linearly decomposed into separate terms resulting from the aberrations and the DM correction, it is not possible for nonlinear controllers to utilize focal-plane E-field estimates, and instead must estimate the E-field in the entrance pupil. This involves solving a phase retrieval problem that is relatively more complicated than focal-plane estimation methods currently in use.

2.1 Stroke minimization: from Lagrange multipliers to penalty method

The SM algorithm finds the smallest DM correction that achieves a desired level of stellar intensity, $I_{\text{target},k}$, integrated over the dark zone.⁴ It is solved by finding the stationary point of the Lagrangian function

$$\mathcal{L}_{\text{SM},k} = \|\mathbf{a}_k\|^2 + \mu_k \left(\|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k)\|^2 - I_{\text{target},k} \right), \quad (6)$$

i.e., a point such that $\partial \mathcal{L}_{\text{SM},k} / \partial \mathbf{a}'_k = 0$, where \mathbf{a}_k is the DM actuator command update, $\mathbf{E}_{\text{DZ},k}$ is the corrected E-field in the dark zone, $\mathbf{a}'_k \triangleq [\mathbf{a}_k^T \ \mu_k]^T$, and μ_k is the Lagrange multiplier. Because this stationary point is a saddle point, it cannot be reached by minimizing $\mathcal{L}_{\text{SM},k}$ directly with respect to \mathbf{a}'_k . Instead, one chooses a fixed starting value for the Lagrange multiplier, μ_k^0 , and minimizes $\mathcal{L}_{\text{SM},k}$ with respect to \mathbf{a}_k to find a corresponding DM solution \mathbf{a}_k^0 . If the constraint $\|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k^0)\|^2 \leq I_{\text{target},k}$ is not satisfied, a larger value $\mu_k^1 > \mu_k^0$ is selected and this procedure is repeated.

The AD-PSM algorithm is based on the same principle, but instead utilizes the cost function

$$J_{\text{PSM},k} = \|\mathbf{a}_k\|^2 + \eta_k \left(\|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k)\|^2 - I_{\text{target},k} \right)^2, \quad (7)$$

where η_k is a penalty parameter. The minimum of $J_{\text{PSM},k}$ with respect to \mathbf{a}_k is coincident with the stationary point of $\mathcal{L}_{\text{SM},k}$.¹⁵

In SM, $\arg \min_{\mathbf{a}_k} \mathcal{L}_{\text{SM},k}$ has an explicit solution in terms of the Jacobian matrix \mathbf{G}_k for any fixed value μ_k^n of the Lagrange multiplier:

$$\mathbf{a}_k^n = - \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \frac{1}{\mu_k^n} \mathbb{I} \right)^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \widehat{\mathbf{E}}_{\text{ab},k} \right\}, \quad (8)$$

where \mathbb{I} is the identity matrix and $\widehat{\mathbf{E}}_{\text{ab},k}$ is the estimate of the aberrated E-field. Finding the SM solution using an iterative optimization algorithm requires $\mathcal{L}_{\text{SM},k}$ to be minimized with respect to \mathbf{a}_k for each value of μ_k^n , which is potentially expensive if many values of μ_k are evaluated before the stationary point is located. On the other hand, minimizing $J_{\text{PSM},k}$ finds the solution directly, with the penalty parameter η_k encoding the relative importance of minimizing actuator stroke and driving the integrated intensity toward the target. The tradeoff is that $J_{\text{PSM},k}$ becomes increasingly ill-conditioned as η_k tends to infinity, which slows down the convergence of the optimization algorithm.¹⁵

2.2 Electric field conjugation

The EFC algorithm attempts to drive the dark zone E-field toward a target, with Tikhonov regularization to mitigate ill-conditioning. Its cost function for a single correction wavelength is given by

$$J_{\text{EFC},k} = \|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k) - \mathbf{E}_{\text{target},k}\|^2 + \|\mathbf{\Gamma}_k \mathbf{a}_k\|^2, \quad (9)$$

where $\mathbf{\Gamma}_k$ is the Tikhonov regularization matrix. In the most common case, one chooses $\mathbf{\Gamma}_k = \alpha_k \mathbb{I}$ and $\mathbf{E}_{\text{target},k} = 0$, making EFC identical to SM with a fixed Lagrange multiplier $\mu_k = 1/\alpha_k^2$. For this case, its solution can be obtained using Eq. (8). In general, the solution for the Jacobian-based formulation of EFC is given by

$$\mathbf{a}_k^* = - \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right)^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \left(\widehat{\mathbf{E}}_{\text{ab},k} - \mathbf{E}_{\text{target},k} \right) \right\}. \quad (10)$$

Here, unlike SM, the regularization constant is fixed, making the problem amenable to an iterative algorithm. Our variant, AD-EFC, is identical to EFC except that the solution is obtained by iteratively minimizing $J_{\text{EFC},k}$ in each WFS&C iteration. The RMAD adjoint model for $J_{\text{EFC},k}$ is provided in Appendix C.

While $\mathbf{E}_{\text{target},k} = 0$ is a common choice for generating high-contrast dark zones, other choices can occasionally be useful. In Section 3.3, we show how one can specify $\mathbf{E}_{\text{target},k}$ to be a plane wave with varying piston phase in order to generate phase diversity for E-field estimation.

2.3 Relationship to Newton’s method

Finding solutions for EFC and SM using the Jacobian matrix is equivalent to minimizing their respective cost functions with respect to \mathbf{a}_k using Newton’s method, as shown in Appendix A. Newton’s method is a second-order optimization technique that utilizes second-derivative information about the cost function, given by the local Hessian matrix $\mathbf{H}_k(\mathbf{a}_k^n)$ at any point \mathbf{a}_k^n in the DM command parameter space. Given an initial guess for the solution \mathbf{a}_k^0 , the full Newton update is given by¹⁵

$$\mathbf{a}_k^1 = \mathbf{a}_k^0 - \mathbf{H}_k^{-1}(\mathbf{a}_k^0) \left. \frac{\partial J_k}{\partial \mathbf{a}_k^T} \right|_{\mathbf{a}_k = \mathbf{a}_k^0}. \quad (11)$$

For cost functions that are exactly quadratic, including EFC and SM, Newton’s method converges in a single iteration.

For general numerical optimization problems, Newton’s method is rarely used in practice because forming the Hessian matrix explicitly is expensive. On the other hand, quasi-Newton methods such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, or the limited-memory BFGS (L-BFGS) variant, can approximate \mathbf{H}_k^{-1} using changes in $\partial J_k / \partial \mathbf{a}_k^T$ over successive optimization iterations. As a consequence, they are substantially less computationally expensive. Although quasi-Newton algorithms do not converge as rapidly as Newton’s method, and in particular can converge slowly for poorly-conditioned problems, they are nonetheless superior to purely first-order methods such as steepest descent.¹⁵

As we show in Appendix A, the Hessian matrix for EFC and SM has an analytic expression in terms of the Jacobian \mathbf{G}_k given by $\mathbf{H}_k = 2 \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{C} \right)$, where \mathbf{C} is a symmetric, positive-definite matrix given by $\mathbf{\Gamma}_k^T \mathbf{\Gamma}_k$ for EFC and \mathbb{I} / μ_k^n for SM. Our approach is to replace this full Newton iteration, requiring computation of the Jacobian, by a series of cheaper quasi-Newton iterations instead, requiring only computation of the gradient $\partial J_k / \partial \mathbf{a}_k^T$, which we achieve using algorithmic differentiation.

3. EXPERIMENTAL SETUP

In this section, we provide an overview of the HiCAT testbed and provide details about the implementation of AD-PSM and AD-EFC, the reference Jacobian-based implementations of SM and EFC, and the E-field estimation algorithm used in the estimation step of the WFS&C loop.

3.1 The HiCAT testbed

HiCAT is a testbed dedicated to technology demonstrations for coronagraphy on segmented-aperture space observatories, with the intent of being directly traceable to a future LUVOIR-like mission. These technologies include Lyot coronagraphy, high-order WFS&C for generating and stabilizing dark zones, and low-order wavefront sensing (LOWFS).¹⁶ HiCAT operates in a mid-contrast regime (10^{-7} to 10^{-8}) which approaches the limit achievable outside of a vacuum environment, and is equipped with two Boston Micromachines Kilo-DMs for high-order sensing and control with 952 actuators each, making it suitable for our proof-of-concept demonstrations. One DM is placed in a plane conjugate to the HiCAT entrance pupil, while the second DM is placed approximately 300 mm farther along the optical axis, corresponding to a Fresnel number $N_F \approx 98$ at a wavelength of 638 nm. This configuration enables simultaneous control of amplitude and phase aberrations over a dark zone that extends over both halves of the image plane. We conducted our experiments using a Thorlabs MCLS1 laser diode source, which emits monochromatic light with a central wavelength $\lambda_0 = 638$ nm.

HiCAT additionally has an IrisAO segmented deformable mirror with 37 hexagonal segments with controllable piston/tip/tilt to act as a telescope pupil simulator and to inform experimental efforts devoted to segment-level tolerancing and stabilization.^{17,18} Figure 5 shows a simplified system layout of HiCAT including the primary imaging beam path as well as several additional beam paths used by the LOWFS and metrology subsystems.

The ultimate planned operational mode for HiCAT is an apodized pupil Lyot coronagraph (APLC)¹⁹ designed specifically to optimize contrast with the IrisAO segmented aperture. However, due to substantial technical challenges associated with properly aligning the APLC pupil-plane masks in the presence of geometrical distortion between the various planes, our experiments on HiCAT utilized a simpler classical Lyot coronagraph (CLC) design with a circularized pupil outer edge. Figure 6 shows an overlay of the CLC pupil masks along with a simulated stellar PSF. Figure 7 shows example experimental PSFs obtained before and after closed-loop WFS&C using SM along with the corresponding DM commands.

For more details about the optical design of HiCAT and recent milestone achievements, we refer elsewhere.²⁰⁻²⁷

3.2 Algorithm implementation

We developed a custom numerical model of HiCAT using Python that included a handwritten adjoint model. To facilitate testing, our model was comprised of several sub-modules each with its own individual forward and adjoint model:

1. A model to compute the DM surface resulting from a given set of actuator commands using a fast convolutional representation. This is described in further detail in Appendix B.
2. A model to apply the phase corrections imparted by the in-pupil DM (DM1) and out-of-pupil DM (DM2), including the free-space propagation between the two DMs.
3. A model to propagate the E-field after DM correction through the HiCAT CLC. We used the semi-analytical Lyot coronagraph model originally described in Ref. 28.

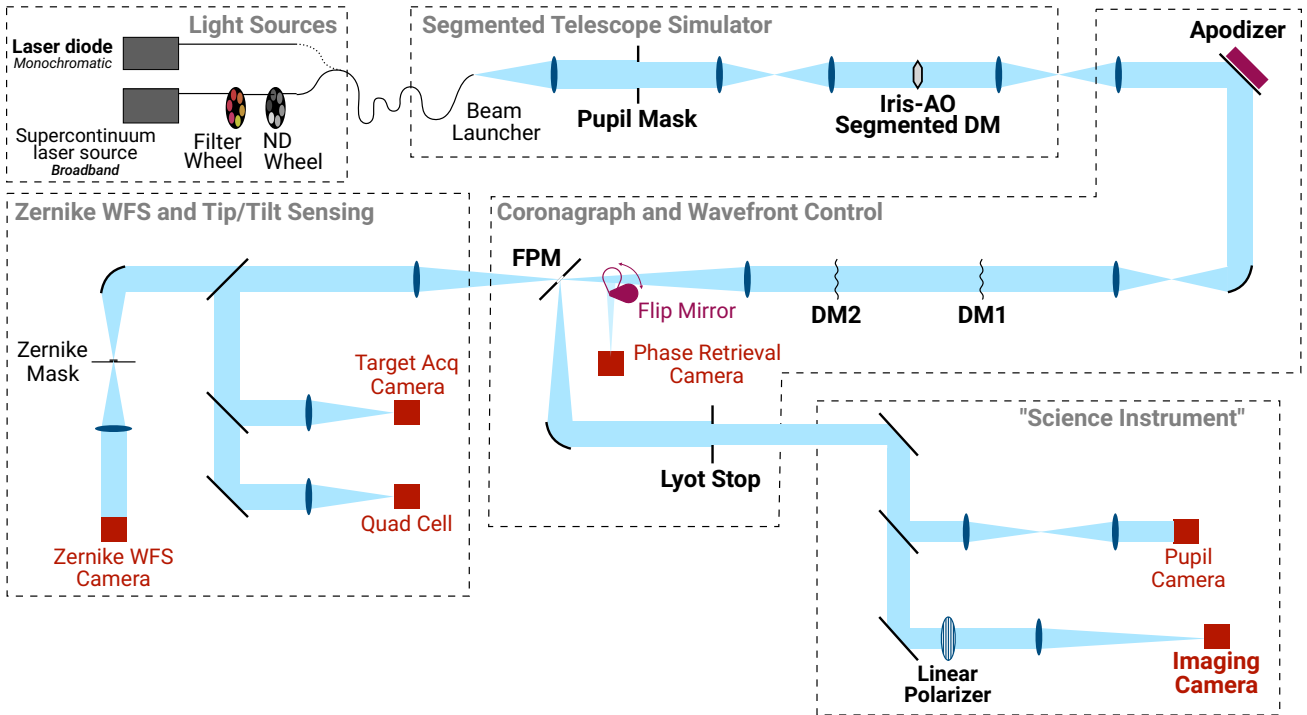


Figure 5. Simplified, partially-unfolded layout of the HiCAT testbed. The elements encountered by the primary imaging beam path are highlighted in bold.

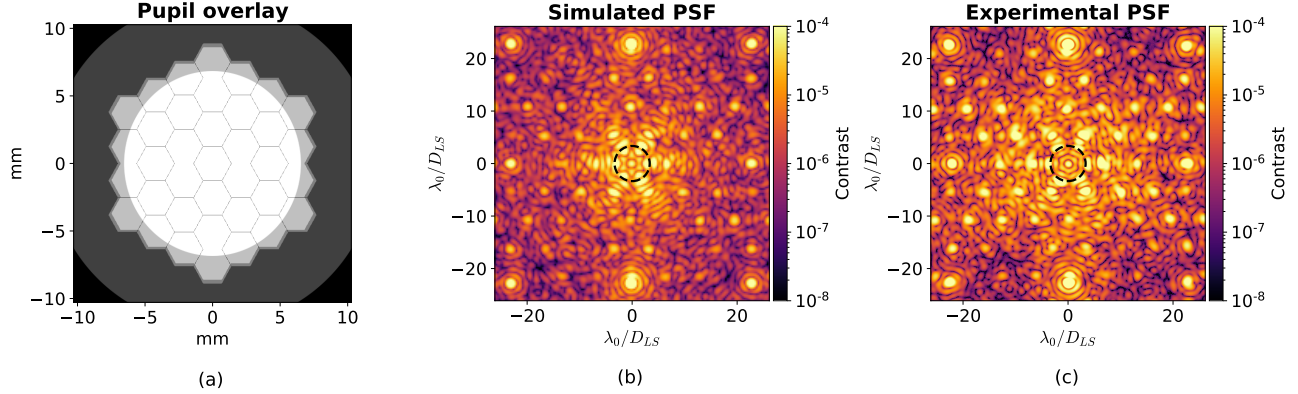


Figure 6. (a) Overlay of HiCAT pupil masks projected onto in-pupil DM plane, including the reflective area of the DM, the entrance pupil mask, the IrisAO segmented aperture, and the Lyot stop. (b) Simulated and (c) experimental coronagraphic PSFs with outline of the geometrical edge of the focal-plane mask.

We refer the reader to our earlier work for detailed descriptions of the operations in the forward and adjoint models for the latter two sub-modules.⁹ A reference numerical model of HiCAT based on the POPPY framework^{29,30} has already been developed and served as a reference for calibrating our custom differentiable model.

3.2.1 Optimization algorithm

We used the limited-memory BFGS (L-BFGS) algorithm,¹⁵ as implemented in the SciPy package,³¹ as the optimization algorithm for AD-PSM and AD-EFC. As discussed in Section 2.3, L-BFGS is a quasi-Newton algorithm, meaning that it uses the gradient vectors collected during optimization to approximate the inverse Hessian matrix \mathbf{H}_k^{-1} of second derivatives, thereby converging to the minimum more rapidly than a purely first-order method such as steepest descent, but with substantially less effort than a full Newton method with an exact Hessian. The ordinary BFGS algorithm forms a dense approximation of \mathbf{H}_k^{-1} using the full collection of gradient vectors, whereas L-BFGS uses an implicit outer-product representation based on only the most recent N_{grad} gradients, where N_{grad} is usually between 10 and 20,¹⁵ making it more memory-efficient for problems where the number of optimization variables is much larger than N_{grad} . For our experiments, we used $N_{\text{grad}} = 10$, the default in SciPy.

All numerical optimization algorithms have a termination criterion that determines when the algorithm has reached a location in the parameter space that is sufficiently close to a local, or preferably global, optimum. For gradient-based algorithms such as L-BFGS, the termination criterion can be defined in terms of the relative improvement in the cost function value between successive optimization iterations, the magnitude of the gradient vector, or both. In the L-BFGS implementation used by SciPy, the termination criterion is set by a tolerance parameter ε defined as the magnitude of the largest element of the gradient:

$$\varepsilon = \max_i \left| \frac{\partial J_k}{\partial \mathbf{a}_k [i]} \right|. \quad (12)$$

As $\varepsilon \rightarrow 0$, the optimization algorithm will expend more effort, in the form of a greater number of optimization iterations, to terminate closer to the local minimum. On the other hand, ε can help to regularize the solution by terminating the optimization algorithm before it converges to an overly aggressive DM correction that would otherwise make the WFS&C loop unstable.

The termination criterion presents an additional consideration not present in SM and EFC, which we recall utilize a closed-form expression for the exact minima of their respective cost functions in any given WFS&C iteration. The value of the tolerance parameter ε should be chosen to optimally trade off between accuracy and computational effort. We discuss this in further detail below.

3.2.2 Scale-invariant cost functions

As the WFS&C loop drives the focal-plane E-field toward zero, the terms in the AD-EFC and AD-PSM cost functions also diminish in magnitude. Consequently, for any fixed ε , the termination criterion will become increasingly loose over time, resulting in solutions that are increasingly suboptimal. The simplest strategy is to choose ε to ensure that the solution is sufficiently close to optimal for the smallest-expected E-field magnitude, with the tradeoff of being overly restrictive in early iterations when the E-field is large. A better strategy is to scale the cost function so that its overall magnitude is invariant to the magnitude of the E-field. For AD-EFC, one such scaling is by the magnitude of the current E-field estimate:

$$J'_{\text{EFC},k} = \frac{1}{\|\hat{\mathbf{E}}_{\text{ab},k}\|^2} J_{\text{EFC},k}, \quad (13)$$

with the scaling factor $\|\hat{\mathbf{E}}_{\text{ab},k}\|^{-2}$ chosen to be independent of the unknown correction \mathbf{a}_k .

For AD-PSM, the situation is slightly more complicated. The intensity penalty $(\|\mathbf{E}_{\text{DZ},k}\|^2 - I_{\text{target},k})^2$ is a fourth-order function of \mathbf{a}_k , whereas the stroke penalty $\|\mathbf{a}_k\|^2$ is only a second-order function. Therefore, even if a scale factor is applied to the entire cost function, the intensity penalty will approach zero more rapidly than the stroke penalty. We found empirically that scaling only the intensity penalty term produced good overall results;

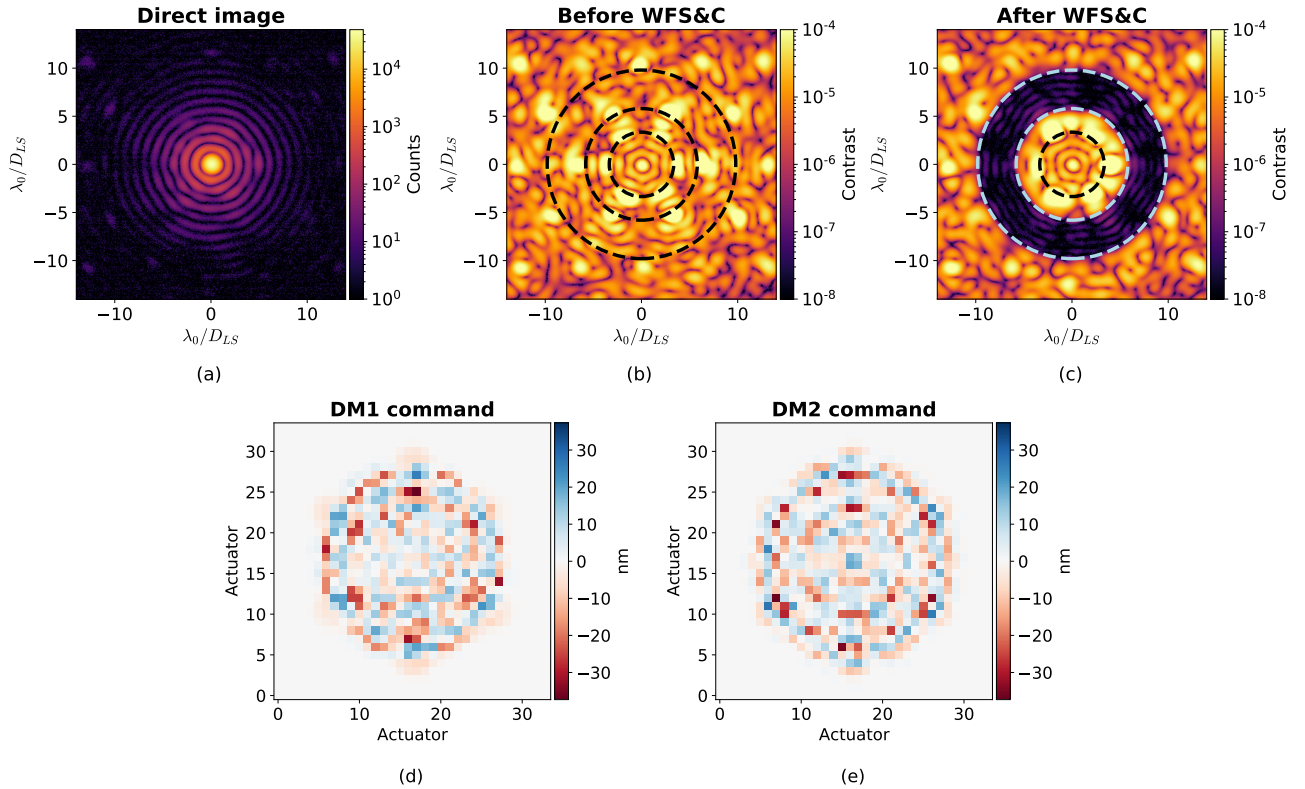


Figure 7. Experimental on-axis images from HiCAT. (a) Non-coronagraphic image, (b) coronagraphic image prior to the WFS&C loop, (c) coronagraphic image after 80 iterations of SM, and corresponding actuator commands for the in-pupil DM (d) and out-of-pupil DM (e). In (b) and (c), the geometrical edge of the focal-plane mask (FPM) with radius $3.34 \frac{\lambda_0}{D_{LS}}$ is shown as well as the inner and outer edges of the dark zone at $5.8 \frac{\lambda_0}{D_{LS}}$ to $9.8 \frac{\lambda_0}{D_{LS}}$, respectively, where D_{LS} is the Lyot stop diameter and $\lambda_0 = 638$ nm.

therefore, the modified AD-PSM cost function is

$$J'_{\text{PSM},k} = \|\mathbf{a}_k\|^2 + \frac{\eta_k}{\left(\|\widehat{\mathbf{E}}_{\text{ab},k}\|^2 - I_{\text{target},k}\right)^2} \left(\|\mathbf{E}_{\text{DZ},k}(\mathbf{a}_k)\|^2 - I_{\text{target},k}\right)^2. \quad (14)$$

As before, the scaling factor depends only on quantities that are fixed in any given WFS&C iteration and independent of the DM correction.

Using the scaled cost functions $J_{\text{EFC},k}$ and $J_{\text{PSM},k}$, we can choose the value of ε as, for instance, a fixed multiple of the value of the cost function with $\mathbf{a}_k = 0$. In Section 4, we explore the performance of AD-PSM and AD-EFC for different combinations of ε and regularization η_k .

3.3 Estimation algorithm

We used the pairwise probe estimator² for the estimation step in all experiments. The pairwise estimator forms a least-squares estimate $\widehat{\mathbf{E}}_{\text{ab},k}$ of the focal-plane E-field $\mathbf{E}_{\text{ab},k}$ by applying a series of P probing DM commands \mathbf{u}_p to generate probing E-fields $\mathbf{E}_{\text{DM},k}(\mathbf{u}_p)$ that interfere with $\mathbf{E}_{\text{ab},k}$. The data vector for the least-squares estimate is formed by differencing the images resulting from $\mathbf{E}_{\text{DM},k}(\mathbf{u}_p)$ and $\mathbf{E}_{\text{DM},k}(-\mathbf{u}_p)$. The estimate of the m th pixel in the dark zone is then found by finding the least-squares solution of the system

$$\begin{bmatrix} \mathbf{I}_{k,1}^+[m] - \mathbf{I}_{k,1}^-[m] \\ \vdots \\ \mathbf{I}_{k,P}^+[m] - \mathbf{I}_{k,P}^-[m] \end{bmatrix} = \begin{bmatrix} \text{Re}\{\mathbf{E}_{\text{DM},k}(\mathbf{u}_1)\}[m] & \text{Im}\{\mathbf{E}_{\text{DM},k}(\mathbf{u}_1)\}[m] \\ \vdots & \vdots \\ \text{Re}\{\mathbf{E}_{\text{DM},k}(\mathbf{u}_P)\}[m] & \text{Im}\{\mathbf{E}_{\text{DM},k}(\mathbf{u}_P)\}[m] \end{bmatrix} \begin{bmatrix} \text{Re}\{\mathbf{E}_{\text{ab},k}\}[m] \\ \text{Im}\{\mathbf{E}_{\text{ab},k}\}[m] \end{bmatrix}, \quad (15)$$

where $\mathbf{I}_{k,p}^\pm \triangleq |\mathbf{E}_{\text{ab},k} + \mathbf{E}_{\text{DM},k}(\pm\mathbf{u}_p)|^2$.

We generated four DM probe functions \mathbf{u}_p that were optimized to generate probing fields of the form

$$\mathbf{E}_{\text{DM},k}(\mathbf{u}_p) = \text{sgn}\{\rho_x\} \exp\left\{i\pi\frac{p-1}{4} \text{sgn}\{\rho_x\}\right\}, \quad (16)$$

where $p \in \{1, 2, 3, 4\}$ and $\text{sgn}\{\rho_x\}$ is the sign of the focal-plane x coordinate. The inclusion of the sign of ρ_x is necessary because in the first-order Taylor series expansion of the E-field used by focal-plane WFS&C algorithms, the in-pupil DM imparts a purely imaginary E-field in the pupil plane, meaning that if the coronagraph masks are all purely real, $\mathbf{E}_{\text{DM},k}(\mathbf{u}_p)$ must have anti-Hermitian symmetry. The probe commands can be found by solving an EFC problem with $\widehat{\mathbf{E}}_{\text{ab},k} = 0$ and $\mathbf{E}_{\text{target},k} = \mathbf{E}_{\text{DM},k}(\mathbf{u}_p)$, which will produce nonzero commands only for the pupil-plane DM in order to satisfy the anti-Hermitian symmetry constraint. For SM and EFC, we generated the probes using Jacobian-based EFC, while for AD-PSM and AD-EFC, we generated the probes using AD-EFC. In all cases, we set the Tikhonov regularization parameter to $\alpha_k = 0.7$. The optimizer tolerance for AD-EFC was set to 10^{-5} . Figure 8 shows the probe commands \mathbf{u}_p along with the corresponding magnitude and phase of $\mathbf{E}_{\text{DM},k}(\mathbf{u}_p)$, while Figure 9 shows the probe commands generated by EFC and AD-EFC side-by-side.

4. EXPERIMENTAL RESULTS AND DISCUSSION

We conducted a series of experiments to compare the contrast performance of AD-PSM and AD-EFC relative to SM and EFC, respectively. All experiments used an annular control region extending from $5.8 \lambda_0/D_{LS}$ to $9.8 \lambda_0/D_{LS}$, where $\lambda_0 = 638$ nm is the central wavelength of the Thorlabs MCLS1 laser diode and D_{LS} is the Lyot stop diameter. Each experiment consisted of 80 WFS&C iterations.

For SM and AD-PSM, we chose $I_{\text{target},k} = 0.5\|\widehat{\mathbf{E}}_{\text{ab},k}\|^2$ as the contrast target, i.e., a factor-of-two improvement in spatially-integrated dark zone contrast iteration-over-iteration. For the Lagrange multiplier line search in SM, we let $\mu_k^{n+1} = 1.3\mu_k^n$, which is a moderately small step size in logarithmically-mapped space. For AD-PSM, we tested combinations of the penalty parameter $\eta_k \in \{10, 100, 1000\}$ and the nonlinear optimization convergence tolerance $\varepsilon \in \{10^{-2}, 10^{-3}, 10^{-4}\}$. We determined these values based on extensive WFS&C simulations and prior experience conducting WFS&C experiments on HiCAT.

For AD-EFC, the Tikhonov regularization matrix was selected as $\mathbf{\Gamma}_k = \alpha_k \mathbb{I}$. We tested combinations of the Tikhonov parameter α_k such that $\alpha_k^2 \in \{10^{-2}, 10^{-3}, 10^{-4}\}$ with the same set of ε values as AD-PSM. For each value of α_k , we also conducted a reference EFC experiment.

For both algorithms, we compared the value of the cost function for two different starting guesses for the DM correction: $\mathbf{a}_k^0 = \mathbf{a}_{k-1}^*$, i.e., the solution of the previous WFS&C iteration, and $\mathbf{a}_k^0 = 0$. The starting guess with the lower of the two cost function values was then selected. In all cases the flat starting guess $\mathbf{a}_k^0 = 0$ was ultimately chosen.

In each experiment, we analyzed the iteration-over-iteration improvement in spatially averaged dark zone contrast to determine the iteration at which the WFS&C loop converged close to its final best contrast. We then collected the spatially averaged dark zone contrast from each WFS&C iteration post-convergence and computed the median and 10-90th percentile contrast values. Figure 10 illustrates this procedure.

Figures 11 and 12 show the spatially averaged dark-zone contrast as a function of WFS&C iteration for example AD-PSM and AD-EFC experiments alongside reference data from SM and EFC, respectively, as well as the PSF from the iteration with deepest contrast. In both cases, the convergence properties of AD-PSM and AD-EFC

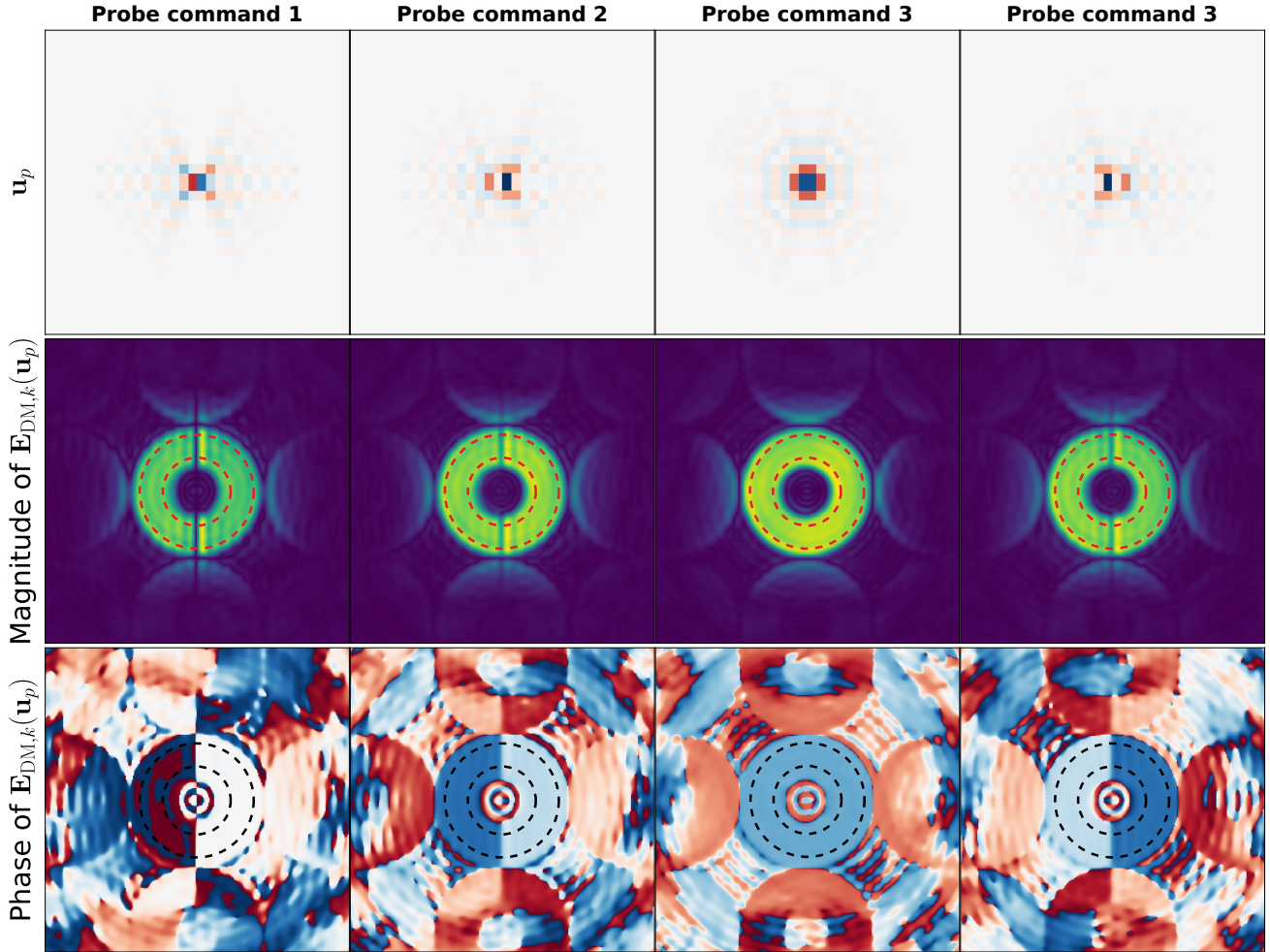


Figure 8. The probe commands \mathbf{u}_p for pairwise estimation were optimized so that the resultant dark-zone E-field $\mathbf{E}_{\text{DM},k}(\mathbf{u}_p) = \text{sgn}\{\rho_x\} \exp\{i\pi \frac{p-1}{4} \text{sgn}\{\rho_x\}\}$, as described in Section 3.3. The inner and outer edges of the dark zone are also shown for reference. The probe commands are close to the inverse Fourier transform of the dark zone geometry (an annulus), modulated by a horizontal sinusoid whose phase angle is proportional to the desired piston phase, and projected onto the pupil-plane DM is modulated. Only the pupil-plane DM is modulated.

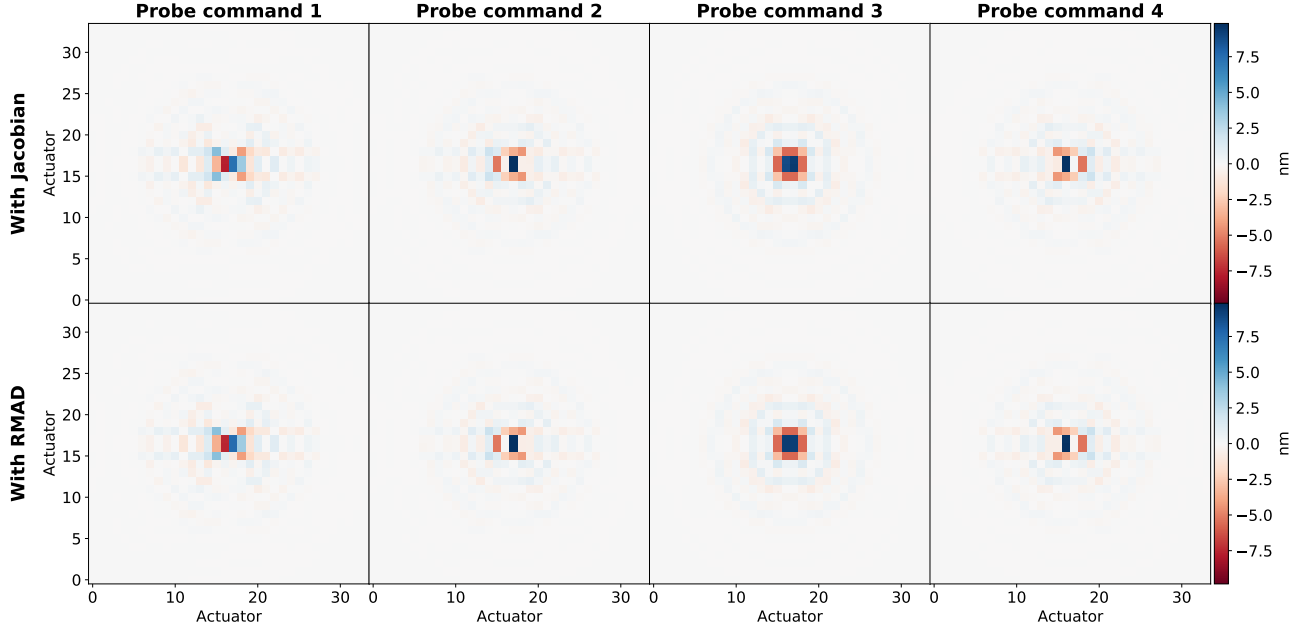


Figure 9. Comparison of probe commands generated by Jacobian-based EFC and AD-EFC as described in Section 3.3.

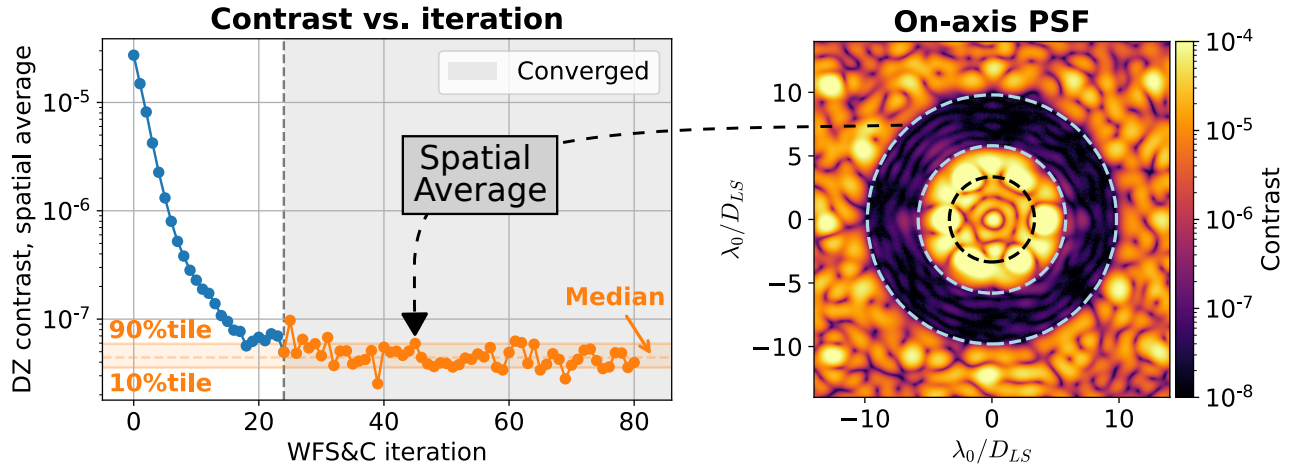


Figure 10. To quantify the contrast performance of each algorithm, we calculate the spatially averaged dark-zone contrast in each WFS&C iteration and analyze it as a function of time to determine the point at which the contrast ceases to continue improving on average. We then collect the spatially averaged dark-zone contrast values from these iterations and calculate their median, 10th percentile, and 90th percentile.

are nearly identical to their Jacobian-based counterparts, validating our proposed approach. Figures 13 and 14 show the statistics of the post-convergence spatially averaged dark-zone contrast achieved with AD-PSM and AD-EFC for each combination of regularization (η_k or α_k) and optimization tolerance ε , respectively, compared to reference experiments with SM and EFC. For all parameter combinations, AD-PSM and AD-EFC equaled the contrast performance of SM and EFC, respectively, within the margin of uncertainty caused by environmental fluctuations and dynamic testbed instabilities.

4.1 Discussion

Our experiments were aimed at exploring a relevant subset of the space of free parameters for each algorithm, namely, the nonlinear optimization convergence tolerance ε , the Tikhonov regularization α_k for AD-EFC, and the

penalty parameter η_k for AD-PSM. In principle, each parameter affects the attainable contrast of the WFS&C loop, but in subtly different ways, which we discuss here.

As discussed in Section 3.2.1, ε determines the effort that the nonlinear optimization algorithm will expend to find a solution close to the true minimum of the cost function. As ε tends toward zero, the solutions obtained using AD-PSM and AD-EFC will asymptotically approach the Jacobian-based solutions from SM and EFC, respectively, but at the cost of a greater number of optimization iterations. This translates to a larger CPU time expenditure, and ultimately a larger time interval between DM updates. On a system such as HiCAT where environmental disturbances cause the aberrated E-field to evolve over time scales on the order of seconds or faster, an excessive delay between the estimation step and application of the DM correction can cause a degradation in achievable contrast. We did not observe a significant degradation in the median spatially averaged contrast for smaller values of ε , but we did observe that for any fixed value of α_k or η_k , smaller values of ε tended to have greater spreads between the 10th and 90th percentile values. In a real spaceborne system, this is unlikely to be a significant consideration because of the much greater E-field stability, and because the total duration of each WFS&C iteration will be dominated by the exposure times needed for the estimation step.

The value ε can also serve as an auxiliary form of regularization, by terminating the optimization algorithm before it reaches an overly aggressive DM correction caused by a noisy E-field estimate, insufficient regularization using α_k or η_k , or both. For instance, in Figure 14, with $\alpha_k^2 = 10^{-4}$, we see that EFC diverged altogether, but AD-EFC did not, and the median spatially averaged contrast improved as ε was reduced. On the other hand, choosing ε too large can impose an effective contrast floor by limiting the ability of the optimization algorithm to converge to appropriately strong corrections. We determined in simulation that this was the case for $\varepsilon > 10^{-2}$.

For EFC and AD-EFC, the Tikhonov regularization α_k trades off between minimizing the energy in the corrected E-field and minimizing the DM actuator stroke in the correction. Smaller values of α_k correspond to more aggressive correction of the E-field. However, because the coronagraph contains actuators that are partially or completely obscured by pupil features, the Hessian matrix $\mathbf{H}_k = 2\left(\text{Re}\left\{\mathbf{G}_k^\dagger \mathbf{G}_k\right\} + \alpha_k^2 \mathbf{I}\right)$ of the EFC cost function (see Section 2.3 and Appendix A) becomes poorly conditioned. This makes the DM correction increasingly sensitive to small perturbations in the estimated E-field, and ultimately makes the WFS&C loop less stable. As we discussed above, for the most aggressive value that we tested, $\alpha_k^2 = 10^{-4}$, the EFC algorithm diverged altogether, but AD-EFC was regularized by the nonzero value of ε . Additionally, we see that the variance of the spatially averaged dark-zone contrast with EFC, as measured by the 10-90th percentile range, was larger for $\alpha_k^2 = 10^{-3}$ than for $\alpha_k^2 = 10^{-2}$. For AD-EFC, for any fixed value of ε , decreasing α_k was also associated with an increase in the variance of the spatially averaged dark-zone contrast, indicating degraded WFS&C loop stability.

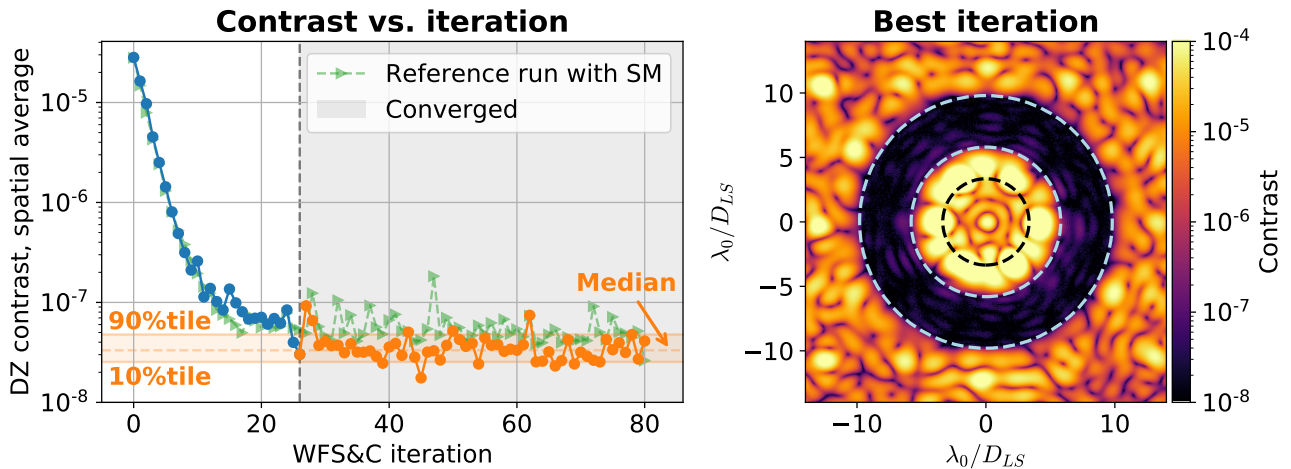


Figure 11. Spatially averaged dark-zone contrast vs. WFS&C iteration for AD-PSM with $\eta_k = 10$ and $\varepsilon = 10^{-4}$. The mean, 10th, and 90th percentile of the converged datapoints in orange, along with a comparison to a reference experiment with SM, are also shown. The on-axis PSF corresponding to the iteration with deepest contrast is displayed to the right.

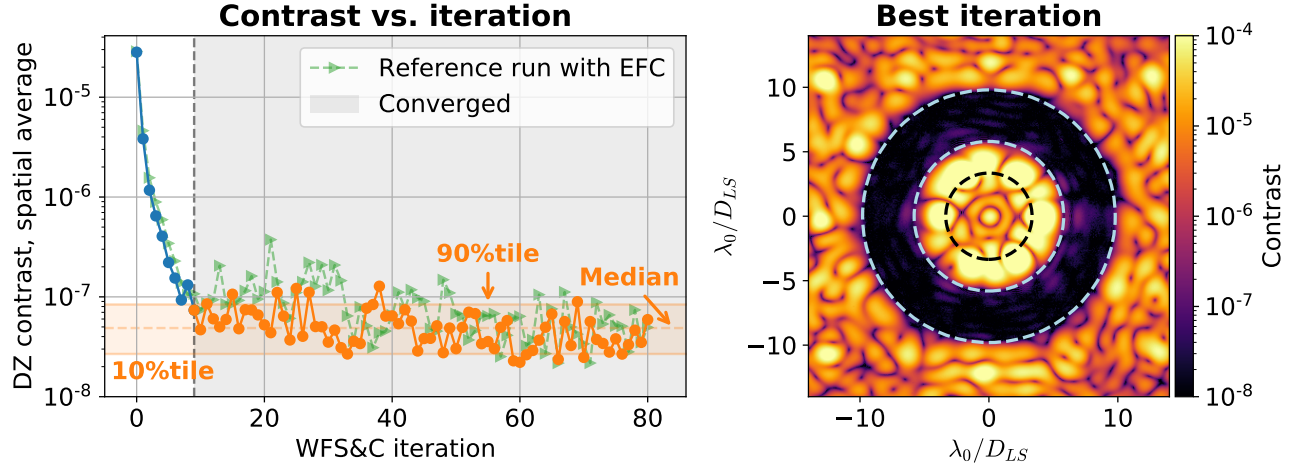


Figure 12. Spatially averaged dark-zone contrast vs. WFS&C iteration for AD-EFC with $\alpha_k^2 = 10^{-2}$ and $\varepsilon = 10^{-4}$. As in Figure 11, the mean, 10th, and 90th percentile of the converged datapoints in orange, a comparison to a reference experiment with EFC, and the on-axis PSF corresponding to the iteration with deepest contrast are also shown.

For AD-PSM and SM, the aggressiveness of the WFS&C control loop is set first and foremost by the targeted energy in the corrected E-field $I_{\text{target},k}$. Any fixed, feasible value of $I_{\text{target},k}$ corresponds to a finite value μ_k^* of the Lagrange multiplier μ_k , which itself corresponds to a fixed value of α_k^2 in EFC. As described in Section 2.1, μ_k^* is found via a line search. On the other hand, the penalty parameter η_k for AD-PSM determines the relative importance of reaching $I_{\text{target},k}$ and minimizing actuator stroke. As η_k tends toward infinity, the correction computed by AD-PSM asymptotically approaches the solution from SM. As such, it has no direct counterpart in SM; however, it can be viewed as a way to tune the aggressiveness of the control loop up to a fixed level imposed by the value of $I_{\text{target},k}$. Our experiments indicated that over the range of values considered, the performance of the WFS&C loop was more or less insensitive to the value of η_k .

In principle, because α_k and ε have similar effects on the performance of AD-EFC, there potentially exists a single combination of the two parameters that is optimal in terms of contrast, or perhaps a continuum of

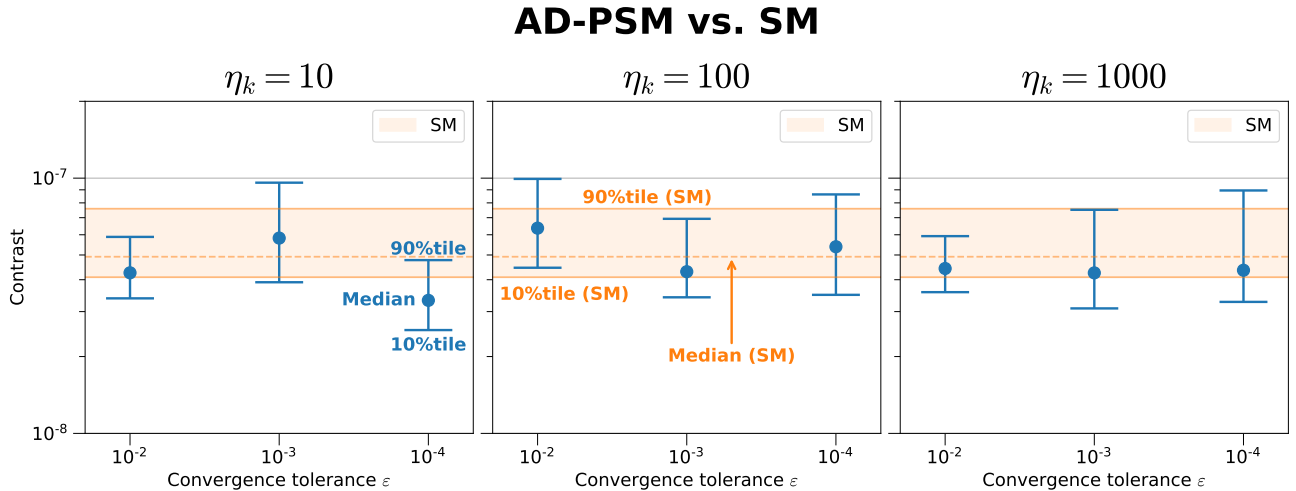


Figure 13. Median, 10th percentile and 90th percentile spatially averaged contrast values achieved by AD-PSM (in blue) as a function of optimizer tolerance, for three different values of the penalty parameter η_k . For comparison, the corresponding contrast values for a reference SM experiment is shown in orange. In all cases, the contrast performance of AD-PSM was equivalent to that of SM within statistical uncertainty.

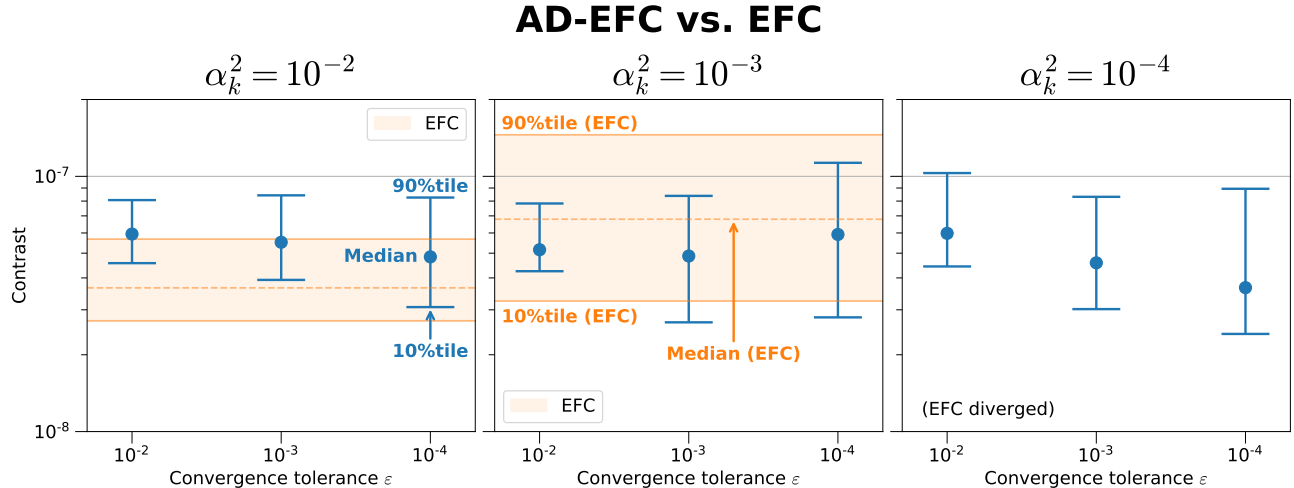


Figure 14. Median, 10th percentile and 90th percentile spatially averaged contrast values achieved by AD-EFC as a function of optimizer tolerance, for three different values of the Tikhonov regularization parameter α_k . For each α_k , we also ran a reference experiment with EFC. The corresponding contrast values for the EFC experiments with $\alpha_k^2 = 10^{-2}$ and $\alpha_k^2 = 10^{-3}$ are shown in orange; the EFC experiment with $\alpha_k^2 = 10^{-4}$ diverged and is not shown. Like AD-PSM in Figure 13, the performance of AD-EFC was comparable to that of EFC for $\alpha_k^2 = 10^{-2}$ and $\alpha_k^2 = 10^{-3}$.

combinations with similar performance, with a change in α_k compensated by a change to ϵ in the opposite direction. The same holds true for AD-PSM. Our experiments exhibited no strongly identifiable trend over the range of values that we considered, which we attribute to the fact that for all parameter contributions, the WFS&C loop was able to reach the contrast floor imposed by environmental, rather than algorithmic, factors.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we reported the first experimental demonstrations of two algorithmic differentiation-based wavefront control algorithms using the HiCAT testbed, AD-PSM and AD-EFC. To within statistical uncertainty, AD-PSM and AD-EFC equaled their Jacobian-based counterparts in dark-zone contrast for all combinations of parameters that we tested. These demonstrations pave the way for future experimental validation at higher contrast.

The analysis in our earlier work indicated that the largest computational gains are realized for DMs with more than 64×64 actuators. The DMs currently in use on HiCAT have 34 actuators across the diameter of the active region, or 952 actuators per DM in total, which is comparatively low. Therefore, our goal was to validate the fundamental capability of AD-PSM and AD-EFC to reach deep contrast, rather than to demonstrate improved computational efficiency. However, we reiterate that at the problem sizes anticipated for future HabEx-like or LUVOIR-like observatories, our approach will be superior in terms of total computing requirements; with the 128×128 DMs baselined for LUVOIR architecture “A”, our analysis indicated more than a tenfold improvement in memory consumption and CPU time compared to Jacobian-based algorithms.

In this work and in our earlier work, we utilized the L-BFGS optimization algorithm to minimize the wavefront control cost function because of its desirable convergence properties compared to first-order optimization methods and low storage requirements. Despite this, L-BFGS is known to converge slowly for poorly-conditioned problems compared to methods that utilize the exact Hessian matrix.¹⁵ In Appendix A, we showed that the Jacobian-based solutions of EFC and SM are equivalent to a single iteration of Newton’s method, with relatively unappealing computational properties apart from convergence rate. However, there exist alternative methods such as truncated Newton algorithms, with excellent convergence properties for quadratic or nearly-quadratic cost functions, that require only the ability to evaluate Hessian-vector products, rather than the Hessian matrix itself.¹⁵ Hessian-vector products can be evaluated using algorithmic differentiation in a similar fashion to gradients. Future work will explore such methods as a potentially faster approach than L-BFGS.

APPENDIX A. EQUIVALENCE OF JACOBIAN-BASED SOLUTIONS AND NEWTON'S METHOD

The cost functions for the EFC algorithm described in Section 2 can be written in the form

$$J_k(\mathbf{a}_k) = \|\mathbf{b}_k\|^2 + \|\mathbf{\Gamma}_k \mathbf{a}_k\|^2, \quad (17)$$

where $\mathbf{\Gamma}_k$ is a regularization matrix. For EFC, $\mathbf{b}_k = \mathbf{E}_{\text{DZ},k}(\mathbf{a}_k) - \mathbf{E}_{\text{target},k}$, while for SM, $\mathbf{b}_k = \mathbf{E}_{\text{DZ},k}(\mathbf{a}_k)$ and $\mathbf{\Gamma}_k = \mathbb{I}/\sqrt{\mu_k}$. For now, we will restrict our attention to the EFC algorithm, but note that the result derived here applies equally as well to minimizing the Lagrangian function for SM with respect to the DM correction \mathbf{a}_k .

Expanding Eq. (17) and recalling that $\mathbf{E}_{\text{DZ},k} = \mathbf{G}_k \mathbf{a}_k + \widehat{\mathbf{E}}_{\text{ab},k}$, the cost function has the form

$$J_{\text{EFC},k}(\mathbf{a}_k) = \mathbf{a}_k^T \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right) \mathbf{a}_k + 2 \mathbf{a}_k^T \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\} + \delta \mathbf{E}_k^\dagger \delta \mathbf{E}_k, \quad (18)$$

where $\delta \mathbf{E}_k \triangleq \widehat{\mathbf{E}}_{\text{ab},k} - \mathbf{E}_{\text{target},k}$ and where we use the fact that \mathbf{a}_k is purely real to discard $\text{Im} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\}$. The Jacobian-based solution is found by finding \mathbf{a}_k such that $\partial J_{\text{EFC},k} / \partial \mathbf{a}_k$ vanishes. We therefore begin by writing down the gradient:

$$\frac{\partial J_{\text{EFC},k}}{\partial \mathbf{a}_k^T} = 2 \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right) \mathbf{a}_k + 2 \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\} = 0. \quad (19)$$

This is a linear system of equations that we can solve for the optimal correction \mathbf{a}_k^* :

$$\mathbf{a}_k^* = - \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right)^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\}. \quad (20)$$

The Hessian matrix is given by

$$\mathbf{H}_k = \frac{\partial^2 J_{\text{EFC},k}}{\partial \mathbf{a}_k \partial \mathbf{a}_k^T} = 2 \left(\text{Re} \left\{ \mathbf{G}_k^\dagger \mathbf{G}_k \right\} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right). \quad (21)$$

Since both terms in \mathbf{H}_k are positive definite, \mathbf{H}_k is positive definite as well, confirming that the solution is a minimum of the cost function.

We will now show that the solution \mathbf{a}_k^* obtained above is the same as the solution obtained by applying a single iteration of Newton's method to the EFC cost function. Let \mathbf{a}_k^0 be an initial guess for the solution. Newton's method produces an iterate of the form¹⁵

$$\mathbf{a}_k^1 = \mathbf{a}_k^0 - \mathbf{H}_k^{-1} \frac{\partial J_{\text{EFC},k}}{\partial \mathbf{a}_k^T} \Big|_{\mathbf{a}_k = \mathbf{a}_k^0}. \quad (22)$$

Combining Eqs. (19) and (21):

$$\frac{\partial J_{\text{EFC},k}}{\partial \mathbf{a}_k^T} \Big|_{\mathbf{a}_k = \mathbf{a}_k^0} = \mathbf{H}_k \mathbf{a}_k^0 + 2 \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\}. \quad (23)$$

Inserting back into Eq. (22):

$$\mathbf{a}_k^1 = \mathbf{a}_k^0 - \mathbf{H}_k^{-1} \left(\mathbf{H}_k \mathbf{a}_k^0 + 2 \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\} \right) \quad (24a)$$

$$= \mathbf{a}_k^0 - \mathbf{H}_k^{-1} \mathbf{H}_k \mathbf{a}_k^0 - 2 \mathbf{H}_k^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\} \quad (24b)$$

$$= \mathbf{a}_k^0 - \mathbf{a}_k^0 - 2 \mathbf{H}_k^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\} \quad (24c)$$

$$= -2 \mathbf{H}_k^{-1} \text{Re} \left\{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \right\}. \quad (24d)$$

Inserting the definition of the Hessian matrix from Eq. (21), we see that the Newton iterate \mathbf{a}_k^1 is identical to the analytical solution in Eq. (20):

$$\mathbf{a}_k^1 = -2 \left[2 \left(\text{Re} \{ \mathbf{G}_k^\dagger \mathbf{G}_k \} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right) \right]^{-1} \text{Re} \{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \} \quad (25a)$$

$$= - \left(\text{Re} \{ \mathbf{G}_k^\dagger \mathbf{G}_k \} + \mathbf{\Gamma}_k^T \mathbf{\Gamma}_k \right)^{-1} \text{Re} \{ \mathbf{G}_k^\dagger \delta \mathbf{E}_k \} \quad (25b)$$

$$= \mathbf{a}_k^* . \quad (25c)$$

As we described earlier, the same result holds if minimizing $\mathcal{L}_{\text{SM},k}$ in Section 2.1 with respect to \mathbf{a}_k .

APPENDIX B. FAST CONVOLUTIONAL DEFORMABLE MIRROR MODEL

Consider a deformable mirror with N_A actuators along each side (i.e., $N_{\text{act}} = N_A^2$) whose surface $s(x, y)$ can be modeled as a linear superposition of identical influence functions $f(x, y)$

$$s(x, y) = \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} f(x - x_m, y - y_n) . \quad (26)$$

For fixed actuator spacing along the horizontal and vertical directions, we can rewrite the above summation as a convolution between a weighted Dirac comb function and the influence function:

$$s(x, y) = f(x, y) * \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} \delta(x - x_m, y - y_n) . \quad (27)$$

Fourier transforming both sides transforms the convolution operation into a multiplication:

$$\mathcal{F}\{s(x, y)\} = \mathcal{F}\{f(x, y)\} \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} \mathcal{F}\{\delta(x - x_m, y - y_n)\} , \quad (28a)$$

$$= \mathcal{F}\{f(x, y)\} \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} \exp \{ -i2\pi(x_m f_x + y_n f_y) \} . \quad (28b)$$

We define $\tilde{s}(f_x, f_y) \triangleq \mathcal{F}\{s(x, y)\}$ and $\tilde{f}(f_x, f_y) \triangleq \mathcal{F}\{f(x, y)\}$:

$$\tilde{s}(f_x, f_y) = \tilde{f}(f_x, f_y) \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} \exp \{ -i2\pi(x_m f_x + y_n f_y) \} . \quad (29)$$

We next define the discretized surface and influence function arrays $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{f}}$ such that

$$\tilde{\mathbf{s}}[p, q] = \tilde{s}(p\Delta f_x, q\Delta f_y) , \quad (30a)$$

$$\tilde{\mathbf{f}}[p, q] = \tilde{f}(p\Delta f_x, q\Delta f_y) , \quad (30b)$$

so that

$$\tilde{\mathbf{s}} = \tilde{\mathbf{f}} \circ \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} a_{m,n} \exp \{ -i2\pi(x_m \mathbf{f}_x + y_n \mathbf{f}_y) \} , \quad (31)$$

where \circ denotes element-wise multiplication. Finally, we define the vectors of actuator center coordinates \mathbf{x}_c and \mathbf{y}_c such that $\mathbf{x}_c[m] = x_m$ and $\mathbf{y}_c[n] = y_n$, as well as the array of actuator commands \mathbf{A} for which $\mathbf{A}[m, n] = a_{m,n}$:

$$\tilde{\mathbf{s}}[p, q] = \tilde{\mathbf{f}}[p, q] \circ \sum_{m=1}^{N_A} \sum_{n=1}^{N_A} \mathbf{A}[m, n] \exp \{ -i2\pi(\mathbf{x}_c[m] \mathbf{f}_x[p] + \mathbf{y}_c[n] \mathbf{f}_y[q]) \} \quad (32a)$$

$$= \tilde{\mathbf{f}}[p, q] \circ \sum_{m=1}^{N_A} \exp \{ -i2\pi \mathbf{x}_c[m] \mathbf{f}_x[p] \} \sum_{n=1}^{N_A} \mathbf{A}[m, n] \exp \{ -i2\pi \mathbf{y}_c[n] \mathbf{f}_y[q] \} . \quad (32b)$$

We can write this more succinctly as the following sequence of element-wise and matrix products:

$$\tilde{\mathbf{s}} = \tilde{\mathbf{f}} \circ (\exp \{-i2\pi \mathbf{f}_x \mathbf{x}_c^T\} \mathbf{A} \exp \{-i2\pi \mathbf{y}_c \mathbf{f}_y^T\}) \quad (33)$$

where exponentiation is performed element-wise and \mathbf{ab}^T denotes the outer product of the vectors \mathbf{a} and \mathbf{b} .

The term in the parentheses is more commonly referred to as the matrix Fourier transform²⁸ or matrix triple product Fourier transform³² of \mathbf{A} , which we denote as follows:

$$\text{MFT}\{\mathbf{A}; \mathbf{x}_c, \mathbf{y}_c, \mathbf{f}_x, \mathbf{f}_y\} \triangleq \exp \{-i2\pi \mathbf{f}_x \mathbf{x}_c^T\} \mathbf{A} \exp \{-i2\pi \mathbf{y}_c \mathbf{f}_y^T\}, \quad (34)$$

yielding

$$\tilde{\mathbf{s}} = \tilde{\mathbf{f}} \circ \text{MFT}\{\mathbf{A}; \mathbf{x}_c, \mathbf{y}_c, \mathbf{f}_x, \mathbf{f}_y\}. \quad (35)$$

The final step is to compute an inverse discrete Fourier transform to obtain the desired discrete DM surface \mathbf{s} , which is carried out most efficiently using the inverse fast Fourier transform, yielding the final result:

$$\mathbf{s} = \text{IFFT}\left\{\tilde{\mathbf{f}} \circ \text{MFT}\{\mathbf{A}; \mathbf{x}_c, \mathbf{y}_c, \mathbf{f}_x, \mathbf{f}_y\}\right\}. \quad (36)$$

For DMs whose active actuators are a subset of the $N_A \times N_A$ grid modeled above, only the elements of \mathbf{A} corresponding to active actuators are set to nonzero values.

B.1 Adjoint model

The algorithm described in the previous section computes the DM surface resulting from a two-dimensional array of actuator commands \mathbf{A} , under the assumptions that the influence function is identical across all actuators and that the surface can be approximated as a linear superposition of the actuator influence functions. In the context of gradient-based nonlinear optimization using RMAD, we can derive an adjoint model for this algorithm that computes the derivative $\overline{\mathbf{A}} \triangleq \partial J / \partial \mathbf{A}^T$ for some scalar cost function J , given the derivative $\overline{\mathbf{s}}$ with respect to the surface \mathbf{s} .

To begin, we break the forward model into the following sequence:

$$\tilde{\mathbf{A}} = \text{MFT}\{\mathbf{A}; \mathbf{x}_c, \mathbf{y}_c, \mathbf{f}_x, \mathbf{f}_y\}, \quad (37a)$$

$$\tilde{\mathbf{s}} = \tilde{\mathbf{f}} \circ \tilde{\mathbf{A}}, \quad (37b)$$

$$\mathbf{s} = \text{IFFT}\{\tilde{\mathbf{s}}\}. \quad (37c)$$

This leads to the following adjoint model, following the RMAD adjoint variable rules in Refs. [9, 33]:

$$\overline{\tilde{\mathbf{s}}} = \text{FFT}\{\overline{\mathbf{s}}\}, \quad (38a)$$

$$\overline{\tilde{\mathbf{A}}} = \tilde{\mathbf{f}}^* \circ \overline{\tilde{\mathbf{s}}}, \quad (38b)$$

$$\overline{\mathbf{A}} = \text{IMFT}\left\{\overline{\tilde{\mathbf{A}}}; \mathbf{f}_x, \mathbf{f}_y, \mathbf{x}_c, \mathbf{y}_c\right\}, \quad (38c)$$

where $*$ denotes element-wise complex conjugation and IMFT denotes the inverse matrix Fourier transform:

$$\text{IMFT}\left\{\overline{\tilde{\mathbf{A}}}; \mathbf{f}_x, \mathbf{f}_y, \mathbf{x}_c, \mathbf{y}_c\right\} \triangleq \exp \{i2\pi \mathbf{x}_c \mathbf{f}_x^T\} \overline{\tilde{\mathbf{A}}} \exp \{i2\pi \mathbf{f}_y \mathbf{y}_c^T\}. \quad (39)$$

Combining these expressions, the adjoint model is then

$$\overline{\mathbf{A}} = \text{IMFT}\left\{\tilde{\mathbf{f}}^* \circ \text{FFT}\{\overline{\mathbf{s}}\}; \mathbf{f}_x, \mathbf{f}_y, \mathbf{x}_c, \mathbf{y}_c\right\}. \quad (40)$$

APPENDIX C. ADJOINT MODEL FOR EFC COST FUNCTION

In Section 2.2, we describe the cost function for the EFC algorithm for a single correction wavelength. Here, we derive its RMAD adjoint model, which computes the derivative $\partial J_{\text{EFC},k}/\partial \mathbf{E}_{\text{DM},k}$.

We begin by writing the cost function as a series of operations evaluated sequentially:

$$\mathbf{E}_{\text{DZ},k} = \mathbf{E}_{\text{DM},k} + \widehat{\mathbf{E}}_{\text{ab},k}, \quad (41\text{a})$$

$$\Delta \mathbf{E}_k = \mathbf{E}_{\text{DZ},k} - \mathbf{E}_{\text{target},k}, \quad (41\text{b})$$

$$J_{\Delta \mathbf{E},k} = \|\Delta \mathbf{E}_k\|^2, \quad (41\text{c})$$

$$\mathbf{c}_k = \mathbf{\Gamma}_k \mathbf{a}_k, \quad (41\text{d})$$

$$J_{\mathbf{c},k} = \|\mathbf{c}_k\|^2, \quad (41\text{e})$$

$$J_{\text{EFC},k} = J_{\Delta \mathbf{E},k} + J_{\mathbf{c},k}. \quad (41\text{f})$$

We now apply the RMAD gradient rules^{9,33} to each step in reverse order to derive the adjoint model, letting $\bar{\mathbf{x}} \triangleq \partial J_k / \partial \mathbf{x}$ for any variable \mathbf{x} :

$$\bar{J}_{\text{EFC},k} = 1., \quad (42\text{a})$$

$$\bar{J}_{\Delta \mathbf{E},k} = \bar{J}_{\mathbf{c},k} = \bar{J}_{\text{EFC},k}, \quad (42\text{b})$$

$$\bar{\mathbf{c}}_k = 2\mathbf{c}_k \bar{J}_{\mathbf{c},k}, \quad (42\text{c})$$

$$\bar{\mathbf{a}}_k = \mathbf{\Gamma}_k^T \bar{\mathbf{c}}_k, \quad (42\text{d})$$

$$\bar{\Delta \mathbf{E}}_k = 2\Delta \mathbf{E}_k \bar{J}_{\Delta \mathbf{E},k}, \quad (42\text{e})$$

$$\bar{\mathbf{E}}_{\text{DZ},k} = \bar{\Delta \mathbf{E}}_k, \quad (42\text{f})$$

$$\bar{\mathbf{E}}_{\text{DM},k} = \bar{\mathbf{E}}_{\text{DZ},k}. \quad (42\text{g})$$

Combining and simplifying, we see that the desired gradient is given by

$$\bar{\mathbf{E}}_{\text{DM},k} = 2\Delta \mathbf{E}_k. \quad (43)$$

This gradient is passed to the next block of the adjoint model, which in this case is the adjoint model for propagation through the coronagraph, which, referring to Figure 4, evaluates the derivatives of $J_{\text{EFC},k}$ with respect to the surfaces $\mathbf{s}_{1,k}$ and $\mathbf{s}_{2,k}$, respectively. We refer to our earlier work for a derivation of the coronagraph propagation adjoint model.⁹

ACKNOWLEDGMENTS

This work was supported in part by the National Aeronautics and Space Administration under Grant 80NSSC19K0120 issued through the Strategic Astrophysics Technology / Technology Demonstration for Exoplanet Missions Program (SAT-TDEM; PI: R. Soummer). E.H.P. was supported by the NASA Hubble Fellowship grant HST-HF2-51467.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NAS5-26555.

HiCAT makes use of the NumPy,^{34,35} Matplotlib,^{36,37} AstroPy,³⁸ SciPy,³⁹ scikit-image,⁴⁰ Pandas,^{41,42} ImageIO,⁴³ Photutils,⁴⁴ HCIPy,⁴⁵ POPPY^{29,30} and CatKit⁴⁶ packages.

REFERENCES

- [1] Perryman, M., [*The Exoplanet Handbook*], Cambridge University Press, 2nd ed. (10 2018).
- [2] Give'on, A., Belikov, R., Shaklan, S., and Kasdin, J., "Closed loop, DM diversity-based, wavefront correction algorithm for high contrast imaging systems," *Opt. Express* **15**(19), 12338–12343 (2007).
- [3] Baudoz, P., Boccaletti, A., Baudrand, J., and Rouan, D., "The self-coherent camera: a new tool for planet detection," *Proc. Int. Astron. Union* **1**(C200), 553–558 (2005).

- [4] Pueyo, L., Kay, J., Kasdin, N. J., Groff, T., McElwain, M., Give'on, A., and Belikov, R., "Optimal dark hole generation via two deformable mirrors with stroke minimization.," *Appl. Opt.* **48**, 6296–6312 (2009).
- [5] The LUVOIR Study Team, "LUVOIR," tech. rep., National Aeronautics and Space Administration (September 2019).
- [6] The Habitable Exoplanet Observatory Study Team, "Habitable Exoplanet Observatory Final Report," tech. rep., Jet Propulsion Laboratory (August 2019).
- [7] Seo, B.-J., Patterson, K., Balasubramanian, K., Crill, B., Chui, T., Echeverri, D., Kern, B. D., Marx, D., Moody, D., Mejia Prada, C., Ruane, G., Shi, F., Shaw, J., Siegler, N., Tang, H., Trauger, J., Wilson, D. W., and Zimmer, R., "Testbed demonstration of high-contrast coronagraph imaging in search for Earth-like exoplanets," *Proc. SPIE* **11117**, 111171V (2019).
- [8] Griewank, A. and Walther, A., [*Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*], Society for Industrial and Applied Mathematics, 2nd ed. (2008).
- [9] Will, S. D., Groff, T. D., and Fienup, J. R., "Jacobian-free coronagraphic wavefront control using nonlinear optimization.," *J. Astron. Telesc. Instrum. Syst.* **7**, 019002 (2021).
- [10] Eldorado Riggs, A. J., Ruane, G., Sidick, E., Coker, C., Kern, B., and Shaklan, S., "Fast linearized coronagraph optimizer (FALCO) I: a software toolbox for rapid coronagraphic design and wavefront correction," *Proc. SPIE* **10698**, 10698V (2018).
- [11] Ruane, G., Wallace, J. K., Steeves, J., Prada, C. M., Seo, B.-J., Bendek, E., Coker, C., Chen, P., Crill, B., Jewell, J., Kern, B., Marx, D., Poon, P. K., Redding, D., Riggs, A. J. E., Siegler, N., and Zimmer, R., "Wavefront sensing and control in space-based coronagraph instruments using Zernike's phase-contrast method," *J. Astron. Telesc. Instrum. Syst.* **6** (Nov. 2020).
- [12] Sun, H., Kasdin, J. N., and Vanderbei, R., "Identification and adaptive control of a high-contrast focal plane wavefront correction system," *J. Astron. Telesc. Instrum. Syst.* **4**(4), 049006 (2018).
- [13] Paul, B., Mugnier, L. M., Sauvage, J. F., Ferrari, M., and Dohlen, K., "High-order myopic coronagraphic phase diversity (COFFEE) for wave-front control in high-contrast imaging systems," *Opt. Express* **21**(26), 31751–31768 (2013).
- [14] Subedi, H. B., Juanola-Parramon, R., Groff, T. D., and Eldorado Riggs, A., "Deformable mirror apodization via nonlinear optimization using algorithmic differentiation," (2021, in preparation).
- [15] Nocedal, J. and Wright, S. J., [*Numerical Optimization*], Springer-Verlag New York (1999).
- [16] Pourcelot, R., N'Diaye, M., Por, E. H., Perrin, M. D., Soummer, R., Luginja, I., Sahoo, A., Carbillet, M., Brady, G. R., Dohlen, K., Maclay, M., McChesney, E., Noss, J., Petrone, P., Pueyo, L., Vigan, A., and Will, S. D., "Experimental validation of active control of low-order aberrations with a Zernike sensor through a Lyot coronagraph," *Proc. SPIE* **11823**, 11823–46 (2021).
- [17] Luginja, I., Sauvage, J.-F., Mugnier, L., Pueyo, L., Perrin, M. D., Noss, J., Will, S. D., Brooks, K. J., Por, E., and Soummer, R., "Wavefront stability of space-based segmented telescopes at very high contrast: experimental validation," *Astron. Astrophys.* (2021, submitted).
- [18] Redmond, S. M., Pueyo, L., Pogorelyuk, L., Soummer, R., Por, E., Luginja, I., Brooks, K., Noss, J., Perrin, M., and Kasdin, J., "Dark hole maintenance results for segmented aperture wavefront error drift in a high contrast space coronagraph," *Proc. SPIE* **11823**, 11823–48 (2021).
- [19] N'Diaye, M., Soummer, R., Pueyo, L., Carlotti, A., Stark, C. C., and Perrin, M. D., "Apodized pupil Lyot coronagraphs for arbitrary apertures. V. Hybrid shaped pupil designs for imaging Earth-like planets with future space observatories," *Astrophys. J.* **818**, 163–171 (2016).
- [20] N'Diaye, M., Choquet, E., Pueyo, L., Elliot, E., Perrin, M. D., Wallace, J. K., Groff, T., Carlotti, A., Mawet, D., Sheckells, M., Shaklan, S., Macintosh, B., Kasdin, N. J., and Soummer, R., "High-contrast imager for complex aperture telescopes (HiCAT): 1. testbed design," *Proc. SPIE* **8864**, 88641K (2013).
- [21] N'Diaye, M., Choquet, E., Egron, S., Pueyo, L., Leboulleux, L., Levecq, O., Perrin, M. D., Elliot, E., Wallace, J. K., Hugot, E., Marcos, M., Ferrari, M., Long, C. A., Anderson, R., DiFelice, A., and Soummer, R., "High-contrast imager for complex aperture telescopes (HiCAT): II. Design overview and first light results," *Proc. SPIE* **9143**, 914327 (2014).

- [22] N’Diaye, M., Mazoyer, J., Choquet, E., Pueyo, L., Perrin, M. D., Egron, S., Leboulleux, L., Levecq, O., Carlotti, A., Long, C. A., Lajoie, R., and Soummer, R., “High-contrast imager for complex aperture telescopes (HiCAT): 3. first lab results with wavefront control,” *Proc. SPIE* **9605**, 96050I–96050I–12 (2015).
- [23] Leboulleux, L., N’Diaye, M., Riggs, A. J. E., Egron, S., Mazoyer, J., Pueyo, L., Choquet, E., Perrin, M. D., Kasdin, J., Sauvage, J.-F., Fusco, T., and Soummer, R., “High-contrast imager for Complex Aperture Telescopes (HiCAT). 4. Status and wavefront control development,” *Proc. SPIE* **9904**, 99043C (2016).
- [24] Soummer, R., Brady, G. R., Brooks, K., Comeau, T., Choquet, E., Dillon, T., Egron, S., Gontrum, R., Hagopian, J., Laginja, I., Leboulleux, L., Perrin, M. D., Petrone, P., Pueyo, L., Mazoyer, J., N’Diaye, M., Riggs, A. J. E., Shiri, R., Sivaramakrishnan, A., Laurent, K. S., Valenzuela, A.-M., and Zimmerman, N. T., “High-contrast imager for complex aperture telescopes (HiCAT): 5. first results with segmented-aperture coronagraph and wavefront control,” *Proc. SPIE* **10698**, 106981O (2018).
- [25] Moriarty, C., Brooks, K., Soummer, R., Perrin, M., Comeau, T., Brady, G., Gontrum, R., and Petrone, P., “High-contrast imager for complex aperture telescopes (HiCAT): 6. software control infrastructure and calibration,” *Proc. SPIE* **10698**, 1069853 (2018).
- [26] Soummer, R., Laginja, I., Will, S., Juanola-Parramon, R., Iii, P. P., Brady, G., Noss, J., Perrin, M. D., Fowler, J., Kurtz, H., Laurent, K. S., Fogarty, K., McChesney, E., Scott, N., Brooks, K., Comeau, T., Ferrari, M., Gontrum, R., Hagopian, J., Hugot, E., Leboulleux, L., Mazoyer, J., Mugnier, L., N’Diaye, M., Pueyo, L., Sauvage, J.-F., Shiri, R., Sivaramakrishnan, A., Valenzuela, A.-M., and Zimmerman, N. T., “High-contrast imager for complex aperture telescopes (HiCAT): 6. Two deformable mirror wavefront control (Conference Presentation),” *Proc. SPIE* **11117**, 111171Y (Sept. 2019).
- [27] Soummer, R., Laginja, I., Will, S. D., Perrin, M. D., Noss, J., Por, E. H., Redmond, S., Pourcelot, R., Sahoo, A., Petrone, P., Brooks, K., Pueyo, L., Nickson, B., Maclay, M., Brady, G. R., Comeau, T., Fowler, J., Glazer, K., Gontrum, R., Hagopian, J., Jolivet, T., Kautz, M., Leboulleux, L., McChesney, E., Mugnier, L., N’Diaye, M., Nguyen, M., Olszewski, H., Pogorelyuk, L., Prothro, J., Sauvage, J.-F., Scott, N., Shiri, R., Sivaramakrishnan, A., Subedi, H., Weinstock, S., Xin, Y., and Zhang, R., “High-contrast imager for complex aperture telescopes (HiCAT): 7. Dark zone demonstration with fully segmented aperture coronagraph,” *Proc. SPIE* **11823**, 11823–27 (2021).
- [28] Soummer, R., Pueyo, L., Sivaramakrishnan, A., and Vanderbei, R. J., “Fast computation of Lyot-style coronagraph propagation,” *Opt. Express* **15**, 15935–15951 (2007).
- [29] Perrin, M. D., Soummer, R., Elliot, E. M., Lallo, M. D., and Sivaramakrishnan, A., “Simulating point spread functions for the James Webb Space Telescope with WebbPSF,” *Proc. SPIE* **84423D**, 8442 – 8442 – 11 (2012).
- [30] Perrin, M., Long, J., Douglas, E., Sivaramakrishnan, A., Slocum, C., and others, “POPPY: Physical Optics Propagation in PYthon,” (Feb. 2016).
- [31] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Jarrod Millman, K., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C., Polat, İ., Feng, Y., Moore, E. W., Vand erPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S. . . , “SciPy 1.0: Fundamental algorithms for scientific computing in Python,” *Nat. Methods* **17**, 261–272 (2020).
- [32] Jurling, A. S., Bergkoetter, M. D., and Fienup, J. R., “Techniques for arbitrary sampling in two-dimensional Fourier transforms,” *J. Opt. Soc. Am. A* **A35**, 1784–1796 (2018).
- [33] Jurling, A. S. and Fienup, J. R., “Applications of algorithmic differentiation to phase retrieval algorithms,” *J. Opt. Soc. Am. A* **31**(7), 1348–59 (2014).
- [34] Oliphant, T. E., [*A guide to NumPy*], vol. 1, Trelgol Publishing USA (2006).
- [35] van der Walt, S., Colbert, S. C., and Varoquaux, G., “The NumPy Array: A Structure for Efficient Numerical Computation,” *Computing in Science and Engineering* **13**, 22–30 (Mar. 2011).
- [36] Hunter, J. D., “Matplotlib: A 2D Graphics Environment,” *Computing in Science and Engineering* **9**, 90–95 (May 2007).

- [37] Caswell, T. A., Droettboom, M., Lee, A., Hunter, J., de Andrade, E. S., Firing, E., Hoffmann, T., Klymak, J., Stansby, D., Varoquaux, N., Nielsen, J. H., Root, B., May, R., Elson, P., Seppänen, J. K., Dale, D., Lee, J.-J., McDougall, D., Straw, A., Hobson, P., Gohlke, C., Yu, T. S., Ma, E., Vincent, A. F., Silvester, S., Moad, C., hannah, Kniazev, N., Ernest, E., and Ivanov, P., “matplotlib/matplotlib: Rel: v3.3.3,” (Nov. 2020).
- [38] Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., Günther, H. M., Lim, P. L., Crawford, S. M., Conseil, S., Shupe, D. L., Craig, M. W., Dencheva, N., Ginsburg, A., VanderPlas, J. T., Bradley, L. D., Pérez-Suárez, D., de Val-Borro, M., Aldcroft, T. L., Cruz, K. L., Robitaille, T. P., Tollerud, E. J., Ardelean, C., Babej, T., Bach, Y. P., Bachetti, M., Bakanov, A. V., Bamford, S. P., Barentsen, G., Barmby, P., Baumbach, A., Berry, K. L., Biscani, F., Boquien, M., Bostroem, K. A., Bouma, L. G., Brammer, G. B., Bray, E. M., Breytenbach, H., Buddelmeijer, H., Burke, D. J., Calderone, G., Cano Rodríguez, J. L., Cara, M., Cardoso, J. V. M., Cheedella, S., Copin, Y., Corrales, L., Crichton, D., D’Avella, D., Deil, C., Depagne, É., Dietrich, J. P., Donath, A., Droettboom, M., Earl, N., Erben, T., Fabbro, S., Ferreira, L. A., Finethy, T., Fox, R. T., Garrison, L. H., Gibbons, S. L. J., Goldstein, D. A., Gommers, R., Greco, J. P., Greenfield, P., Groener, A. M., Grollier, F., Hagen, A., Hirst, P., Homeier, D., Horton, A. J., Hosseinzadeh, G., Hu, L., Hunkeler, J. S., Ivezić, Ž., Jain, A., Jenness, T., Kanarek, G., Kendrew, S., Kern, N. S., Kerzendorf, W. E., Khvalko, A., King, J., Kirkby, D., Kulkarni, A. M., Kumar, A., Lee, A., Lenz, D., Littlefair, S. P., Ma, Z., Macleod, D. M., Mastropietro, M., McCully, C., Montagnac, S., Morris, B. M., Mueller, M., Mumford, S. J., Muna, D., Murphy, N. A., Nelson, S., Nguyen, G. H., Ninan, J. P., Nöthe, M., Ogaz, S., Oh, S., Parejko, J. K., Parley, N., Pascual, S., Patil, R., Patil, A. A., Plunkett, A. L., Prochaska, J. X., Rastogi, T., Reddy Janga, V., Sabater, J., Sakurikar, P., Seifert, M., Sherbert, L. E., Sherwood-Taylor, H., Shih, A. Y., Sick, J., Silbiger, M. T., Singanamalla, S., Singer, L. P., Sladen, P. H., Sooley, K. A., Sornarajah, S., Streicher, O., Teuben, P., Thomas, S. W., Tremblay, G. R., Turner, J. E. H., Terrón, V., van Kerkwijk, M. H., de la Vega, A., Watkins, L. L., Weaver, B. A., Whitmore, J. B., Woillez, J., Zabalza, V., and Astropy Contributors, “The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package,” *Astrophys. J.* **156**, 123 (Sept. 2018).
- [39] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods* **17**, 261–272 (2020).
- [40] van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Goullart, E., Yu, T., and scikit-image contributors, t., “scikit-image: Image processing in Python,” *PeerJ* (June 2014).
- [41] Wes McKinney, “Data Structures for Statistical Computing in Python,” in [*Proceedings of the 9th Python in Science Conference*], Stéfan van der Walt and Jarrod Millman, eds., 56 – 61 (2010).
- [42] Reback, J., McKinney, W., jbrockmendel, den Bossche, J. V., Augspurger, T., Cloud, P., gyoung, Sinhrks, Hawkins, S., Roeschke, M., Klein, A., Petersen, T., Tratner, J., She, C., Ayd, W., Naveh, S., Garcia, M., Schendel, J., Hayden, A., Saxton, D., Jancauskas, V., McMaster, A., Battiston, P., Seabold, S., chris b1, h vetinari, Dong, K., Hoyer, S., Overmeire, W., and Gorelli, M., “pandas-dev/pandas: Pandas 1.1.4,” (Oct. 2020).
- [43] Silvester, S., Tanbakuchi, A., Müller, P., Nunez-Iglesias, J., Harfouche, M., McCormick, M., Ladegaard, A., Rai, A., OrganicIrradiation, Smith, T. D., Konowalczyk, M., Lee, A., Klein, A., rreilink, Nises, J., jackwalker64, Vaillant, G. A., Barnes, C., Zulko, Hsieh, P.-C., NiklasRosenstein, Komarčević, M., Michael Hirsch, P., Schambach, M., Singleton, J., van Kemenade, H., Inngs, G., Yang, G., Kohlgrüber, F., and Dusold, C., “imageio/imageio v2.8.0,” (Feb. 2020).
- [44] Bradley, L., Sipőcz, B., Robitaille, T., Tollerud, E., Vinícius, Z., Deil, C., Barbary, K., Wilson, T. J., Busko, I., Günther, H. M., Cara, M., Conseil, S., Bostroem, A., Droettboom, M., Bray, E. M., Bratholm, L. A., Lim, P. L., Barentsen, G., Craig, M., Pascual, S., Perren, G., Greco, J., Donath, A., de Val-Borro, M., Kerzendorf, W., Bach, Y. P., Weaver, B. A., D’Eugenio, F., Souchereau, H., and Ferreira, L., “astropy/photutils: 1.0.1,” (Sept. 2020).

- [45] Por, E. H., Haffert, S. Y., Radhakrishnan, V. M., Doelman, D. S., van Kooten, M., and Bos, S. P., “High Contrast Imaging for Python (HCIPy): an open-source adaptive optics and coronagraph simulator,” *Proc. SPIE* **10703**, 1070342 (2018).
- [46] Noss, J., Fowler, J., Moriarty, C., Brooks, K. J., Laginja, I., Perrin, M. D., Soummer, R., Comeau, T., and Olszewski, H., “spacetelescope/catkit: v0.36.1,” (July 2021).