applied optics

Subaperture translation estimation accuracy in transverse translation diversity phase retrieval

DUSTIN B. MOORE* AND JAMES R. FIENUP

The Institute of Optics, University of Rochester, 275 Hutchinson Road, Rochester, New York 14627, USA
**Corresponding author: moored@secac.com*

Received 12 October 2015; revised 17 January 2016; accepted 9 February 2016; posted 10 February 2016 (Doc. ID 251720); published 22 March 2016

For optical metrology by transverse translation diversity phase retrieval (or ptychography), information theoretic limits on the ability to estimate subaperture translation, essential for accurate metrology, are assessed as a function of the optical aberrations of the system being measured. Special attention is given to the case that an unknown linear phase aberration, or equivalent detector or target motion, is present that varies with each point spread function in the measured data.

OCIS codes: (010.7350) Wave-front sensing; (100.5070) Phase retrieval; (120.3940) Metrology; (120.5050) Phase measurement.

http://dx.doi.org/10.1364/AO.55.002526

1. INTRODUCTION

Transverse translation diversity phase retrieval (TTDPR) [1–4] is an image-based wavefront sensing method similar to the coherent diffractive imaging technique known as ptychography [5–10]. In TTDPR, point spread functions (PSFs) of an aberrated optical system are typically acquired as a subaperture moves in a plane perpendicular to the optical axis somewhere inside the system. This plane is usually conjugate to a pupil plane of the system and for simplicity this discussion will be restricted to that case. The subaperture motion allows each PSF to come from the light passing through different regions of the plane but with some overlap. The wavefront over the union of these regions is then reconstructed by a phase retrieval algorithm. The measured PSF data, knowledge of the transmission function of the subaperture, and the subaperture translation for each PSF can be used in a TTDPR algorithm to find the exit pupil field. In this discussion, we assume it is just the phase of the complex field that is unknown, but it is also possible to estimate the amplitude and thus retrieve the complex field. If there is error in the subaperture translation knowledge, estimates of these values can be refined by including them as unknowns in the phase retrieval process [1]. In this case, the PSFs themselves must supply the information necessary to determine the missing subaperture translation knowledge, which is the central concern of this paper. We assume here that it is necessary to estimate subaperture translation from the contents of each PSF. For brevity, we will just refer to "translation" whenever translation of the subaperture is referenced.

Our motivation to understand translation estimation in TTDPR comes from monitoring the aberrations of the long-wave channel on the near-infrared camera (NIRCam) [11] of the James Webb Space Telescope. In this TTDPR application, a Lyot stop in a long-wave pupil plane serves as the subaperture. The Lyot stops are intended for coronographic imaging [12] but also enable TTDPR wavefront sensing [2,4]. The long-wave channels of NIRCam's two modules lack the weak defocus lenses present in the short-wave channels that permit conventional focus-diverse phase retrieval [11,13]. TTDPR provides an opportunity to track the aberrations of NIRCam (separate from the rest of the telescope) during ground testing and on orbit. However, hardware and test specifics prevent the exact nature of the translation from being known to the necessary precision and must be estimated from the PSF data. Other test constraints require that the target being imaged be allowed to move by an uncalibrated amount between each PSF acquisition in some test applications. Provided such target motion is within the isoplanatic patch of the system, it can be modeled as a phase aberration linear in subaperture plane coordinates under the paraxial assumptions of Section 2. This additional unknown linear phase component must be estimated by the phase retrieval algorithm separately for each PSF. We refer to TTDPR measurements where these linear phase terms must be estimated separately for each PSF as the unshared linear phase (USLP) case. When the linear phase is the same for all PSFs, we refer to it as the shared linear phase (SLP) case.

For unobscured systems, like NIRCam imaging its internal source, most PSFs in a TTDPR measurement will derive from exit pupils lacking significant amplitude variations besides the imposed subaperture. Similarly, in the experimental results of [1], the subaperture did not overlap the edge of the aperture stop. Assuming other amplitude variations can be neglected,

we will show that estimating translation from PSF data is equivalent to estimating the phase retrieved from the PSF. We will parameterize that phase by coefficients of a Zernike polynomial description [14], though other basis functions may be similarly useful. We then calculate the Cramer–Rao lower bound for unbiased estimators of translation as a function of the pupil phase aberrations of the system under test. We show how they differ significantly in the SLP and USLP cases. In the SLP case, the ability to assume that each PSF shares the same linear phase contributes subaperture translation information that is unavailable in the USLP case. Translation estimation in the USLP case is entirely dependent on the existence of aberrations higher in order than defocus and astigmatism (those higher than second radial degree) in the optical system under test. Through Monte Carlo simulation, we evaluate translation estimation for a specific population of aberrations to quantify the error in translation estimation for the two cases. In higher fidelity simulations, we performed TTDPR to determine the actual error in translation estimation. Both simulations show that large defocus aberrations will significantly decrease translation estimation error in the SLP case but not for the USLP case, which must rely on the higher-order phase errors for translation estimation. Translation estimation is still possible in the USLP case when the higher-order aberrations are sufficiently large in magnitude and the inherent wavefront sensing error due to noise is sufficiently low.

2. SUBAPERTURE TRANSLATION ESTIMATION ACCURACY

In this section, we will estimate the accuracy of translation estimation using PSFs derived from an aberrated optical system using a simplified model suitable for Cramer–Rao lower bounds. For translation estimation methods useful for actual translation computation, see [1,4,7–10].

It is presumed that the optical system under test obeys the scalar linear-systems theory for isoplanatic imaging with the aberrations. If the target is an unresolved point, the observed noise-free irradiance for the kth PSF in the detector plane is

$$I_k(x,y) = |G_k(x,y)|^2, \tag{1}$$

where $G_k(x,y)$ is the field in the detector. Let $g_k(u,v)$ be a generalized pupil function [15] whose phase, $\phi_k(u,v)$, is the phase aberration of the system. The phase of the USLP case will depend on the conditions in place while measuring the PSF, so $\phi_k(u,v)$ is subscripted with PSF index k. We will assume that $G_k(x,y)$ is well approximated by a Fraunhofer propagation $\mathcal{P}$ of $g_k(u,v)$ with [15]

$$\begin{aligned} G_k(x,y) &= \mathcal{P}[g_k] \\ &= \frac{A}{\lambda L}\iint_{-\infty}^{\infty} g_k(u,v)\exp\left[-\frac{i2\pi}{\lambda L}(xu+yv)\right]\mathrm{d}u\mathrm{d}v, \end{aligned} \tag{2}$$

where A is a constant amplitude, λ is the wavelength of light, and L is the distance between the exit pupil and detector planes. Any quadratic phase in the detector field due to the propagation is omitted from Eq. (2) as it would not alter the PSF intensity. A derivation similar to the following can be done if $\mathcal{P}$ is a Fresnel propagation but the result is complicated by the need to track the additional quadratic phase in u and v. For simplicity, this derivation will be in the Fraunhofer limit, or equivalently, the quadratic phase term within the Fresnel integral is made part of the phase of $g_k(u,v)$.

Let $A_F(u,v)$ be the amplitude of the fixed field that would have been in the exit pupil had the subaperture not been introduced into the system. This fixed aperture arises from (i) the aperture stop of the system, (ii) the illumination provided by the unresolved target being imaged to yield PSFs, and (iii) modifications to the illumination made by vignetting surfaces of the system which may not be conjugate to the pupil. Further, assume for simplicity that the subaperture is in a plane conjugate to the exit pupil and that it is perfectly reimaged to the exit pupil yielding a real, nonnegative, transmission function $A_S(u,v)$ in the plane of the exit pupil. We model the amplitude of the exit pupil as the product of the translated subaperture and the amplitude of the fixed field, making the complex field of the generalized exit pupil

$$g_k(u,v) = A_S(u-s, v-t)A_F(u,v)\exp[i\phi_k(u,v)], \tag{3}$$

where (s,t) is the translation in exit pupil coordinates. In general, s and t will depend on k but in this discussion we consider the estimation of translation for an arbitrary but specific PSF, so the index k is suppressed for brevity.

We will require that $\phi_k(u,v)$ be a slowly spatially varying phase function that can be expressed by a Zernike polynomial expansion,

$$\phi_k(u,v) = \sum_{j=1}^{J} a_{j,k}Z_j(u,v), \tag{4}$$

where the coefficients $a_{k,j}$ of the jth Zernike polynomial $Z_j(u,v)$ are indexed by PSF number k as necessary for the SLP and USLP cases. The assumption that $\phi_k(u,v)$ is slowly varying is usually true in a well-corrected optical system. We use the Noll-ordered polynomials [14,16], the first 11 terms of which are listed in Table 1. We will require that u and v are normalized such that the region of the exit pupil accessed by all subaperture translations is encompassed by the circle $u^2+v^2 \le 1$.

Table 1. Zernike Polynomials in Cartesian Coordinates with j, the Index of the Polynomial, and n, the Radial Degree

j	n	$Z_j(u,v)$	Aberration
1	0	1	Piston
2	1	$2u$	u linear
3	1	$2v$	v linear
4	2	$\sqrt{3}(2u^2+2v^2-1)$	Defocus
5	2	$2\sqrt{6}uv$	Astig. 45°
6	2	$\sqrt{6}(u^2-v^2)$	Astig. 0°
7	3	$\sqrt{8}(3u^2v+3v^3-2)$	v coma
8	3	$\sqrt{8}(3u^3+3uv^2-2)$	u coma
9	3	$\sqrt{8}(3u^2v-v^3)$	v trefoil
10	3	$\sqrt{8}(u^3-3uv^2)$	u trefoil
11	4	$\sqrt{5}[6(u^2+v^2)^2-6(u^2+v^2)+1]$	Spherical

In the SLP case, the only parameters of the system changing between PSF acquisitions are the subaperture translations (s, t). The phase coefficients $a_{j,k}$ will all be independent of k and thus the same values are shared among all PSFs. In the USLP case, the linear phase terms $a_{2,k}$ and $a_{3,k}$ will be different for each PSF. We next consider two experimental conditions that lead to the linear phase terms becoming unshared. First, consider a detector that shifts in its plane between PSF acquisitions. This motion is equivalent to a modification of the linear Zernike coefficients by the Fourier shift theorem. If the detector plane is shifted by a vector (ϵ_k, η_k), then from Eq. (2) the field in the detector plane is

$$G_k(x-\epsilon_k, y-\eta_k) = \mathcal{P}\left\{g_k(u,v)\exp\left[\frac{i2\pi}{\lambda L}(u\epsilon_k + v\eta_k)\right]\right\} \equiv \mathcal{P}[h_k(u,v)], \tag{5}$$

where $h_k(u, v)$ is an instance of $g_k(u, v)$ with the linear phase coefficients $a_{2,k}$ and $a_{3,k}$ increased by $\pi\epsilon_k/\lambda L$ and $\pi\eta_k/\lambda L$, respectively. The PSF detected by a shifted detector is identical to the PSF due to $h_k(u, v)$ with modified linear phase terms since by Eq. (5),

$$|G_k(x-\epsilon_k, y-\eta_k)|^2 = |\mathcal{P}[h_k]|^2. \tag{6}$$

Thus, a detector that moves between PSF acquisitions is equivalent to each PSF having arisen from its own unique linear phase coefficients giving rise to a USLP situation. Second, suppose instead that the detector was fixed and it was the unresolved target being imaged that was translating. Also, let the amount of the target translation be bounded so that the target stays within an isoplanatic patch of the optical system under test. The effect of target translation inside the isoplanatic patch is a translation of the PSF but by an amount scaled by the magnification of the system. The target translation can be modeled as an exit pupil phase that has appropriately varying linear phase terms. Consequently, an unknown and varying target translation also yields a USLP situation.

The subaperture positions for the kth PSF can, in principle, be recovered if the PSF for that subaperture position is different than a PSF derived from identical conditions but with an erroneous subaperture position. Since the PSF arises exclusively from the field in the exit pupil, these differences must be due to changes in the amplitude and/or phase of $g_k(u, v)$. If the fixed aperture has hard-edged features, these variations act as fiducials against which translation can be estimated. This is the case when the moving subaperture overlaps a hard-edged aperture stop in $A_F(u, v)$ during one or more PSF acquisitions. However, if the subaperture is smaller than the aperture stop and vignetting of the subaperture by the aperture stop is mild or nonexistent, the utility of amplitude variation is limited since those PSFs will come from regions of the fixed aperture not overlapping any such amplitude fiducials. For this reason we focus our attention on the more difficult case that the effects of $A_F(u, v)$ on each PSF are negligible and

$$A_S(u-s, v-t)A_F(u,v) \approx A_S(u-s, v-t). \tag{7}$$

The fixed amplitude for the experiment in [1] was consistent with Eq. (7) aside from some very small imperfections, referred to as "digs," in the surface under test.

The reduced role of fixed aperture features differentiates translation estimation in TTDPR from the established position refinement methods in ptychography for coherent diffractive imaging [9,10]. In coherent diffractive imaging, specimens generally have a multitude of well-defined edges in amplitude, in phase, or in both. These edges yield far-field diffraction patterns that change robustly with small variations in translation or "probe position." For TTDPR wavefront sensing, there are typically few such amplitude fiducials contributing to the subaperture positions, and it is variations in the phase that must be relied upon for translation estimation.

Assuming Eq. (7), Eq. (3) can be approximated as

$$g_k(u,v) = A_S(u-s, v-t)\exp[i\phi_k(u,v)]. \tag{8}$$

Next, consider the PSF intensity that would be expected from an incorrect subaperture translation. Let the subaperture be translated by an additional error vector (q, r), also implicitly dependent on k, as in the exit pupil field:

$$g'_k(u,v) = A_S(u-s-q, v-t-r)\exp[i\phi_k(u,v)]. \tag{9}$$

The PSF intensity is

$$I'_k(x,y) = |\mathcal{P}\{A_S(u-s-q, v-t-r)\exp[i\phi_k(u,v)]\}|^2. \tag{10}$$

Making the variable substitutions $u \to u' + q$ and $v \to v' + r$ in the integral implicit in Eq. (10), one gets

$$\begin{aligned} I'_k(x,y) &= \left|\frac{A}{\lambda L}\iint_{-\infty}^{\infty} A_S(u'-s, v'-t)\exp[i\phi_k(u'+q, v'+r)]\right. \\ &\quad \times \exp\left[\frac{-i2\pi}{\lambda L}(xq+yr)\right] \\ &\quad \left.\times \exp\left[\frac{-i2\pi}{\lambda L}(xu'+yv')\right]du'dv'\right|^2 \\ &= \left|\frac{A}{\lambda L}\iint_{-\infty}^{\infty} A_S(u'-s, v'-t)\exp[i\phi_k(u'+q, v'+r)]\right. \\ &\quad \left.\times \exp\left[\frac{-i2\pi}{\lambda L}(xu'+yv')\right]du'dv'\right|^2, \end{aligned} \tag{11}$$

where the complex exponential in $xq + yr$ from the first line disappears because it does not affect the intensity. Since inside Eq. (11) is a propagation,

$$I'_k(x,y) = |\mathcal{P}\{A_S(u-s, v-t)\exp[i\phi_k(u+q, v+r)]\}|^2. \tag{12}$$

From Eq. (12), we observe that the change in the PSF due to a misestimation of the translation by (q, r) is equivalent to a shift in the exit pupil phase ϕ_k by $(-q, -r)$. Evaluating the shift in terms of the Zernike polynomials gives

$$\phi_k(u+q, v+r) = \sum_{j=1}^{J} a_{j,k}Z_j(u+q, v+r). \tag{13}$$

A shifted Zernike polynomial term, $Z_j(u + q, v + r)$, can be exactly expressed as a weighted sum of unshifted Zernike polynomials $Z_j(u, v)$ by endnote 36 in [17]. That proof involves expanding $Z_j(u + q, v + r)$ as a Taylor series about the point (u, v) in the shift variables (q, r) and various partial derivatives of the unshifted Zernike polynomial [17]. Since derivatives of the unshifted Zernike polynomial of an order

higher than the polynomial's radial degree are zero [14], the Taylor series is finite in length. Of the terms in the Taylor series, those involving first derivatives of Zernike polynomials are themselves each finite series of Zernike polynomials of lower radial degree [14,18]. The Taylor series terms involving higher-order partial derivatives can be found by repeatedly differentiating a first derivative series to find the finite series of Zernike polynomials that equals the higher-order partial derivative. Since every term in the Taylor series can be expressed as a finite series in Zernike polynomials, the shifted Zernike polynomial can be written exactly as a weighted sum of unshifted Zernike polynomials of lesser radial degree. We use this to write the apparent change in a Zernike polynomial as a function of translation estimation error (q, r) as

$$Z_j(u+q, v+r) = Z_j(u, v) + \sum_{j'=1}^{J} \gamma_{j,j'}(q, r) Z_{j'}(u, v), \quad \textbf{(14)}$$

where $\gamma_{j,j'}(q, r)$ is a matrix of polynomials in q and r. The polynomials in the matrix $\gamma_{j,j'}(q, r)$ for the significant terms up to spherical aberration were evaluated algebraically and are listed in Table 2. It is possible to evaluate $\gamma_{j,j'}(q, r)$ for shifted-Zernike polynomials of arbitrarily high radial degree but we stopped at spherical aberration for simplicity. Terms that when shifted yield only a constant phase are invisible to phase retrieval, so terms of $\gamma_{j,j'}(q, r)$ where $j' = 1$ are omitted from Table 2. For example, shifted versions of the linear phase term are the same phase unshifted but with an additional constant:

$$\begin{aligned} Z_2(s+q, t+r) &= 2(s+q) = 2s + 2q \\ &= Z_2(s, t) + qZ_1(s, t). \end{aligned} \quad \textbf{(15)}$$

Substituting Eq. (14) into Eq. (13) yields

$$\begin{aligned} \phi_k(u+q, v+r) &= \sum_{j=1}^{J} a_{j,k} Z_j(u, v) + \sum_{j=1}^{J} a_{j,k} \sum_{j'=1}^{J} \gamma_{j,j'}(q, r) Z_{j'}(u, v) \\ &= \sum_{j=1}^{J} a_{j,k} Z_j(u, v) + \sum_{j'=1}^{J} \sum_{j=1}^{J} a_{j',k} \gamma_{j',j}(q, r) Z_j(u, v) \\ &= \sum_{j=1}^{J} \left[a_{j,k} + \sum_{j'=1}^{J} a_{j',k} \gamma_{j',j}(q, r) \right] Z_j(u, v), \end{aligned} \quad \textbf{(16)}$$

where we note that the labels j and j' have been permuted in the right-hand part of the sum in the second step to yield the simplification. From Eq. (12) it is observed that the change in the PSF due to a translation misestimation is equivalent to altering the phase that contributed to the PSF. According to Eq. (16), the apparent alteration in phase is equivalent to modifying the Zernike coefficients $a_{j,k}$ by adding an erroneous contribution of $\sum_{j'=1}^{J} a_{j',k} \gamma_{j',j}(q, r)$ to each coefficient.

Suppose there is some calculation involving the PSFs with indices not equal to k that estimates $a_{j,k}$. For instance, if we perform a successful TTDPR phase retrieval operation on all of the PSFs that are not the kth PSF, and the kth PSF comes from the same optical system, then we would expect the TTDPR solution from the other PSFs to be a good estimate of $a_{j,k}$. This is not possible for $a_{2,k}$ and $a_{3,k}$ in the USLP case and we address this later. Also suppose that it is possible to perform a single-PSF phase retrieval on the kth PSF assuming erroneous values for (s, t) and that the results of this successful phase retrieval solution are polynomial coefficients $c_{j,k}$ for the jth Zernike. Note that performing this single-PSF phase retrieval may be ill-posed if the subaperture extent does not allow unique estimation of each of the Zernike coefficients of interest. For this argument, we assume this single-PSF phase retrieval accesses enough of the field to uniquely estimate the necessary Zernike phase coefficients.

In this case, Eq. (16) suggests a way to solve for the unknown translation as follows: the error (q, r) in the translations is governed by the nonlinear system of equations,

$$c_{j,k} - a_{j,k} = \sum_{j'=1}^{J} a_{j',k} \gamma_{j',j}(q, r) = \sum_{j'=2}^{J} a_{j',k} \gamma_{j',j}(q, r), \quad \textbf{(17)}$$

according to Eq. (16) and the fact that $\gamma_{1,j}(q, r) = 0$ for all j, since $Z_1(u+q, v+r) = 1 = Z_1(u, v) + 0$. If this system of equations is solvable for (q, r), the resulting estimates of (q, r) can be added to the erroneous values of (s, t) to re-estimate the true (s, t).

Incidentally, we suspect that the translation estimation problem as posed in this section is similar to translation estimation in some forms of subaperture stitching interferometry (SSI) where (i) the interferometer aperture is confined to translate in the plane of the acquired interferograms, (ii) the translation must be estimated from acquired interferograms, and (iii) it is not possible to track the change in constant piston phase between acquisitions. A particular interferogram with unknown translation could then be processed to yield an estimate for $c_{j,k}$. If $a_{j,k}$ were estimable from the ensemble of interferograms, the system of equations represented by Eq. (17) would exist for

Table 2. Nonzero Terms of the Zernike Translation Matrix $\gamma_{j'j}(q,r)$ for $j < 12$

j'	$j = 2$	3	4	5	6	7	8
4	$2\sqrt{3}q$	$2\sqrt{3}r$					
5	$\sqrt{6}r$	$\sqrt{6}q$					
6	$\sqrt{6}q$	$-\sqrt{6}r$					
7	$6\sqrt{2}qr$	$3\sqrt{2}q^2 + 9\sqrt{2}r^2$	$2\sqrt{6}r$	$2\sqrt{3}q$	$-2\sqrt{3}r$		
8	$9\sqrt{2}q^2 + 3\sqrt{2}r^2$	$6\sqrt{2}qr$	$2\sqrt{6}q$	$2\sqrt{3}r$	$2\sqrt{3}q$		
9	$6\sqrt{2}qr$	$3\sqrt{2}(q^2 - r^2)$		$2\sqrt{3}q$	$2\sqrt{3}r$		
10	$3\sqrt{2}(q^2 - r^2)$	$-6\sqrt{2}qr$		$-2\sqrt{3}r$	$2\sqrt{3}q$		
11	$2\sqrt{5}q + 12\sqrt{5}(q^3 + qr^2)$	$2\sqrt{5}r + 12\sqrt{5}(q^2r + r^3)$	$4\sqrt{15}(q^2 + r^2)$	$4\sqrt{30}qr$	$2\sqrt{30}(q^2 - r^2)$	$2\sqrt{10}r$	$2\sqrt{10}q$

Table 3. Nonzero Terms of the Linearized Zernike Translation Matrix $\gamma^q_{j'j}$ for $j < 12$

j'	$j=2$	3	4	5	6	8
4	$2\sqrt{3}$					
5		$\sqrt{6}$				
6	$\sqrt{6}$					
7				$2\sqrt{3}$		
8			$2\sqrt{6}$		$2\sqrt{3}$	
9				$2\sqrt{3}$		
10					$2\sqrt{3}$	
11	$2\sqrt{5}$					$2\sqrt{10}$

the SSI translation estimation problem as it does in the TTDPR case.

Since $\gamma_{j,j'}(q, r)$ involves nonlinear terms, it is difficult to predict the invertability of the system of equations in Eq. (17) for a general aberration. When the expected translation estimation errors (q, r) are small, however, it is possible to make a linear approximation that is more easily analyzed. Suppose that the guesses for s and t used to find $c_{j,k}$ are sufficiently accurate that $|q|$ and $|r|$ are much smaller than unity. The linear terms will then dominate the terms in $\gamma_{j,j'}(q, r)$ that are quadratic and higher in q and r. In this case, the matrix of polynomials $\gamma_{jj'}(q, r)$ can be split into the portion listed in Table 3 that is linear in q and the portion listed in Table 4 that is linear in r, which we label $\gamma^q_{jj'}$ and $\gamma^r_{jj'}$, respectively. Neglecting the higher-order terms in q and r in Eq. (17), we can produce a linear approximation for small q and r,

$$c_{j,k} - a_{j,k} \approx \sum_{j'=2}^{J} a_{j',k}(\gamma^q_{j',j}q + \gamma^r_{j',j}r) = qQ_j + rR_j, \tag{18}$$

where

$$Q_j = \sum_{j'=2}^{J} a_{j',k}\gamma^q_{j',j}, R_j = \sum_{j'=2}^{J} a_{j',k}\gamma^r_{j',j} \tag{19}$$

can be seen as vectors in the space of the differences of $c_{j,k} - a_{j,k}$ that depend on the aberrations of the system.

Let $\boldsymbol{c}$, $\boldsymbol{a}$, $\boldsymbol{Q}$, and $\boldsymbol{R}$ be column vectors representing the values of $c_{j,k}$, a_j, Q_j, and R_j, respectively, starting at $j = 2$ in the SLP case and $j = 4$ in the USLP case. The SLP case column vectors begin at $j = 2$ because a_0 in Eq. (18) is unknowable. In the USLP case, the vectors begin with $j = 4$ because $a_{2,k}$ and $a_{3,k}$ vary with each PSF and are thus unknowable *a priori*. For simplicity, we have terminated $\gamma^q_{j'j}$, $\gamma^r_{j'j}$, $c_{j,k}$, a_j, Q_j, R_j

Table 4. Nonzero Terms of the Linearized Zernike Translation Matrix $\gamma^r_{j'j}$ for $j < 12$

j'	$j=2$	3	4	5	6	7
4		$2\sqrt{3}$				
5	$\sqrt{6}$					
6		$-\sqrt{6}$				
7			$2\sqrt{6}$		$-2\sqrt{3}$	
8				$2\sqrt{3}$		
9					$2\sqrt{3}$	
10				$-2\sqrt{3}$		
11		$2\sqrt{5}$				$2\sqrt{10}$

and their vector equivalents at spherical aberration but, as with $\gamma_{j,j'}(q, r)$, they could be extended to include aberrations with arbitrarily high Zernike polynomial index.

Given Eq. (18), one can approximate the process of estimating q and r given noisy measurements of $c_{j,k}$ using a linear model,

$$\boldsymbol{c} = \boldsymbol{H\theta} + \boldsymbol{a} + \boldsymbol{\eta}, \tag{20}$$

where $\boldsymbol{H} = [\boldsymbol{Q}\ \boldsymbol{R}]$, the values to be estimated are $\boldsymbol{\theta} = [q\ r]^{\mathrm{T}}$, and $\boldsymbol{\eta}$ represents noise in the measurement $\boldsymbol{c}$ as a result of the single-PSF phase retrieval. We will approximate the error in the phase retrieval measurement as an additive, white, Gaussian noise with variance σ_j^2 for the jth Zernike coefficient and with zero covariance between terms. The results of a phase retrieval measurement for $c_{j,k}$ are not really going to be distributed in this manner [19], but it suffices for the purposes of understanding the basic sensitivity of estimators of q and r to aberrations $a_{j,k}$. A minimum-variance unbiased estimator [20] can be calculated for Eq. (20) yielding an estimator $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$. The estimator will have a covariance matrix,

$$\boldsymbol{C}_{\hat{\theta}} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{H})^{-1}, \tag{21}$$

where $\boldsymbol{C}_{\eta}$ is the covariance matrix of the noise $\boldsymbol{\eta}$. Since we assume zero covariance between terms, $\boldsymbol{C}_{\eta}$ and $\boldsymbol{C}_{\eta}^{-1}$ are diagonal matrices with values of σ_j^2 and σ_j^{-2} along their diagonals, respectively. Equation (21) involves a 2-by-2 matrix inverse and a closed-form evaluation gives

$$\boldsymbol{C}_{\hat{\theta}} = \frac{1}{\det(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{H})}\begin{bmatrix} \boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R} & -\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R} \\ -\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q} & \boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q} \end{bmatrix}, \tag{22}$$

from which the individual variances for estimators of q and r can be extracted and written more simply as

$$\mathrm{var}(\hat{q}) = \left[\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q} - \frac{(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R})^2}{\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R}}\right]^{-1},$$
$$\mathrm{var}(\hat{r}) = \left[\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R} - \frac{(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R})^2}{\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q}}\right]^{-1}. \tag{23}$$

If the variances are small for a particular aberration and wavefront sensing accuracy, then translation estimation is expected to be robust. Since the coordinates u and v have been normalized as fractions of the radius of the circle $u^2 + v^2 \le 1$ and are hence unitless, $\mathrm{var}(\hat{q})$ and $\mathrm{var}(\hat{r})$ are similarly unitless. A minimum-variance unbiased estimator of a linear model is efficient [20], so, assuming Eq. (20) and our simplified model for wavefront sensing error are true, $\mathrm{var}(\hat{q})$ and $\mathrm{var}(\hat{r})$ are Cramer–Rao lower bounds for any unbiased estimator of translation.

The scalar values $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q}$, $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R}$, and $(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R})^2$ are always positive, so Eq. (23) has the lower bounds:

$$\mathrm{var}(\hat{q}) \ge (\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q})^{-1},$$
$$\mathrm{var}(\hat{r}) \ge (\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R})^{-1}. \tag{24}$$

Thus a necessary condition for robust estimation of translations along the u and v coordinates is that the optical system have aberrations that yield $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q}$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R}$ large in comparison to 1. To ensure robust estimation it is also required that $(\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R})^2$ is small in comparison to the product of $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{Q}$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\eta}^{-1}\boldsymbol{R}$, according to Eq. (23).

3. SUBAPERTURE TRANSLATION ESTIMATION WITH SHARED LINEAR PHASE

In the SLP case for which the terms $a_{2,k}$ and $a_{3,k}$ are inferable from PSFs with all indices but k, evaluating the components of Eq. (19) for arbitrary aberrations a_j yields

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}} &= \Big[2\sqrt{3}a_4 + \sqrt{6}a_6 + 2\sqrt{5}a_{11}, \sqrt{6}a_5, 2\sqrt{6}a_8,\\ &\quad 2\sqrt{3}(a_7 + a_9), 2\sqrt{3}(a_8 + a_{10}), 0, 2\sqrt{10}a_{11}\Big],\\ \boldsymbol{R}^{\mathrm{T}} &= \Big[\sqrt{6}a_5, 2\sqrt{3}a_4 - \sqrt{6}a_6 + 2\sqrt{5}a_{11}, 2\sqrt{6}a_7,\\ &\quad 2\sqrt{3}(a_8 - a_{10}), 2\sqrt{3}(a_9 - a_7), 2\sqrt{10}a_{11}, 0\Big]. \qquad (25)\end{aligned}$$

The components of Eq. (24) in the SLP case are therefore

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{Q} &= \sigma_2^{-2}\left(2\sqrt{3}a_4 + \sqrt{6}a_6 + 2\sqrt{5}a_{11}\right)^2 + 6\sigma_3^{-2}a_5^2\\ &\quad + 24\sigma_4^{-2}a_8^2 + 12\sigma_5^{-2}(a_7 + a_9)^2 + 12\sigma_6^{-2}(a_8 + a_{10})^2\\ &\quad + 40\sigma_8^{-2}a_{11}^2,\\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R} &= 6\sigma_2^{-2}a_5^2 + \sigma_3^{-2}\left(2\sqrt{3}a_4 - \sqrt{6}a_6 + 2\sqrt{5}a_{11}\right)^2\\ &\quad + 24\sigma_4^{-2}a_7^2 + 12\sigma_5^{-2}(a_8 - a_{10})^2 + 12\sigma_6^{-2}(a_9 - a_7)^2\\ &\quad + 40\sigma_7^{-2}a_{11}^2. \qquad (26)\end{aligned}$$

Equation (26) indicates that the lower bounds for the variance, Eq. (24), are inversely proportional to a series of sums depending on second and higher radial degree aberrations adding in quadrature. If there are no phase aberrations in the exit pupil, $\boldsymbol{Q} = \boldsymbol{R} = \boldsymbol{0}$ and the lower bounds in Eq. (24) are infinite. Hence, subaperture translation estimation is not practical for systems with completely unaberrated exit pupils or those with only piston or linear phase terms. If the translations were uncertain prior to TTDPR in these cases, no additional information about what region of the subaperture plane was unaberrated could be gained through TTDPR. Phase retrieval would indicate just that an unspecified region of the subaperture plane was aberration-free.

The expected variance and Cramer–Rao lower bound in Eq. (23) also require that the square of

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R} &= (\sigma_2^{-2} + \sigma_3^{-2})a_5\left(6\sqrt{2}a_4 + 2\sqrt{30}a_{11}\right)\\ &\quad + 6(\sigma_2^{-2} - \sigma_3^{-2})a_5a_6 + 24\sigma_4^{-2}a_7a_8\\ &\quad + 12(\sigma_5^{-2} + \sigma_6^{-2})(a_8a_9 - a_7a_{10})\\ &\quad + 12(\sigma_5^{-2} - \sigma_6^{-2})(a_7a_8 - a_9a_{10}) \qquad (27)\end{aligned}$$

be small compared to the product of $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{Q}$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R}$ in Eq. (26) for a low variance of the bound. If the wavefront sensing errors in linear phase and astigmatism are independent of rotation, then $\sigma_3 = \sigma_2$ and $\sigma_6 = \sigma_5$, and Eq. (27) simplifies to

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R} &= 4\sigma_2^{-2}a_5\left(3\sqrt{2}a_4 + \sqrt{30}a_{11}\right)\\ &\quad + 24\sigma_4^{-2}a_7a_8 + 24\sigma_5^{-2}(a_8a_9 - a_7a_{10}). \qquad (28)\end{aligned}$$

For highly aberrated systems, which are described by many Zernike coefficients that are large in magnitude, $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{Q}$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R}$ are likely large as well, and thus the lower bounds in Eq. (24) on translation estimation error are small. These bounds are more problematic on nearly unaberrated systems where $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{Q}$ and $\boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R}$ will be much smaller; thus, we concentrate on that case.

Cramer–Rao analysis of defocus-diverse phase retrieval shows that wavefront sensing in the presence of noise performs better when defocus has been induced in the system [21,22], so it can be expected that some defocus has been induced to improve wavefront sensing. When this intentionally added defocus dominates the other aberrations, the expected translation estimation error in the SLP case is approximately

$$\begin{aligned}\sigma_2^{-2}\,\mathrm{var}(\hat{q}) &\approx \frac{1}{12a_4^2},\\ \sigma_3^{-2}\,\mathrm{var}(\hat{r}) &\approx \frac{1}{12a_4^2}, \qquad (29)\end{aligned}$$

according to Eqs. (26), (27), and (23), under the assumption that $a_j \ll 1$ for $j \neq 4$ and that the variance of wavefront sensing error is comparable for all aberrations. For example, if the aberrations are dominated by a defocus of $a_4 = 4.62$ rad RMS, equivalent to 2.55 waves center-to-edge of the circle $u^2 + v^2 \le 1$, Eq. (29) is

$$\sigma_2^{-2}\,\mathrm{var}(\hat{q}) = \sigma_3^{-2}\,\mathrm{var}(\hat{r}) \approx 3.91 \times 10^{-3}\ \mathrm{rad}^{-2}. \qquad (30)$$

In general, for any single aberration that dominates all of the others, the variance of translation is inversely proportional to the square of the dominating aberration with the constant of proportionality determined by the factors in Eq. (26).

If there is not just one dominating aberration such as defocus, Eqs. (26), (27), and (23) must be evaluated for the specific phase aberrations of the system and wavefront sensing error variances to predict translation estimation performance. If the specific aberration of interest is unknown, these equations can be evaluated over a specified population of aberrations to establish expected behavior. This is done in Section 5 for a specific population of aberrations in a Monte Carlo experiment.

4. SUBAPERTURE TRANSLATION ESTIMATION WITH UNSHARED LINEAR PHASE

In the USLP case, $\boldsymbol{c}$, $\boldsymbol{a}$, $\boldsymbol{Q}$, and $\boldsymbol{R}$ become column vectors representing the values of $c_{j,k}$, a_j, Q_j, and R_j for $4 \le j \le 11$ since $a_{2,k}$ and $a_{3,k}$ are unshared and unknown. For this case, evaluating Eq. (19) yields

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}} &= \Big[2\sqrt{6}a_8, 2\sqrt{3}(a_7 + a_9), 2\sqrt{3}(a_8 + a_{10}), 0, 2\sqrt{10}a_{11}\Big],\\ \boldsymbol{R}^{\mathrm{T}} &= \Big[2\sqrt{6}a_7, 2\sqrt{3}(a_8 - a_{10}), 2\sqrt{3}(a_9 - a_7), 2\sqrt{10}a_{11}, 0\Big]. \qquad (31)\end{aligned}$$

Therefore,

$$\begin{aligned}\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{Q} &= 24\sigma_4^{-2}a_8^2 + 12\sigma_5^{-2}(a_7 + a_9)^2\\ &\quad + 12\sigma_6^{-2}(a_8 + a_{10})^2 + 40\sigma_8^{-2}a_{11}^2,\\ \boldsymbol{R}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R} &= 24\sigma_4^{-2}a_7^2 + 12\sigma_5^{-2}(a_8 - a_{10})^2\\ &\quad + 12\sigma_6^{-2}(a_9 - a_7)^2 + 40\sigma_7^{-2}a_{11}^2, \qquad (32)\end{aligned}$$

and

$$\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{C}_{\boldsymbol{\eta}}^{-1}\boldsymbol{R} = 24\sigma_4^{-2}a_7a_8 + 12(\sigma_5^{-2}+\sigma_6^{-2})(a_8a_9 - a_7a_{10}) + 12(\sigma_5^{-2}-\sigma_6^{-2})(a_7a_8 - a_9a_{10}). \quad (33)$$

The aberrations that contribute to improved subaperture estimation in the USLP case are those of third radial degree and above. Notably missing are the second radial degree terms: defocus and the two astigmatisms. This is due to the first two entries in $\boldsymbol{Q}$ and $\boldsymbol{R}$ in the SLP case, which include the second radial degree terms, being absent from the $\boldsymbol{Q}$ and $\boldsymbol{R}$ of the USLP case. These terms are excluded from Eq. (20) in the USLP case because $a_{2,k}$ and $a_{3,k}$ are unknown and thus $c_{2,k} - a_2$ and $c_{3,k} - a_{3,k}$ are unusable equations of Eq. (18). In less mathematical terms, an errant subaperture translation samples a mean wavefront slope different than the mean slope across the subaperture in the correct translation. If the second radial degree aberrations are nonzero, the difference in mean wavefront slope has a component linear in pupil coordinates and that yields a translation of the PSF relative to where it would be expected, given $\boldsymbol{a}$. This PSF shift constitutes information about the subaperture translation and it is due to the existence of the second radial degree aberrations. However, this shift is indistinguishable from unknown detector or target motion since that also produces a translation of the PSF. Thus, the second radial degree aberrations in the optical system contribute no subaperture translation information if the detector or target are also translating. We observe that it is necessary to have aberrations of the exit pupil higher in order than defocus and astigmatism to estimate translation in a system for which the linear phase terms vary unpredictably between PSFs or, equivalently, when the detector or target are undergoing unknown translation. Similarly, in the USLP case with a dominant defocus or astigmatism aberration, it is much more difficult to estimate the variance of translation estimation without specifying the smaller aberrations of higher radial degree. In the next section, we turn to Monte Carlo analysis to evaluate the behavior of the variance of translation estimation with a particular population of aberrations.

5. MONTE CARLO EXPERIMENTS

We compared the Cramer–Rao lower bounds on translation estimation in the SLP and USLP cases for a specific aberration population using a Monte Carlo experiment. Five hundred random aberration coefficient vectors a_j were composed by drawing independent and normally distributed random numbers from distributions having standard deviations of 10, 0.2, 0.1, 0.1, and 0.3 rad RMS for the linear, astigmatism, coma, trefoil, and spherical aberration terms, respectively. For each vector, the defocus term was assumed to be the arbitrary constant $a_4 = 4.62$ rad RMS used in Section 3 and the piston term was assumed zero. The mean wavefront aberration averaged across the population of aberration vectors was 4.64 rad RMS with linear phase excluded, but just 0.429 rad RMS with defocus phase also excluded. An example exit pupil phase synthesized from one such aberration vector is plotted in Fig. 1(a) and in Fig. 1(b) with the defocus subtracted. In both figures, the phase has been plotted over the circle $u^2 + v^2 \le 1$. In addition to an aberration coefficient vector, evaluation of Eq. (23)

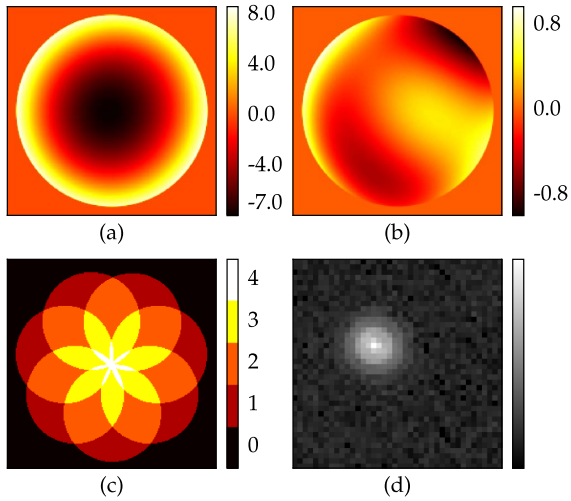

Fig. 1. (a) Sample exit pupil phase with linear phase subtracted and a color bar in radians, (b) sample wavefront phase having also had the defocus subtracted with a color bar in radians, (c) map of the subaperture/pupil plane with a color bar indicating number of subaperture positions overlapping the pupil, and (d) simulated noisy PSF from the sample exit pupil phase and one of the subaperture positions.

requires the variance of wavefront sensing error by Zernike coefficient for a single-PSF retrieval. Such variances will depend on the particular aberration coefficient vector but we assume that the variance of wavefront sensing error for a given Zernike coefficient averaged over the ensemble of coefficient vectors is adequate. In Section 6, a particular simulated wavefront sensing experiment is described for which ensemble-averaged variances have been estimated by a Monte Carlo process described in that section. In this section, we take these mean wavefront sensing error variances as a given and they are listed in Table 5. To the extent that the astigmatism variances are larger than that of defocus and the defocus variance is larger than the coma terms, these mean variances are consistent with the variances calculated numerically in [19], though the exact values and the experiment simulated differ.

For each of the coefficient vectors, the value of $\mathrm{var}(\hat{q})$ was evaluated in the SLP and USLP cases using Eq. (23) assuming the wavefront sensing error variances in Table 5. A log histogram of $\mathrm{var}(\hat{q})$ values over the population of 500 vectors is shown in Fig. 2. Since the aberration statistics chosen do not have an orientation bias, a log histogram of $\mathrm{var}(\hat{r})$ would

Table 5. Mean Wavefront Sensing Error Variances Estimated from the Simulated Phase Retrieval Experiment in Section 6

Term	Mean Value (rad^2)
σ_2^2	9.44×10^{-5}
σ_3^2	9.49×10^{-5}
σ_4^2	9.15×10^{-5}
σ_5^2	1.58×10^{-4}
σ_6^2	1.55×10^{-4}
σ_7^2	1.84×10^{-5}
σ_8^2	1.83×10^{-5}

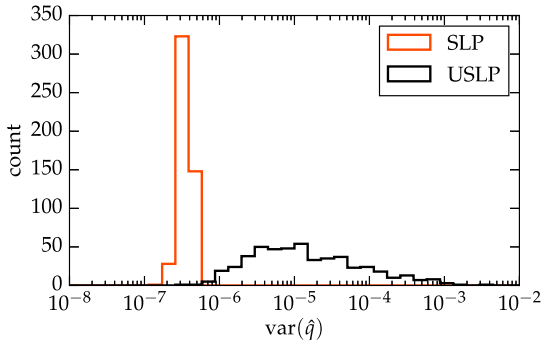

Fig. 2. Log histogram of var($\hat{q}$) over the population of coefficient vectors in the SLP and USLP cases.

yield nearly identical distributions. Of particular note, the SLP cases are clustered around an average value of 3.58×10^{-7}. Substituting σ_4^2 from Table 5 into the defocus-dominant limit in Eq. (30) yields a prediction for var($\hat{q}$) of 3.57×10^{-7}. This confirms that Eq. (30) is useful for predicting the expected variance of translation estimation in this defocus-dominated SLP case without requiring a Monte Carlo experiment, just as Section 3 noted. The maximum variance of the population is about 1.3 times the mean variance and the mean variance is about 2.9 times the minimum. The square root of the mean variance yields a standard deviation of the subaperture translation estimation error of 0.03% of the width of the reference circle described by $u^2 + v^2 \le 1$.

In contrast, the histogram for the USLP case in Fig. 2 is far more spread out and denotes variances much larger than the SLP case. The USLP variances range over almost 4 orders of magnitude with a maximum 44 times that of the mean and the mean variance is 214 times the minimum. The mean variance of 6.75×10^{-5} in the USLP case is about 189 times larger than the mean in the SLP case. The standard deviation of subaperture translation estimation error corresponding to the mean variance is 0.41% of the width of the circle $u^2 + v^2 \le 1$ while the maximum variance corresponds to a standard deviation of 2.7% of the width of the circle. A worse-case translation error of 2.7% could be unacceptably large for some demanding applications.

Since Eqs. (26)–(28) and Eqs. (32) and (33) are linear in the inverse of the wavefront sensing error variances, Eq. (23) would be linear in any constant factor change that was applied to all wavefront sensing error variances equally. Supposing that the wavefront sensing errors in Table 5 were each 100 times larger, the variance statistics calculated in this section would also be 100 times larger and the standard deviations larger by a factor of 10. In this case, the worst aberration vector would yield a translation error of 27%, which is almost certainly unacceptably large, since it would be very unclear to what part of the exit pupil each PSF's phase measurements corresponded.

6. COMPARISON WITH PHASE RETRIEVAL SIMULATION

Next, we used Monte Carlo analysis to examine whether the bound suggested by Eq. (23) is consistent with the variance of translation estimation error of phase retrieval in simulation. For each of the aberration coefficient vectors, 30 TTDPR experiments were simulated for the SLP and USLP cases with varying noise realizations. In each of these TTDPR experiments, seven PSFs were simulated from varying positions of a circular subaperture defined by $u^2 + v^2 \le 0.25$. The subaperture positions were arranged as indicated by the map of subaperture overlap in Fig. 1(c). This arrangement of the subapertures just inscribes the circle $u^2 + v^2 \le 1$. To each of the PSFs, different realizations of simulated Poisson- and Gaussian-distributed detector noise were applied. It was assumed that the peak pixel received 7,000 photoelectrons and there was a read noise of 30 photoelectrons. The simulated detector pixels were Nyquist sampled in intensity for a notional circular aperture defined by $u^2 + v^2 \le 1$. No such circular aperture stop was simulated in the fixed aperture, though, to be consistent with the assumptions of Section 2 that specify no amplitude fiducials. TTDPR computation using the method of [1] was applied to each of the simulated data to estimate exit pupil phase and subaperture translation assuming the SLP and USLP cases.

TTDPR is a nonlinear optimization technique that locates a minimum of an error metric given a starting point for the unknown phase aberrations and translations. For this analysis, we wished to assess the translation estimation accuracy of the TTDPR algorithm limited primarily by detector noise and independent of stagnation issues associated with converging to the correct exit pupil phase estimate. Therefore, we chose as starting aberration coefficient values the actual aberration coefficients used in the simulated data. Doing the same with the translation estimates, however, does not yield representative translation estimation errors in our testing. If we start the minimization process at the true subaperture translations, it has a tendency to find a minimum of the error metric near the true solution even when deeper minima of the metric exist somewhat further away from the true solution. In metrology where the true translation estimation is unknown, it is the translation estimation errors associated with these deeper minima that are of interest since these are likely to be found by the minimization process. To increase the chance of locating these deeper minima, 20 phase retrievals were done for each PSF noise realization with different initial subaperture translation estimates perturbed from the true values by a Gaussian-distributed random error having a variance of 2.25×10^{-2}. Of these 20 retrievals, the retrieval with the lowest final error metric was selected as the result for a given PSF noise realization.

We evaluated MSE(s), the mean squared difference between the retrieved translations and the actual translations present in the simulated PSF data, over the selected phase retrievals from each of the 30 noise realizations for each aberration vector. If the subaperture estimation in the TTDPR performed in simulation is unbiased, it would be expected that

$$\mathrm{MSE}(s) \gtrsim \mathrm{var}(\hat{q}), \tag{34}$$

where var($\hat{q}$) is calculated using Eq. (23) since it is a Cramer–Rao lower bound. The validity of this bound is contingent on the assumptions of Section 2 including the approximation used to model the effects of detector noise on single-PSF

phase retrieval. In particular, it was assumed that the noise in the PSF induces errors in the retrieved Zernike coefficients equivalent to uncorrelated, independently distributed Gaussian noise. Though TTDPR applied to PSFs corrupted by Gaussian and Poisson noise yield different error statistics, we take the variance of actual single-PSF phase retrieval to be an estimate for the variances σ_j^2 in $\boldsymbol{C}_{\boldsymbol{\eta}}$. We assessed the variance of single-PSF phase retrieval through additional Monte Carlo simulation experiments involving one PSF arising from an untranslated subaperture centered in the pupil. For each of the σ_j^2 in $\boldsymbol{C}_{\boldsymbol{\eta}}$, a Monte Carlo experiment was performed to estimate the variance of phase retrieval wavefront sensing error with respect to wavefront sensing error in just the jth Zernike term in the presence of PSF noise. Thirty noisy PSFs were synthesized for each of the randomly drawn aberration coefficient vectors used in the main simulation described earlier. For each noisy PSF and selected aberration type, phase retrieval was performed to estimate the jth Zernike coefficient while the other aberrations had their coefficients assumed to be the true value. Subaperture translation was not estimated and the phase retrieval algorithm used the correct translation. The squared errors of the wavefront sensing error for each Zernike term were averaged over the population of aberration coefficient vectors and noise realizations. These results are listed in Table 5. By estimating the error in each Zernike term individually, we excluded from our variance estimates the effects of the covariance of errors between different Zernike terms. This made our experimentally estimated variances consistent with the assumption of no covariance made in Section 2.

In Fig. 3, the value of MSE(s) for each aberration coefficient vector is plotted versus var($\hat{q}$) assuming the mean covariances in Table 5 for both the SLP and USLP cases. Also plotted is a line representing the Cramer–Rao lower bound in Eq. (34). Each of the 500 aberration coefficient vectors yielded a MSE(s) consistent with the Cramer–Rao lower bound inequality in Eq. (34) for the SLP case. In the USLP case, three of the vectors out of the 500 had MSE(s) slightly less than the bound predicts, the least of which had a MSE(s) 12% smaller

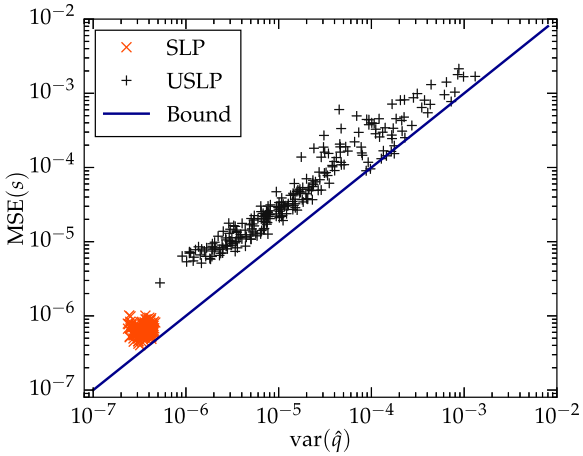

Fig. 3. Log–log scatter plot of MSE(s) of simulated TTDPR experiments for a given aberration vector versus the σ^{-2} var($\hat{q}$) for the aberration for both the SLP and USLP cases. Includes line representing the predicted Cramer–Rao lower bound of Eq. (34).

than the bound. Still, the vast majority of aberration coefficient vectors had estimation errors consistent with the bound. This is remarkable given the significant approximations made in Section 2 and simplification of the wavefront sensing noise to a Gaussian model that neglects covariance. It is also impressive that the translation errors of the simulated TTDPR are correlated to the simplified variance model in Eq. (23) as they are over the 3.8 orders of magnitude encompassed by Fig. 3. We found that 89% of the aberration vectors had a MSE(s) between var($\hat{q}$) and 4 × var($\hat{q}$) suggesting that Eq. (23) is a useful predictor of translation error for the actual phase retrieval simulated and not just an approximate lower bound. Since the TTDPR algorithm was estimating both the subaperture translations and the phase aberrations shared between PSFs, it would not be expected to achieve the Cramer–Rao lower bound for estimating translation alone. That it approximately follows the bound Eq. (23) but with a constant scale factor is unexpected but potentially useful for predicting the behavior of real measurements that involve maximum-likelihood estimators rather than minimum-variance estimators.

Next, we compared MSE(s) values in the experiment involving simulated TTDPR between the USLP and SLP cases. The MSE(s) across the population in the USLP case was 1.51×10^{-4} and 230 times that of the MSE in the SLP case. The ratio predicted by comparing the mean values of var($\hat{q}$) between USLP and SLP cases in Section 5 was 189 so the actual ratio of 230 is only slightly more than expected. The error between 189 and 230 of just 21% suggests that the theory outlined in Section 5 is useful for comparing the behavior of SLP and USLP cases of real systems when averaged across an ensemble of potential aberrations.

Finally, we compared the wavefront sensing accuracy by finding the mean of the RMS error between the retrieved aberrations and the true aberrations for each simulated phase retrieval. These RMS errors were averaged to estimate the expected wavefront sensing error over the population of aberrations. In the SLP case, the mean RMS error was 7.6×10^{-3} rad while in the USLP case it was 1.1×10^{-2} rad. If the wavefront sensing requirements for the simulated wavefront sensing experiment were one-hundredth wave RMS, or 6.2×10^{-2} rad RMS, TTDPR in both the SLP and USLP cases could be expected to succeed for the selected population of aberrations.

7. CONCLUSION

In TTDPR, PSFs of an optical system are acquired for use in a phase retrieval algorithm to estimate the aberrations of that system. A subaperture in a plane conjugate to the exit pupil is translated to a different position for each PSF. In cases where the knowledge of this translation is insufficiently accurate or unavailable, it must be estimated from the PSF data. Misestimation of the translation of the subaperture leads to misestimation of the aberrations, so the accuracy of this translation estimation is critical. In Section 2, we developed analytic expressions for the variance of a minimum-variance unbiased estimator and a Cramer–Rao lower bound for subaperture translation estimation assuming various approximations. Notable among these approximations, we assumed that the exit pupil does not have fixed amplitude features that would aide

subaperture translation estimation. For unobscured imaging systems, most PSFs will derive from subaperture translations that do not overlap an amplitude feature that would aide translation estimation. The Cramer–Rao lower bound depends on wavefront sensing quality and the composition of phase aberrations of the system. Which aberrations contribute to the variance varies according to whether each PSF arose from identical linear phase terms or dissimilar linear phase terms due to unknown translations of the detector acquiring the PSF or of the target being imaged. We refer to these two cases as the shared linear phase and unshared linear phase cases, respectively.

The expressions for the bound on translation estimation vary significantly between the SLP case evaluated in Section 3 and USLP cases evaluated in Section 4. The aberrations with radial degree of 2 contribute to lowering the variance of translation estimation in the SLP case but do not directly contribute in the USLP case according to the model of Section 2. In the common situation that a significant defocus has been introduced to improve wavefront sensing, USLP is at a significant disadvantage because this additional defocus does not directly aid translation estimation as it does in the SLP case. We note, though, that the model in Section 2 does not take into account the variation of wavefront sensing error σ due to the composition of aberrations. In practice, increasing the magnitude of defocus and astigmatism will likely decrease wavefront sensing error for all aberrations and thus aide translation estimation even in the USLP case.

Evaluating whether translation estimation is viable for a particular experiment requires knowledge of the noise inherent in the wavefront sensing process and predictions for the actual aberrations expected in the subaperture plane. For instance, if the aberrations are dominated by large coma or spherical aberration terms, translation estimation may perform acceptably well in both SLP and USLP cases, depending on the wavefront sensing error. In Section 5, we chose a specific large population of randomly generated aberrations and evaluated the predictions of the theory to understand and evaluate the differences between the SLP and USLP cases assuming one had an efficient estimator that met the Cramer–Rao lower bound. From this analysis, we observed that the mean value of translation variance over this population of aberrations should be about 189 times larger in the USLP case than in the SLP case. In Section 6, we simulated TTDPR using simulated PSFs and actual TTDPR algorithm evaluations. Translation estimate errors from this actual phase retrieval algorithm would not be expected to meet the Cramer–Rao lower bound but they should not be lower than the Cramer–Rao bound derived in Section 2. This was the case for all of the aberrations examined in the SLP case and virtually all of the aberrations in the USLP case, suggesting that the bound was plausible. The 0.6% of the aberrations that violated the bound in the USLP case did so by at most 12%, which is understandable given the nature of the approximations in the bound. For this set of experiments, the mean squared error of translation was 1.51×10^{-5} in the USLP case where the units of translation are radii of the circle $u^2 + v^2 \leq 1$. This mean squared error was 230 times larger than the average SLP case or, alternately, more than 15 times larger if the RMS error values were compared. Our comparison between theory-predicted Cramer–Rao bound for the variance and the actual mean squared error achieved by TTDPR suggested a strong correlation, as shown by Fig. 3. We found that 89% of the aberrations yielded translation variances between the predicted Cramer–Rao bound and 4 times the bound over a range of bound values encompassing 3.8 orders of magnitude. This level of agreement is remarkable given that the model for the Cramer–Rao bound does not take into account the variation of wavefront sensing error with aberration composition or the covariance of the wavefront sensing errors between the sensed wavefront coefficients. The simulation in Section 5 demonstrated a phase aberration sensing error better than one-hundredth wave RMS in both the SLP and USLP cases. On average, the USLP cases exhibited slightly higher wavefront sensing error than the SLP cases due to the effects of additional subaperture translation misestimation.

It is likely that these results derived for TTDPR have an analog in conventional subaperture stitching interferometry when the position of the subaperture has errors that must be estimated from interferogram data.

Finally, we return to evaluating TTDPR for use with the long-wave channel of the NIRCam on the James Webb Space Telescope. The USLP case arises in ground testing involving external pupil illumination as well as when the internal light sources of NIRCam are used for ground testing or on-orbit testing of NIRCam. The results of Section 4 indicate that additional defocus does not directly aid translation estimation in the USLP case, which must rely solely on higher-order aberrations. Consequently, the lack of weak defocus lenses in the long-wave channel does not significantly hinder translation estimation for that case. If weak lenses were present, the improvement would only be that due to the indirect effect of increasing wavefront sensing accuracy. Instead, it is the third radial degree aberrations like coma and trefoil in NIRCam that directly assist TTDPR. Fortunately, the NIRCam system develops such aberrations when the Lyot stops are introduced. This is because the Lyot stop is patterned on a prism substrate that shifts the field of view of NIRCam when selected in the filter wheel [12]. NIRcam involves a collimating lens assembly and a camera lens assembly whose individual aberrations cancel most ideally when the prism is not introduced between them. We have observed that introduction of the prism element unbalances the overall aberration correction in NIRCam enough for successful USLP TTDPR. In cases where NIRCam is presented with an input beam having well-defined fixed amplitude features, such as the segment boundaries of the telescope's primary mirror, estimation is likely to be even better than that achieved using pupil phase alone.

In conclusion, we have computed a Cramer–Rao lower bound for subaperture translation estimation in TTDPR aberration retrieval when amplitude fiducials are unavailable in the pupil. This estimation is essential to phase measurements when the translation must be estimated using the PSF data. The bound shows that Zernike coma, trefoil, and those terms of higher radial degree are critical to estimation when the linear phase terms of the pupil are effectively unknown and also vary from PSF to PSF. Monte Carlo experiments involving simulations of TTDPR were consistent with the lower bound and also

correlated with the bound, suggesting that it could approximate the average MSE expected. We believe that the insights provided by this approximate but analytic bound will be invaluable for designing robust optical metrology experiments using TTDPR in difficult situations.

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

Acknowledgment. We would like to thank NASA's Goddard Space Flight Center for their support under grant NNX15AE94A.

REFERENCES

1. 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).
2. T. P. Zielinski and J. R. Fienup, "Phase retrieval with a translating Lyot stop coronagraph mask in the JWST," in *Signal Recovery and Synthesis*, OSA Technical Digest Series (Optical Society of America, 2009), paper SWA2.
3. D. B. Moore and J. R. Fienup, "Fast linear approximation for phase retrieval of partially coherently illuminated objects," in *Wavefronts and Aberrations*, OSA Technical Digest Series (Optical Society of America, 2012), paper FTu2F.4.
4. D. B. Moore and J. R. Fienup, "Transverse translation diversity wavefront sensing with limited position and pupil illumination knowledge," Proc. SPIE **9143**, 91434F (2014).
5. H. M. L. Faulkner and J. M. Rodenburg, "A phase retrieval algorithm for shifting illumination," Appl. Phys. Lett. **85**, 4795–4797 (2004).
6. H. M. L. Faulkner and J. M. Rodenburg, "Movable aperture lensless transmission microscopy: a novel phase retrieval algorithm," Phys. Rev. Lett. **93**, 023903 (2004).
7. M. Guizar-Sicairos and J. R. Fienup, "Phase retrieval with transverse translation diversity: a nonlinear optimization approach," Opt. Express **16**, 7264–7278 (2008).
8. M. Guizar-Sicairos and J. R. Fienup, "Measurement of coherent x-ray focused beams by phase retrieval with transverse translation diversity," Opt. Express **17**, 2670–2685 (2009).
9. F. Zhang, I. Peterson, J. Vila-Comamala, A. Diaz, F. Berenguer, R. Bean, B. Chen, A. Menzel, I. K. Robinson, and J. M. Rodenburg, "Translation position determination in ptychographic coherent diffraction imaging," Opt. Express **21**, 13592–13606 (2013).
10. A. Tripathi, I. McNulty, and O. G. Shpyrko, "Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods," Opt. Express **22**, 1452–1466 (2014).
11. L. W. Huff, "NIRCam instrument optics," Proc. SPIE **5904**, 590404 (2005).
12. J. E. Krist, C. A. Beichman, J. T. Trauger, M. J. Rieke, S. Somerstein, J. J. Green, S. D. Horner, J. A. Stansberry, F. Shi, M. R. Meyer, K. R. Stapelfeldt, and T. L. Roellig, "Hunting planets and observing disks with the JWST NIRCam coronagraph," Proc. SPIE **6693**, 66930H (2007).
13. 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).
14. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **68**, 207–211 (1967).
15. J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts and Company, 2005).
16. D. Malacara, *Optical Shop Testing*, 3rd ed. (Wiley, 2007).
17. A. Guirao, D. R. Williams, and I. G. Cox, "Effect of rotation and translation on the expected benefit of an ideal method to correct the eye's higher-order aberrations," J. Opt. Soc. Am. A **18**, 1003–1015 (2001).
18. P. C. L. Stephenson, "Recurrence relations for the Cartesian derivatives of the Zernike polynomials," J. Opt. Soc. Am. A **31**, 708–715 (2014).
19. L. Meynadier, V. Michau, M.-T. Velluet, J.-M. Conan, L. M. Mugnier, and G. Rousset, "Noise propagation in wave-front sensing with phase diversity," Appl. Opt. **38**, 4967–4979 (1999).
20. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice Hall, 1993).
21. J. R. Fienup, J. C. Marron, T. J. Schulz, and J. H. Seldin, "Hubble space telescope characterized by using phase retrieval algorithms," Appl. Opt. **32**, 1747–1767 (1993).
22. J. R. Fienup, B. J. Thelen, R. G. Paxman, and D. A. Carrara, "Comparison of phase diversity and curvature wavefront sensing," Proc. SPIE **3353**, 930–940 (1998).