Phase-error estimation and image reconstruction from digital-holography data using a Bayesian framework

CASEY J. PELLIZZARI,[1,*] MARK F. SPENCER,[2] AND CHARLES A. BOUMAN[1]

[1]*School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907, USA*
[2]*Air Force Research Laboratory, Directed Energy Directorate, KAFB, Albuquerque, New Mexico 87111, USA*
Corresponding author: cpelliz@purdue.edu

The estimation of phase errors from digital-holography data is critical for applications such as imaging or wave-front sensing. Conventional techniques require multiple i.i.d. data and perform poorly in the presence of high noise or large phase errors. In this paper, we propose a method to estimate isoplanatic phase errors from a single data realization. We develop a model-based iterative reconstruction algorithm that computes the maximum *a posteriori* estimate of the phase and the speckle-free object reflectance. Using simulated data, we show that the algorithm is robust against high noise and strong phase errors.   © 2017 Optical Society of America

1. INTRODUCTION

Digital holography (DH) uses coherent-laser illumination and detection to gain significant improvements in sensitivity for remote-sensing applications. Detection involves measuring the modulation of a strong reference field by a potentially weak signal field. This modulation allows for the detection of signals with energies equivalent to a single photon or less [1].

In practice, DH systems are sensitive to phase errors imparted by atmospheric perturbations or flaws in the optical system. For imaging applications, we must estimate and remove these phase errors to form a focused image [2–10]. Similarly, for wavefront-sensing applications, estimation of the phase errors represents the desired sensor output [11–14]. In either case, accurately estimating the phase errors is a necessary and critical task.

State-of-the-art techniques for estimating phase errors from DH data involve maximizing an image sharpness metric with respect to the phase errors [3–5,7–10]. Typically, multiple data realizations are obtained for which the image-to-image speckle variation is decorrelated but the phase errors are identical. For applications using pupil-plane detection, the data realizations are corrected using an initial estimate of the phase errors and then inverted using a fast Fourier transform (FFT). The inversion produces estimates of the complex-valued reflection coefficient, g, which are incoherently averaged to reduce speckle variations. The sharpness of this speckle-averaged image is then maximized to obtain the final phase-error estimate.

Image sharpening (IS) algorithms are designed to use multiple data realizations. However, Thurman and Fienup demonstrated

that in favorable conditions, the algorithms can still be used when only one realization is available [3]. For a high signal-to-noise ratio (SNR), with several hundred detected photons per detector element, their algorithm was able to estimate atmospheric phase functions with residual root-mean-square (RMS) errors as low as 1/8 *wave*.

While IS algorithms have been successfully demonstrated for estimating both isoplanatic and anisoplanatic phase errors, there remains room for improvement. First, IS algorithms use relatively simple inversion techniques to reconstruct the magnitude squared of the reflection coefficient, $|g|^2$, rather than the real-valued reflectance, r. The reflectance, given by $r = E[|g|^2]$, where $E[\cdot]$ indicates the expected value, is a smoother quantity with higher spatial correlation between elements in the image. We are accustomed to seeing the reflectance in conventional images and it is of greater interest for many imaging applications. Conversely, reconstructing $|g|^2$ leads to images with high-spatial-frequency variations known as speckle. By reconstructing r, we can leverage its higher spatial correlation to better constrain the estimation process and potentially produce more accurate estimates of the phase errors [15].

Another limitation of IS algorithms is that the process of estimating the phase errors is not tightly coupled to image reconstruction. Information about the object can be further leveraged during the phase-estimation process. For example, this information can help rule out contributions to the phase caused by noise. By jointly estimating both the phase errors and the reflectance, we can incorporate additional information into our estimates, which helps reduce uncertainty [15].

Lastly, in many practical applications it may not be possible to obtain multiple data realizations. Particularly, in cases where there is not enough relative movement between the object and the receiver to decorrelate the speckle variations between images, the phase errors are rapidly changing, or when timing requirements prohibit multiple images from being obtained. Without multiple realizations, IS algorithms can struggle to produce accurate phase-error estimates [3] and an alternate method is needed.

In this paper, we propose an improved method of DH phase recovery based on a framework of joint image reconstruction and phase-error estimation. Our approach builds on the work of [15] by reconstructing the spatially correlated reflectance, r, rather than the reflection coefficient, g, which allows for better estimates of the phase errors from less data. We focus on reconstructing from a single data realization under isoplanatic atmospheric conditions for the off-axis pupil-plane recording geometry (PPRG) [13]. Our major contributions include:

1. We jointly compute the maximum *a posteriori* (MAP) estimates of the image reflectance, r, and phase errors, ϕ from a single data realization. Joint estimation reduces uncertainty in both quantities. Additionally, estimating r rather than g helps to further constrain the problem and produce images without speckle.

2. We derive the forward model for a DH system using the off-axis PPRG. The model ensures our estimates are consistent with the physics and statistics of the remote-sensing scenario.

3. We develop a technique to compute the 2D phase errors that occur in DH. Our approach allows for the estimation of phase errors on a lower resolution grid to help reduce the number of unknowns.

4. We model the phase errors as a random variable using a Gaussian Markov random field (GMRF) prior model. This approach allows us to compute the MAP estimate of the phase errors, which constrains the estimate to overcome high noise and strong turbulence.

5. We compare the proposed algorithm to the image sharpening approach in [3] over a range of SNRs and atmospheric turbulence strengths.

Overall, our experimental results using simulated, single-frame, isoplanatic data demonstrate that the proposed MBIR algorithm can reconstruct substantially higher-quality phase and image reconstructions compared to IS algorithms.

2. ESTIMATION FRAMEWORK

Figure 1 shows an example scenario for a DH system using the off-axis PPRG. Our goal is to compute the MAP estimates of the reflectance, $r \in \mathbb{R}^N$, and the phase errors, $\phi \in \mathbb{R}^N$, from the noisy data, $y \in \mathbb{C}^N$. Note that we are using vectorized notation for these 2D variables. The joint estimates are given by

$$(\hat{r}, \hat{\phi}) = \underset{(r,\phi)\in\Omega}{\mathrm{argmin}}\{-\log p(r, \phi|y)\}$$

$$= \underset{(r,\phi)\in\Omega}{\mathrm{argmin}}\{-\log p(y|r, \phi) - \log p(r) - \log p(\phi)\}, \quad (1)$$

where Ω represents the jointly feasible set and the quantities r and ϕ are assumed to be independent.
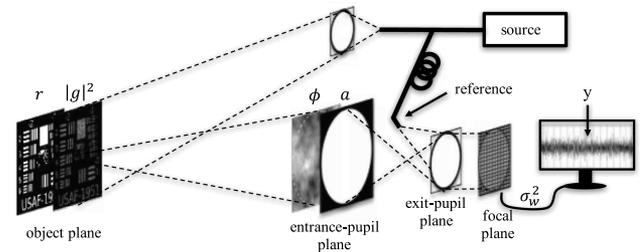

Fig. 1. Example DH system using an off-axis pupil-plane recording geometry. A coherent source is used to flood illuminate an object, which has a reflectance function, r, and corresponding reflection coefficient, g. The return signal is corrupted by atmospheric phase errors, ϕ, and passes through the entrance pupil to the imaging system, a. Both a and ϕ are considered to be located in the same plane. The entrance-pupil plane is then imaged onto a focal-plane array where it is mixed with a reference field. For simplicity, only the entrance and exit pupils of the optical system are shown. Finally, the noisy data, y (with noise power σ_w^2), is processed to form an image and/or an estimate of ϕ.

To evaluate the cost function for Eq. (1), we must derive the model for $p(y|r, \phi)$. The reader is referred to Appendix A for a more detailed derivation. In short, we can represent the data using an additive noise model given by

$$y = Af + w, \quad (2)$$

where $w \in \mathbb{C}^N$ is the measurement noise, $A \in \mathbb{C}^{N\times N}$ is the linear forward model operator, and $f \in \mathbb{C}^N$ is the complex field in the object plane. For an object with reflectance function $g \in \mathbb{C}^N$, f is defined as

$$f = \Gamma g, \quad (3)$$

where $\Gamma \in \mathbb{C}^{N\times N}$ is a diagonal matrix that applies the object-plane, quadratic phase factor from the Fresnel propagation integral [16]. Note that $|f_i| = |g_i|$, and therefore $r_i = E[|g_i|^2] = E[|f_i|^2]$ for all i. By representing the data as a function of f, rather than g, we avoid having to explicitly define Γ in our model.

In Eq. (2), the matrix A accounts for the propagation and measurement geometry. If we ignore the blurring effects caused by sampling the signal with finite-sized pixels, A can be decomposed as

$$A = \mathcal{D}(a)\mathcal{D}(\exp\{j\phi\})F. \quad (4)$$

Here $\mathcal{D}(\cdot)$ denotes an operator that produces a diagonal matrix from its vector argument, $a \in \mathbb{R}^N$ is the entrance-pupil transmission function, and ϕ is the phase-error function previously defined. For our purposes, a is a binary circular function defining the entrance pupil of the imaging system, as shown in Fig. 1. Finally, we choose the reconstruction parameters such that $F \in \mathbb{C}^{N\times N}$ is a two-dimensional discrete Fourier transform (DFT) matrix scaled so that $F^H F = I$.

For surfaces that are rough relative to the illumination wavelength, the vector g can be modeled as a conditionally complex Gaussian random variable. Given the underlying reflectance, r, of the scene, the conditional distribution is given by

$$p(g|r) \sim CN(0, \mathcal{D}(r)), \quad (5)$$

where for any random vector $z \in \mathbb{C}^M$, $CN(\mu, C)$ is the multivariate complex normal distribution with mean, μ, and covariance matrix, Σ, such that

$$CN(\mu, C) = \frac{1}{\pi^M |\Sigma|} \exp\{-(z - \mu)^H \Sigma^{-1}(z - \mu)\}. \quad (6)$$

Here, superscript H indicates the Hermitian transpose. The vector, f, then has an identical distribution given by

$$p(f|r) \sim CN(0, \Gamma \mathcal{D}(r)\Gamma^H)$$
$$= CN(0, \mathcal{D}(r)). \quad (7)$$

The equality in Eq. (7) results from the commutative property of diagonal matrices and from the fact that $\Gamma\Gamma^H = I$.

It is common for coherent detection systems to use a reference beam that is much stronger than the return signal. For such cases, shot noise, driven by the power of the reference beam, is the dominate source of measurement noise and can be modeled as additive, zero-mean, complex Gaussian white noise [17,18]. The distribution of w is therefore given by

$$p(w) \sim CN(0, \sigma_w^2 I), \quad (8)$$

where σ_w^2 is the noise variance.

From Eqs. (2)–(8), the likelihood function of the data, given the reflectance and the phase errors, is distributed according to [19]

$$p(y|r, \phi) \sim CN(0, A\mathcal{D}(r)A^H + \sigma_w^2 I). \quad (9)$$

Therefore, the MAP cost function is given by

$$c(r, \phi) = -\log p(y|r, \phi) - \log p(r) - \log p(\phi)$$
$$= \log|A\mathcal{D}(r)A^H + \sigma_w^2 I| + y^H(A\mathcal{D}(r)A^H + \sigma_w^2 I)^{-1}y$$
$$- \log p(r) - \log p(\phi). \quad (10)$$

Unfortunately, the determinate and inverse make direct optimization of Eq. (10) an extremely computationally expensive task.

Equation (10) is similar to the MAP cost function in [15], except that the forward model has changed and we include the distribution of ϕ in the cost function. Despite these differences, we use a similar optimization approach, which leverages the expectation maximization (EM) algorithm and allows us to replace Eq. (10) with a more tractable surrogate function.

In order to use the EM algorithm, we need to first introduce the concept of surrogate functions. For any function $c(x)$, we define a surrogate function, $Q(x; x')$, to be an upper-bounding function, such that $c(x) \leq Q(x; x') + \kappa$, where x' is the current value of x that determines the functional form of Q and κ is a constant that ensures the two functions are equal at x'. Surrogate functions have the property that minimization of Q implies minimization of f. That is,

$$\{Q(x; x') < Q(x'; x')\} \Rightarrow \{c(x) < c(x')\}. \quad (11)$$

Surrogate functions are useful because in some cases it is possible to find a simple surrogate function, Q, that we can iteratively minimize in place of a more difficult to compute function, c. In fact, the EM algorithm provides a formal framework for constructing such a surrogate function. More specifically, for our problem, the surrogate function of the EM algorithm takes the following form:

Repeat{
$$\hat{r} \leftarrow \underset{r}{\operatorname{argmin}}\, Q(r, \phi'; r', \phi')$$
$$\hat{\phi} \leftarrow \underset{\phi}{\operatorname{argmin}}\, Q(r', \phi; r', \phi')$$
$$r' \leftarrow \hat{r}$$
$$\phi' \leftarrow \hat{\phi}$$
}

Fig. 2. EM algorithm for joint optimization of the MAP cost surrogate function.

$$Q(r, \phi; r', \phi') = -\mathrm{E}[\log p(y, f|r, \phi)|Y = y, r', \phi']$$
$$- \log p(r) - \log p(\phi), \quad (12)$$

where r' and ϕ' are the current estimates of r and ϕ, respectively. A core idea of the EM algorithm is that there is missing or unobserved data that we can only indirectly observe through the measured data, y [20]. Since knowledge of this unobserved data would simplify the problem, we conjecture what that data must be, given our knowledge of the observed data [21]. Here we choose f to be the unobserved data.

Evaluation of the expectation in Eq. (12), with respect to f, constitutes the E-step of the EM algorithm. After the surrogate function is specified, we conduct the M-step to jointly minimize the Q function according to

$$(\hat{r}, \hat{\phi}) = \underset{(r,\phi)}{\operatorname{argmin}}\{Q(r, \phi; r', \phi')\}. \quad (13)$$

Figure 2 shows the alternating minimization approach used for implementing Eq. (13). Note that we jointly minimize Q with respect to both variables before we update its functional form by updating r' and ϕ'. Updating r' prior to computing $\hat{\phi}$ leads to a slightly different algorithm; however, both approaches monotonically reduce the MAP cost function and we observed that both approaches produce very similar results.

The proposed algorithm of Fig. 2 shares the same convergence and stability properties as the algorithm in [15]. More specifically, the EM algorithm converges in a stable manner to a local minima. Since both the MAP cost function and its surrogate are nonconvex, the algorithm is likely to converge to a local minimum that is not a global minimum. We cannot directly evaluate the MAP cost function to compare different local minima. Therefore, to ensure we find a good local minima, in the sense that it produces a focused and useful image, we pay close attention to our initial conditions. In Section 4.C, we present a multistart algorithm that we empirically found to robustly compute good local minima.

3. PRIOR MODELS

In this section, we present the distributions used to model both the reflectance, r, and the phase errors, ϕ. In both cases, we choose to use Markov random fields (MRFs) with the form

$$p(r) = \frac{1}{z} \exp\left\{ -\sum_{\{i,j\} \in \mathcal{P}} b_{i,j} \rho(\Delta) \right\}, \quad (14)$$

where z is the partition function, $b_{i,j}$ is the weight between neighboring pixel pairs (r_i and r_j, or ϕ_i and ϕ_j), Δ is the

difference between those pixel values, $\mathcal{P}$ is the set of all pairwise cliques falling within the same neighborhood, and $\rho(\cdot)$ is the potential function [21]. We obtain different models for the reflectance and phase errors by choosing different potential functions.

A. Reflectance

For the reflectance function, r, we used a Q-generalized Gaussian Markov random field (QGGMRF) potential function with the form

$$\rho_r\left(\frac{\Delta_r}{\sigma_r}\right) = \frac{|\Delta_r|^p}{p\sigma_r^p}\left(\frac{\left|\frac{\Delta_r}{T\sigma_r}\right|^{q-p}}{1 + \left|\frac{\Delta_r}{T\sigma_r}\right|^{q-p}}\right), \qquad (15)$$

where T is a unitless threshold value that controls the transition of the potential function from having the exponent q to having the exponent p [22]. The variable, σ_r, controls the variation in $\hat{r}$. The parameters of the QGGMRF potential function affect its shape and therefore the influence neighboring pixels have on one another.

While more sophisticated priors exist for coherent imaging, the use of a QGGMRF prior is relatively simple and computationally inexpensive yet effective [15,23]. The low computational burden is important for DH applications where we desire near-real-time, phase-error estimations, such as with wavefront sensing. In this work, we focus on reconstructing relatively simple objects. The reader is directed to [23] for information about how these models affect reconstruction of more complicated objects.

B. Phase Errors

In practice, the phase-error function may not change much from sample to sample or we may wish to correct the phase at a defined real-time frame rate using a deformable mirror with fewer actuators than sensor measurements. Therefore, to reduce the number of unknowns, we allow the phase-error function, ϕ, to be modeled on a grid that has lower resolution than the measured data. We denote the subsampled function as $\bar{\phi}$.

Using $\bar{\phi}$ allows us to estimate only a single phase value for multiple data samples. This technique reduces the number of unknowns when estimating r and ϕ from $2N$ to $N(1 + 1/n_b^2)$, where n_b is the factor of subsampling used in both dimensions. To scale $\bar{\phi}$ to the resolution of ϕ for use in the forward model, we use a nearest-neighbor-interpolation scheme given by

$$\phi = P\bar{\phi}. \qquad (16)$$

Here, P is an $N \times N/n_b^2$ interpolation matrix with elements in the set [0,1]. While a more sophisticated interpolation method could be used, nearest neighbor requires no additional computation since we simply copy each sample, $\bar{\phi}_s$, to the corresponding samples in ϕ. This low computational cost is attractive for applications like wavefront sensing where speed is important.

To model the distribution of unwrapped phase-error values in $\bar{\phi}$, we use a Gaussian potential function given by

$$\rho_{\bar{\phi}}\left(\frac{\Delta_{\bar{\phi}}}{\sigma_{\bar{\phi}}}\right) = \frac{|\Delta_{\bar{\phi}}|^2}{2\sigma_{\bar{\phi}}^2}, \qquad (17)$$

where $\sigma_{\bar{\phi}}$ controls the variation in $\hat{\bar{\phi}}$. The Gaussian potential function enforces a strong influence between neighboring

samples, which generates a more smoothly varying output when compared to QGGMRF. Since the unwrapped phase function will not have edges or other sharp discontinuities, the GMRF is more appropriate than QGGMRF in this case. An added benefit is that it is less computationally expensive than the QGGMRF model, which allows for fast evaluation.

4. ALGORITHM

In this section, we present the details for how we execute the proposed EM algorithm shown in Fig. 2. The discussion describes how we evaluate the surrogate function in the E-step and how we optimize it in the M-step.

A. E-Step—Surrogate Function Evaluation

Using the models given by Eqs. (7)–(9) and (14)–(17), along with the conditional posterior distribution of f given y and r, we can evaluate the surrogate function of Eq. (12). Its final form is given by

$$Q(r, \phi; r', \phi') = -\frac{1}{\sigma_w^2}2\,\mathrm{Re}\{y^H A_\phi \mu\} + \log|\mathcal{D}(r)|$$
$$+ \sum_{i=1}^{N}\frac{1}{r_i}(C_{i,i} + |\mu_i|^2) + \sum_{\{i,j\}\in\mathcal{P}}b_{i,j}\rho_r\left(\frac{\Delta_r}{\sigma_r}\right)$$
$$+ \sum_{\{i,j\}\in\mathcal{P}}b_{i,j}\rho_{\bar{\phi}}\left(\frac{\Delta_{\bar{\phi}}}{\sigma_{\bar{\phi}}}\right). \qquad (18)$$

Here, A_ϕ indicates that the matrix A is dependent on ϕ, μ_i is the ith element of the posterior mean given by

$$\mu = C\frac{1}{\sigma_w^2}A_{\phi'}^H y, \qquad (19)$$

and $C_{i,i}$ is the ith diagonal element of the posterior covariance matrix given by

$$C = \left[\frac{1}{\sigma_w^2}A_{\phi'}^H A_{\phi'} + \mathcal{D}(r)^{-1}\right]^{-1}. \qquad (20)$$

We simplify the evaluation of Eq. (18) by approximating the posterior covariance as

$$C \approx \mathcal{D}\left(\frac{\sigma_w^2}{1 + \frac{\sigma_w^2}{r}}\right), \qquad (21)$$

which requires that

$$A_\phi^H A_\phi = F^H \Phi^H \Lambda^H \Lambda \Phi F \approx I. \qquad (22)$$

For a square entrance-pupil function filling the detector's field of view, $\Lambda = I$ and the approximations of Eqs. (21) and (22) become exact as both F and Φ are unitary matrices. When Λ represents a circular aperture, the relationship of Eq. (22) is approximate. However, since it dramatically simplifies computations, we will use this approximation and the resulting approximation of Eq. (21) in all our simulations when computing the Q function for the EM update. In practice, we have found this to work well.

B. M-Step—Optimization of the Surrogate Function

The goal of the M-step is to minimize the surrogate function according to Eq. (13), using the alternating optimization shown

in Fig. 2. Specifically, we use iterative coordinate descent (ICD) to update both r and ϕ [21]. ICD works by sequentially minimizing the entire cost function with respect to a single sample, r_s or ϕ_s.

By considering just the terms in Eq. (B6), which depend on a pixel r_s, the cost function for the reflectance update is given by

$$q_s(r_s; r', \phi') = \log r_s + \frac{C_{s,s} + |\mu_s|^2}{r_s} + \sum_{j \in \partial s} b_{s,j} \rho_r \left(\frac{r_s - r_j}{\sigma_r} \right), \tag{23}$$

where the sum over $j \in \partial s$ indicates a sum over all neighbors of the pixel r_s. We carry out the minimization of Eq. (23) with a 1D line search over $\mathbb{R}^+$.

To minimize Eq. (18) with respect to the phase errors, we consider just the terms that depend on $\bar{\phi}$. The phase-error cost function becomes

$$q(\phi; r', \phi') = -\frac{1}{\sigma_w^2} 2 \operatorname{Re}\{y^H A_\phi \mu\} + \frac{1}{2\sigma_{\bar{\phi}}^2} \sum_{\{i,j\} \in \mathcal{P}} b_{i,j} |\Delta_{\bar{\phi}}|^2. \tag{24}$$

We use ICD to sequentially minimize the cost with respect to each element of the subsampled phase-error function, $\bar{\phi}$. For element $\bar{\phi}_p$, corresponding to a $n_b \times n_b$ group of data samples with indices in the set B, the cost function becomes

$$q_p(\bar{\phi}_p; r', \phi') = -|\chi| \cos(\phi_\chi - \bar{\phi}_p) + \frac{1}{2\sigma_{\bar{\phi}}^2} \sum_{j \in \partial p} b_{i,j} |\bar{\phi}_p - \bar{\phi}_j|^2, \tag{25}$$

where

$$\chi = \frac{2}{\sigma_w^2} \sum_{i=1}^{n_b^2} y_{B(i)}^* [F\mu]_{B(i)} \tag{26}$$

and $j \in \partial p$ is an index over neighboring phase samples on the lower resolution grid. We minimize Eq. (25) using a 1D line search over $\bar{\phi}_p \in [\bar{\phi}^{**} - \pi, \bar{\phi}^{**} + \pi]$, where $\bar{\phi}^{**}$ minimizes the prior term. By minimizing Eq. (25), we obtain an estimate of the unwrapped phase errors. Finally, we obtain the full-resolution estimate of the unwrapped phase errors, ϕ, according to Eq. (16).

C. Initialization and Stopping Criteria

Figure 3 summarizes the steps of the EM algorithm. To determine when to stop the algorithm, we can use either a set number of iterations, N_K, or a metric such as

$$\epsilon = \frac{\|r^k - r^{k-1}\|}{\|r^{k-1}\|}, \tag{27}$$

where k is the iteration index, and we stop the algorithm when ϵ falls below a threshold value of ϵ_T.

The EM algorithm in Fig. 3 is initialized according to

$$r \leftarrow |A^H y|^{\circ 2}, \quad \sigma_r \leftarrow \frac{1}{\gamma} \sqrt{s^2(r)}, \tag{28}$$

where $|\cdot|^{\circ 2}$ indicates the elementwise magnitude square of a vector, γ is a unitless parameter introduced to tune the amount of regularization in r, and $s^2(\cdot)$ computes the sample variance of a vector's elements [24].

```
EM {
Inputs: y, r', φ', σ_r, σ_φ̄, σ²_w, q, p, T, b, (either N_K or ε_T)
Outputs: r̂, φ̂
while k < N_K or ε > ε_T  do
    C ← D ( σ²_w / (1 + σ²_φ̄/r') )
    μ ← C (1/σ²_w) A^H_φ' y,
    for all s ∈ S do
        r_s ← argmin {q_s(r_s; r', φ')}
              r_s∈ℝ⁺
    for all p ∈ P do
        φ_p ← argmin {q_p(φ̄_p; r', φ')}
               φ
    φ ← Fφ̄
}
```

Fig. 3. EM algorithm for the MAP estimates of r and ϕ. Here, S is the set of all pixels and P is the set of all phase-error elements.

```
MBIR Algorithm {
Inputs: y, γ, σ_w, q, p, T, b, N_K, N_L, ε_T
Outputs: r̂, φ̂
φ ← 0
for i = 1 : N_L do
    r ← |A^H y|^{∘2}, σ_r ← (1/γ)√var(r)
    φ ← EM {y, r, φ, σ_r, σ_φ̄, σ²_w, 2, 2, 1, G(0.8), N_K}
r ← |A^H y|^{∘2}, σ_r ← (1/γ)√var(r)
r, φ ← EM {y, r, φ, σ_r, σ_φ̄, σ²_w, q, p, T, b, ε_T}
}
```

Fig. 4. Algorithm that initializes and runs the EM algorithm. An iterative process is used to initialize the phase-error vector ϕ.

We use a heuristic similar to that used in [15] to iteratively initialize the phase-error estimate. Figure 4 details the steps of this iterative process. The initial estimate of the phase-error vector is simply $\phi \leftarrow 0$. Next, for a set number of outer-loop iterations, N_L, we allow the EM algorithm to run for N_K iterations. At the beginning of each outer-loop iteration, we reinitialize according to Eq. (28).

After the outer loop runs N_L times, we again reinitialize according to Eq. (28) and run the EM algorithm until it reaches the stopping threshold ϵ_T. A Gaussian prior model for r works best in the outer loop for the initialization of ϕ. Specifically, we use $q = 2$, $p = 2$, $T = 1$, $\gamma = 2$, and $b = G(0.8)$, where $G(\sigma)$ indicates a 3×3 Gaussian kernel with standard deviation, σ. Once the initialization process is complete, we can use different prior-model parameters for the actual reconstruction.

5. METHODS

To compare the proposed algorithm to an IS algorithm from [3], we generated simulated data using an approach similar to that of [3]. Figure 5(a) shows the binary 256×256 reflectance function that we multiplied by an incident power, I_0, to get the

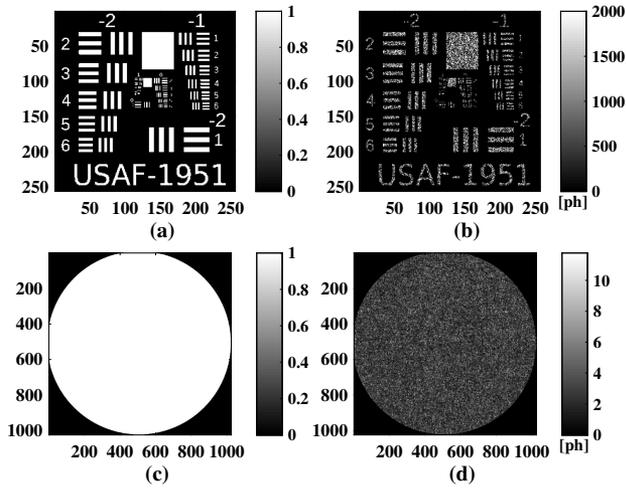**Fig. 5.** (a) 256 × 256 reflectance function used for generating simulated data. (b) Intensity of the reflected field in the object plane, $|f|^{\circ 2}$. (c) Entrance-pupil transmission function, $a(m, n)$. (d) Intensity of the field passing through the entrance pupil.

Fig. 6. Atmospheric phase errors for D_{ap}/r_0 values of (a) 10, (b) 20, (c) 30, and (d) 40. Values shown are wrapped to $[-\pi, \pi)$.

object intensity, $I(p, q)$. Figure 5(b) shows the magnitude squared of the corresponding reflection coefficient generated according to $f \sim CN(0, \mathcal{D}(I_0 r))$.

In accordance with the geometry shown in Fig. 1, we used an FFT to propagate the field, f, from the object plane to the entrance-pupil plane of the DH system. The field was padded to 1024 × 1024 prior to propagation. Next, we applied a phase error, $\phi(m, n)$, and the entrance-pupil transmission function, $a(m, n)$, shown in Fig. 5(c). Figure 5(d) is an example of the intensity of the field passing through the entrance pupil, which has a diameter, D_{ap}, equal to the grid length.

To generate the atmospheric phase errors, we used an FFT-based technique described in [25]. Using a Kolmogorov model for the refractive-index power spectral density (PSD), we generated random coefficients for the spatial-frequency components of the atmospheric phase, and then we used an inverse FFT to transform to the spatial domain. To set the turbulence strength, we parameterized the Kolmogorov PSD by the coherence length, r_0, also known as the Fried parameter [25,26]. The value of r_0 specifies the degree to which the phase of a plane wave passing through the turbulent medium is correlated. Two points in the phase function, which are separated by a distance greater than r_0, will typically be uncorrelated. In this paper, we report turbulence strength using the ratio D_{ap}/r_0, which is related to the degrees of freedom in the atmospheric phase. We simulated data for D_{ap}/r_0 values of 10, 20, 30, and 40. Figure 6 shows examples of the wrapped phase errors for each case.

After we added phase errors to the propagated field and applied the aperture function shown in Fig. 5, we mixed the signal with a modulated reference beam and detected the resultant power. Following [3], the reference-beam power was set at approximately 80% of the well depth per detector element, i.e., 80% of 5 × 10⁴ photoelectron (pe). We modeled the Gaussian read noise with a standard deviation of 40 pe and digitized the output to 12 bits. After detection, we demodulated the signal to
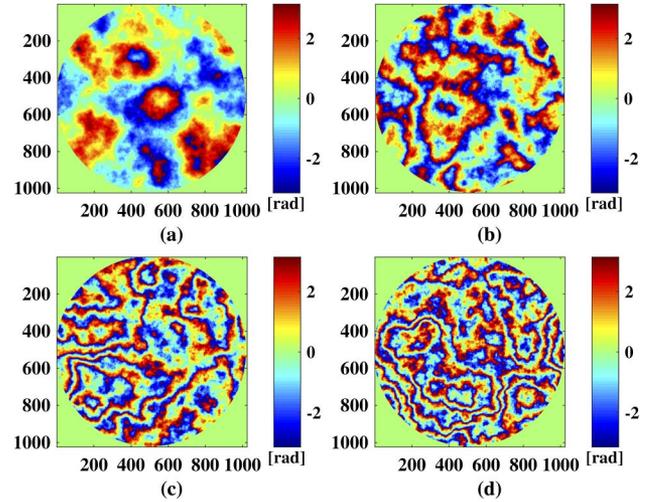
remove the spatial-frequency offset from the reference beam, low-pass filtered to isolate the signal of interest, and decimated to obtain a 256 × 256 data array. (It is typical for this process to be carried out by taking an FFT, windowing a small region around the desired signal spectrum, and taking an inverse FFT.) The resulting data were represented by Eq. (2) after vectorization.

We generated data over a range of SNRs, which we define as

$$\text{SNR} = \frac{s^2(Af)}{s^2(w)}, \tag{29}$$

where $s^2(\cdot)$ is the sample-variance operator used in Eq. (28). For optically coherent systems, SNR is well approximated by the average number of detected signal photons per pixel [17,27]. At each turbulence strength, and at each SNR, we generated 18 i.i.d. realizations of the data. We then processed each i.i.d. realization independently and computed the average performance over the 18 independent cases.

To measure the distortion between the reconstructed images, $\hat{r}$, and the simulation input, r, we used the normalized root mean square error (NRMSE) given by

$$\text{NRMSE} = \sqrt{\frac{\|\alpha^* \hat{r} - r\|^2}{\|r\|^2}}, \tag{30}$$

where

$$\alpha^* = \underset{\alpha}{\text{argmin}}\{\|\alpha \hat{r} - r\|^2\} \tag{31}$$

is the least-squares fit for any multiplicative offset between r and $\hat{r}$.

To measure distortion between the reconstructed phase error, $\hat{\phi}$, and the actual phase error, ϕ, we calculated the Strehl ratio according to

$$S = \frac{[|\text{FFT}\{a(m, n)e^{j[\hat{\phi}(m,n) - \phi(m,n)]}\}|^2]_{\text{max}}}{[|\text{FFT}\{a(m, n)\}|^2]_{\text{max}}}, \tag{32}$$

where FFT$\{\cdot\}$ is the FFT operator and $[\cdot]_{\text{max}}$ indicates that we take the maximum value of the argument. The function, $a(m, n)$, is a binary function that represents the aperture in

the observation plane. It takes on the value of 1 inside the white dotted circle shown in Fig. 5 and 0 outside.

The digital phase correction of DH data is analogous to using a piston-only, segmented deformable mirror (DM) in adaptive optics. Therefore, using the DM fitting error from [28] and Maréchal's approximation, we can compute the theoretical limit of the Strehl ratio as

$$S_{\max} = e^{-1.26\left(\frac{d}{r_0}\right)^{\frac{5}{3}}}, \tag{33}$$

where d is the spacing between correction points in $\hat{\phi}$.

Using the NRMSE and the Strehl ratio, we compared performance of the proposed algorithm to the IS approach presented in [3] using a point-by-point estimate of ϕ and the M_2 sharpness metric. The algorithm computes the phase-error estimate according to
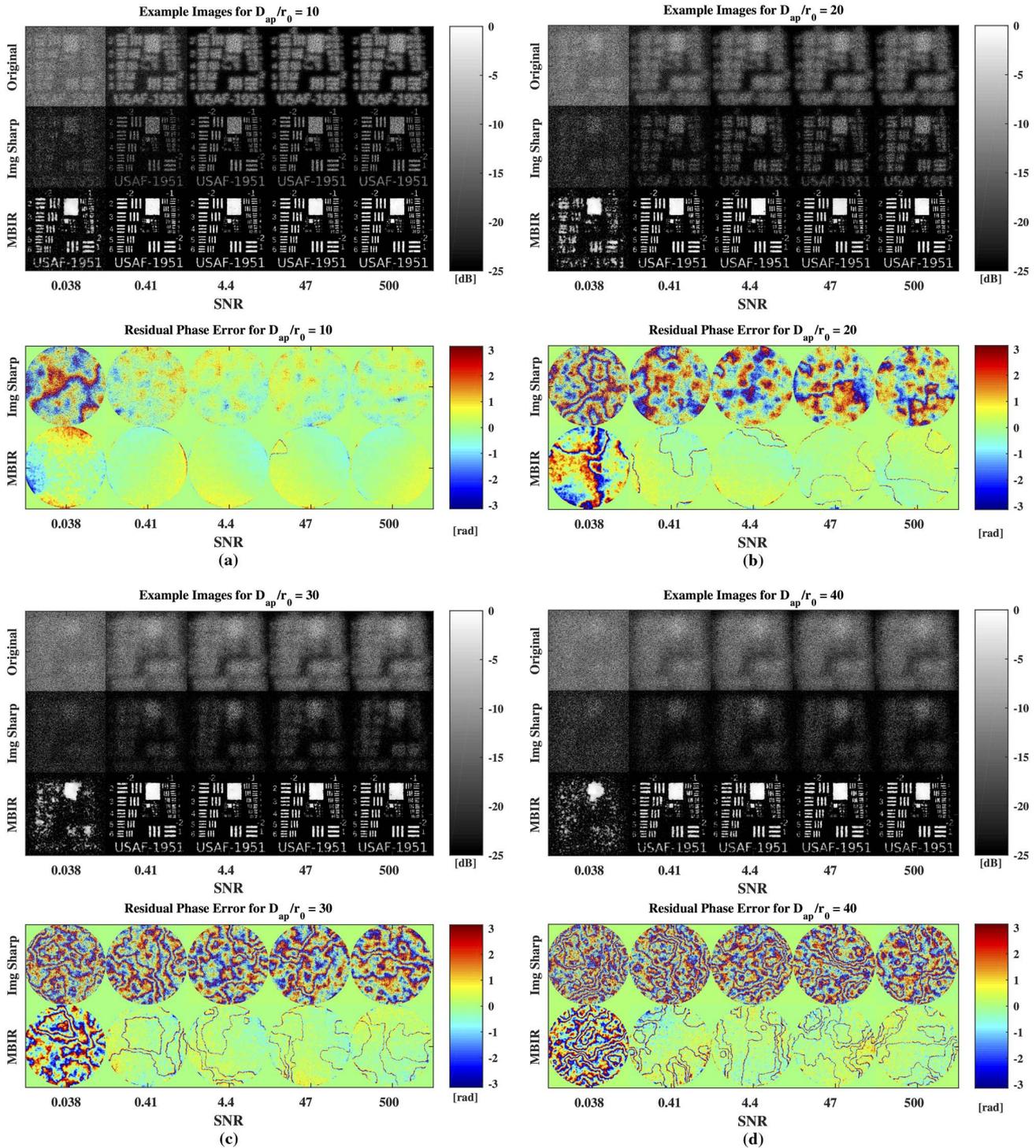

Fig. 7. Example images and residual phase errors for D_{ap}/r_0 of (a) 10, (b) 20, (c) 30, and (d) 40. The top row of each image is the original blurry image. These examples represent the reconstructions with the median Strehl ratio at the chosen SNRs. Note that we only show five of the 20 SNR values. The images are shown using a log-based dB scale and the residual phase errors are wrapped to $[-\pi, \pi]$.

$$\hat{\phi} = \underset{\phi}{\mathrm{argmax}} \left\{ -\sum_{p,q} (|F^H \mathcal{D}(\exp\{j\phi\})^H y|^{\circ 2})^{\circ 0.5} \right\}, \qquad (34)$$

where $\circ$ indicates the application of an exponent to each element. Following the process described in [3], we used 20 iterations of conjugate gradient to optimize Eq. (34) and the algorithm was initialized using a 15th-order Zernike polynomial estimate of the phase errors. Following [3], we obtained the polynomial estimate using an iterative method to estimate only up to the 3rd-order terms, then up to the 4th, and so on, continuing up to 15th order.

For the proposed MBIR algorithm, we allowed the outer initialization loop to run $N_L = 2 \times 10^2$ times, with $N_K = 10$ EM iterations each time. We kept N_L constant for all reconstructions. Once the iterative initialization process was complete, we set a stopping criteria of $\epsilon_T = 1 \times 10^{-4}$ and let the EM algorithm run to completion. We used $q = 2$, $p = 1.1$, $T = 0.1$, $\gamma = 2$, and $b = G(0.1)$ as the QGGMRF prior parameters for image reconstruction. Additionally, we used $n_b = 2$, $\sigma_{\dot{\phi}} = 0.1$, and $b = G(0.1)$ for the phase error prior parameters. Using $n_b = 2$ gives a total number of unknowns of $5/4N$. Ideally, we would adjust $\sigma_{\dot{\phi}}$ in Eq. (17) to fit the turbulence strength. However, in an attempt to reduce the number of parameters that must be adjusted in the MBIR algorithm, and since we may not know the turbulence strength in practical applications, we fix $\sigma_{\dot{\phi}}$ to a value that robustly produces good results. We found our parameters to work well over a wide range of conditions.

6. EXPERIMENTAL RESULTS

Figure 7 shows example reconstructions for a subset of the results. Each block of images shows the reconstructions corresponding to the median Strehl ratio of the 18 i.i.d. data sets. Note that we only show five of the 20 SNR levels for each turbulence strength. The top row of each image block shows the original blurry images, the middle shows the IS reconstructions, and the bottom shows the MBIR reconstruction. To compress the large dynamic range that occurs in coherent imagery, we present the results using a log-based decibel (dB) scale given by $r_{\mathrm{dB}} = 10\log_{10}(\tilde{r})$, where $\tilde{r} \in (0, 1]$ is the normalized reflectance function. The residual phase errors, wrapped to $[-\pi, \pi)$, are shown below each image block. To aid in viewing the higher-order residual phase errors in Fig. 7, we removed the constant and linear phase components, piston, tip, and tilt. The values were estimated according to

$$(\hat{p}, \hat{t}_x, \hat{t}_y) = \underset{(p,t_x,t_y)}{\mathrm{argmin}} \sum_{m,n} |\mathrm{W}\{\phi_r(m, n) - (p + t_x m + t_y n)\}|^2,$$

$$(35)$$

where $\mathrm{W}\{\cdot\}$ is an operator that wraps its argument to $[-\pi, \pi)$, ϕ_r is the residual phase error, p is a constant, and t_x and t_y are the linear components in the x and y dimensions, respectively.

The examples in Fig. 7 show that the IS algorithm is able to correct most of the phase errors for $D_{ap}/r_0 = 10$ and some of the errors at $D_{ap}/r_0 = 20$. For $D_{ap}/r_0 = 30$ and $D_{ap}/r_0 = 40$, the IS images are blurred beyond recognition. In contrast, the proposed MBIR algorithm produced focused images for all but the lowest SNR reconstructions. In addition, we see how the MBIR reconstruction has significantly less speckle variation. The MBIR results also show that in many cases, the residual phase errors contain large-scale patches separated by wrapping cuts. However, as Fig. 7 shows, the patches are approximately modulo-2π equivalent and still produce focused images.
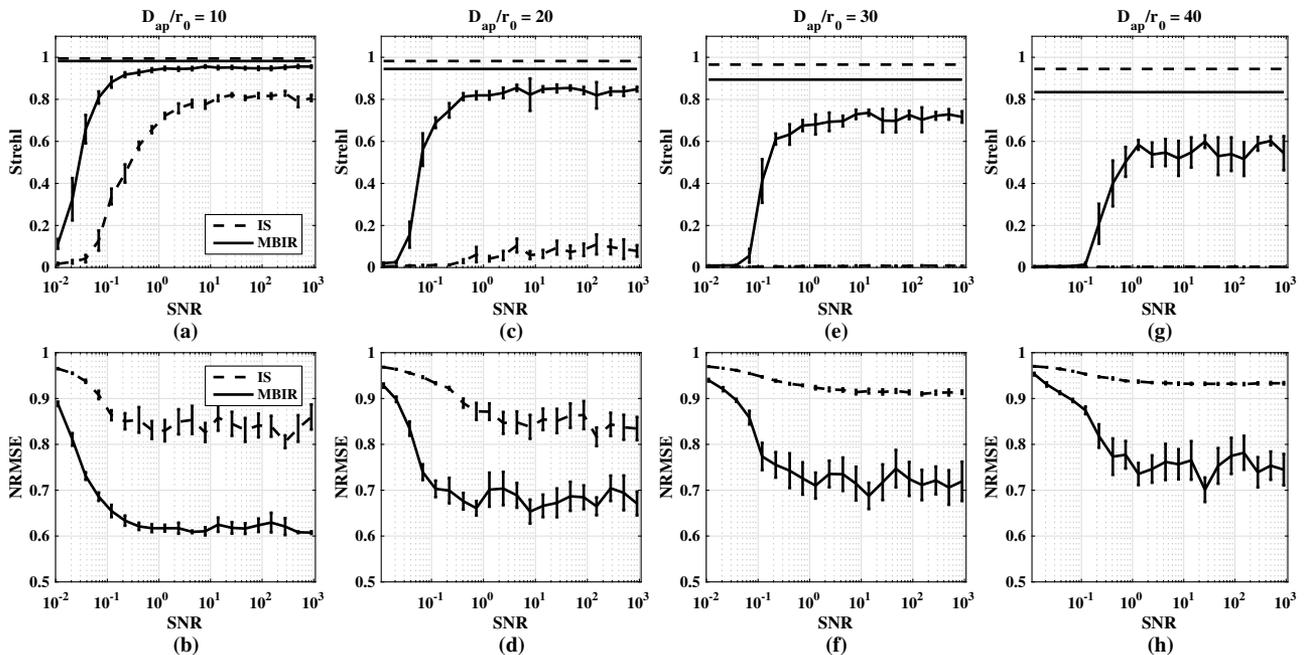

Fig. 8. Strehl ratio (top row) and reconstruction error (bottom row) versus SNR for D_{ap}/r_0 of (a), (b) 10, (c), (d) 20, (e), (f) 30, and (g), (h) 40. The dashed curves shows the IS results and the solid curves show the MBIR results. The horizontal lines in the top row show the corresponding Strehl ratio limit given by Eq. (33).

Figure 8 shows the resulting Strehl ratio and NRMSE for each algorithm as a function of SNR and D_{ap}/r_0. The curves in Fig. 8 show the average results for all 18 i.i.d. realizations along with error bars that span the standard deviation. We also plot the Strehl ratio limits given by Eq. (33) where we used $d = D_{ap}/256$ for the IS algorithm and $d = D_{ap}/128$ for the MBIR algorithm.

The results show that the IS algorithm works in a limited range of conditions for single-shot data. For $D_{ap}/r_0 = 10$, IS peaks at a Strehl ratio of 0.8 when the SNR exceeds 10. For lower SNRs or stronger turbulence, the IS algorithm's performance tapers off quickly. Conversely, the proposed MBIR algorithm is able to obtain Strehl ratios much higher than IS, even for low SNRs and strong turbulence. In the $D_{ap}/r_0 = 10$ case, MBIR reaches the IS Strehl limit of 0.8 with about 92% less SNR.

While Figs. 7 and 8 show that the MBIR algorithm produces reconstructions with higher Strehl ratios and lower NRMSEs compared to the IS algorithm, we must also consider how long they take to run. To compare the algorithm runtimes, we generated 20 i.i.d. data realizations at $D_{ap}/r_0 = 10$ and SNR $= 100$. We then computed the IS and MBIR reconstructions for each of the 20 data sets. Figure 9 shows the average Strehl ratio as a function of computer runtime. We used a computer with a 4 GHz Intel Core i7 processor and executed the reconstructions in MATLAB using mex files to run the optimization functions for both algorithms.

The runtime results in Fig. 9 show that the MBIR algorithm can reach high Strehl ratios much faster than the IS algorithm. The average Strehl ratio of the IS algorithm reached its peak of 0.81 after approximately 153 s, while the MBIR algorithm achieved an average Strehl ratio of 0.81 after approximately 6 s. MBIR reached average Strehl ratios of 0.93 after approximately 13 s and its peak of 0.95 after approximately 28 s. When IS transitions from the polynomial estimate to the
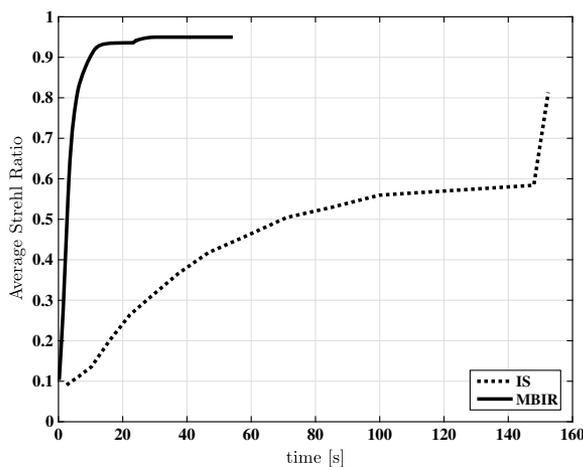
point-by-point estimate at ≈148 s, there is a sharp increase in the average Strehl ratio. Note that running the point-by-point IS algorithm longer does not continue this trend.

7. CONCLUSION

In this paper, we presented an inverse, model-based approach to estimating phase errors and reconstructing images from digital holography (DH) data. We designed the algorithm for cases where only a single data realization is available and the phase errors are isoplanatic. Rather than estimating the spatially uncorrelated reflection coefficient, g, we estimate the highly correlated reflectance function, r. This allows us to better constrain the underdetermined system and produce speckle-reduced images and accurate phase-error estimates with less data. Using first principals, we derived a discrete forward model for use in the MAP cost function. To obtain a more tractable surrogate function, we used the EM algorithm. Additionally, we introduced a GMRF prior model for the phase-error function modeled on a lower resolution grid and presented optimization schemes for the joint estimates.

We compared the proposed algorithm to a leading image sharpness approach over a range of conditions. The results showed that the MBIR algorithm produced phase estimates with higher Strehl ratios and lower image distortions than the IS technique. For cases of high noise and large phase errors, IS was not effective, while MBIR was still able to produce accurate estimates. Furthermore, we showed that the MBIR algorithm is much faster than IS at reaching high Strehl ratios. In conclusion, we showed that the proposed algorithm is an effective alternative to IS algorithms for estimating phase errors from single-shot DH data and reconstructing images.

APPENDIX A: FORWARD MODEL DERIVATION

In this section, we derive the linear forward model expressed by Eq. (2). Following the scenario depicted in Fig. 1, we derive a continuous model for the data and then a discrete representation.

Assume that we use a monochromatic plane wave to illuminate an object that has a complex-valued reflection coefficient, $\tilde{g}(\xi, \eta)$. We denote continuous functions using tildes. The return field passing through the entrance pupil can be represented using the Fresnel diffraction integral [16] given by

$$\tilde{U}_s(x, y) = \alpha \tilde{a}(x, y) e^{j\tilde{\psi}(x,y)} e^{j\frac{k}{2z}(x^2+y^2)}$$

$$\times \iint_{-\infty}^{\infty} \left\{ \tilde{g}(\xi, \eta) e^{j\frac{k}{2z}(\xi^2+\eta^2)} \right\} e^{-j\frac{k}{z}(x\xi+y\eta)} \mathrm{d}\xi \mathrm{d}\eta, \quad \textbf{(A1)}$$

where α is a complex constant, $\tilde{a}(x, y)$ represents the circular transmission function for the system's entrance pupil, $\tilde{\psi}(x, y)$ is the atmospheric phase-error function assumed to be concentrated in a single layer in the entrance-pupil plane (i.e., an isoplanatic phase error), $k = 2\pi/\lambda$ is the wave number, and λ is the wavelength.

To simplify notation, we define the function

$$\tilde{f}(\xi, \eta) = \tilde{g}(\xi, \eta) e^{j\frac{k}{2z}(\xi^2+\eta^2)} \quad \textbf{(A2)}$$

and its corresponding continuous space Fourier transform given by

Fig. 9. Average Strehl ratio as a function of algorithm runtime for 20 i.i.d. realizations at $D_{ap}/r_0 = 10$ and SNR $= 100$. The jump in Strehl ratio for the IS algorithm occurs at approximately 148s when we switch from the 15th-order polynomial estimate to the point-by-point estimate. Here we limited the number of outer-loop iterations for the MBIR algorithm to $N_K = 50$ since the Strehl ratio did not change significantly after ≈30 s.

$$\tilde{F}(u, v) = \iint_{-\infty}^{\infty} \tilde{f}(\xi, \eta) e^{-j2\pi(u\xi+v\eta)} \, d\xi \, d\eta. \tag{A3}$$

We can also combine the pupil-plane quadratic phase function of Eq. (A1) with the atmospheric phase errors to get a single unknown phase term given by

$$\tilde{\phi}(x, y) = \tilde{\psi}(x, y) + \frac{k}{2z}(x^2 + y^2). \tag{A4}$$

Given the definitions in Eqs. (A2) through (A4), the return signal of Eq. (A1) can be written as

$$\tilde{U}_s(x, y) = \alpha \tilde{a}(x, y) e^{j\tilde{\phi}(x,y)} \iint_{-\infty}^{\infty} \tilde{f}(\xi, \eta) e^{-j2\pi\left(\frac{x}{\lambda z}\xi + \frac{y}{\lambda z}\eta\right)} \, d\xi \, d\eta$$

$$= \alpha \tilde{a}(x, y) e^{j\tilde{\phi}(x,y)} \tilde{F}\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right), \tag{A5}$$

where $\tilde{F}(\frac{x}{\lambda z}, \frac{y}{\lambda z})$ implies $\tilde{F}(u = \frac{x}{\lambda z}, v = \frac{y}{\lambda z})$.

Optical heterodyne detection is performed by mixing the returned signal with a local reference beam given by

$$\tilde{U}_r(x, y) = R_0 e^{j2\pi(\nu_x x + \nu_y y)}, \tag{A6}$$

where R_0 is the amplitude and ν_x and ν_y are factors that control the spatial-frequency modulation. We combine the signal and reference onto the detector, which measures the intensity given by

$$\tilde{I}(x, y) = |\tilde{U}_r(x, y) + \tilde{U}_s(x, y)|^2$$

$$= |\tilde{U}_r(x, y)|^2 + |\tilde{U}_s(x, y)|^2$$

$$\quad + \tilde{U}_r(x, y)\tilde{U}_s^*(x, y) + \tilde{U}_r^*(x, y)\tilde{U}_s(x, y)$$

$$= R_0^2 + \alpha^2 \tilde{a}(x, y)^2 \left|\tilde{F}\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right)\right|^2$$

$$\quad + \alpha \tilde{a}(x, y) R_0 e^{-j\tilde{\phi}(x,y)} \tilde{F}^*\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right) e^{j(\nu_x x + \nu_y y)}$$

$$\quad + \alpha \tilde{a}(x, y) R_0 e^{j\tilde{\phi}(x,y)} \tilde{F}\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right) e^{-j(\nu_x x + \nu_y y)}, \tag{A7}$$

where $*$ indicates the complex conjugate.

After detection, we demodulate the detector output to remove the spatial-frequency offset of the reference and low-pass filter to remove the unwanted terms. Modulation by the reference requires that we sample the data at a higher rate than would otherwise be necessary for the nonmodulated signal. After demodulation, and low-pass filtering, the signal bandwidth is only a fraction of the sampling bandwidth. We therefore decimate the signal to reduce the computational requirements. Note that this is mathematically equivalent to taking an FFT of the intensity data, extracting a smaller region around the offset signal and then taking an inverse FFT. In either case, we isolate the last term of Eq. (A7), which gives us our signal of interest,

$$\tilde{y}(x, y) = \tilde{a}(x, y) e^{j\tilde{\phi}(x,y)} \tilde{F}\left(\frac{x}{\lambda z}, \frac{y}{\lambda z}\right). \tag{A8}$$

(In practice, the output of the detection process is only proportional to the right-hand side of Eq. (A8) by some unknown constant.) Equation (A8) is continuous in both the x–y and ξ–η coordinate systems and does not include noise. However, we wish to represent the signal as discrete noisy measurements,

$y(m, n)$, generated from a discrete-space signal, $f(p, q)$. We start by representing the discrete field in the object plane as

$$f(p, q) = \tilde{f}(\xi, \eta)|_{\substack{\xi=pT_\xi \\ \eta=qT_\eta}}, \tag{A9}$$

where T_ξ and T_η are the spatial-sampling periods in the object plane. Furthermore, if we sample the signal with a focal-plane array and ignore the blurring effects of finite-sized pixels, then the discrete measurements can be represented as

$$y(m, n) = \tilde{y}(x, y)|_{\substack{y=mT_y \\ x=nT_x}} + w(m, n), \tag{A10}$$

where T_x and T_y are the spatial-sampling periods in the measurement plane. We assume the system is shot-noise limited with noise $w(m, n)$, driven by the power of the reference beam [18], and modeled as additive, zero-mean, white Gaussian noise [17].

Combining Eqs. (A8)–(A10), we get

$$y(m, n) = a(m, n) e^{j\phi(m,n)} \sum_{p,q} f(p, q) F_{m,n;p,q} + w(m, n), \tag{A11}$$

where a and ϕ are discrete versions of $\tilde{a}$ and $\tilde{\phi}$, respectively, and

$$F_{m,n;p,q} = \frac{1}{\sqrt{N}} \exp\left\{-j2\pi\left(\frac{T_\xi T_x}{\lambda z} mp + \frac{T_\eta T_y}{\lambda z} nq\right)\right\}. \tag{A12}$$

Equation (A11) can be represented more compactly using matrix-vector notation as

$$y = Af + w, \tag{A13}$$

where $w \in \mathbb{C}^M$ is the vectorized measurement noise and $f \in \mathbb{C}^M$ is the vectorized field in the object plane. The matrix A can be decomposed as

$$A = \mathcal{D}(a)\mathcal{D}(\exp\{j\phi\})F. \tag{A14}$$

Here $\mathcal{D}(\cdot)$ denotes an operator that produces a diagonal matrix from its vector argument, $a \in \mathbb{R}^N$ is the vectorized entrance-pupil transmission function, and $\phi \in \mathbb{C}^N$ is the vectorized phase-error function. The matrix F is defined in Eq. (A12).

Since the sum in Eq. (A11) represents the forward propagation of the field f, we have scaled F by $1/\sqrt{N}$ so that $F^H F = I$. This ensures we conserve energy when propagating between the object and entrance-pupil planes. Furthermore, we choose our reconstruction parameters such that $T_\xi T_x/\lambda z = N_x$ and $T_\eta T_y/\lambda z = N_y$, where N_x and N_y are the grid sizes in the x and y dimensions, respectively. Thus, F is exactly a DFT kernel and can be efficiently implemented using a FFT.

APPENDIX B: DERIVATION OF THE EM SURROGATE FUNCTION

In this section, we derive the EM surrogate for the MAP cost function. We start by writing Eq. (12) as

$$Q(r, \phi; r', \phi') = -E[\log p(y, f|r, \phi)|Y = y, r', \phi']$$

$$\quad - \log p(r) - \log p(\phi)$$

$$= -E[\log p(y|f, \phi) + \log p(f|r)|Y = y, r', \phi']$$

$$\quad - \log p(r) - \log p(\phi), \tag{B1}$$

where we have used Bayes' theorem inside the expectation and the fact that $p(y|f, r, \phi) = p(y|f, \phi)$. Next, we substitute in the forward and prior models specified in Sections 2 and 3. This gives

$$
Q(r, \phi; r', \phi') = \mathrm{E}\left[\frac{1}{\sigma_w^2}\|y - A_\phi f\|^2 \,\middle|\, Y = y, r', \phi'\right] + \log|\mathcal{D}(r)|
$$
$$
+ \sum_{i=1}^{N} \frac{1}{r_i} \mathrm{E}[|f_i|^2 \,|\, Y = y, r', \phi'] + \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\rho_r\left(\frac{\Delta_r}{\sigma_r}\right)
$$
$$
+ \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\rho_{\tilde{\phi}}\left(\frac{\Delta_{\tilde{\phi}}}{\sigma_{\tilde{\phi}}}\right) + \kappa, \tag{B2}
$$

where A_ϕ indicates the matrix A is dependent on ϕ and the variable κ is a constant with respect to r and ϕ.

To evaluate the expectations in Eq. (B2), we must specify the conditional posterior distribution of f. Using Bayes' theorem,

$$
p(f|y, r, \phi) = \frac{p(y|f, r, \phi)p(f|r)}{p(y|r, \phi)}
$$
$$
= \frac{1}{z} \exp\left\{-\frac{1}{\sigma_w^2}\|y - A_\phi f\|^2 - f^H \mathcal{D}(r)^{-1} f\right\}, \tag{B3}
$$

where z is the partition function which absorbs any exponential terms that are constant with respect to f. By completing the square, we can show that the posterior distribution is a complex Gaussian with mean

$$
\mu = C \frac{1}{\sigma_w^2} A_{\phi'}^H y \tag{B4}
$$

and covariance

$$
C = \left[\frac{1}{\sigma_w^2} A_{\phi'}^H A_{\phi'} + \mathcal{D}(r)^{-1}\right]^{-1}. \tag{B5}
$$

Using the posterior distribution specified by Eq. (B3), we can evaluate the expectations in Eq. (B2) to get the final form of our EM surrogate function given by

$$
Q(r, \phi; r', \phi') = -\frac{1}{\sigma_w^2} 2\,\mathrm{Re}\{y^H A_\phi \mu\} + \log|\mathcal{D}(r)|
$$
$$
+ \sum_{i=1}^{N} \frac{1}{r_i}(C_{i,i} + |\mu_i|^2) + \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\rho_r\left(\frac{\Delta_r}{\sigma_r}\right)
$$
$$
+ \sum_{\{i,j\}\in\mathcal{P}} b_{i,j}\rho_{\tilde{\phi}}\left(\frac{\Delta_{\tilde{\phi}}}{\sigma_{\tilde{\phi}}}\right) + \kappa, \tag{B6}
$$

where μ_i is the ith element of the posterior mean and $C_{i,i}$ is the ith diagonal element of the posterior covariance.

REFERENCES

1. P. J. Winzer and W. R. Leeb, "Coherent LIDAR at low signal powers: basic considerations on optical heterodyning," J. Modern Opt. **45**, 1549–1555 (1998).
2. J. C. Marron and K. S. Schroeder, "Holographic laser radar," Opt. Lett. **18**, 385–387 (1993).
3. S. T. Thurman and J. R. Fienup, "Phase-error correction in digital holography," J. Opt. Soc. Am. A **25**, 983–994 (2008).
4. S. T. Thurman and J. R. Fienup, "Correction of anisoplanatic phase errors in digital holography," J. Opt. Soc. Am. A **25**, 995–999 (2008).
5. J. C. Marron, R. L. Kendrick, N. Seldomridge, T. D. Grow, and T. A. Höft, "Atmospheric turbulence correction using digital holographic detection: experimental results," Opt. Express **17**, 11638–11651 (2009).
6. J. Marron, R. Kendrick, S. Thurman, N. Seldomridge, T. Grow, C. Embry, and A. Bratcher, "Extended-range digital holographic imaging," Proc. SPIE **7684**, 76841J (2010).
7. A. E. Tippie and J. R. Fienup, "Phase-error correction for multiple planes using a sharpness metric," Opt. Lett. **34**, 701–703 (2009).
8. A. E. Tippie and J. R. Fienup, "Multiple-plane anisoplanatic phase correction in a laboratory digital holography experiment," Opt. Lett. **35**, 3291–3293 (2010).
9. A. E. Tippie, "Aberration correction in digital holography," Ph.D. thesis (University of Rochester, 2012).
10. J. R. Fienup, "Phase error correction in digital holographic imaging," in *Digital Holography and Three-Dimensional Imaging* (Optical Society of America, 2014), p. DM1B-1.
11. R. A. Muller and A. Buffington, "Real-time correction of atmospherically degraded telescope images through image sharpening," J. Opt. Soc. Am. **64**, 1200–1210 (1974).
12. M. F. Spencer, I. V. Dragulin, D. S. Cargill, and M. J. Steinbock, "Digital holography wave-front sensing in the presence of strong atmospheric turbulence and thermal blooming," Proc. SPIE **9617**, 961705 (2015).
13. M. T. Banet, M. F. Spencer, R. A. Raynor, and D. K. Marker, "Digital holography wavefront sensing in the pupil-plane recording geometry for distributed-volume atmospheric aberrations," Proc. SPIE **9982**, 998208 (2016).
14. M. F. Spencer, R. A. Raynor, M. T. Banet, and D. K. Marker, "Deep-turbulence wavefront sensing using digital-holographic detection in the off-axis image plane recording geometry," Opt. Eng. **56**, 031213 (2017).
15. C. Pellizzari, R. Trahan III, H. Zhou, S. Williams, S. Williams, B. Nemati, M. Shao, and C. A. Bouman, "Synthetic aperture LADAR: a model based approach," IEEE Trans. Comput. Imaging **PP**, 1 (2017).
16. J. W. Goodman, *Introduction to Fourier Optics* (Roberts & Company, 2005).
17. R. L. Lucke and L. J. Rickard, "Photon-limited synthetic aperture imaging for planet surface studies," Appl. Opt. **41**, 5084–5095 (2002).
18. V. V. Protopopov, *Laser Heterodyning* (Springer, 2009), Vol. **149**.
19. S. M. Kay, *Fundamentals of Statistical Signal Processing, Detection Theory*, Prentice Hall Signal Processing Series (Prentice-Hall, 1993), Vol. **II**.
20. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. R. Stat. Soc. Ser. B **39**, 1–38 (1977).
21. C. A. Bouman, "Model based image processing," (2015) https://engineering.purdue.edu/~bouman/publications/pdf/MBIP-book.pdf.
22. J. B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multi-slice helical CT," Med. Phys. **34**, 4526–4544 (2007).
23. C. Pellizzari, R. Trahan III, H. Zhou, S. Williams, S. Williams, B. Nemati, M. Shao, and C. A. Bouman, "Optically coherent image formation and denoising using plug and play inversion framework," Appl. Opt. **56**, 4735–4744 (2017).
24. C. Forbes, M. Evans, N. Hastings, and B. Peacock, *Statistical Distributions* (Wiley, 2011).
25. J. D. Schmidt, *Numerical Simulation of Optical Wave Propagation with Examples in MATLAB* (SPIE, 2010).
26. L. C. Andrews and R. L. Phillips, *Laser Beam Propagation through Random Media* (SPIE, 2005), Vol. **52**.
27. H. A. Haus, *Electromagnetic Noise and Quantum Optical Measurements* (Springer, 2012).
28. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford University, 1998).