

**Compressive ptychographical coherent diffractive
imaging based on Poissonian maximum likelihood and
sparse approximations for phase and magnitude**

Vladimir Katkovnik and Jaakko Astola

Department of Signal Processing, Tampere University of Technology,

P. O. Box 553, Tampere, Finland, vladimir.katkovnik@tut.fi.

Ptychography is a lensless coherent diffractive imaging that uses intensity measurements of multiple diffraction patterns collected with a localized illumination probe from overlapping regions of an object. A novel iterative algorithm is proposed targeted on optimal processing noisy measurements. The noise suppression is enabled by two instruments. First, by the maximum likelihood technique formulated for Poissonian (photon counting) measurements, and second, by sparse approximation of phase and magnitude of object and probe. It is shown in particular, that for noisy data the maximum likelihood estimate of the wavefield at the sensor plane is essentially different from the famous Gerchberg-Saxton-Fienup solution, where the magnitude of the estimate is replaced by the square root of the intensity measurement. The simulation experiments demonstrate the state-of-the-art performance of the proposed algorithm both numerically and visually. © 2012 Optical Society of America

OCIS codes: 070.2025, 100.2000, 100.3010, 100.3190

1. Introduction

The ptychography is a method of lensless coherent diffractive imaging that uses multiple overlapping diffraction patterns in order to obtain high-resolution reconstructions of phase and magnitude of an object. It is envisaged in the original formulation by Hoppe [1] that such patterns could be obtained by placing a very narrow probe-aperture in the plane of the object so that each diffraction pattern would overlap with one another. The name ptychography, from the Greek for *fold*, derives from this optical configuration: each object area is modulated by the probe and convolved with some wavefield propagation function, and thus made to interfere with its neighbors. Measuring only the intensities of diffraction patterns leads to the phase retrieval formulation of the ptychographical imaging. Only recently this principal was implemented. It has been demonstrated that phase and magnitude of the object and the probe can be extracted from ptychographical data with a good accuracy (e.g. [2–5]). Ptychographical imaging has resulted in increased spatial resolution microscopy without the need for lenses at both optical and X-ray wavelengths [6, 7].

First of all, as the recent state-of-the art we wish to mention the *ptycho-*

graphical iterative engine (PIE) algorithm by Rodenburg and Faulkner [8] developed for the complex-valued object reconstruction assuming that the illumination probe is known. In the further development the requirement that an accurate model of the probe function is known is removed. The *extended ptychographical iterative engine* (e-PIE) proposed by Maiden and Rodenburg in the paper [9] reconstructs both the object and the probe complex-valued transmissive functions. Another type of the iterative algorithm (known as *parallel ptychographical iterative engine* (*p-PIE*)) for reconstruction of the object and the probe is proposed by Thibault et al. [10]. This latter algorithm is based on projections of successive iterations on restrictions in the object and sensor planes using the so-called *difference map*. *Superresolution ptychographical imaging* proposed by Maiden et al. [11] is a further development of the e-PIE algorithm.

All these iterative algorithms are based on methods developed for the phase retrieval problem and exploit the ideas of the *Fienup's hybrid input-output algorithm* [12, 13]. The algorithms cycle through all data sets sequentially to update both the object and probe wave front reconstructions. In these iterations modules of the diffraction patterns are replaced by the square roots

of the observed intensity measurements and phase values obtained from the forward propagation procedure are preserved. Overall these algorithms can be treated as sequential projections of reconstructions on convex sets defined by the measurements and forward propagation models. General representation and discussion of these techniques can be seen in [14, 15].

The application of maximum-likelihood principle to the ptychographical image reconstruction is demonstrated by Thibault and Guizar-Sicairos [16]. The minus Poissonian log-likelihood function is used as the criterion. The proposed iterative algorithm is based on the unconstrained conjugate gradient minimization and includes preconditioning, regularization and some typical modifications of the statistical noise model.

In this paper we are focussed on a variational formalization of the ptychographical imaging and a development of an optimal algorithm. Our approach is original in two aspects. First, starting from Poissonian modeling of intensity measurements we propose a constraint maximum likelihood technique for reconstruction both object and probe. In this way the approach is targeted on optimal suppression of noise in intensity measurements. Second, a sparse modeling is used for object and probe functions of interest. This parametric

modeling being flexible and universal enables improved quality reconstructions. It is used separately for phase and magnitude characteristics of object and probe. Thus, the reconstruction is formulated as a variational problem incorporating both the noisy measurement modeling and the sparse modeling for object and probe. The solution (algorithm) is iterative designed for optimization of the introduced objective functions.

The term *compressive imaging* in the title of the paper refers to the very popular recent development in imaging known also as *compressive (or compressed) sensing (CS)* (e.g. [17–19]). The main constructive idea of this technique, as it is used in our paper, to build shortest (simplest) models for object and probe. It is done by using l_0 or l_1 norms as a prior for functions to be reconstructed. While CS is targeted on minimization of number of used observations, in our approach we are focussed on optimization of accuracy and quality of imaging assuming that the number of observation is fixed.

The proposed approach is quite flexible allowing multiple further development: compensation of errors in probe locations, ambiguity in support of the probe, various prior information of phase and magnitude of object and probe, superresolution imaging, etc.

For object and probe sparse approximation and compressive imaging we use the technique developed in our previous papers, where they are exploited for: complex-valued transmissive object reconstruction from complex-valued observations [20], the phase retrieval in 4f optical setup with SLM phase modulation [21] and for the ghost imaging of the magnitude object [22].

The rest of the paper is organized as follows. In Section 2 the image formation for ptychographical digital imaging and Poissonian observation model are presented. The sparse modeling for phase and magnitude are discussed in subsection 2.C. The decoupled constrained variational approach for the image retrieval from intensity observations are presented in Section 3. The proposed algorithm is obtained in Section 4. In Section 5 the proposed algorithm is discussed in comparison with the most popular modern p-PIE and e-PIE techniques. Simulation experiments are shown and discussed in Section 6. Algorithm derivation from the variational formulation of the problem is presented in Appendix.

2. Image formation and sparse object and probe modeling

We consider a microscope ptychographical imaging in the setup illustrated in Fig. 1. A tested transmissive object (O) is illuminated by a narrow laser

beam through a pinhole (P). A phase diffuser in front of the object generates a speckle wavefield enabling an improved high-frequency sensing of the object. The laser beam propagating through the pinhole and the diffuser at the object plane forms a probe wave front. The object is mounted on a motorized x/y translation stage. The interaction of the object with the probe beam produces the exit wavefield impinging on the digital sensor registering the intensity of this wavefield. Stage x/y movements are used for scanning of the whole object by the overlapping probe beams. A spacial light modulator (SLM) can be used as a computer programmable diffuser.

We use a notation $\psi_j(r)$ for the product of the object $o(r)$ and probe $p(r)$ waves at the object plane. It is convenient to represent $\psi_j(r)$ in the form

$$\psi_j(r) = o(r + r_j)p(r). \quad (1)$$

Here $r_j = (x_j, y_j)$ is a vector defining a location of the probe beam with respect to the object.

For the far field Fraunhofer's approximation, the wavefield at the sensor plane is calculated (within a factor of an invariant magnituder) as

$$u_j(\omega) = \mathcal{F}\{\psi_j(r)\} = \mathcal{F}\{o(r + r_j)p(r)\}, \quad (2)$$

where $r \in \mathbb{R}^2$ and $\omega \in \mathbb{R}^2$ are object/probe and reciprocal (Fourie) space

coordinate vectors, and $\mathcal{F}\{\cdot\}$ stands for the $2D$ integral Fourier transform.

Then the intensity of the wavefield incident on the sensor is calculated as

$$I_j(\omega) = |u_j(\omega)|^2 = |\mathcal{F}\{o(r + r_j)p(r)\}|^2, \quad (3)$$

The ptychography imaging problem is to reconstruct both the object $o(r)$ and probe $p(r)$ wavefields from a set of measured diffraction patterns $\{I_j(\omega)\}_{j=1,\dots,J}$. Both unknowns $o(r)$ and $p(r)$ are complex-valued thus the phase as well as the magnitude of these wavefields should be reconstructed.

If $p(r)$ is known the reconstruction of $o(r)$ is from the class of *inverse imaging problems*. If $p(r)$ is unknown the reconstruction of $o(r)$ is from the class of *blind inverse imaging problems*.

2.A. Discrete image formation

In discrete modeling of $u_j(\omega)$ and $\psi_j(r)$, the continuous arguments ω and r are replaced by digital one with a corresponding replacement of all continuous functions by their sampled and finite support counterparts: $u_j(\omega) \rightarrow u_j[l]$, $\psi_j(r) \rightarrow \psi_j[k]$ with a $2D$ integer arguments $l = [l_1, l_2]$ and $k = [k_1, k_2]$.

Let the discrete sensor of the camera be of $M_s \times N_s$ square pixels of the size $\Delta_l \times \Delta_l$, and a discrete model for the object and probe has, in general rectangular, pixels of the size $\Delta_x \times \Delta_y$.

Assume the following link between pixel's size in the object/probe and sensor domains

$$[\Delta_x, \Delta_y] = \frac{\lambda z}{\Delta_l} \left[\frac{1}{M_s}, \frac{1}{N_s} \right], \quad (4)$$

where λ is the wave length and z is a distance between the sensor and the object planes.

Then the discrete model for the Fraunhofer wavefield propagation defined by Eq.(2) takes the form

$$u_j[l] = \mathcal{DFT}\{\psi_j[k']\}, \quad (5)$$

$$\psi_j[k'] = o[k' + k_j]p[k'], \quad (6)$$

$$k = k' + k_j, \quad (7)$$

where k' are coordinates in probe, k are coordinates in object.

In this model the Discrete Fourier transform (DFT) is used instead of the integral Fourier transform in Eq. (2) due to the conditions (4), and $I_j[l]$, $o[k]$, $p[k']$ are the sampled versions of the corresponding $I_j(u)$, $o(r + r_j)$, $p(r)$. We assume for simplicity that r_j are such that the corresponding k , k' and k_j in (5)-(7) are integer.

The support of the probe p is much smaller than that for the object, thus the support of $\psi_j[k]$ is defined by the support of the probe. In DFT the pixel

sizes for the signal and frequency domains can be different, but the number of the pixels in these domains is assumed to be equal.

In order to exploit the full capacity of the sensor and to improve the accuracy of numerical modeling for forward wavefield propagation the dimension of $u_j[l]$ in the DFT domain is taken equal or close to the size of the sensor. If the size of the sensor in pixels is larger than that pixel size of the probe $\psi_j[k']$ is zeropadded to the size of the sensor.

In what follows we use a vector-matrix notations for images and DFTs. These vectors are obtained from the standard matrix representation for images by concatenating the columns of these matrices. We use bold lower case characters for these vectors.

Let $\mathbf{u}_j \in \mathbb{C}^{n_s}$, $n_s = M_s N_s$, and $\boldsymbol{\psi}_j \in \mathbb{C}^{n_p}$, $n_p = M_p N_p$, be vectorial representations for the matrixes $u_j[l]$, $M_s \times N_s$, and $\psi_j[k]$, $M_p \times N_p$. Then the forward propagation of the wavefield $\boldsymbol{\psi}_j$ from the object to the sensor plane can be presented as a linear transformation given in the matrix form

$$\mathbf{u}_j = \mathbf{A}\boldsymbol{\psi}_j, \tag{8}$$

where \mathbf{A} is a generic notation for any discrete forward propagation operator (complex-valued $n_s \times n_p$ matrix, $\mathbf{A} \in \mathbb{C}^{n_s \times n_p}$).

For the DFT model in Eq. (5) the columns of the matrix \mathbf{A} are orthonormal and $\mathbf{A}^T \mathbf{A} = \mathbf{I}_{n_p \times n_p}$. If the conditions (4) are not fulfilled more general Fresnel and Fraunhofer models could be used instead of Eq.(5). All these models allow representation in the generic form (8).

In order to express ψ_j through $o[k]$ and $p[k]$ we introduce the object-vector $\mathbf{o} \in \mathbb{C}^{n_o}$ corresponding to the $M_o \times N_o$ matrix-image $o[k]$, $n_o = M_o N_o$, and the probe-vector $p \in \mathbb{C}^{n_p}$, corresponding to the $M_p \times N_p$ matrix-image $p[k]$, $n_p = M_p N_p$.

Let $\mathbf{o}_j \in \mathbb{C}^{n_p}$ be a vector corresponding to a part of the vector-object $\mathbf{o} \in \mathbb{C}^{n_o}$ covered by the probe \mathbf{p} in $\psi_j[k'] = o[k' + k_j]p[k']$. The links between the vectors \mathbf{o}_j and \mathbf{o} can be given in the forms

$$\mathbf{o}_j = \mathbf{G}_j^o \mathbf{o},$$

where \mathbf{G}_j^o is an incidence matrix of the size $n_p \times n_o$. The items of these matrices equal to 1 show a location of the corresponding items of \mathbf{o} in the vector \mathbf{o}_j . All other items of \mathbf{G}_j^o are equal to zero. Each line of \mathbf{G}_j^o has only one element equal to 1.

Then, the vectorized input-output model for Eqs. (5)-(6) takes the following

equivalent forms

$$\mathbf{u}_j = \mathbf{A} \cdot (\mathbf{o}_j \circ \mathbf{p}) = \quad (9)$$

$$\mathbf{A} \cdot \text{diag}(\mathbf{p}) \mathbf{G}_j^o \mathbf{o} = \quad (10)$$

$$\mathbf{A} \cdot \text{diag}(\mathbf{G}_j^o \mathbf{o}) \mathbf{p}. \quad (11)$$

Here 'o' stands in Eq. (9) for the Hadamard element wise product of two vectors \mathbf{o}_j and p . Operation $\text{diag}(\cdot)$ means that the vector-argument is transformed in the diagonal matrix $(n_p \times n_p)$ with diagonal elements equal to the items of the corresponding vector. Eq. (10) gives $\boldsymbol{\psi}_j$ as a product of the vector \mathbf{o} and a fixed matrix, provided that p is given. Contrary to it Eq.(11) gives $\boldsymbol{\psi}_j$ as a product of the vector p and a fixed matrix provided that \mathbf{o} is given. We use these relations for minimization of some objective functions with respect to \mathbf{o} and p .

The resolution of the object imaging is defined by the sampling intervals Δ_x and Δ_y , respectively, for the coordinates x and y . Accordingly to the formulas (4) for a smaller z and larger M_s, N_s these sampling intervals are smaller and the resolution is higher.

The noiseless intensity observations for the introduced modeling of the pty-

chographical system has the following form

$$I_j[l] = |u_j[l]|^2, \quad j = 1, \dots, J. \quad (12)$$

Thus, the observations are given for intensities obtained at all sensor pixels at various J locations of the probe. In vector notations the set of observations (12) is given in the form

$$\mathbf{I}_j = |\mathbf{u}_j|^2, \quad \mathbf{I}_j \in \mathbb{R}^{n_s}, \quad j = 1, \dots, J. \quad (13)$$

2.B. Poissonian observations

We consider Poissonian random observations of the intensities (12). In Poissonian modeling observations take random integer values interpreted as a counted number of photons detected by the sensor. This discrete distribution has a single parameter μ and defined by the formula $p(I_j[l] = L) = \exp(-\mu) \frac{\mu^L}{L!}$. Here $p(I_j[l] = L)$ is the probability that a random $I_j[l]$ takes value L , where $L \geq 0$ is an integer. The parameter μ is the intensity flow of Poissonian random events.

For the considered system the parameter μ is different for different experiment j and different l with values $\mu = I_j[l]$. Then the probabilistic Poissonian

observation model is given by the formula

$$p(I_j[l] = L) = \exp(-I_j[l]\chi) \frac{(I_j[l]\chi)^L}{L!}. \quad (14)$$

According to the properties of the Poissonian distribution, we have for the mean value and the variance of the observed $\hat{I}_j[l]$, $E\{\hat{I}_j[l]\} = var\{\hat{I}_j[l]\} = I_j[l]\chi$. In these formulas χ is a scaling parameter which can be interpreted as an exposure time. A larger χ means a larger exposure time, and a larger number of the photons with the intensity $I_j[l]$ is recorded.

The probabilistic observation model corresponding to (3) can be written in the form

$$\hat{I}_j[l] = \mathcal{Poisson}\{I_j[l]\chi\}, \quad (15)$$

where $\mathcal{Poisson}\{\cdot\}$ stands for a generator of random Poissonian numbers with the corresponding intensity.

Instead of (15) the normalized intensity (per sec.) is measured as

$$\hat{I}_j[l] = \mathcal{Poisson}\{I_j[l]\chi\}/\chi. \quad (16)$$

It is not difficult to realize that $E\{\hat{I}_j[l]\} = I_j[l]$ and $var\{\hat{I}_j[l]\} = I_j[l]/\chi$. Thus, larger χ (larger observation time) results in a smaller value of the variance, and always $\hat{I}_j[k]$ is an unbiased estimate of $I_j[k]$.

Note that the distribution of $\hat{I}_j[l]$ defined by Eq. (16) is not Poissonian anymore, in particular just because $E\{\hat{I}_j[l]\} \neq var\{\hat{I}_j[l]\}$.

For a sufficiently large $I_j[l]\chi$, $I_j[l]\chi > 1000$ (even for much smaller $I_j[l]\chi$), the Poissonian distribution (14) is well approximated by the Gaussian distribution with the mean and the variance equal to $I_j[l]\chi$ [23]. Then, $\hat{I}_j[l]$ in Eq. (16) is also Gaussian with the distributions $\mathcal{N}(I_j[l], I_j[l]/\chi)$.

With this Gaussian approximation the observed random $\hat{I}_j[l]$ can be represented in the form with an additive random noise

$$\hat{I}_j[l] = I_j[l] + \sigma_{j,l}\varepsilon_{j,l}, \quad (17)$$

$$\sigma_{j,l}^2 = I_j[l]/\chi. \quad (18)$$

where $\varepsilon_{j,l}$ is the standard zero-mean Gaussian white random noise, $\varepsilon_{j,l} \sim \mathcal{N}(0, 1)$.

In the vector notations the observation model (17)-(18) is of the form

$$\hat{\mathbf{I}}_j = \mathbf{I}_j + \boldsymbol{\sigma}_j \circ \boldsymbol{\varepsilon}_j, \quad (19)$$

$$\boldsymbol{\sigma}_j^2 = \mathbf{I}_j/\chi, \quad j = 1, \dots, J, \quad (20)$$

where $\boldsymbol{\varepsilon}_j \sim \mathcal{N}(0, \mathbf{I}_{n_s \times n_s})$. A thermal Gaussian random noise usual for digital sensors can be included in the models (17) and (19) as Gaussian random

summands.

2.C. *Sparse object and probe modeling*

One of the fundamental assumptions and facts of CS is that *natural* signals are *sparse*. The sparsity assumes that there exists a basis consisting of a small number of items where the image can be represented exactly or approximately with high accuracy. This *ideal sparse* basis is prior unknown and selected from a given set of potential bases (dictionary or dictionaries). The popularity and success of *CS* are due to the attractive theory, the efficient algorithms and the evidence that the developed formalism fits perfectly to many important applications. In CS the image is reconstructed from subsampled data. A total number of available observations can be smaller (much smaller) than size of the image. CS can be treated as a special regularization technique for inverse imaging where regularization is introduced through a *sparse* image modeling.

For sparse imaging we exploit two complementary ideas. First, an overcomplete dictionary (set of potential bases) is designed. Second, a suitable basis actually used for modeling of object and probe are selected from this dictionary. The overcomplete dictionary is composed from linear dependent functions (known as frames) and should be reach enough in order indeed

include the desirable sparse basis. The overcomplete frame representations are discussed further in this subsection. The suitable sparse basis selection is obtained by solving special optimization problems considered in the next section.

Overcomplete approximation for real-valued signal \mathbf{Y} can be given in two different *synthesis* and *analysis* forms, respectively, as follows

$$\mathbf{Y} = \mathbf{\Psi}\boldsymbol{\theta}, \boldsymbol{\theta} = \mathbf{\Phi}\mathbf{Y}. \quad (21)$$

Here, $\mathbf{Y} \in \mathbb{R}^n$, $\boldsymbol{\theta} \in \mathbb{R}^m$, and $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are transform matrices of $n \times m$ and $m \times n$, respectively. The vector $\boldsymbol{\theta}$, usually called *spectrum*, gives the parameters for the parametric approximation of the image \mathbf{Y} as $\mathbf{Y} = \mathbf{\Psi}\boldsymbol{\theta}$. Thus, $\mathbf{Y} = \sum_{j=1}^m \mathbf{\Psi}_j \cdot \boldsymbol{\theta}_j$, where $\mathbf{\Psi}_j$ are the columns of the matrix $\mathbf{\Psi}$, and $\boldsymbol{\theta}_j$ are the items of the vector $\boldsymbol{\theta}$. It is recognized that overcomplete representations for \mathbf{Y} with $m \gg n$ and linearly dependent $\mathbf{\Psi}_j$ form a much more powerful tool for advanced imaging than the classical orthonormal bases, where $m = n$. The concept of *frame* is a generalization of the classical bases developed for overcomplete (synthesis and analysis) representations with linearly dependent approximating functions (e.g. [24]). There are special links between the analysis and synthesis frames. The requirement, $\mathbf{\Psi} \cdot \mathbf{\Phi} = \mathbf{I}_{n \times n}$, where $\mathbf{I}_{n \times n}$ is

the $n \times n$ identity matrix, enables a perfect reconstruction of any \mathbf{Y} from the corresponding spectrum $\boldsymbol{\theta}$, indeed $\mathbf{Y} = \boldsymbol{\Psi}\boldsymbol{\theta} = \boldsymbol{\Psi} \cdot \boldsymbol{\Phi}\mathbf{Y} = \mathbf{Y}$. For the so-called *tight* frames, $\boldsymbol{\Phi}^T \cdot \boldsymbol{\Phi} = \mathbf{I}_{n \times n}$ and $\boldsymbol{\Psi} = \boldsymbol{\Phi}^T$.

For details and applications of overcomplete, in particular, frame based modeling for imaging we refer to the recent book [25].

Modeling of the object and probe functions \mathbf{o} and p lies at the core of the variational approach to ptychographical imaging proposed in this paper. Following to Katkovnik and Astola [20] separate sparse models are used for magnitude and phase. The following equations link magnitude and phase of the object function with their corresponding transform (spectrum) representations:

$$\text{mod}(\mathbf{o}) = \boldsymbol{\Psi}_{a,o}\boldsymbol{\theta}_{a,o}, \text{ angle}(\mathbf{o}) = \boldsymbol{\Psi}_{\varphi,o}\boldsymbol{\theta}_{\varphi,o}, \quad (22)$$

$$\boldsymbol{\theta}_{a,o} = \boldsymbol{\Phi}_{a,o} \cdot \text{mod}(\mathbf{o}), \boldsymbol{\theta}_{\varphi,o} = \boldsymbol{\Phi}_{\varphi,o} \cdot \text{angle}(\mathbf{o}), \quad (23)$$

where $\boldsymbol{\theta}_{a,o}$ and $\boldsymbol{\theta}_{\varphi,o}$ are vectors of the magnitude and phase spectra for the object. The modulus and angle operations applied to vectors in (22)-(23) are elementwise. Thus $\text{abs}(\mathbf{o})$ and $\text{angle}(\mathbf{o})$ are the vectors of magnitude and phase values.

The *frame* synthesis and analysis matrices $\boldsymbol{\Psi}_{a,o}$, $\boldsymbol{\Phi}_{a,o}$, $\boldsymbol{\Psi}_{\varphi,o}$, $\boldsymbol{\Phi}_{\varphi,o}$ are shown

with the indices a and φ for magnitude and phase, respectively.

Eqs. (22) define the synthesis giving magnitude ($\text{abs}(\mathbf{o})$) and phase ($\text{angle}(\mathbf{o})$) from the magnitude and phase spectra $\boldsymbol{\theta}_{a,o}$ and $\boldsymbol{\theta}_{\varphi,o}$. On the other hand, the analysis Eqs. (23) give the spectra for magnitude and phase of \mathbf{o} .

The l_0 -norm of the vector $\boldsymbol{\theta}$, denoted by $\|\boldsymbol{\theta}\|_{l_0}$, is defined as a number of nonzero elements of the vector. The l_1 -norm of $\boldsymbol{\theta}$ is defined as the sum of the absolute values of all items of the vector, $\|\boldsymbol{\theta}\|_{l_1} = \sum_s |\boldsymbol{\theta}_s|$. Both these norms are used in order to characterize *sparsity* of approximation. A smaller value of the norm means a higher sparsity of approximation.

In our approach the sparsity is evaluated separately for phase and magnitude by the l_p -norms, $\|\boldsymbol{\theta}_{\varphi,o}\|_{l_p}$ and $\|\boldsymbol{\theta}_{a,o}\|_{l_p}$, where $l_p = l_0, l_1$ for the l_0 - and l_1 -norms, respectively. Overall, the main intention is to find sparsest (shortest) models for phase and magnitude with smallest values of the l_p -norms.

The modeling of the probe function p is used in the form similar to (22)-(23) with the transforms matrices where the index ‘ o ’ is replaced by the index ‘ p ’:

$$\text{mod}(\mathbf{p}) = \boldsymbol{\Psi}_{a,p} \boldsymbol{\theta}_{a,p}, \quad \text{angle}(\mathbf{p}) = \boldsymbol{\Psi}_{\varphi,p} \boldsymbol{\theta}_{\varphi,p}, \quad (24)$$

$$\boldsymbol{\theta}_{a,p} = \boldsymbol{\Phi}_{a,p} \cdot \text{mod}(\mathbf{p}), \quad \boldsymbol{\theta}_{\varphi,p} = \boldsymbol{\Phi}_{\varphi,p} \cdot \text{angle}(\mathbf{p}), \quad (25)$$

where $\boldsymbol{\theta}_{a,p}$ and $\boldsymbol{\theta}_{\varphi,p}$ are vectors of the magnitude and phase spectra for the

probe p .

The above formulas represent general and formal aspects of overcomplete modeling for CS. The problem is how to obtain frames (dictionary) which are indeed so rich that they can be quite universal and efficient in applications. A family of the BM3D algorithms for various imaging problems were recently proposed within the framework of nonlocal patch-wise image modeling by Dabov et al. [26]. The corresponding analysis and synthesis frames, formalizing BM3D image modeling, are developed by Danielyan et al. [27]. We exploit these BM3D frames in this paper for implementation of our algorithm. The success of BM3D image processing for various applications works in favor of this choice. We wish to note that the analysis and synthesis frame operations are implemented algorithmically without explicit calculation of the analysis and synthesis matrices $\Psi_{a,o}$, $\Phi_{a,o}$, $\Psi_{\varphi,o}$, $\Phi_{\varphi,o}$, which may have extremely large dimensions.

3. Decoupled compressive imaging technique

Conventionally the compressive imaging assumes that reconstructed images are obtained as solutions of optimization problems with objective functions including two kind of terms. First ones are obtained from the probabilistic

observation models. Second ones give evaluation of the image complexity calculated from the spectra θ using their l_p -norms.

In this paper instead of the standard single-objective minimizations, we consider an algorithm based on multi-objective optimization proposed for real- and complex-valued imaging in the papers [27] and [20], respectively.

The motivations of this approach are as follows:

(A) The algorithm produces a better imaging and better accuracy than those of the algorithms using the fidelity and regularization terms together in single-objective optimization.

(B) The algorithm is simple in implementation because the inversion and the filtering are decoupled. An efficient procedures based on DFT are developed for the inversion steps, and hard- and soft- thresholdings for the phase and magnitude filtering.

Let us introduce the following two objective function:

$$\mathcal{L}_1(\mathbf{o}, \mathbf{p}, \Theta) = \frac{\chi}{2} \sum_{j=1}^J \|(\hat{\mathbf{I}}_j - |\mathbf{u}_j|^2) ./ |\mathbf{u}_j|\|_2^2 + \quad (26)$$

$$\frac{1}{\gamma_o} \|\mathbf{o} - \mathbf{v}_o\|_2^2 + \frac{1}{\gamma_p} \|\mathbf{p} - \mathbf{v}_p\|_2^2, \quad (27)$$

$$\mathbf{v}_o = (\Psi_{a,o} \boldsymbol{\theta}_{a,o}) \circ \exp(j\Psi_{\varphi,o} \boldsymbol{\theta}_{\varphi,o}), \quad (28)$$

$$\mathbf{v}_p = (\Psi_{a,p} \boldsymbol{\theta}_{a,p}) \circ \exp(j\Psi_{\varphi,p} \boldsymbol{\theta}_{\varphi,p}); \quad (29)$$

$$\mathcal{L}_2(\mathbf{o}, \mathbf{p}, \Theta) = \tau_{a,o} \cdot \|\boldsymbol{\theta}_{a,o}\|_{l_p} + \tau_{\varphi,o} \cdot \|\boldsymbol{\theta}_{\varphi,o}\|_{l_p} + \quad (30)$$

$$\tau_{a,p} \cdot \|\boldsymbol{\theta}_{a,p}\|_{l_p} + \tau_{\varphi,p} \cdot \|\boldsymbol{\theta}_{\varphi,p}\|_{l_p} + \quad (31)$$

$$\frac{1}{2} \|\boldsymbol{\theta}_{a,o} - \Phi_{a,o} \text{mod}(\mathbf{o})\|_2^2 + \frac{1}{2} \|\boldsymbol{\theta}_{\varphi,o} - \Phi_{\varphi,o} \text{angle}(\mathbf{o})\|_2^2 + \quad (32)$$

$$\frac{1}{2} \|\boldsymbol{\theta}_{a,p} - \Phi_{a,p} \text{mod}(\mathbf{p})\|_2^2 + \frac{1}{2} \|\boldsymbol{\theta}_{\varphi,p} - \Phi_{\varphi,p} \text{angle}(\mathbf{p})\|_2^2. \quad (33)$$

The first summand in Eq. (26) is the fidelity term obtained due to the maximum likelihood method as the minus logarithm of the joint probabilistic distribution of the Gaussian observations (19)-(20). The second (27), third (28) and fourth summands (29) are the synthesis constraints (22) and (24) given for complex-valued \mathbf{o} and p through quadratic penalties. γ_o and γ_p are parameters of this penalization.

Summands in Eqs.(30)-(31) define the complexity (sparsity) of the models for phase and magnitude. The summands in Eqs.(32)-(33) are the analysis constraints (25) and (23) given using the quadratic penalties. The quadratic penalization of constraint-equations is the usual practice in order to reduce the constrained optimization to an unconstrained one. Smaller values of the corresponding penalization parameters ($\gamma_{a,o}, \gamma_{\varphi,o}, \gamma_{a,p}, \gamma_{\varphi,p}$) mean that in the solution the corresponding equations will be fulfilled more accurately.

The objective functions \mathcal{L}_1 and \mathcal{L}_2 include the observation model as well as the all constraints linking the variables of the problem. Following to the approach developed in [20] we use the alternating minimization of \mathcal{L}_1 and \mathcal{L}_2 , provided that \mathcal{L}_1 is minimized with respect to \mathbf{o} , p and \mathcal{L}_2 is minimized with respect to $\Theta = (\theta_{a,o}, \theta_{\varphi,o}, \theta_{a,p}, \theta_{\varphi,p})$.

It results in the following type of the iterative algorithm:

$$(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}) = \arg \min_{\mathbf{o}, \mathbf{p}} \mathcal{L}_1(\mathbf{o}, \mathbf{p}, \Theta^t), \quad (34)$$

$$\Theta^{t+1} = \arg \min_{\Theta} \mathcal{L}_2(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}, \Theta). \quad (35)$$

In Eq. (34) minimization of \mathcal{L}_1 on (\mathbf{o}, p) inverses the nonlinear forward propagation operators from \mathbf{o}, p to $\{\mathbf{I}_j\}$ and, filters noise in the observations $\{\hat{\mathbf{I}}_j\}$. Minimization of \mathcal{L}_2 on Θ filters phase and magnitude of the object \mathbf{o} and the probe p from observation noise appeared in $(\mathbf{o}^{t+1}, p^{t+1})$ and looks for the suitable basis for reconstruction of \mathbf{o} and p .

Thus, in the proposed algorithm the inverse and model selection for \mathbf{o} and p are decoupled. Instead of minimization of a single objective function the algorithm (34)-(35) is looking for a quite different solution which is known as

a *fixed-point* $(\mathbf{o}^*, \mathbf{p}^*, \Theta^*)$ defined by the equations:

$$(\mathbf{o}^*, \mathbf{p}^*) = \arg \min_{\mathbf{o}, \mathbf{p}} \mathcal{L}_1(\mathbf{o}, \mathbf{p}, \Theta^*), \quad (36)$$

$$\Theta^* = \arg \min_{\Theta} \mathcal{L}_2(\mathbf{o}^*, \mathbf{p}^*, \Theta).$$

This fixed-point balances two different intentions: minimization of the fidelity term (accuracy of observation fitting) and the complexity of the used model (defined by the norms $\|\boldsymbol{\theta}_{a,o}\|_{l_p}$, $\|\boldsymbol{\theta}_{\varphi,o}\|_{l_p}$, $\|\boldsymbol{\theta}_{a,p}\|_{l_p}$, $\|\boldsymbol{\theta}_{\varphi,p}\|_{l_p}$).

In terms of the game theory the problem (36) can be interpreted as a game of two players identified, respectively, with two groups of variables (\mathbf{o}, \mathbf{p}) and Θ [28]. An interaction between the players is noncooperative because minimization of \mathcal{L}_1 with respect to (\mathbf{o}, \mathbf{p}) in general results in increase of \mathcal{L}_2 and minimization of \mathcal{L}_2 with respect to Θ increases \mathcal{L}_1 . The equilibrium of this game called *Nash equilibrium* defines the *fixed point* $(\mathbf{o}^*, \mathbf{p}^*, \Theta^*)$ of this optimization. For $p = 1$, the problem (36) is convex.

Hence the Nash equilibrium provides a balance between the fit of the reconstructions \mathbf{o} and p to the observations $\hat{\mathbf{I}}_j$ and the complexity of the models for the magnitude and phase for the object and probe.

In the algorithm development we modify the objective function $\mathcal{L}_1(\mathbf{o}, p, \Theta)$

by replacing it for

$$\begin{aligned} \mathcal{L}_1(\mathbf{o}, \mathbf{p}, \mathbf{w}, \Theta) &= \frac{\lambda}{2} \sum_{j=1}^J \|(\hat{\mathbf{I}}_j - |\mathbf{w}_j|^2) ./ |\mathbf{w}_j|\|_2^2 + \\ &\frac{1}{\gamma_1} \sum_{j=1}^J \|\mathbf{u}_j - \mathbf{w}_j\|_2^2 + \frac{1}{\gamma_o} \|\mathbf{o} - \mathbf{v}_o\|_2^2 + \frac{1}{\gamma_p} \|\mathbf{p} - \mathbf{v}_p\|_2^2, \\ \mathbf{v}_o &= (\Psi_{a,o} \boldsymbol{\theta}_{a,o}) \circ \exp(j \Psi_{\varphi,o} \boldsymbol{\theta}_{\varphi,o}), \quad \mathbf{v}_p = (\Psi_{a,p} \boldsymbol{\theta}_{a,p}) \circ \exp(j \Psi_{\varphi,p} \boldsymbol{\theta}_{\varphi,p}). \end{aligned} \quad (37)$$

In the fidelity term of this new \mathcal{L}_1 the variable \mathbf{u}_j is replaced by its estimate \mathbf{w}_j . This additional variable serves as a splitter simplifying minimization. This estimate of \mathbf{u}_j is obtained due to the quadratic penalization term $\frac{1}{\gamma_1} \sum_{j=1}^J \|\mathbf{u}_j - \mathbf{w}_j\|_2^2$.

With this new variable the optimization procedure (34) is complemented by minimization on \mathbf{w}_j . Thus Eqs.(34)-(35) take the form

$$(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}, \mathbf{w}^{t+1}) = \arg \min_{\mathbf{o}, \mathbf{p}, \mathbf{w}} \mathcal{L}_1(\mathbf{o}, \mathbf{p}, \mathbf{w}, \Theta^t), \quad (38)$$

$$\Theta^{t+1} = \arg \min_{\Theta} \mathcal{L}_2(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}, \Theta). \quad (39)$$

We replace optimization of \mathcal{L}_1 simultaneous with respect to variables \mathbf{o}, p ,

\mathbf{w} on the successive one thus the algorithm becomes

$$\mathbf{o}^{t+1} = \arg \min_{\mathbf{o}} \mathcal{L}_1(\mathbf{o}, \mathbf{p}^t, \mathbf{w}^t, \Theta^t), \quad (40)$$

$$\mathbf{p}^{t+1} = \arg \min_{\mathbf{p}} \mathcal{L}_1(\mathbf{o}^{t+1}, \mathbf{p}, \mathbf{w}^t, \Theta^t), \quad (41)$$

$$\mathbf{w}^{t+1} = \arg \min_{\mathbf{w}} \mathcal{L}_1(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}, \mathbf{w}, \Theta^t), \quad (42)$$

$$\Theta^{t+1} = \arg \min_{\Theta} \mathcal{L}_2(\mathbf{o}^{t+1}, \mathbf{p}^{t+1}, \Theta). \quad (43)$$

4. Algorithm development

Solution of the problems (40)-(43) are detailed in Appendix. In this section we present the corresponding results and discuss how they are used in the successive steps of the developed algorithm.

4.A. Solution of optimization problems

1. The problem (40) gives as a solution the following expression for the object o :

$$o[k] = \frac{v_o[k]/\gamma_o + \frac{1}{\gamma_1} \sum_{j=1}^J \hat{\psi}_j[k - k_j] p^*[k - k_j]}{1/\gamma_o + \frac{1}{\gamma_1} \sum_{j=1}^J |p[k - k_j]|^2}, \quad (44)$$

$$\hat{\psi}_j[k] = \mathcal{DFT}^{-1}\{w_j\},$$

where k is an integer $2D$ argument variable for the object function.

Let us check the scaling of the estimate of $o[k]$ in Eq.(44). In this solution, $v_o[k]$ is an estimate of $o[k]$, and $\hat{\psi}_j[k]$ is an estimate of $\psi_j[k] = o[k +$

$k_j]p[k]$. Assuming for a moment that $v_o[k] \simeq o[k]$ and $\hat{\psi}_j \simeq \psi_j[k]$ and inserting these functions in Eq.(44) we obtain $o[k] \simeq o[k]$. Thus, the solution (44) enables an accurate scaling for the estimate of $o[k]$.

2. The problem (41) gives as a solution for the probe p :

$$p[k'] = \frac{v_p[k']/\gamma_p + \frac{1}{\gamma_1} \sum_{j=1}^J \hat{\psi}_j[k'] o^*[k' + k_j]}{1/\gamma_p + \frac{1}{\gamma_1} \sum_{j=1}^J |o[k' + k_j]|^2}. \quad (45)$$

Note, that the argument variable k' is different from the one used in Eq.(44). In Eq.(45) it takes values in the area defined by the probe support $s_p(k')$. The scaling of the solution $p[k']$ can be checked assuming that $v_p[k'] \simeq p[k']$ and $\hat{\psi}_j[k'] \simeq o[k' + k_j]p[k']$. Then, we can see that $p[k'] \simeq p[k']$. Thus, the scaling in Eq.(45) as the estimate of $p[k']$ is quite accurate.

3. A solution of the problem (42) has a form of the nonlinear operator

$$w_j[l] = \mathcal{G}(\hat{I}_j[l], u_j[l]) \triangleq |\hat{\mathbf{w}}_j[l]| \exp(j\text{angle}(u_j[l])), \quad (46)$$

where the phase is equal to the phase (angle) of $u_j[l]$, and where the magnitude is a solution of the optimization problem

$$|\hat{\mathbf{w}}_j[l]| = \arg \min_{|\mathbf{w}_j[l]|} \frac{\chi}{2} (\hat{I}_j[l] - |w_j[l]|^2)^2 / |w_j[l]|^2 + \frac{1}{\gamma_1} (|u_j[l]| - |\mathbf{w}_j[l]|)^2. \quad (47)$$

The solution $|\hat{\mathbf{w}}_j[l]|$ belongs to the interval $[|u_j[l]|, \sqrt{\hat{I}_j[l]}]$. For a small noise level, when $\chi\gamma_1$ is comparatively large, $\chi\gamma_1 \rightarrow \infty$, this solution takes the form

$$w_j[l] = \frac{u_j[l]}{|u_j[l]|} \sqrt{\hat{I}_j[l]}. \quad (48)$$

Thus, we arrive to the well known rule commonly used in the phase retrieval algorithms: the variable $w_j[l]$ has the absolute value equal to the square root of the observation, and the phase is equal to the one predicted by the wavefield propagation model. The conventional solution (48) is optimal for noiseless data but for essentially noisy observations can be very different from the optimal solution (46).

4. The problem (43) is additive with respect to the components of Θ . It follows that the solution can be given in the form independent for each component of Θ . In the explicit form these solutions are as follows:

$$\boldsymbol{\theta}_{a,o} = \mathfrak{Ih}_{\tau_{a,o}}(\Phi \bmod(\mathbf{o})), \quad \boldsymbol{\theta}_{\varphi,o} = \mathfrak{Ih}_{\tau_{\varphi,o}}(\Phi_{\varphi,o} \text{angle}(\mathbf{o})), \quad (49)$$

$$\boldsymbol{\theta}_{a,p} = \mathfrak{Ih}_{\tau_{a,p}}(\Phi_{a,p} \bmod(\mathbf{p})), \quad \boldsymbol{\theta}_{\varphi,p} = \mathfrak{Ih}_{\tau_{\varphi,p}}(\Phi_{\varphi,p} \text{angle}(\mathbf{p})), \quad (50)$$

where the vectors $\Phi_{a,o} \bmod(\mathbf{o})$, $\Phi_{\varphi,o} \text{angle}(\mathbf{o})$, $\Phi_{a,p} \bmod(p)$, and $\Phi_{\varphi,p} \text{angle}(p)$ are the analysis spectra calculated for magnitude and phase

of object and probe defined accordingly to the analysis operations given in Eq.(23).

The threshold operator \mathfrak{Th}_τ with a parameter τ is applied independently to each elements of the corresponding vectors. For an arbitrary scalar variable b it is defined according to the generic formulas

$$\hat{b} = \mathfrak{Th}_\tau(b) = \begin{cases} \mathfrak{Th}_{\sqrt{2\tau}}^{hard}(b) = b \cdot 1(|b| \geq \sqrt{2\tau}), \text{ if } l_p = l_0, \\ \mathfrak{Th}_\tau^{soft}(b) = \text{sign}(b) \cdot \max(|b| - \tau, 0), \text{ if } l_p = l_1, \end{cases} \quad (51)$$

where *hard* and *soft* indicate the type of the thresholdings: hard or soft.

It is emphasized that these different thresholding operations obtained as solutions for the problem (43). The hard-thresholding is an operation obtained for the l_0 -norm and the soft-thresholding is an operator obtained for the l_1 -norm in the criterion \mathcal{L}_2 . Note that in Eqs.(51) the threshold parameters for the soft- and hard thresholdings are different.

The thresholding enables a strong filtering for magnitude and phase of the object and probe functions performed in the spectral domain. The hard thresholding keeps only large spectral components in such way that $\hat{b} = b$ for $|b| \geq \sqrt{2\tau}$ and $\hat{b} = 0$ otherwise. For the soft thresholding $\hat{b} = 0$

if $|b| \leq \tau$. When the spectral components are not zero, $|b| > \tau$, they have absolute values which are smaller than the absolute values of the original b , these nonzero spectral components are calculated as $\text{sign}(b) (|b| - \tau)$.

4.B. Proposed C-PI algorithm

Combining the solutions (44)-(51) we developed the following iterative algorithm.

Algorithm C-PI

Input: $o^{init}, p^{init}, \{\hat{I}_j\}$

Set $t = 0, o^0 = v_o^0 = o^{init}, p^0 = v_p^0 = p^{init},$

Repeat until convergence

1. Forward propagation: $u_j^t[l] = \mathcal{DFT}\{o^t[k + k_j]p^t[k]\};$
 2. Observation filtering: $w_j^t[l] = \mathcal{G}(\hat{I}_j[l], u_j^t[l]);$
 3. Back propagation: $\hat{\psi}_j^t[k - k_j] = \mathcal{DFT}^{-1}\{w_j^t\}, j = 1, \dots, J;$
 4. Object update:
$$o^{t+1}[k] = \frac{v_o^t[k]/\gamma_o + \frac{1}{\gamma_1} \sum_{j=1}^J \hat{\psi}_j^t[k - k_j](p^t[k - k_j])^*}{1/\gamma_o + \frac{1}{\gamma_1} \sum_{j=1}^J |p^t[k - k_j]|^2};$$
 5. Probe update:
$$p^{t+1}[k'] = \frac{v_p^t[k']/\gamma_p + \frac{1}{\gamma_1} \sum_{j=1}^J \hat{\psi}_j^t[k'](o^{t+1}(k' + k_j))^*}{1/\gamma_p + \frac{1}{\gamma_1} \sum_{j=1}^J |o^{t+1}[k' + k_j]|^2};$$
 6. Probe relaxation: $p^{t+1}[k'] = (1 - \alpha)p^{t+1}[k'] + \alpha p^t[k'];$
 7. Object filtering: $v_o^{t+1} = \text{BM3D-filter}(o^{t+1});$
 8. Probe filtering: $v_p^{t+1} = \text{BM3D-filter}(p^{t+1});$
- $t = t + 1.$

The input variables of the algorithm are: the intensity observations $\{\hat{I}_j\}_1^J;$

initial guesses for object o^{init} and probe p^{init} .

Step 1 corresponds to the wavefield propagation from the object/probe plane to the sensor (Eq.(5)). It gives the complex wavefields $u_j^{t+1}[l]$. Step 2 based on the solution (46)-(47) is used in order to correct the sensor wavefield predictions by observed intensities $\{\hat{I}_j\}_1^J$. This correction is done in the DFT space and returned to the image domain by the inverse DFT (step 3).

Step 4 and Step 5 update the object and probe wavefields using the solutions (44) and (45).

Step 6 slows down changes in the probe update. This relaxation of the probe variations stabilizes the blind inverse imaging and improve the accuracy of reconstruction for both the object and the probe. In our simulation $\alpha = .9$.

BM3D-filter shown in Step 7 and Step 8 combines the following four operations. First, the design of the analysis and synthesis frames Φ and Ψ . Second, the analysis of the object and probe images o^{t+1} and p^{t+1} (see Eqs.(23) and (25)). Third, the filtering of these spectral representations of these images using the thresholding solutions (49)-(50). Fourth, inverse the filtered spectral representations into the image domains (see Eqs.(22) and (24)). At this final step we obtain the the images v_o^{t+1} and v_p^{t+1} . These wavefields are filtered

versions of the object and probe images o^{t+1} and p^{t+1} . The analysis, filtering and return of images from the spectral representations are produced for both the magnitude and phase of the wavefields o^{t+1} and p^{t+1} . The filtered v_o^{t+1} and v_p^{t+1} are used in Step 3 and Step 4, where they are aggregated with the estimates of the object and probe images obtained from noisy observations $\{\hat{I}_j\}_1^J$.

BM3D-filter allows two different modes: first, with fixed analysis and synthesis frames obtained on the previous iterations and, second, with the design of the analysis and synthesis frames for varying o^{t+1} and p^{t+1} .

In our experiments we assume that the first $t_0 - 1$ iterations are produced with no BM3D-filtering. It means that Step 6 and Step 7 are dropped, and $v_o^t = v_p^t = 0$ and $1/\gamma_o = 1/\gamma_p = 0$ in Steps 4 and 5.

The mode of BM3D-filter with the frame design is compulsory for the iteration $t = t_0$. The BM3D frames are data dependent and for this step they are obtained for o^{t_0} and p^{t_0} . These frames (different for object and probe) may be fixed up to the end of reconstruction or updated by BM3D-filter after a number of iterations. It was noticed that only a few updates are able to improve the reconstruction.

We use the name **C-PI** for the proposed algorithm as the abbreviation for **C**ompressive **P**tychographical **I**maging.

4.C. Basic (no filtering) C-PI algorithm

As a simplified version of the C-PI algorithm we present it in “*no-filtering*” mode where BM3D-filtering is dropped and the optimal observation filtering is replaced by the conventional operator (48). This simplified algorithm which

we call **Basic C-PI** takes the form.

Algorithm: Basic C-PI (no filtering)

Input: $o^{init}, p^{init}, \{\hat{I}_j\}$

Set $t = 0, o^0 = o^{init}, p^0 = p^{init},$

Repeat until convergence:

1. Forward propagation: $u_j^t[l] = \mathcal{DFT}\{o^t[k + k_j]p^t[k]\};$
2. Observation processing, Eq.(72): $w_j^t[l] = \frac{u_j^t[l]}{|u_j^t[l]|} \sqrt{\hat{I}_j[l]};$
3. Back propagation: $\hat{\psi}_j^t[k - k_j] = \mathcal{DFT}^{-1}\{w_j^t\}, j = 1, \dots, J;$

4. Object update:

$$o^{t+1}[k] = \frac{\sum_{j=1}^J \hat{\psi}_j^t[k - k_j] (p^t[k - k_j])^*}{\sum_{j=1}^J |p^t[k - k_j]|^2};$$

5. Probe update:

$$p^{t+1}[k'] = \frac{\sum_{j=1}^J \hat{\psi}_j^t[k'] (o^{t+1}[k' + k_j])^*}{\sum_{j=1}^J |o^{t+1}[k' + k_j]|^2};$$

6. Probe relaxation:

$$p^{t+1}[k'] = (1 - \alpha)p^{t+1}[k'] + \alpha p^t[k']$$

$t = t + 1.$

5. Comparison of the algorithms

In this section we compare the proposed algorithm versus two algorithms which are the most popular in the field. For reader's convenience and more transparent discussion we briefly present these algorithms using our notation for variables.

5.A. Parallel Ptychographical Iterative Engine (p-PIE)

This algorithm is proposed by Thibaut et al. [10]. The algorithm iterates the object and probe reconstructions between the image and Fourier transform spaces and “*projects*” these reconstructions on constraints given in these spaces. In our notation the object and probe are updated according to the formulas:

$$\begin{aligned} o^{t+1}[k] &= \frac{\sum_{j=1}^J \psi_j^t[k - k_j](p^t[k - k_j])^*}{\sum_{j=1}^J |p^t[k - k_j]|^2 s_p[k - k_j]}, \\ p^{t+1}[k'] &= \frac{\sum_{j=1}^J \psi_j^t[k'](o^{t+1}[k' + k_j])^*}{\sum_{j=1}^J |o^{t+1}[k' + k_j]|^2}. \end{aligned} \quad (52)$$

The main iterations produced for the variables ψ_j^{t+1} are as follows (see Eq. (9) in [10]):

$$\psi_j^{t+1}[k] = \psi_j^t[k] + \Pi_F\{2 \tilde{\psi}_j^{t+1}[k] - \psi_j^{t+1}[k]\} - \tilde{\psi}_j^{t+1}[k], \quad j = 1, \dots, J, \quad (53)$$

where $\tilde{\psi}_j^t$ are calculated according to Eq.(1) using the estimates o^t and p^t

from Eqs.(52):

$$\tilde{\psi}_j^t[k] = o^t[k + k_j]p^t[k], \quad (54)$$

and $\Pi_F\{\}$ is the “*projection operator*” corresponding to the rule (48). It consisting from three successive step: DFT, replacement of the DFT magnitude by $\sqrt{\hat{I}_j[k]}$ and the inverse DFT.

The algorithm given by Eqs.(52)-(54) is referred in [9] as *p-PIE* algorithm due to the fact that all views ψ_j^{t+1} are used in parallel simultaneously in the updates (52).

5.B. *Extended Ptychographical Iterative Engine (e-PIE)*

The e-PIE algorithm by Maiden and Rodenburg [9] is different from p-PIE first of all by sequential use of the observed diffraction patterns one-by-one. The patterns are addressed in a random sequence $q(i)$, $i = 1, \dots, J$, and random $q(i)$ takes one of the values in the set $\{1, \dots, J\}$. The object and probe updates are calculated according to the formulas

$$o^{i+1}[k] = o^i[k] + \quad (55)$$

$$\alpha \frac{(\tilde{\psi}_{q(i)}^i[k - k_{q(i)}] - \psi_{q(i)}^i[k - k_{q(i)}])(p^i[k - k_{q(i)}])^*}{|p^i[k - k_{q(i)}]|_{\max}^2},$$

$$p^{i+1}[k'] = p^i[k'] + \beta \frac{(\tilde{\psi}_{q(i)}^i[k'] - \psi_{q(i)}^i[k'])(o^{s(i)}[k' + k_{q(i)}])^*}{|o^{s(i)}[k' + k_{q(i)}]|_{\max}^2}. \quad (56)$$

Here $\psi_{q(i)}^i$ are calculated using Eq.(54), and $\tilde{\psi}_{q(i)}^i$ is calculated from $\psi_{q(i)}^i$ according to the rule (48) by replacing the magnitudes of its DFT by $\sqrt{\hat{I}_{q(i)}}$ and then inverse this modified DFT, the step parameters $\alpha = \beta = 1$.

The iterations (55) can be treated as a gradient descent (gradient with respect to o) for a sequence of the objective functions $L_o^i[k] = \frac{|\tilde{\psi}_i[k - k_i] - \psi_i[k - k_i]|^2}{|p^i[k - k_i]|_{\max}^2}$, $j = 1, \dots, J$, taken in random order. It is a sort of vectorial optimization where the successive steps are done with respect to different objective functions.

Similarly, the iterations (56) can be treated as a gradient descent (gradient with respect to p) for the objective functions $L_p^i[k] = \frac{|\tilde{\psi}_i[k'] - \psi_i[k']|^2}{|o^i[k + k_i]|_{\max}^2}$, $j = 1, \dots, J$, taken in the random order used for the iterations (55).

It is not difficult to realize that the sum $\sum_{i=1}^J \sum_k |\tilde{\psi}_i[k'] - \psi_i[k']|^2$ gives an indirect estimate of $\sum_{i=1}^J \sum_k (\sqrt{\hat{I}_i[k]} - \sqrt{I_i[k]})^2$ where \hat{I}_i is an estimate obtained from ψ_i for the observations I_i . Thus, minimizations of L_o^i and L_p^i are targeted on minimization of $\sum_{i=1}^J \sum_k (\sqrt{\hat{I}_i[k]} - \sqrt{I_i[k]})^2$.

5.C. C-PI versus p-PIE and e-PIE

The C-PI algorithm differs from the p-PIE and e-PIE algorithms in three principal aspects. First of all, the C-PI algorithm is derived for noisy Pois-

sonian data. Minimization of the minus log-likelihood summand $\frac{\chi}{2} \sum_{j=1}^J \|(\hat{\mathbf{I}}_j - |\mathbf{u}_j|^2) \cdot \mathbf{u}_j\|_2^2$ in \mathcal{L}_1 enables the ability of C-PI to filter Poissonian signal dependent noise. The problem of noise filtering is not formulated in derivation of the p-PIE and e-PIE algorithms. Minimization of the criteria $\sum_{i=1}^J (\sqrt{\hat{I}_i} - \sqrt{I_i})^2$ discussed in Subsection 5.B actually means that the p-PIE algorithm tries to fit the observed noisy data as accurately as possible. The noise filtering effects in the p-PIE and e-PIE algorithms appear only due to the overcompleteness of observations.

Second, the C-PI algorithm uses the sparse modeling for phase and magnitude. It enables the principal advantage of this algorithm in the accuracy and quality of imaging. p-PIE and e-PIE belong to the class of nonparametric algorithms which are oriented on reconstruction of arbitrary object and probe without any approximation of their variations.

Third, contrary to p-PIE and e-PIE the C-PI algorithm is derived from variational formulation of the reconstruction problem as a result of minimization of special objective functions. Thus the design of the algorithm is highly formalized. This variational approach is quite general and flexible. It can be extended in order to include, for instance, extra priory information

on object and probe, or uncertainties in modeling of image formation and experimental parameters.

The C-PI algorithm similar to p-PIE has a parallel structure using all observations simultaneously (Step 3 and Step 4). However, these parallel updates of object and probe are different in these algorithms.

6. Simulation experiments

6.A. Settings and parameters

In the presented results we use a complex valued transmissive object $o = a_O \exp(j \cdot \varphi_O)$, where the modulus a_O is given by the gray scale test-image “lena” and the phase φ_O by the binary test-image “chessboard”. Both images are of the size 256×256 . The modulus is scaled to range $[0.1, 1.1]$ and the phase to interval $[0, \pi/2]$.

The sensor $M_S \times N_S$ ($M_S = N_S = 128$) has square pixels $\Delta_S \times \Delta_S$, $\Delta_S = 7.4 \mu m$. A pixelation (discretization) of the object plane is defined by the object pixels $\Delta_x = \Delta_y = \Delta_O$, $\Delta_O = \Delta_S/4$. The wavelength $\lambda = .532 \mu m$ corresponds to the green laser.

The object scan is produced by the 21×21 probe array. The step size of the scan is equal to 10 object’s pixels, i.e. $10 \cdot \Delta_O = 18.5 \mu m$, with array

overlapping of about 70%.

The probe positions are known and given but integer random offsets from a regular square grid. These offsets have uniform distributions over the interval $[-2, 2]$ and introduced in order to suppress periodic artefacts in the reconstruction [9]. The probe beam is generated by propagating the pinhole wavefield a short distance $z/5$ from the pinhole to the object.

The observations o_r are Poissonian generated according to Eq. (16). Parameter χ in Eq. (15) defines the level of random noise in observations. In our experiments $\chi \in [.01, 10^4]$. For the introduced object the mean count of photons in each observation pattern is varying from $3.4 \cdot 10^4$ to $3.4 \cdot 10^{10}$, respectively. Thus, $\chi = 0.01$ corresponds to the noisiest case, while $\chi = 10^4$ corresponds to the nearly noiseless observations.

In what follows we show the results obtained for the hard-shresholding filtering in BM3D only, corresponding using in the C-PI algorithm the l_0 -norm in (51). Our experiments show that the l_0 -norm usually results in a better image reconstruction than the l_1 -norm.

According to Eq. (1) the observations depend on the products $o(r+r_j)p(r)$. It follows that the object and probe can be reconstructed only within invariant

complex-valued factors. In order to avoid this ambiguity in evaluation of the algorithm accuracy, following to [9], we use the corrected reconstructions o_h^t and p_h^t instead of the reconstructions o^t and p^t given by the algorithm iterations:

$$\begin{aligned} o_h^t &= o^t h_o^t, \quad p_h^t = p^t h_p^t, \\ h_o^t &= (o^t)^* o / \|o^t\|_2^2, \quad h_p^t = (p^t)^* p / \|p^t\|_2^2. \end{aligned} \tag{57}$$

The correction factors h_o and h_p , calculated for the object and the probe, are found from minimizations with respect to h of two criteria, respectively $\|o - o^t h\|_2^2$ and $\|p - p^t h\|_2^2$. We will use the term *h-corrected* for these reconstructions.

The phase of the factors h corrects the phase of the reconstructions and the modulus of the factors enables a proper scaling of the modulus reconstructions. The quality of fit for the complex-valued object and probe are calculated using the relative root-mean-squared error criteria [9]:

$$E_{0,\text{object}}(t) = \sqrt{\frac{\|o - o_h^t\|_2^2}{\|o\|_2^2}}, \quad E_{0,\text{probe}}(t) = \sqrt{\frac{\|p - p_h^t\|_2^2}{\|p\|_2^2}}. \tag{58}$$

We calculate also the standard root-mean-squared-error (*RMSE*) separately for the magnitude and phase reconstructions. It is done for *h-corrected* reconstructions. It is emphasized that the *h-corrected* reconstructions are

used only for the accuracy criteria calculation and for visualization of the reconstructions. In the algorithm the original uncorrected estimates of the variables are used.

In our experiments we use BM3D filtering only for the object reconstruction and omit it for the probe. It is done because the data for the probe are highly overcomplete and sufficient for non-parametric reconstruction without any special modeling for the phase and magnitude of the probe.

The computational complexity of the algorithm is characterized by the time required for 100 iterations. This time depends on the configuration of the algorithm. In the all filtering operations are excluded (Basic C-PI) the C-PI reconstruction takes about 110 sec. If *BM3D* filtering is on in C-PI but the filtering of observations is excluded (step 3 of the algorithm) the time of calculations is about 120 sec.

The observation filtering is the most demanding operation. If both observation and BM3D filtering are activated in C-PI the time of calculation is increased to 280 seconds.

100 iterations of the e-PIE algorithm require about 160 sec. This data are obtained for the computer and the software used in the experiments:

Intel Core 2Duo E8400 @ 3GHz, RAM 4GB, Windows Xp SP3, Matlab 7.9.0 (R2009b).

We make our MATLAB programs for the demo version of our C-PI and e-PIE algorithms are publicly available for testing:

<http://www.cs.tut.fi/~lasip/DDT/>.

6.B. Results and algorithm comparison

For the noiseless observations ($\chi = 10000$) the both e-PIE and C-PI algorithms works well and reconstruct the object and probe wavefields being initialized by the free space models for the object and the probe ($O = 1$ and $P = 1$).

However, for noisy data a much more accurate initial guess for the probe is important. In the further shown results this initial guess is obtained by applying C-PI in the noiseless scenario, $\chi = 10000$, with 200 iterations. Thus, a noiseless experiment with blind reconstruction of the probe and the object is used as a special test-experiment done for reconstruction of the probe, produced provided a larger exposure time corresponding $\chi = 10000$. In this calibration experiment we can use any transmissive object and the algorithm reconstructs both the unknown object and probe.

It is noticed that the free space object recommended in some publications for the probe reconstruction is a poor choice. Much better quality reconstructions can be obtained using the objects with strongly varying modulus/phase enabling a reach diversity of the intensities registered in diffraction patterns.

Fig. 2 shows $E_{0,\text{object}}$ and $E_{0,\text{probe}}$ curves obtained after 100 iterations for χ varying from 0.01 to 10000 (photon count from $3.4 \cdot 10^4$ to $3.4 \cdot 10^{10}$).

First, we compare the results corresponding to the e-PIE and Basic C-PI (without filtering) algorithms. Starting from about count 10^6 the curves for $E_{0,\text{object}}$ and $E_{0,\text{probe}}$ are more or less horizontal. It means that the level of the noise is so low that the reconstructions only slightly depend on it.

For this χ interval the values of $E_{0,\text{object}}$ and $E_{0,\text{probe}}$ for e-PIE are lower than those for C-PI. Thus, the e-PIE algorithm performs better than the C-PI algorithm without filtering.

For the higher noise level, count smaller than 10^6 , the situation is changed, and the C-PI algorithm performs better with the lower values of $E_{0,\text{object}}$ and $E_{0,\text{probe}}$, in particular for counts of the order $10^4 - 10^5$.

The best results are demonstrated by the C-PI algorithm where the both BM3D and observation filtering are exploited. The most essential advantage

is demonstrated for the high level noise observations with photon counts of the order $10^4 - 10^5$.

This advantage of the C-PI algorithm with filtering is supported by the results shown in Fig. 3 for RMSE criteria shown for the modulus and phase reconstructions. We wish to note that the observation filtering for the interval of low noise level is not compulsory and can be dropped. It allows essentially accelerate the C-PI algorithm.

The comparative converge rate of the algorithms is demonstrated in Fig. 4. The C-PI algorithm is faster than the e-PIE and Basic C-PI algorithms.

The filtering properties of C-PI are defined mainly by the following parameters: γ_1/γ_o (in the inverse formula (44)), γ_1/γ_p (in the inverse formula (45)), $\gamma_1\chi$ (in the observation filtering formula (47)), and $(\tau_{a,o}, \tau_{\varphi,o})$. The last two parameters define the thresholdings in BM3D thresholding Eqs. (49)-(50).

For the results given by the C-PI algorithm with BM3D and observation filtering these parameters have been selected for the heavy noise case, $\chi = .02$, with the following values: $\gamma_1/\gamma_o = .25$, $\gamma_1/\gamma_p = 0$, $\gamma_1\chi = 10\chi$ and $\tau_{a,o} = .02$, $\tau_{\varphi,o} = 1.0$. These parameters for filtering are unchanged for all other values of $\chi \in [.01, 10000]$. The curves for C-PI with filtering in Figs 2 and 4 are

obtained for these fixed values of the parameters.

The observation filtering is started from the iteration 10. The BM3D filtering starts from iteration 25. At this iterations BM3D frames (transforms Φ and Ψ) are calculated for the corresponding reconstructions of the modulus and phase. These frames are fixed up to iteration 50, where they are updated accordingly to the current reconstructions of the modulus and phase. The third and the last update is produced for iteration 75.

Further improvement of reconstructions can be achieved using different strategies of BM3D filtering and selection of various thresholds $\tau_{a,o}$, $\tau_{\varphi,o}$ and the algorithm parameters for different noise level and different experiment parameters.

The visual results for reconstruction of the modulus and the phase for very noisy data $\chi = .01$ (photon count 3×10^4) are shown in Figs. 5 and 6. In these figures, the first row shows the reconstruction images of the modulus and the phase from left-to-right. In the second row one can see the cross-sections of these reconstructions where also the true values of the modulus and the phase are given. The reconstructions shown in Fig. 5 are obtained by e-PIE. Both the magnitude and phase images are damaged severely by noise

and diffraction artifacts. The e-PIE algorithm fails to separate the modulus (lena) and phase (chessboard) images. We can see the chessboard squares in the modulus reconstruction as well as some features of lena's hat in the phase reconstruction. The imaging obtained by the Basic C-PI (without filtering) looks quite similar (images are not shown) despite of some numerical advantage shown by lower values of RMSEs for the modulus and phase.

The reconstruction by the C-PI algorithm equipped by both BM3D and observation filtering is shown Fig. 6. The advantage over the data shown in Fig. 5 is obvious. The modulus and phase information are perfectly separated, noise and diffraction effects are suppressed.

In Fig. 7 we show the nearly perfect reconstructions obtained by C-PI for the low noise level, $\chi = 10000$. These results are obtained by using the above BM3D filtering where the threshold $\tau_{a,o} = .02$ is replaced for $\tau_{a,o} = .01$, and the observation filtering is dropped. For comparison in Fig. 8 we show also very good reconstructions given by the e-PIE algorithm. However, RMSE values are slightly worse than those for C-PI in Fig. 7, and visually the phase reconstruction is not so perfect. One can see slight traces of lena's hat in the phase image and the cross-section of the phase is not smooth.

Finally, in Fig. 9 we show the probe reconstruction obtained by the C-PI algorithm with BM3D filtering for $\chi = 10000$ after 200 iterations and the initial guess $P = 1$. This probe reconstruction is used as the initial guess for the all results discussed above.

7. Conclusion

A novel iterative algorithm targeted on optimal processing noisy measurements is proposed for ptychographical reconstruction of complex-valued object and probe. The noise suppression is enabled by two instruments. First, by the maximum likelihood technique formulated for Poissonian (photon counting) measurements, and second, by sparse approximation of phase and magnitude of the object and probe. It is shown, in particular, that for noisy data the maximum likelihood estimate of the wavefield at the sensor plane is essentially different from the famous Gerchberg-Saxton-Fienup solution, where the magnitude of the estimate is replaced by the square root of the intensity measurement. The simulation experiments demonstrate the state-of-the-art performance of the developed algorithm both numerically and visually.

The proposed approach to the algorithm design is quite flexible and can be extended for various developments, in particular, such as a compressive

super-resolution ptychographical imaging and compensation of unknown object positions. Our experiments show that using a phase SLM as a diffuser as it is shown in Fig. 1 allows significantly improve the reconstruction. However, the problem of a proper selection of this phase mask is out of scope of this paper.

Appendix: Derivation of the C-PI algorithm

Minimization of \mathcal{L}_1 on \mathbf{o} (problem 40)

$$\mathcal{L}_1(\mathbf{o}, \mathbf{p}, \mathbf{w}, \Theta) = \frac{\chi}{2} \sum_{j=1}^J \|(\hat{\mathbf{I}}_j - |\mathbf{w}_j|^2) ./ |\mathbf{w}_j|\|^2 + \quad (59)$$

$$\frac{1}{\gamma_1} \sum_{j=1}^J \|\mathbf{u}_j - \mathbf{w}_j\|^2 + \frac{1}{\gamma_o} \|\mathbf{o} - \mathbf{v}_o\|_2^2 + \frac{1}{\gamma_p} \|\mathbf{p} - \mathbf{v}_p\|_2^2,$$

$$\mathbf{v}_o = (\Psi_{a,o} \boldsymbol{\theta}_{a,o}) \circ \exp(j\Psi_{\varphi,o} \boldsymbol{\theta}_{\varphi,o}), \quad \mathbf{v}_p = (\Psi_{a,p} \boldsymbol{\theta}_{a,p}) \circ \exp(j\Psi_{\varphi,p} \boldsymbol{\theta}_{\varphi,p}).$$

A necessary minimum condition with respect to complex-valued \mathbf{o} has a form:

$$\partial \mathcal{L}_1 / \partial \mathbf{o}^* = \frac{1}{\gamma_o} (\mathbf{o} - \mathbf{v}_o) + \frac{1}{\gamma_1} \sum_{j=1}^J \frac{\partial \mathbf{u}_j^H}{\partial \mathbf{o}^*} (\mathbf{u}_j - \mathbf{w}_j) = 0. \quad (60)$$

Using the formula (10) we obtain

$$\partial \mathcal{L}_1 / \partial \mathbf{o}^* = \frac{1}{\gamma_o} (\mathbf{o} - \mathbf{v}_o) + \frac{1}{\gamma_1} \sum_{j=1}^J \frac{\partial (\mathbf{A} \cdot \text{diag}(\mathbf{p}) \mathbf{G}_j^o \mathbf{o})^H}{\partial \mathbf{o}^*} (\mathbf{u}_j - \mathbf{w}_j) = \quad (61)$$

$$\frac{1}{\gamma_o} (\mathbf{o} - \mathbf{v}_o) + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{A} \cdot \text{diag}(\mathbf{p}) \mathbf{G}_j^o)^H (\mathbf{A} \cdot \text{diag}(\mathbf{p}) \mathbf{G}_j^o \mathbf{o} - \mathbf{w}_j) = 0.$$

It follows

$$\left[\frac{1}{\gamma_o} \mathbf{I}_{n_o \times n_o} + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \text{diag}(\mathbf{p}^*) \mathbf{A}^H \mathbf{A} \cdot \text{diag}(\mathbf{p}) \mathbf{G}_j^o \right] \mathbf{o} = \quad (62)$$

$$\frac{1}{\gamma_o} \mathbf{v}_o + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \cdot \text{diag}(\mathbf{p}^H) \mathbf{A}^H \mathbf{w}_j.$$

For the matrix \mathbf{A} corresponding to DFT $\mathbf{A}^H \mathbf{A} = \mathbf{I}_{n_s \times n_s}$ and the linear equation for \mathbf{o} is of the form:

$$\left[\frac{1}{\gamma_o} \mathbf{I}_{n_o \times n_o} + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \text{diag}(|\mathbf{p}|^2) \mathbf{G}_j^o \right] \mathbf{o} = \frac{1}{\gamma_o} \mathbf{v}_o + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \cdot \text{diag}(\mathbf{p}^H) \mathbf{A}^H \mathbf{w}_j. \quad (63)$$

It is easily to check that $\frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \text{diag}(|p|^2) \mathbf{G}_j^o$ is a diagonal matrix, then Eq.(63) can be easily solved by calculating the elements of the vector \mathbf{o} independently by inverting the diagonal elements of the matrix $\left[\frac{1}{\gamma_o} \mathbf{I}_{n_o \times n_o} + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{G}_j^o)^T \text{diag}(|p|^2) \mathbf{G}_j^o \right]$. This solution can be presented in the image domain by going back from the vectorized to the image domain variables. It can be verified that this passage gives the solution of the problem (40) in the form (44).

Minimization of \mathcal{L}_1 on p (problem (41))

Similar transformations starting from the minimum condition $\partial \mathcal{L}_1 / \partial p^* = 0$ give the linear equation for p :

$$\partial \mathcal{L}_1 / \partial \mathbf{p}^* = \frac{1}{\gamma_p} (\mathbf{p} - \mathbf{v}_p) + \frac{1}{\gamma_1} \sum_{j=1}^J \frac{\partial \mathbf{u}_j^H}{\partial \mathbf{p}^*} (\mathbf{u}_j - \mathbf{w}_j) = 0.$$

Using Eq.(11) it gives

$$\begin{aligned}
\partial\mathcal{L}_1/\partial\mathbf{p}^* &= \frac{1}{\gamma_p}(\mathbf{p} - \mathbf{v}_p) + \frac{1}{\gamma_1} \sum_{j=1}^J \frac{\partial(\mathbf{A}\cdot\text{diag}(\mathbf{G}_j^o\mathbf{o})\mathbf{p})^H}{\partial\mathbf{p}^*}(\mathbf{u}_j - \mathbf{w}_j) = \\
&\frac{1}{\gamma_p}(\mathbf{p} - \mathbf{v}_p) + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{A}\cdot\text{diag}(\mathbf{G}_j^o\mathbf{o}))^H (\mathbf{A}\cdot\text{diag}(\mathbf{G}_j^o\mathbf{o})\mathbf{p} - \mathbf{w}_j) \implies \\
&\left[\frac{1}{\gamma_p} \mathbf{I}_{n_p \times n_p} + \frac{1}{\gamma_1} \sum_{j=1}^J \text{diag}(\mathbf{G}_j^o\mathbf{o}^H) \mathbf{A}^H \mathbf{A} \text{diag}(\mathbf{G}_j^o\mathbf{o}) \right] \mathbf{p} = \\
&\frac{1}{\gamma_p} \mathbf{v}_p + \frac{1}{\gamma_1} \sum_{j=1}^J (\mathbf{A}\cdot\text{diag}(\mathbf{G}_j^o\mathbf{o}))^H \mathbf{w}_j.
\end{aligned}$$

For $\mathbf{A}^H \mathbf{A} = \mathbf{I}_{n_p \times n_p}$ we obtain

$$\left[\frac{1}{\gamma_p} \mathbf{I}_{n_p \times n_p} + \frac{1}{\gamma_1} \sum_{j=1}^J \text{diag}(\mathbf{G}_j^o|\mathbf{o}|^2) \right] \mathbf{p} = \frac{1}{\gamma_p} \mathbf{v}_p + \frac{1}{\gamma_1} \sum_{j=1}^J \text{diag}(\mathbf{G}_j^o\mathbf{o}^H) \mathbf{A}^H \mathbf{w}_j. \quad (64)$$

The matrix in the left-hand side of this equation is diagonal. Thus, the solution of this equation is very simple and can be given in the image domain variables in the form of Eq. (45).

Minimization of \mathcal{L}_1 on $\mathbf{w}_j \in \mathbb{C}^{M_s N_s}$ (problem (42))

The only part of \mathcal{L}_1 depending on \mathbf{w}_j is

$$\tilde{\mathcal{L}}_1 = \frac{\chi}{2} \sum_{j=1}^J \|(\hat{\mathbf{I}}_j - |\mathbf{w}_j|^2) ./ |\mathbf{w}_j|\|^2 + \frac{1}{\gamma_1} \sum_{j=1}^J \|\mathbf{u}_j - \mathbf{w}_j\|^2 \quad (65)$$

This criterion is separable with respect to the components of the vector

necessary \mathbf{w}_j . Thus, the problem is reduced to multiple scalar nonquadratic optimization problems

$$\min_{\mathbf{w}_j[l]} \frac{\chi}{2} (\hat{\mathbf{I}}_j[l] - |\mathbf{w}_j[l]|^2)^2 ./ |\mathbf{w}_j[l]|^2 + \frac{1}{\gamma_1} |\mathbf{u}_j[l] - \mathbf{w}_j[l]|^2, \quad (66)$$

where $\mathbf{w}_j[l]$ is complex-valued. The only term of this criterion depending on the phase of $\mathbf{w}_j[l]$ is the second one. It is easy to conclude that the minimum with respect to the phase angle is achieved for

$$\text{angle}(\hat{\mathbf{w}}_j[l]) = \text{angle}(\mathbf{u}_j[l]). \quad (67)$$

Then the criterion (66) takes the form

$$\frac{\chi}{2} (\hat{\mathbf{I}}_j[l] - |\mathbf{w}_j[l]|^2)^2 ./ |\mathbf{w}_j[l]|^2 + \frac{1}{\gamma_1} \left| |\mathbf{u}_j[l]| - |\mathbf{w}_j[l]| \right|^2, \quad (68)$$

and depends on the magnitude of $\mathbf{w}_j[l]$ only. This optimal magnitude is a solution of the real valued problem

$$|\hat{\mathbf{w}}_j[l]| = \arg \min_{|\mathbf{w}_j[l]|} \frac{\chi}{2} (\hat{\mathbf{I}}_j[l] - |\mathbf{w}_j[l]|^2)^2 ./ |\mathbf{w}_j[l]|^2 + \frac{1}{\gamma_1} \left| |\mathbf{u}_j[l]| - |\mathbf{w}_j[l]| \right|^2. \quad (69)$$

Then the optimal solution of minimization of (65) with respect to the complex valued \mathbf{w}_j takes the form

$$\hat{\mathbf{w}}_j[l] = |\hat{\mathbf{w}}_j[l]| \exp(j \text{angle}(\mathbf{u}_j[l])), \quad (70)$$

where the magnitude $|\hat{\mathbf{w}}_j[l]|$ is found numerically from Eq.(69). It can be shown that this solution belongs to the interval $[\sqrt{\hat{\mathbf{I}}_j[l]}, |\mathbf{u}_j[l]|]$.

For brevity we denote the nonlinear operator giving the solution (70) as

$$w_j[l] = \mathcal{G}(\hat{I}_j[l], u_j[l]). \quad (71)$$

The following comment is of importance. It can be seen easily that $|w_j[l]| \rightarrow \sqrt{\hat{I}_j[l]}$ for larger $\gamma_1\chi$. Then

$$w_j[l] = \mathcal{G}(\hat{I}_j[l], u_j[l]) = \frac{u_j[l]}{|u_j[l]|} \sqrt{\hat{I}_j[l]}. \quad (72)$$

Thus, we arrive to the well known and popular rule commonly used in the phase retrieval algorithms: the variable $\mathbf{w}_j[l]$ has the absolute value equal to the square root of observation and the phase predicted by the wavefield model. Of course, it is not an optimal solution of $\min_{\mathbf{w}_j} \mathcal{L}_1$ and but it is close to the optimal one if the noise level in the observations is small, it has a place when $\gamma_1\chi \rightarrow \infty$.

Minimization of \mathcal{L}_2 on Θ (problem (43))

The objective function \mathcal{L}_2 is additive with respect to the elements of the

spectral vectors Θ . The optimization with respect to these variables can be solved independently for each one. This optimization is reduced to consideration the following scalar problem:

$$\boldsymbol{\theta}_i = \arg \min_{\boldsymbol{\theta}_i} \tau \cdot \|\boldsymbol{\theta}_i\|_p + \frac{1}{2} (\boldsymbol{\theta}_i - \mathbf{B}_i)^2. \quad (73)$$

There is an analytical solution for (73) known as the hard- and soft thresholding (or hard- and soft shrinkage) operators, respectively for l_0 and l_1 .

Let us denote this operator as

$$\boldsymbol{\theta} = \mathfrak{Th}_\tau(\mathbf{B}), \quad (74)$$

where \mathbf{B} is an input-vector and $\boldsymbol{\theta}$ is a solution-vector returned by the operator.

For l_0 and l_1 this elementwise operator is specified as (e.g. [25])

$$\boldsymbol{\theta} = \mathfrak{Th}_\tau(\mathbf{B}) = \quad (75)$$

$$\begin{cases} \mathfrak{Th}_{\sqrt{2\tau}}^{hard}(\mathbf{B}) = \mathbf{B} \circ 1(|\mathbf{B}| \geq \sqrt{2\tau}), \text{ if } l_p = l_0, \\ \mathfrak{Th}_\tau^{soft}(\mathbf{B}) = \text{sign}(\mathbf{B}) \circ \max(|\mathbf{B}| - \tau, 0), \text{ if } l_p = l_1, \end{cases}$$

where the indexes '*soft*' and '*hard*' indicate the type of the solution. The threshold parameters for the soft- and hard thresholdings are different.

With this notation the solution for the problem (34) is of the form (49)-(50).

A. Acknowledgments

This work was supported by the Academy of Finland: project no. 213462, 2006-2011 (Finnish Programme for Centres of Excellence in Research) and project no. 138207, 2011-2014.

References

1. W. Hoppe, “ Trace structure analysis, ptychography, phase tomography, ” *Ultramicroscopy* **10**, 187 – 198 (1982).
2. J. M. Rodenburg, “Ptychography and related diffractive imaging methods,” in *Advances in Imaging and Electron Physics*, P.W. Hawkes, ed. (Elsevier, 2008), **150**, pp. 87–184.
3. J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, “Hard-x-ray lensless imaging of extended objects,” *Phys. Rev. Lett.* **98**, 034801 (2007).
4. M. Dierolf, A. Menzel, P. Thibault, P. Schneider, C. M. Kewish, R. Wepf, O. Bunk, and F. Pfeiffer, “Ptychographic x-ray computed tomography at the nanoscale,” *Nature* **467**, 436–439 (2010).
5. Hüe, F., Rodenburg, J.M., Maiden, A.M., Sweeney, F., and Midgley, P.A. ‘Wave-front phase retrieval in transmission electron microscopy via ptychography,’ *Physical Review B* **82** ,Article Number: 121415 (2010).
6. J. M. Rodenburg, A. C. Hurst, and A. G. Cullis, “Transmission microscopy without lenses for objects of unlimited size,” *Ultramicroscopy* **107**, 227–231 (2007).
7. A. Schropp, P. Boye, A. Goldschmidt, S. Hönig, R. Hoppe, J. Patommel, C. Rakete, D.

- Samberg, S. Stephan, S. Schöder, M. Burghammer, and C. G. Schroer, “Non-destructive and quantitative imaging of a nano-structured microchip by ptychographic hard x-ray scanning microscopy,” *J. Microsc.* **241**, 9–12 (2011).
8. J.M.Rodenburg,H.M.L.Faulkner, "A phase retrieval algorithm for shifting illumination," *Appl. Phys. Lett.* **85** (20) 4795–4797 (2004).
 9. A. M. Maiden and J. M. Rodenburg, “An improved ptychographical phase retrieval algorithm for diffractive imaging,” *Ultramicroscopy* **109**, 1256–1262 (2009).
 10. P. Thibault, M. Dierolf, O. Bunk, A. Menzel, F. Pfeiffer, “Probe retrieval in ptychographic coherent diffractive imaging,” *Ultramicroscopy*, **109**, 338–343 (2009).
 11. A. M. Maiden, M. J. Humphry, F. Zhang and J. M. Rodenburg, “Superresolution imaging via ptychography,” *J. Opt. Soc. Am. A* **28**, 604-612 (2011).
 12. J.R.Fienup, “Phase retrieval algorithms: a comparison,” *Appl. Opt.* **21** (15), 2758–2769 (1982).
 13. M.Guizar-Sicairos,J.R.Fienup, “Phase retrieval with transfer set translation diversity: a nonlinear optimization approach,” *Opt.Express* **16** (10), 7264–7278 (2008).
 14. V.Elser, Phase retrieval by iterated projections, *J. Opt. Soc. Am. A* **20**(1), 40–55 (2003).
 15. Y.Censor, N.Cohen, T.Kotzer, J.Shamir, Summed squared distance error reduction by simultaneous multiprojections and applications, *Appl. Math. Comput.* **126**, 157–179 (2002).
 16. P Thibault and M. Guizar-Sicairos, “Maximum-likelihood refinement for coherent diffractive imaging,” *New Journal of Physics* **14**, 063004 (20 pp) (2012).
 17. E. J. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: exact signal

- reconstruction from highly incomplete frequency information,” *IEEE Trans. Inf. Theory* **52**, 489–509 (2006).
18. D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
 19. S. Gazit, A. Szameit, Y. C. Eldar, M. Segev, "Super-resolution and reconstruction of sparse sub-wavelength images, " *Optics Express* **17**, pp. 23920-23946 (2009).
 20. V. Katkovnik and J. Astola, "High-accuracy wavefield reconstruction: decoupled inverse imaging with sparse modeling of phase and amplitude," *JOSA A* **29**, pp. 44–54 (2012).
 21. V. Katkovnik and J. Astola, “Phase retrieval via spatial light modulator phase modulation in 4f optical setup: numerical inverse imaging with sparse regularization for phase and amplitude,” *J. Opt. Soc. Am. A* **29**, 105-116 (2012).
 22. V. Katkovnik and J. Astola, “Compressive sensing computational ghost imaging,” *J. Opt. Soc. Am. A*, **29**, 1556–1567 (2012).
 23. A. Papoulis, *Probability, random variables, and stochastic processes* (New York: McGraw-Hill, 1991).
 24. D. Han, K. Kornelson, D. Larson, E. Weber, *Frames for undergraduates* (Student Mathematical Library, AMS, 2007).
 25. M. Elad, *Sparse and redundant representations: from theory to applications in signal and image processing*, Springer (2010).
 26. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3D transform-domain collaborative filtering,” *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
 27. A. Danielyan, Katkovnik, V. and Egiazarian, K., "BM3D frames and variational image

deblurring," *IEEE Trans. Image Processing* **21** (2012).

28. F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *A Quarterly Journal of Operations Research* **5**, 173–210 (2007).

List of Figure Captions

Figure 1. Schematic of ptychographical imaging.

Figure 2. E_0 criteria calculated for object (solid) and probe (dashed) versus the photon count. Results for three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best fitting object and probe complex-valued functions is enabled by C-PI algorithm with BM3D and observation filtering.

Figure 3. $RMSE$ calculated for modulus (solid) and phase (dashed) of object (solid). Three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best accuracy is achieved by C-PI algorithm with BM3D and observation filtering.

Figure 4. E_0 criteria calculated for object versus iteration number. Three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best convergence rate is obtained by C-PI algorithm with BM3D and observation filtering.

Figure 5. Complex object reconstruction. Very noisy observations, $\chi = .01$. e-PIE algorithm: First row: images of modulus and phase reconstructions;

second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color).

Figure 6. Complex object reconstruction. Very noisy observations, $\chi = .01$. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color).

Figure 7. Complex object reconstruction. Noiseless observations, $\chi = 10000$. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color). Visually perfect reconstruction.

Figure 8. Complex object reconstruction. Noiseless observations, $\chi = 10000$. e-PIE algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color). Phase reconstruction is a bit corrupted by noise as well as some features of lena's image.

Figure 9. Complex probe reconstruction used for initialization of all algorithms. Noiseless observations, $\chi = 10000$, 200 iterations. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-

sections of the true image (solid, red in color) and the reconstruction (dash, blur in color).

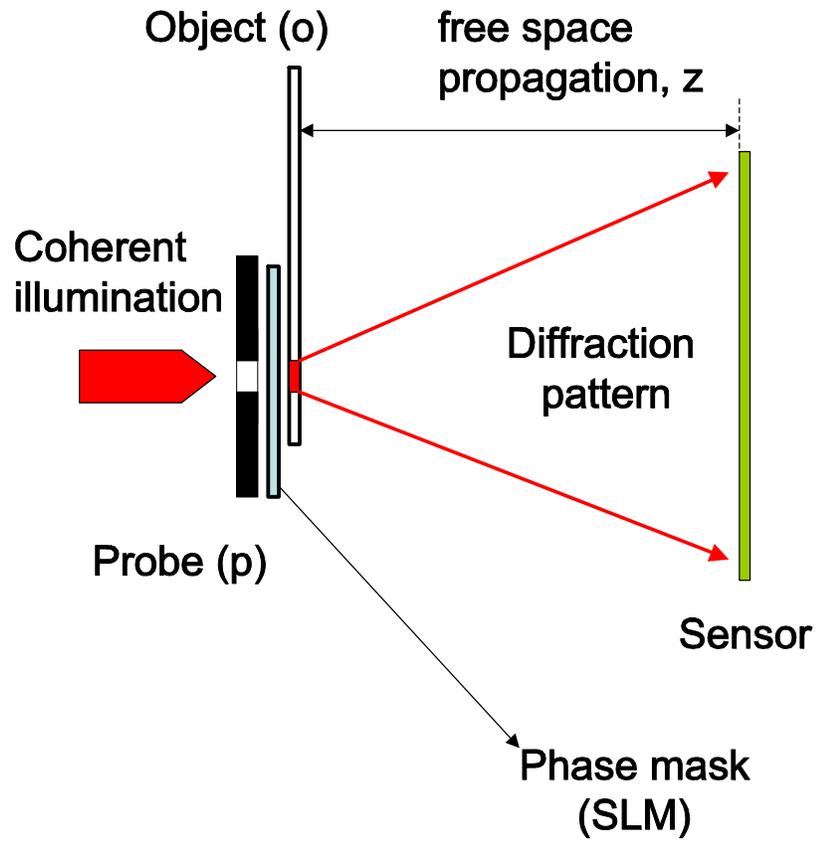


Fig. 1. Schematic of ptychographical imaging.

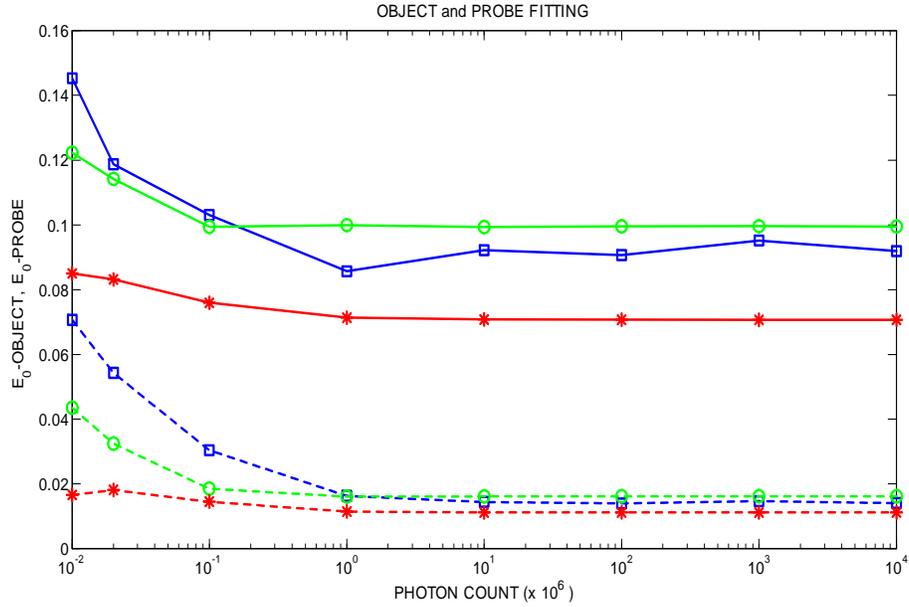


Fig. 2. E_0 criteria calculated for object (solid) and probe (dashed) versus the photon count. Results for three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best fitting object and probe complex-valued functions is enabled by C-PI algorithm with BM3D and observation filtering.

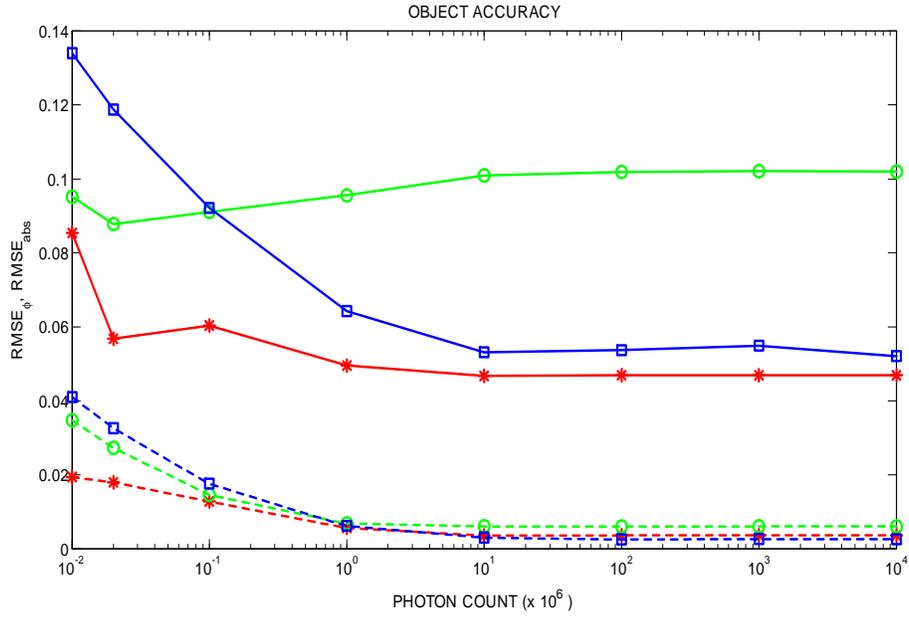


Fig. 3. *RMSE* calculated for modulus (solid) and phase (dashed) of object (solid). Three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best accuracy is achieved by C-PI algorithm with BM3D and observation filtering.

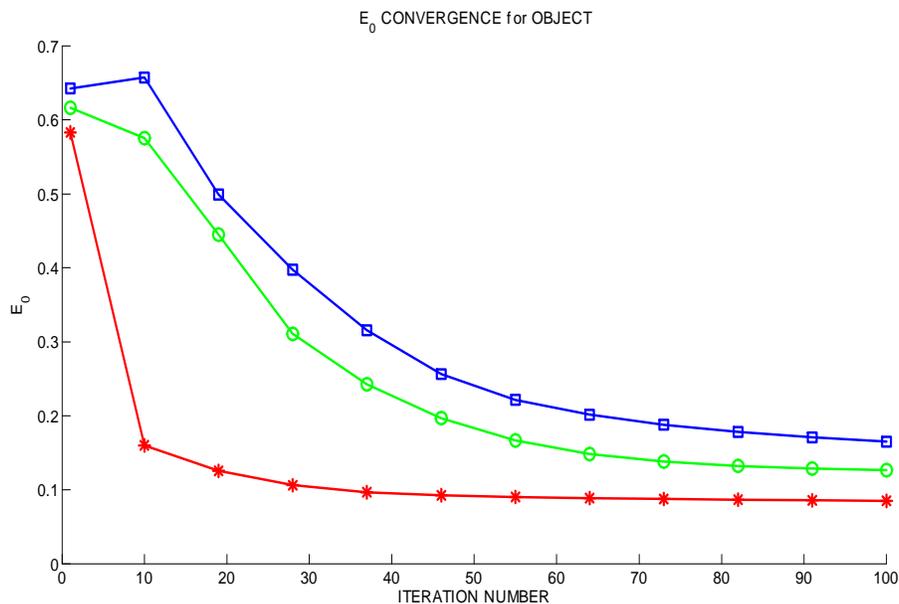


Fig. 4. E_0 criteria calculated for object versus iteration number. Three algorithms are compared: e-PIE ('squares', blue in color); basic (no filtering) C-PI ('o', green in color); C-PI ('stars', red in color). The best convergence rate is obtained by C-PI algorithm with BM3D and observation filtering.

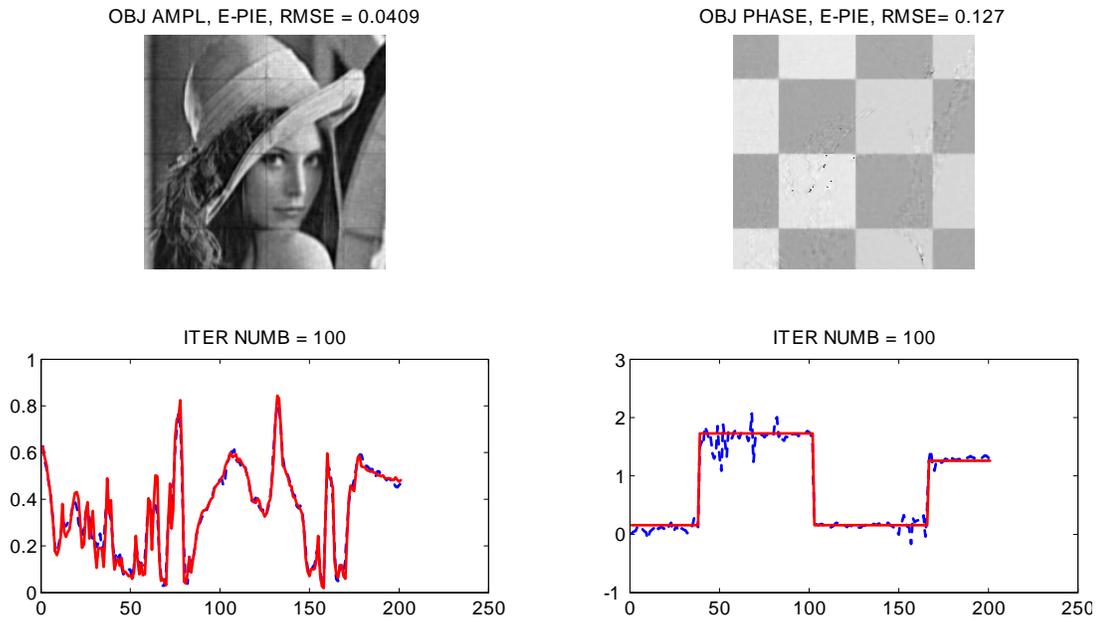


Fig. 5. Complex object reconstruction. Very noisy observations, $\chi = .01$. e-PIE algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color).

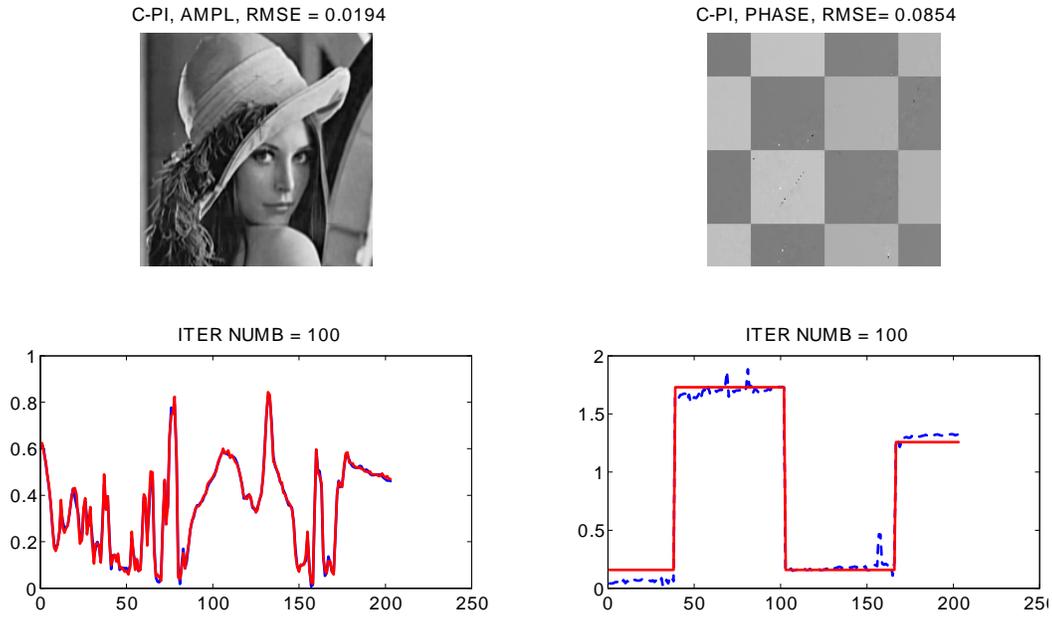


Fig. 6. Complex object reconstruction. Very noisy observations, $\chi = .01$. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color).

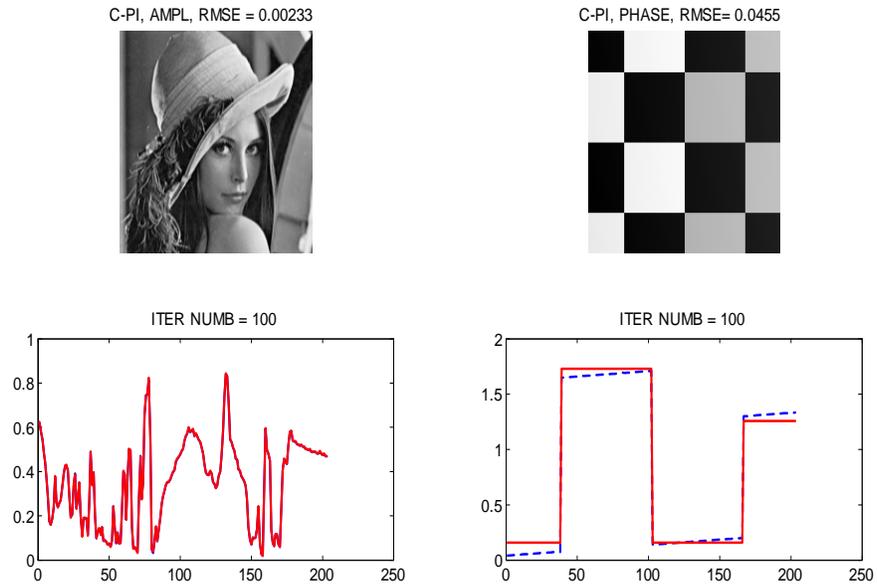


Fig. 7. Complex object reconstruction. Noiseless observations, $\chi = 10000$. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color). Visually perfect reconstruction.

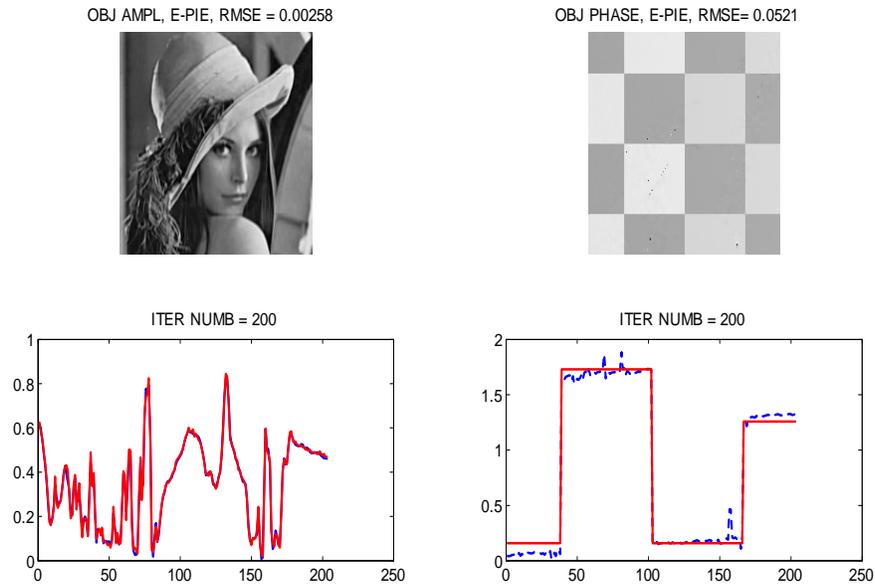


Fig. 8. Complex object reconstruction. Noiseless observations, $\chi = 10000$. e-PIE algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color). Phase reconstruction is a bit corrupted by noise as well as some features of lena's image.

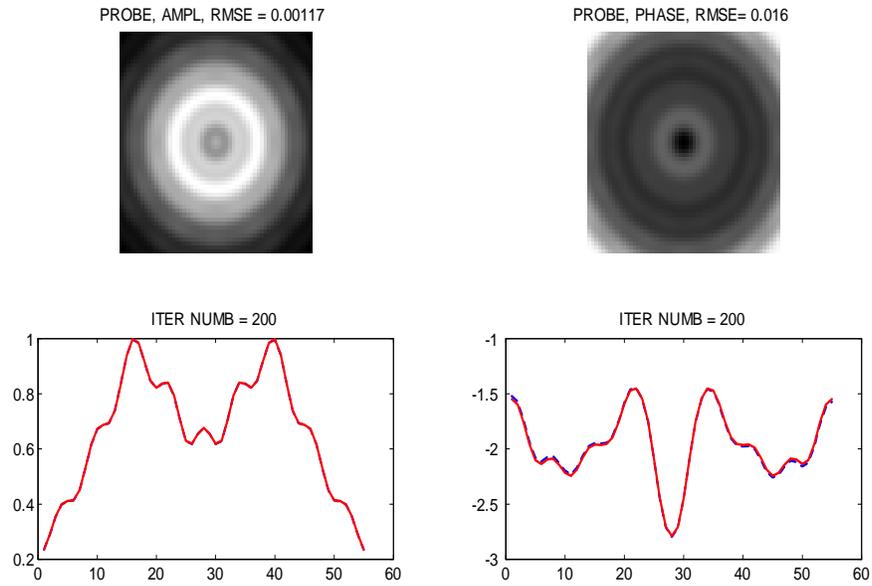


Fig. 9. Complex probe reconstruction used for initialization of all algorithms. Noiseless observations, $\chi = 10000$, 200 iterations. C-PI algorithm: First row: images of modulus and phase reconstructions; second row: cross-sections of the true image (solid, red in color) and the reconstruction (dash, blue in color).