Sparse phase retrieval from noisy data

Vladimir Katkovnik
Department of Signal Processing, Technology University of Tampere,
Tampere, Finland. E-mail: vladimir.katkovnik@tut.fi

A variational approach to object phase and amplitude reconstruction from multiple noisy Poissonian intensity observations is developed for the typical phase retrieval scenario. Sparse modeling of amplitude and absolute phase of the object is one of the key elements of the derived SPAR algorithm. The efficiency of this algorithm is demonstrated by simulation experiments for the coded diffraction pattern scenario. The comparison is produced versus the truncation Wirtingling flow (TWF) algorithm (Y. Chen and E. J. Candès, 2015, http://statweb.stanford.edu/~candes/papers/TruncatedWF.pdf), which is the state-of-the-art in the field. For noisy observations SPAR demonstrates a definite advantage over TWF. For the low noise level the performance of SPAR as well as its simplified version the GS algorithm, where the sparse modeling of the object is omitted, is nearly identical to the performance of TWF. The GS algorithm is faster than TWF while SPAR computationally much more demanding is slower than both TWF and GS. © 2015 Optical Society of America

OCIS codes: 070.2025, 100.3010, 100.3190, 100.5070

1. Introduction

1.1. Phase retrieval formulation

Transparent specimens, for instance biological cells and some tissues, do not change the intensity of passing light but introduce phase delays caused by variations in thickness, density and refractive indices. Visualization of phase variations by transforming them in light intensity is one of the important problems in optics. The revolutionary phase contrast imaging (Frits Zernike 1930s, Nobel prize 1953) solves the problem by introducing a modulation of the wavefront in the focal (Fourier) plane behind the first principal lens. In this way qualitative phase visualization is achieved where the observed light intensity is linked with phase variations, however, there is no proportional relation of the light intensity with the phase and contrast inversions may disrupt a proper phase interpretation. Despite this drawback the phase contrast microscopy is one of the most frequently applied optical methods in medical and biological research.
Quantitative phase visualization is targeted on precise phase imaging. In the modern development it is fundamentally based on digital data processing. One of the popular mathematical formulations and computational techniques relevant to the quantitative phase visualization is phase retrieval.

The phase retrieval problem is formulated as finding a complex-valued vector $\mathbf{x} \in \mathbb{C}^n$ from real-valued observations $\mathbf{y}_s \in \mathbb{R}^m$:

$$\mathbf{y}_s = |\mathbf{A}_s \mathbf{x}|^2, \ s = 1, ..., L. \quad (1)$$

In terms of the coherent diffractive imaging the model (1) allows the following interpretation. Provided the unit intensity of a laser beam the vector $\mathbf{x}$ is an object (specimen) transfer function and the complex-valued wavefront just behind the object (see Fig.1a); $\mathbf{A}_s \in \mathbb{C}^{m \times n}$ is an $m \times n$ matrix of the wavefront propagation operator from the object to the sensor plane and the vector $\mathbf{y}_s$ is an intensity of the wavefront registered by the sensor. The squared absolute value in (1) is an element-wise operation. Thus, the items of the vector $\mathbf{y}_s$ are squared absolute values of the corresponding items of the vector $\mathbf{A}_s \mathbf{x} \in \mathbb{C}^m$. The model (1) corresponds to a multiple observation scenario of $L$ experiments.

![Fig. 1. Examples of optical setups for lensless (a) and 4f lens (b) scenarios for phase retrieval.](image)

Imaging that phases of the complex-valued $\mathbf{A}_s \mathbf{x}$ denoted as $\mathbf{u}_s$ are known then the quadratic
equations (1) can be replaced by the linear ones $A_s x = u_s$ and finding of $x$ is reduced to the linear algebra problem. Thus, these phases of $A_s x$ killed by the modulus in (1) transform the linear problem in the much more complex quadratic one.

Conventionally, the term *phase retrieval problem* is addressed to reconstruction of the missing phases of the vectors $A_s x$. However in Eqs.(1) the phase of the complex-valued object $x \in \mathbb{C}^n$ is also unknown and actually reconstruction (retrieval) of this phase is the main goal of the problem at hand.

In this paper we refocus the phase retrieval problem by treating the missed phases of $A_s x$ as auxiliary variables and the phase and the amplitude of the object $x$ as the main unknowns of interest. Accordingly, the sparse modeling mentioned in the title of this paper is applied to both the phase and the amplitude of $x$. We show that this sparse modeling applied to $x$ leads to an efficient algorithm solving the problem.

In this interpretation the phase retrieval becomes the quantitative phase imaging problem with the real-valued observations $y_s \in \mathbb{R}^m$ given by Eqs.(1).

The paper is focussed on 2D imaging. The vectors and matrixes in the above equations correspond to the vectorized representations conventionally applied for this kind of the 2D problems.

Design of the image formation operators $A_s$ in Eqs.(1) is a crucial moment in order to gain observation diversity sufficient for reliable reconstruction of both object phase as well as object amplitude. Pragmatically, this diversity means that the set $\{A_s\}_{s=1}^{L}$ consists of the operators different in such way that the observations $y_s$ are very different for different $s$ while being obtained for the same $x$. Image defocussing is one of the popular ways to get this kind of the diverse observations. First results based on the defocus approach are demonstrated in 1973 by Misell [1] and in 1980 by Saxton [2] for two defocussed images obtain using optical lens.

In the coherent lensless imaging (Fig.1a) the laser beam goes through the object and after free-space propagation the intensity of the diffracted wavefield is registered by the sensor array. The corresponding operators $A_s$ depends on the object-sensor distance $d_s$ and the wavelength $\lambda_s$, $A_s = \mathcal{A}(d_s, \lambda_s)$, where the operator $\mathcal{A}$ can be modelled by the rigorous Rayleigh-Sommerfeld integral or by the Fresnel and Franhofer approximations of this integral (Goodman [3]). The latter modeling is equivalent to the discrete Fourier transform $F_s$ for the operator $A_s$ because the phase term of the Franhofer transform is cancelled by the modulus operation in Eq.(1).

Experiments for a set of distances $d_s$ (displaced sensor planes) is a popular mean to get the sufficient observation diversity. This approach is developed and studied by Pedrini, et al. [4], Almor, et al. [5] and [6]. In the recent development, in particular we refer to Kohler, et al. [7] and Camacho, et al. [8], a spatial light modulator (SLM) is exploited in order to get a set of differently defocussed images.

A phase modulation of the wavefront near the lens plane is a popular and general tool
to get diverse observations. The phase modulation at the object plane plus the Fraunhofer wavefront propagation result in the observation model known as a coded diffraction pattern:

$$y_s = |FD_s x|^2, \ s = 1, \ldots, L,$$

where \( F \in \mathbb{C}^{n \times n} \) denotes the Fourier transform and \( D_s \in \mathbb{C}^{n \times n} \) is the diagonal matrix of complex exponents, \( D_s = \text{diag}\{\exp(j\phi_1(s)), \exp(j\phi_2(s)), \ldots, \exp(j\phi_n(s))\} \).

This phase modulation can be implemented by a special phase mask inserted just behind the object plane in the setup shown in Fig.1a.

The phase \( \phi_k(s) \) in \( D_s \) can be generated as random. Let \( \phi_k(s) \) be i.i.d. zero mean Gaussian, \( \phi_k(s) \sim N(0, \sigma) \). Then, \( d_{ks} = \exp(j\phi_k(s)) \) are random elements of \( D_s \), such that \( E\{d_{ks}\} = 0 \) and the matrices \( A_s = FD_s \) are zero-mean random. This random phase modulation changes the spectrum of \( FD_s x \) in a radical way extending a distribution of the intensity from low to high frequency components. The independence of \( \phi_k(s) \) for all \( k \) and \( s \) enables a strong diversity of observations in (2).

The phase modulation (coded aperture imaging) is applied in various optical setups. For instance, consider the coherent two lens 4f optical system shown in Fig.1b, where \( f \) is a focal length of the lenses. The wavefront in the focal plane behind the first lens is the Fourier transform of the wavefront going through the object [3]. Further, the wavefront at the sensor plane, which locates at the focal length of the second lens, is the Fourier transform of the wavefront in the Fourier plane of the first lens. The phase modulation is produced using phase mask or SLM in the focal (Fourier) plane of the first lens. The intensity registered by the sensor in this system can be represented as

$$y_s = |FD_s Fx|^2, \ s = 1, \ldots, L.$$  

It is demonstrated in Falldorf, et al. [9] that the phases \( \phi_k(s) \) in \( D_s \) can be selected in such way that the distributions of the wavefront at the sensor plane imitates desired displacements \( \{d_s\} \) of the sensor plane with respect to the object plane. The phase retrieval for phase imaging in the 4f optical setup with the random phase modulation in the Fourier plane is studied in Katkovnik, et al. [10].

Developments of the phase modulation techniques with various applications can be seen in the book by Glückstad and Palima [11].

1.B. Phase retrieval algorithms

Let us start from the popular Gerchberg-Saxton (GS) algorithms (Gerchberg and Saxton [12], Fienup [13]). These iterative algorithms are based on alternating projections between the object plane with a complex-valued \( x \) and the Fourier (diffraction) plane \( Ax \) with a given (measured) amplitude \( z \). At the Fourier plane the amplitudes in the vectors \( Ax \) are replaced by the corresponding items of \( z \). The backprojection of this result to the object plane is modified according to the prior information on the object, e.g. support size and shape,
amplitude value of \( x \), etc. The GS algorithms exist in various modifications. The review and analysis of these algorithms as well as further developments can be seen in the recent paper by Guo, et al. [14].

The algorithms, known as a single-beam multiple-intensity reconstruction (SBMIR), are targeted on reconstruction of 3D wavefield covered by 2D intensity measurement planes (e.g. Ivanov, et al. [15], Pedrini et al. [4], [16] and Almoro, et al. [6], [5]). An existence of a single unknown object-source of radiation \( x \) is not assumed. The SBMIR algorithm starts from an initial guess of the complex amplitude at the first measurement plane. Then, this initial guess propagates numerically forward from the one measurement plane to the next following one successively through the all sequence of the recordings. At each plane the calculated modulus of the wave field is replaced by the square root of the intensity measured for this plane according to the GS algorithm. When the last measurement plane is reached the wave field estimate at this plane is propagated back to the first plane. This iterative process is repeated until convergence.

Contrary to the intuitively clear heuristic GS algorithms the variational approaches to phase retrieval usually have a stronger mathematical background starting from an image formation modeling, further going to formulation of the objective function (criterion) and finally to numerical techniques solving corresponding optimization tasks. Here we wish refer to the recent overview by Shechtman et al. [17] concentrated on the algorithms for the phase retrieve models of the form (1). The constrains sufficient for uniqueness of the solution are presented in detail. Beyond the alternating projection GS a few novel mathematical methods are discussed: semidefinite programming phase lifting using matrix completion (PhaseLift algorithm) by Candes et al. [18] and greedy sparse phase retrieval (GESPAR algorithm) by Shechtman et al. [19]. Some optical applications of the phase retrieval algorithms are considered in this overview paper.

Many publications concern variational techniques as well as revisions of the intuitive GS algorithms by using optimization formulations. In particular, the links between the conventional GS and variational techniques are studied by Fienup [20] and by Bauschke et al. [21]. A sophisticated variational formulation for the phase retrieval is demonstrated by Irwan and Lane [22], where the criterion corresponding to Poissonian observations and the prior defining the smoothness of the phase are proposed. The problem is formalized as a penalized likelihood optimization. The conjugate gradient iterative algorithm for this setting is proposed by Lane [23].

Especially we wish to note the recent Wirtingling flow (WF) algorithms presented in Candes et al. [24] and Chen and E. J. Candès [25]. These algorithms are iterative complex domain gradient descents. Specific features of these algorithms are as follows: a special spectral initialization, a non-trivial growing step-size parameter and truncation of the gradient in the truncation Wirtingling flow (TWF) version of the algorithms [24]. Meticulous mathematical analysis is produced for the algorithm design, parameter selection and performance evalua-
tion. It is stated that the solution of the quadratic equations (1) can be done "nearly as easy as solving linear equation". In this mathematical analysis the elements of the matrices $A_s$ in (1) are random independent and subject of a complex-valued Gaussian distribution. Simulation experiments demonstrate that the TWF algorithm works and works well provided a small level of the random noise.

1.C. Contribution and structure of this paper

The sparsity hypothesis is a hot topic in phase imaging for optics. In the GESPAR algorithm [19] the sparsity in the signal domain is exploited: the length of the vector-solution $x$ is minimized. The transform domain sparsity developed for amplitude and phase of $x$ is a base of the Sparse Phase Amplitude Reconstruction (SPAR) algorithm by Katkovnik and Astola [10]. This transform domain sparsity formulation has been applied for high-accuracy phase imaging in various setups by Katkovnik and Astola [27], [26], [28].

In this paper the SPAR technique is developed for the phase retrieval problem with the intensities defined in the form (1) for Poissonian observations. Data adaptive non-local frames (BM3D frames by Danielyan et al. [29]) are used for the transform domain representations and sparse approximations of phase and amplitude.

The SPAR phase retrieval algorithm derived from the variational formulation of the problem incorporates two types of filtering: filtering of Poissonian noise at the sensor plane and filtering of phase and amplitude at the object plane. If both these filters are omitted the SPAR algorithm becomes quite similar to the conventional GS algorithm. We use the term GS algorithm for this simplified version of the SPAR algorithm.

Surprisingly, this novel GS algorithm demonstrates the performance nearly identical to the advanced TWF algorithm [25]. Both algorithms enable similar accuracy for the phase and amplitude reconstruction as well as the similar computational complexity.

The complete SPAR algorithm computationally more demanding than the GS algorithm for noisy data demonstrates much higher accuracy as compared versus both TWF and GS algorithms. The phase unwrapping included in the iterations of the SPAR algorithm, when the phase variation overcomes $2\pi$ range, enables an efficient noise suppression and accurate absolute phase reconstruction in situations when the TWF algorithm fails.

The paper is organized as follows. In Section 2 we consider different variational formulations of phase retrieval including derivation of the new GS algorithm and the sparsity modelling for phase and amplitude. The complete SPAR algorithm development is a subject of Section 3, where step-by-step solutions of the variational problems are discussed and the SPAR algorithm is composed. Section 4 concerns the experimental study of the proposed GS and SPAR algorithms and their comparison versus the TWF algorithm.
2. Problem formulation

2.A. Sparse wavefront modeling

It is recognized that many natural images (and signals) admit sparse representations in
the sense that they can be well approximated by linear combinations of a small number of
functions. This is a consequence of the self-similarity of these images: it is very likely to
find in them many similar patches in different locations and at different scales. The topic of
sparse and redundant representations is of tremendous interest in the last ten years. This
interest stems from the role that the low dimensional models play in many signal and image
areas such as compression, restoration, classification, and design of priors and regularizers,
just to name a few [30].

Let $x \in \mathbb{C}^n$ be a complex-valued wavefront. Denote $b = \text{abs}(x)$ and $\varphi = \text{angle}(x) \in [-\pi, \pi)$ as, respectively, the corresponding images of amplitude (modulus) and the wrapped phase, $\varphi$. Then we have $x = b \times \exp(j\varphi)$. Herein, all functions applied to vectors are to be understood in the component-wise sense; the same applies to multiplications (denoted as $\times$) and divisions of vectors.

With the objective of formulating treatable phase imaging problems, most approaches
follow a two-step procedure: in the first step, an estimate of the so-called principal (wrapped, interferometric) phase in the interval $[-\pi, \pi)$ is determined; in the second step, termed phase unwrapping, the absolute phase is inferred by adding of an integer number of $2\pi$ multiples to the estimated interferometric phase [31]. In what follows, we denote the principal phase as $\varphi$ and the absolute phase as $\varphi_{\text{abs}}$. We introduce the phase-wrap operator $\mathcal{W} : \mathbb{R} \mapsto [-\pi, \pi)$, linking the absolute and principal phase as $\varphi = \mathcal{W}(\varphi_{\text{abs}})$. We also define the unwrapped phase as $\varphi_{\text{abs}} = \mathcal{W}^{-1}(\varphi)$. Notice that $\mathcal{W}^{-1}$ is not the inverse function of $\mathcal{W}$ because the latter is not one-to-one and thus is does not have inverse.

In sparse coding for complex valued $x$, we may think in two different directions: either we
use a complex valued sparse representation to model directly the complex image $x$, as recently
proposed in [32] and [33], or we use sparse real valued representations for the amplitude $b$ and absolute phase $\varphi_{\text{abs}}$ images of $x$. To some extend the choice of the type of the sparse
modeling depends on the application. The former is suited to wavefront reconstruction,
where the interferometric (wrapped) phase carries all necessary phase information, whereas
the latter is suited to applications requiring the inference of the absolute phase, herein termed phase reconstruction.

In this paper, we follow the second type of the wavefront modeling. We introduce formally
this sparse wavefront modeling as the following matrix operations:

\begin{align}
    b &= \Psi_b \theta_a, \\
    \varphi &= \Psi_\varphi \theta_\varphi, \\
    \theta_a &= \Phi_a b, \\
    \theta_\varphi &= \Phi_\varphi \varphi_{\text{abs}},
\end{align}

where $\theta_a \in \mathbb{R}^p$ and $\theta_\varphi \in \mathbb{R}^p$ are, respectively, the amplitude and absolute phase spectra of the
object \(x\). In Eqs.(4), the amplitude \(b \in \mathbb{R}^n\) and absolute phase \(\varphi \in \mathbb{R}^n\) are synthesized from the amplitude and phase spectra \(\theta_a\) and \(\theta_\varphi\). On the other hand, the analysis Eqs.(5) give the spectra for amplitude and phase of the wavefront \(x\). In Eqs.(4)-(5) the synthesis \((n \times p)\) and analysis \((p \times n)\) matrices are denoted as \(\Psi_a, \Psi_\varphi, \Phi_a, \Phi_\varphi,\) respectively.

Following the sparsity rationale we assume that amplitude and phase spectra, \(\theta_a\) and \(\theta_\varphi\), respectively, are sparse; i.e., most elements thereof are zero. In order to quantify the level of sparsity of \(\theta_a\) and \(\theta_\varphi\), i.e., their number of non-zero (active) elements, we use the pseudo \(l_0\)-norm \(\| \cdot \|_0\) defined as a number of non-zero elements of the vector-argument. Therefore, in the ensuing formulations, we will design estimation criteria promoting low values of \(\| \theta_a \|_0\) and \(\| \theta_\varphi \|_0\).

Usually, the spectral dimensions are much higher than the dimensions of the image \(x\), \(p \gg n\), while the number of the active elements, i.e. the pseudo \(l_0\)-norms of spectra, are much smaller than \(p\) and smaller than \(n\). The sparse approximations in the form of Eqs.(4)-(5) are initiated from our works [26] and [27].

It is obvious that for the complex exponent there is no difference between the principal and absolute phase, \(\exp(j \varphi_{abs}) = \exp(j \varphi)\), and the angle operator in \(\varphi = \text{angle}(x)\) gives the principal phase. However, there is a great deal of difference between the sparsity for the absolute and interferometric phases, because in many cases the absolute phase can be smooth or piece-wise smooth function easily allowing sparsification while provided \(\max(\abs(\varphi_{abs})) > \pi\) the corresponding wrapped phase may experience multiple heavy discontinuities and be quite difficult for direct sparse approximations. Nevertheless note, that an efficient sparsification of the wrapped phase can be achieved through approximation of the complex exponent \(\exp(j \varphi)\). Here we wish to mentioned the windowed Fourier transform developed for fringe processing by Kemao [34] as well as different forms of the Gabor transform which are definitely good candidates for this problem.

Another styles of the data adaptive efficient approximators for the complex exponent are proposed in the recent papers [32] and [33] based on the leaning dictionary techniques and high-order SVD non-local complex domain filtering.

2.B. Noisy observation modeling

The measurement process in optics amounts to count the photons hitting the sensor’s elements and is well modeled by independent Poisson random variables: the probability that a random Poissonian variable \(z_s[l]\) of the mean value \(y_s[l]\) takes a given non-negative integer \(k\), is given by

\[
p(z_s[l] = k) = \exp(-y_s[l]) \frac{(y_s[l])^k}{k!},
\]

where \(y_s[l]\) is the intensity of the wavefront at pixel \(l\) defined by Eq.(1).

The parameter \(\chi > 0\) in (6) is a scaling factor, which can be interpreted as an exposure time or as a sensitivity of the sensor. Recall that the mean and the variance of Poisson random
variable \( z_s[l] \) are equal and are given by \( y_s[l] \), i.e., \( E\{z_s[l]\} = \text{var}\{z_s[l]\} = y_s[l] \). Defining the observation signal-to-noise ratio (SNR) as the ratio between the square of the mean and the variance of \( z_s[l] \), we have \( SNR = \frac{E^2\{z_s[l]\}}{\text{var}\{z_s[l]\}} = y_s[l] \). Thus, the noisiness of observations approaches infinite when \( \chi \to 0 \) \((SNR \to 0)\) and approaches zero when \( \chi \to \infty \) \((SNR \to \infty)\). The latter case corresponds to the noiseless scenario, i.e. \( z_s[l]/\chi \to y_s[l] \) with the probability 1.

The scale parameter \( \chi \) is of importance for modeling as it allows to control a level of randomness in observations. For real data processing one usually can take \( \chi = 1 \) assuming image scaling such that \( E\{z\} = y \).

3. Algorithm development

We formulate reconstruction of the wavefront as a variational problem with estimation of the amplitude and phase of \( x = b \times \exp(j\varphi) \) from noisy Poissonian observations \( z_s \). This problem is a rather challenging mainly due to nonlinearity of \( x \) with respect to the amplitude and phase and the periodic nature of \( x \) with respect to the phase.

3.A. GS algorithm

Let us start from a simplified setting of the problem. Assume that the sparsity hypotheses for the amplitude and phase of \( x \) are not imposed. Then, the maximum likelihood concept for the observations (6) gives the criterion

\[
L(\{u_s\}) = \sum_{s=1}^{L} \sum_{l=1}^{n} [||u_s[l]||^2 \chi - z_s[l] \log(||u_s[l]||^2 \chi)],
\]

where \( u_s = A_s x \).

The WF and TWF algorithms in [24] and [25] implement a direct minimization of \( L(\{u_s\}) \) based on straightforward calculations of the gradient of \( L(\{u_s\}, x) \) with respect to \( x \in \mathbb{C}^n \).

Contrary to this approach we reformulate the problem as a constrained optimization \( \min_{u_s, x} L(\{u_s\}) \) subject to \( u_s = A_s x, s = 1, \ldots, L \). The quadratic penalization of the constraints leads to the criterion

\[
L_1(\{u_s\}, x) = \sum_{s=1}^{L} \sum_{l=1}^{n} [||u_s[l]||^2 \chi - z_s[l] \log(||u_s[l]||^2 \chi)] + \frac{1}{\gamma_1} \sum_{s=1}^{L} ||u_s - A_s x||_2^2,
\]

where the weight parameter \( \gamma_1 > 0 \).

The iterative alternative minimization is used for optimization of (7) with respect to \( u_s \in \mathbb{C}^m \) and \( x \in \mathbb{C}^n \).

\[
\{\hat{u}^t_s\} = \arg \min_{u_s} L_1(\{u_s\}, \hat{x}^t), \tag{8}
\]

\[
\hat{x}^{t+1} = \arg \min_x L_1(\{\hat{u}^t_s\}, x). \tag{9}
\]
The solution for \((8)\) is of the form \([35]\):

\[ u_s[l] = b_s[l] \exp(j \cdot \text{angle}(v_s[l])), \]

where

\[ b_s[l] = \frac{|v_s[l]|/(\gamma_1 \chi) + \sqrt{|v_s[l]|^2/(\gamma_1 \chi)^2 + 4|z_s[l]|(1 + 1/(\gamma_1 \chi))/\chi}}{2(1 + 1/(\gamma_1 \chi))} \]

\[ v_s = A_s x. \]

In this solution the amplitude \(b_s[l]\) depends on both the observation \(z_s\) and the amplitude of \(v_s[l]\).

Note, that for large \(\gamma_1 \chi \to \infty\) (noiseless case)

\[ u_s[l] \to \sqrt{|z_s[l]|/\chi} \exp(j \cdot \text{angle}(v_s[l])), s = 1, \ldots, L. \]

Optimization of \(\mathcal{L}_1(\{u_s^k\}, x)\) with respect to \(x \in \mathbb{C}^n\) (for the problem \((9)\)) leads to the minimum condition of the form \(\partial \mathcal{L}_1(\{u_s^k\}, x)/\partial x^* = 0\) and to the normal least-squares equation for \(x\)

\[ \sum_{s=1}^{L} A_s^H A_s x = \sum_{s=1}^{L} A_s^H u_s \]

and to the solution

\[ x = (\sum_{s=1}^{L} A_s^H A_s)^{-1} \sum_{s=1}^{L} A_s^H u_s. \]

For the Fraunhofer approximation of the forward wavefront propagation diffraction operator \(A_s = F_s\), where \(F_s\) is the discrete Fourier transform, and \(A_s^H A_s = I_{n \times n}\) provided \(n = m\). Then Eq.(15) takes the form

\[ x = \frac{1}{L} \sum_{s=1}^{L} A_s^H u_s. \]

In general, the situation can be much more complex, in particular, because \(A_s^H A_s\) are ill-conditioned due to the fact that the operators \(A_s\) are a low-path filters suppressing high frequency components of the object \(x\) (e.g. [3]) and in many cases \(m < n\).

Then the solution of Eq.(14) can be found using iterations

\[ e^{k+1} = e^k \beta (\sum_{s=1}^{L} A_s^H A_s e^k - \sum_{s=1}^{L} A_s^H u_s), k = 0, 1, \ldots, \]

where \(e^k\) is an estimate of \(x\) and \(\beta > 0\) is a step-size parameter.

Note, that in this modeling \(A_s e^k\) is the forward propagation of the wavefront \(e^k\) and \(A_s^H e^k\) is the backward propagation of the wavefront \(A_s e^k\). The required for phase retrieval the sufficient phase diversity means that \(\sum_{s=1}^{L} A_s^H A_s\) is a positive definite matrix. Then, Eq.(14) has an unique solution and exists a small enough \(\beta\) such that the iterations in Eq.(17) converge.
Combining the solutions (13) and (15) for Eqs.(8)-(9) we arrive to the iterative GS algorithm shown in Table 1.

At Step 1 the object wavefront estimate $\hat{x}^t$ propagates using the operators $A_s$ and defines the wavefront $\hat{v}_s^t$ at the sensor plane. At Step 2 this wavefront is updated to the variable $\hat{u}_s^t$ by changing the amplitude of $\hat{v}_s^t$ according to the given observations $z_s$, while the phase of $\hat{v}_s^t$ is preserved in $\hat{u}_s^t$. At Step 3 the estimates $\{\hat{u}_s^t\}$ backpropagate to the object plane and update the object wavefront $\hat{x}^{t+1}$.

We have here both the typical features of the GS algorithms. First, iterative forward and backward propagations, and second, update of the amplitudes in the transform domain accordingly to the given observations. Respectively, we use the name GS for this phase retrieval algorithm.

Contrary to the usual heuristic design of this type of the algorithms this one is derived from optimization formulation. Remind that the assumption $\gamma_1 \chi \to \infty$ is used in this design. It means that this algorithm is optimal for noiseless observations only. For noisy observations ($\chi$ is not large) the algorithm can be optimized accordingly using in Step 2 the amplitudes defined by Eq.(11) instead of the standard GS rule (13).

<table>
<thead>
<tr>
<th>Table 1. GS Phase Retrieval Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> ${z_s}$, $s = 1,...,L$ , $x^1$;</td>
</tr>
<tr>
<td>For $t = 1,...,N$;</td>
</tr>
<tr>
<td>1. <strong>Forward propagation:</strong></td>
</tr>
<tr>
<td>$\hat{v}_s^t = A_s \hat{x}^t$, $s = 1,...,L$;</td>
</tr>
<tr>
<td>2. <strong>Observation constrains:</strong></td>
</tr>
<tr>
<td>$\hat{u}_s^t[l] = \sqrt{z_s[l]/\chi} \exp(j \cdot \angle (\hat{v}_s^t[l])), s = 1,...,L$;</td>
</tr>
<tr>
<td>3. <strong>Backward propagation:</strong></td>
</tr>
<tr>
<td>$\hat{x}^{t+1} = (\sum_{s=1}^{L} A_s^\dagger A_s)^{-1} \sum_{s=1}^{L} A_s^\dagger \hat{u}_s^t$;</td>
</tr>
<tr>
<td><strong>Output:</strong> $\hat{x}^{N+1}$.</td>
</tr>
</tbody>
</table>

With $\gamma_1 \to \infty$ the criterion $L_1(\{u_s\}, x) \to \sum_{s=1}^{L} \sum_{l=1}^{n} [u_s[l]^2 \chi - z_s[l] \log(|u_s[l]|^2 \chi)]$, thus the GS algorithm minimizes this minus loglikelihood written for Poissonian observations. Note that the TWF algorithm in [25] is designed for minimization of the latter criterion. Thus, both TWF and GS algorithms being completely different are intended for minimization of the same minus loglikelihood criterion.

It seems that the goal of the optimal algorithm design is achieved because the solution if found is optimal for Poissonian observations at least for large $L$. However, a large $L$ is not practical and for a small $L$ randomness of observations is revealed in randomness of the object wavefront $x$, in particular in the object phase. The random errors in $x$ can be quite
strong due to the fact that the problem at hand is usually ill-posed and noise amplification for the variables of interest is a typical effect.

This point motivates us to move further and to develop an algorithm with improved filtering properties for \( x \). In what follows we show that the sparsity hypotheses for amplitude and phase of \( x \) is a relevant instrument to achieve this goal.

3.B. Sparse phase retrieval

Two principally different variational formulations classified as the analysis and synthesis approaches can be viewed for sparse modelling. In the synthesis approach, the relations between the signal and spectrum variables are given by the synthesis equations (4), while in the analysis approach these relations are given by the analysis equations (5).

In the synthesis approach the variational setup is of the form

\[
\min_{\theta_a, \theta_\varphi} \sum_{s=1}^{L} \sum_{l=1}^{n} ||u_s[l]||^2 \chi - z_s[l] \log(||u_s[l]||^2 \chi) + \alpha_1 ||\theta_a||_0 + \alpha_2 ||\theta_\varphi||_0,
\]

subject to: \( u_s = A_s x, \quad x = (\Psi_a \theta_a) \times \exp(j \Psi_\varphi \theta_\varphi) \).

The first summand in Eq.(18) is the minus loglikelihood corresponding to the Poissonian distribution (6), where \( y_s[l] \) are replaced by \( ||u_s[l]||^2 \). The pseudo \( l_0 \)-norms with the coefficients \( \alpha_1 \) and \( \alpha_2 \) are included in order to enable the sparsity of the amplitude and phase in the spectrum domain. According to the constrains (19) \( u_s = A_s [ (\Psi_a \theta_a) \times \exp(j \Psi_\varphi \theta_\varphi) ] \). Substituting this expression in Eq.(18) one can see that the minimization criterion depends only the spectrum variables \( \theta_a \) and \( \theta_\varphi \). This is a specific feature of the synthesis formulation: minimization is produced with respect to the spectrum variables and only the synthesis operators \( \Psi_a \) and \( \Psi_\varphi \) are used. The amplitude and the phase are calculated due to Eqs.(4). Thus, \( x \) is calculated using only the synthesis operators.

In the analysis approach the variational setup is of the form

\[
\min_{x} \sum_{s=1}^{L} \sum_{l=1}^{n} ||u_s[l]||^2 \chi - z_s[l] \log(||u_s[l]||^2 \chi) + \alpha_1 ||\Phi_a b||_0 + \alpha_2 ||\Phi_\varphi \varphi_{abs}||_0.
\]

subject to: \( u_s = A_s x, \quad b = \text{abs}(x), \quad \varphi_{abs} = \mathcal{W}^{-1}(\text{angle}(x)) \).

Substituting \( u_s, b \) and \( \varphi_{abs} \) from Eq.(21) into Eq.(20) one can see that this criterion depends only on the spatial complex-valued variable \( x \). Thus, contrary to the synthesis approach the optimization is produced in the signal domain and only the analysis operators \( \Phi_a \) and \( \Phi_\varphi \) are used.

It is clear from Eqs. (18)-(21) that both the synthesis and analysis setups lead to quite complex optimization problems.

Herein, we adopt a different Nash equilibrium approach. The constrained optimization with a single criterion function, as it is in (18) and (20), is replaced by a search for the Nash equilibrium balancing two criteria. Details of this approach, links with the game theory and
demonstrations of its efficiency for the synthesis-analysis sparse inverse imaging can be seen in [29], where it is done for linear observation modeling. Applications of the Nash equilibrium technique for optical problems with nonlinear objects and nonlinear observations can be seen in [26], [27], [35]. In what follows we use the approaches developed in these papers.

3.C. SPAR algorithm

The following two criteria are introduced for formalization of the algorithm design:

\[
L_1(\{u_s\}, x) = \sum_{s=1}^{L} \sum_{l=1}^{n} ||u_s[l]||^2 \chi - z_s[l] \log(||u_s[l]||^2 \chi) + \frac{1}{\gamma_1} \sum_{s=1}^{L} ||u_s - A_s x||_2^2, \tag{22}
\]

\[
L_2(\theta_\varphi, \theta_a, \varphi_{abs}, b) = \tau_a \cdot ||\theta_a||_0 + \tau_\varphi \cdot ||\theta_\varphi||_0 + \frac{1}{2} ||\theta_a - \Phi_a b||_2^2 + \frac{1}{2} ||\theta_\varphi - \Phi_\varphi \varphi_{abs}||_2^2, \tag{23}
\]

where \( b = \text{abs}(x), \varphi_{abs} = W^{-1}(\text{angle}(x)) \).

The criterion (22) is identical to (7) and already discussed in Subsection 3.A. As it is emphasized in Subsection 2.A we use the separate sparse modeling for the absolute phase \( \varphi_{abs} \) and the amplitude \( b \) of the wavefront \( x \). The criterion (23) promote this sparsity in the analysis transform domain. The regularization terms \( \frac{1}{2} ||\theta_a - \Phi_a b||_2^2 \) and \( \frac{1}{2} ||\theta_\varphi - \Phi_\varphi \varphi_{abs}||_2^2 \) are squared Euclidean norms calculated for differences between spectra \( \theta_a \) and \( \theta_\varphi \) and their predictors \( \Phi_a b \) and \( \Phi_\varphi \varphi_{abs} \).

Minimization of \( L_1(\{u_s\}, x) \) with respect to \( u_s \) and \( x \) gives the solutions shown in Eqs.(11) and (15). Minimization of \( L_2 \) on \( \theta_a \) and \( \theta_\varphi \) results in the well known hard-thresholding solutions:

\[
\hat{\theta}_a = (\Phi_a b) \times 1 \left[ \text{abs}(\Phi_a b) \geq \sqrt{2\tau_a} \right], \tag{24}
\]

\[
\hat{\theta}_\varphi = (\Phi_\varphi \varphi_{abs}) \times 1 \left[ \text{abs}(\Phi_\varphi \varphi_{abs}) \geq \sqrt{2\tau_\varphi} \right],
\]

where \( 1[w], w \in \mathbb{R}^p \), is an element-wise vector function: \( \mathbb{R}^p \mapsto \mathbb{R}^p, 1[w_k] = 1 \) if \( w_k \geq 0 \) and \( 1[w_k] = 0 \) if \( w_k < 0 \).

Here \( t_{\theta_a} = \sqrt{2\tau_a} \) and \( t_{\theta_\varphi} = \sqrt{2\tau_\varphi} \) are thresholds for the amplitude and the phase, respectively. The items of the spectral coefficients \( \text{abs}(\Phi_a b) \) and \( \text{abs}(\Phi_\varphi \varphi_{abs}) \), which are smaller than the corresponding thresholds are zeroed in Eq.(24).

According to the idea of the Nash equilibrium balancing multiple penalty functions (e.g. [37]) the proposed algorithm is composed of alternating optimization steps performed for the
criteria $\mathcal{L}_1$ and $\mathcal{L}_2$. It leads to the iterative procedure:

$$\{\hat{u}^t_s\} = \arg \min_{\{u_s\}} \mathcal{L}_1(\{u_s\}, \hat{x}^t), \quad (25)$$

$$\hat{x}^t = \arg \min_{x} \mathcal{L}_1(\{\hat{u}^t_s\}, x), \quad (26)$$

$$(\hat{\theta}_{\varphi}^t, \hat{\theta}_a^t) = \arg \min_{\theta_{\varphi}, \theta_a} \mathcal{L}_2(\theta_{\varphi}, \theta_a, \hat{\varphi}_{abs}^t, \hat{b}^t), \quad (27)$$

$$\hat{b}^{t+1} = \Psi_a \hat{\theta}_a^t; \hat{\varphi}_{abs}^{t+1} = \Psi_{\varphi} \hat{\theta}_{\varphi}^t \quad (28)$$

$$\hat{x}^{t+1} = \hat{b}^{t+1} \times \exp(j \hat{\varphi}_{abs}^{t+1}), \quad (29)$$

where $b^{t+1}$ and $\varphi_{abs}^{t+1}$ are updates of the amplitude and the absolute phase for $x$.

The success of any sparse imaging depends on how reach and redundant are the transforms/dictionaries used for analysis and synthesis. In our algorithm for the analysis and synthesis operations we use the BM3D frames, where BM3D is the abbreviation for Block-Matching and 3D filtering [29]. Let us recall some basic ideas of this advanced technique. At the first stage the image is partitioned into small overlapping square patches and further the vector corresponding to each patch is modeled as a sparse linear combination of vectors taken from a given orthonormal bases. For each patch a group of similar patches is collected which are stacked together and form a 3D array (group). This stage is called grouping. The entire 3D group-array is projected onto a 3D transform basis. The obtained spectral coefficients are hard-thresholded and the inverse 3D transform gives the filtered patches, which are returned to the original position of these patches in the image. This stage is called collaborative filtering. This process is repeated for all pixels of the entire wavefront and obtained overlapped filtered patches are aggregated in the final image estimate. This last stage is called aggregation. The details of BM3D as an advanced image filter can be seen in [36].

It follows from [29], that the steps (27)-(28) including the grouping operations defining the analysis $\Phi$ and synthesis $\Psi$ frames can be combined in a single algorithm. In what follows we use the notation BM3D for this algorithm. Note, that the standard BM3D algorithm as it is presented in the original paper [36] is composed from two successive steps: thresholding and Wiener filtering. In this paper BM3D corresponding to the procedures (27) and (28) consists of only the first thresholding (hard-thresholding) step.

The criterion $\mathcal{L}_2$ is separable on the variables $\theta_{\varphi}$ and $\theta_a$. It follows that the corresponding solutions can be calculated independently for the amplitude and the phase. Using the BM3D algorithm for implementation of the steps (27)-(28) we obtain:

$$\hat{b}^{t+1} = BM3D_{ampl}(\hat{b}^t, th_a), \quad (30)$$

$$\hat{\varphi}_{abs}^{t+1} = BM3D_{phase}(\hat{\varphi}_{abs}^t, th_{\varphi}).$$

In Eq.(30) we use BM3D with different subscripts because different parameters can be used in BM3D for amplitude and phase processing.

Combining the solutions obtain in Subsection 3.A for Eqs.(25)-(26) and Eqs.(30) for the
steps (27)-(28) we arrive to the phase retrieval algorithm shown in Table 2.

Table 2. SPAR Phase Retrieval Algorithm

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Forward propagation:</td>
</tr>
<tr>
<td></td>
<td>( \hat{v}_s^t = A_s \hat{x}_s^t ), ( s = 1, ..., L );</td>
</tr>
<tr>
<td>2.</td>
<td>Poissonian noise suppression:</td>
</tr>
<tr>
<td></td>
<td>( \hat{u}_s^t = b_s^t \times \exp(j \cdot \text{angle}(\hat{v}_s^t)) ), Eq.(11) for ( b_s^t );</td>
</tr>
<tr>
<td>3.</td>
<td>Backward propagation:</td>
</tr>
<tr>
<td></td>
<td>( \hat{x}<em>s^t = (\sum</em>{s=1}^{L} A_s^H A_s)^{-1} \sum_{s=1}^{L} A_s^H \hat{u}_s^t );</td>
</tr>
<tr>
<td>4.</td>
<td>Phase unwrapping:</td>
</tr>
<tr>
<td></td>
<td>( \hat{\phi}_{abs}^t = \mathcal{W}^{-1}(\text{angle}(\hat{x}_s^t)) );</td>
</tr>
<tr>
<td>5.</td>
<td>Phase and amplitude filtering:</td>
</tr>
<tr>
<td></td>
<td>( \hat{\phi}<em>{abs}^{t+1} = BM3D</em>{\text{phase}}(\hat{\phi}<em>{abs}^t, th</em>\phi) );</td>
</tr>
<tr>
<td></td>
<td>( \hat{b}^{t+1} = BM3D_{\text{ampl}}(\text{abs}(\hat{x}_s^t), th_a) );</td>
</tr>
<tr>
<td>6.</td>
<td>Object wavefront update:</td>
</tr>
<tr>
<td></td>
<td>( \hat{x}<em>s^{t+1} = \hat{b}^{t+1} \times \exp(j \hat{\phi}</em>{abs}^{t+1}) );</td>
</tr>
<tr>
<td></td>
<td>Output: ( \hat{\phi}_{abs}^{N+1}, \hat{b}^{N+1} ).</td>
</tr>
</tbody>
</table>

The first three steps of this algorithm are identical to the corresponding three steps of the GS algorithm with the only difference that in Step 2 instead of the standard GS rule we use a more general procedure (11) for calculation of the amplitude taking into consideration the noisiness of observations.

The sparsification (filtering on the base of sparse approximation) is produced in Step 5. The unwrapping of the phase in Step 4 is necessary in order to use the sparsity hypothesis imposed on the absolute phase. In the GS algorithm update of \( x \) is produced in complex domain. Contrary to it the update of \( x \) in Step 6 is produced through the updates of the amplitude and the absolute phase calculated in parallel.

The proposed GS algorithm can be treated as a special case of the SPAR algorithm. Indeed, if \( th_\phi, th_a \to 0 \) the BM3D filter does not filter input signal. Then, we do not need the absolute phase, Step 4 as well as Step 5 can be dropped and the SPAR algorithm becomes identical to the GS algorithm within the difference in Step 2.

4. Numerical experiments

For our simulations we select the coded diffraction pattern scenario Eq.(2):

\[
y_s = |\mathbf{F} \mathbf{D}_s \mathbf{x}|^2, \ s = 1, ..., L.
\]  

(31)
Following to the publications \cite{24} and \cite{25} the wavefront modulation is enabled by the random phases $\phi_k$ in $D_s$ with equal probabilities taking four values $[0, \pi/2, -\pi/2, \pi]$ and the number of experiments $L = 12$. To some extend, the choice for experiments of the model (31) with these four random phase values for phase modulation is caused by our intention to consider TWF as the main counterpart to our algorithm. The rigorous mathematical background of TWR developed for the Poissonian data makes this algorithm one of the best in the area \cite{25}. The MATLAB codes of TWF provided for the model (31) make the comparative analysis simple for implementation. While the results in \cite{25} are presented mainly for noiseless data or small level noise herein we are concentrated on noisy data and show that in this case the developed sparse modeling allows to achieve a dramatic improvement in the accuracy of phase and amplitude imaging. For the model (31) $A_s = FD_s$ and

$$
\sum_{s=1}^{L} A_s^H A = \sum_{s=1}^{L} D_s^H F^H F D_s = \sum_{s=1}^{L} D_s^H D_s = LI_{n \times n}, \tag{32}
$$

and Step 2 of our algorithm is simplified to the form

$$\hat{x}' = \frac{1}{L} \sum_{s=1}^{L} D_s^H F^H \hat{u}_s'. \tag{33}$$

All results presented in this section can be reproduced by running the publicly available MATLAB demo-codes \footnote{http://www.cs.tut.fi/~lasip/DDT/index3.html}.

4.A. Processing without phase unwrapping

For experiments we use $256 \times 256$ MATLAB test-images: Lena and Cameraman. For the phase they are scaled to the interval $[0, \pi/2]$. In this case the phase unwrapping is not required and Step 4 is omitted in the SPAR algorithm.

The accuracy of the wavefront reconstruction is characterized by \textit{RMSE} criteria calculated independently for amplitude and phase. In the phase retrieval problem the object phase image can be estimated within an invariant phase-shift only. Following \cite{24} and \cite{25} the estimated phase image is corrected by an invariant phase-shift $\varphi_{shift}$ defined as

$$\varphi_{shift} = \arg\left( \min_{\varphi \in [0,2\pi]} \| \exp(-j\varphi) \hat{x} - x \|_2^2 \right), \tag{34}$$

here $x$ and $\hat{x}$ are the true phase and the estimate of the wavefront, respectively. This correction of the phase is done only for calculation of the criteria and for result imaging and not used in the algorithm iterations.

In what follows we compare the results obtained by TWF (MATLAB codes available at http://web.stanford.edu/~yxchen/TWF/) and ours GS and SPAR algorithms. The results are shown as functions of the Poissonian scale parameter $\chi$, $1 \leq \chi \leq 0.25 \cdot 10^{-5}$. The smallest
\( \chi \) results in the noisiest data. The corresponding Signal-to-Noise Ratio (SNR) is calculated as

\[
SNR = 10 \log_{10}(\chi^2 \sum_{s=1}^{L} ||y_s||^2_2 / \sum_{s=1}^{L} ||y_s \chi - z_s||^2_2) \ dB.
\] (35)

As an interesting parameter of the observed data we calculate also the mean values of \( \{z_s\} \), \( N_{\text{photon}} = \sum_{s=1}^{L} \sum_{k=1}^{n} z_s(k)/Ln \), i.e. the mean number of photons per pixel of the sensor. Variations of SNR and \( N_{\text{photon}} \) naturally depend on amplitude and phase images. In our experiments for \( \chi \in [0.25 \cdot 10^{-5}, 1.0] \) these variations approximately take values from 1dB to 60dB for SNR and from 1 photon to \( 6 \cdot 10^4 \) photons per pixel for \( N_{\text{photon}} \).

The achieved accuracies in RMSE values are shown in Figs. 2 and 4, respectively, for the phase images of Lena and Cameraman with the amplitude equal to 1.

![Fig. 2. Lena phase image: RMSE for phase and amplitude reconstructions versus the parameter \( \chi \). Comparison of the TWF, GS and SPAR algorithms.](image)

Visualization of these phase reconstructions is presented in Figs.(3) and (5). The advantage of the SPAR algorithm is obvious both visually and numerically.

Now let us consider the more demanding scenario when both phase and amplitude are space varying. Lena is used for the phase and Cameraman for the amplitude. In these experiments Cameraman is scaled to the interval \( [0.1, 1.1] \). The corresponding RMSE values are shown in Fig.6.

It can be seen in comparison with the previous curves for RMSE that the accuracy of the reconstruction is significantly lower for the high-level noise, \( \chi < 0.01 \). Images in Fig.7 for noisy case, \( \chi = 0.001 \), show a visible degradation of the phase reconstruction due to the leakage from the space varying Cameraman amplitude. Some features of Cameraman can be seen in the Lena phase reconstruction. For the lower nose level, Fig.8, \( \chi = 0.01 \), the reconstruction of the phase is of the quality close to the considered above cases when only
phase is spatially varying and the amplitude is invariant. The serious advantage of the SPAR algorithm is obvious in these results.

4.B. Super-resolution (compressed sensing) imaging

Consider the case when the size of the sensor measured in number of pixels is smaller than the size of the image to be reconstructed. It is a typical scenario of the super-resolution or compressed sensing, when sampling in the Fourier domain is produced only for the lower frequency components. The problem is to reconstruct the high-resolution image from these sub-sampled data.

In this case the criterion $\mathcal{L}_1$ in Eq.(22) takes the form

$$
\mathcal{L}_1(u_s, v_s) = \sum_{s=1}^{L} \sum_{l \in Z} |u_s[l]|^2 \chi - z_s[l] \log(|u_s[l]|^2 \chi) + \frac{1}{\gamma_1} \sum_{s=1}^{L} ||u_s - v_s||^2_2,
$$

where $Z$ denotes a set of the sensor pixels.

Step 2 of the SPAR algorithm takes the form

$$
\hat{u}_s[l] = \begin{cases} 
    b_s[l] \exp(j \cdot \text{angle}(v_s[l])), & \text{if } l \in Z, \\
    v_s[l] & \text{if } l \notin Z. 
\end{cases}
$$

Here, the vectors $u$ and $v$ have the same dimension, $u, v \in \mathbb{C}^n$, and only the pixels of $u_s$ from $Z$ are subjects of filtering while others are equal to the corresponding elements of $v_s$. The corresponding changes in Step 2 in the GS algorithm take the form

$$
\hat{u}_s[l] = \begin{cases} 
    \sqrt{z_s[l]/\chi} \exp(j \cdot \text{angle}(\hat{v}_s[l])) & \text{if } l \in Z, \\
    v_s[l] & \text{if } l \notin Z. 
\end{cases}
$$

Simulation experiments show that the SPAR algorithm works very well provided a subsampling up to 25% of the initial image size. The dependence of RMSE on $\chi$ for the Lena
phase modelling and invariant amplitude imaging are shown in Fig.9. Visualization of the phase imaging by SPAR and GS algorithms for $\chi = 0.01$ are shown Fig.10. The advantage of the SPAR algorithm versus the GS algorithm visually and numerically in RMSE values is obvious. These results demonstrate the serious advantage of the SPAR algorithm versus the GS algorithm explained by using the sparse modeling of the wavefront. The SPAR algorithm not only filter noise in the phase and amplitude imaging but also allows to reconstruct fine details of the images missed in subsampled observations. Note the TWF algorithm in the form presented by the authors is not applicable for the considered compressed sensing scenario.

4.C. Absolute phase imaging with phase unwrapping

Here we simulate three complex-valued data sets of the size $100 \times 100$ with the invariant amplitude equal to 1 and spatially varying absolute phase: Gaussian (phase range 44 radians), truncated Gaussian (phase range 44 radians), and Shear Plane (phase range 149 radians) exploited, in particular, in [32].

For the phase imaging we apply all three considered algorithms: TWF, GS and SPAR. The first two algorithms give the wrapped phase reconstructions which are unwrapped by the PUMA algorithm [38], identical to the one used in SPAR iterations.

The results are demonstrated for very noisy Poissonian observations obtained with $\chi = 2.5 \cdot 10^{-4}$. The 3D surfaces in Fig.11 are reconstructions of the absolute Gaussian phase. From left-to-right we can see three reconstructions TWF (failed), GS (quite noisy), SPAR and the true absolute phase surface. The SPAR reconstruction is slightly noisy and of the best quality and accuracy. In Fig.12 one can see the corresponding wrapped phases. RMSEs
for the phase images in Figs. 11-12 and in what follows are calculated for the absolute phase and wrapped phase, respectively.

Similar images in Figs. 13-14 are shown for the truncated Gaussian surface. This discontinuous function is much more difficult for reconstruction and for unwrapping as compared with the continuous Gaussian surface. Again we can see that TWF fails, GS gives a low quality reconstruction while SPAR is able to show much better results.

Images in Figs. 15-16 are shown for the discontinuous shear surface. Again we can see that TWF fails, GS gives a noisy reconstruction while the best results are demonstrated by the SPAR algorithm.

For lower noise level, $\chi > 10^{-3}$, all algorithms demonstrate a much better performance. TWF and GS show quite close results with a clear advantage of SPAR. For nearly noiseless cases, $\chi > 10^{-1}$, all algorithms enable a perfect reconstruction of all three absolute phase surfaces.

The advantage of SPAR versus GS reconstructions demonstrates the principal importance of the sparse phase modeling in iterations of the SPAR algorithm. The unwrapping of the GS reconstruction applied only for the final estimate is not able to produce results comparable with the ones obtained by SPAR. For small noise level the accuracy of TWF and GS is nearly identical.

4.D. Parameters of the SPAR algorithm

The performance of the SPAR algorithm essentially depends on its parameters. Optimization can be produced for each magnitude/phase distribution and the noise level. However, in our experiments the parameters are fixed for all our experiments and enabling a good quality of reconstruction. The image patches used in BM3D are always square $8 \times 8$. The group size is limited by the number 25. The step size between the neighboring patches is equal to 3. The transforms DCT (for patches) and Haar (for the group length) are used for 3D group
The parameters defining the iterations of the algorithm are as follows: $\gamma_1 = 1/\chi$; $\theta_a = 1.4$; $\theta_{\varphi} = 1.4$. The number of the iterations is fixed to 50.

For our experiments we used MATLAB R2014a and the computer with the processor Intel(R) Core(TM) i7-4800MQ@2.7 GHz. The complexity of the algorithm is characterized by the time required for processing. For 50 iterations and $256 \times 256$ images this time is as follows: TWF→10 sec.; GS→5 sec.; SPAR→70 sec. (no unwrapping); SPAR→90 sec. (with unwrapping).

5. Conclusion

This paper introduces a variational approach to object phase and amplitude reconstruction from noisy Poissonian intensity observations in the typical phase retrieval scenario. The maximum likelihood criterion used in the developed multiobjective optimization (Nash equilibrium approach) defines the tendency to get statistically optimal estimates. Sparse modeling of amplitude and absolute phase is one of the key elements of the developed SPAR algorithm used for modeling spatially varying amplitude and phase. This sparse modeling enables regularization in general ill-posed inverse imaging problems. The complexity of the algorithm is defined by the built-in BM3D filters generating data adaptive synthesis and analysis frames varying in iterations. The efficiency of the algorithm is demonstrated by simulation experiments for the coded diffraction pattern scenario. The comparison is produced versus the state-of-the-art TWF algorithm. For noisy observations the SPAR algorithm demonstrates a definite advantage over the TWF. For the low noise level the accuracy of
the SPAR algorithm as well as its simplified version the GS algorithm is nearly identical to the accuracy of the TWF algorithm. The GS algorithm is essentially faster than TWF while SPAR computationally much more demanding is slower than both TWF and GS.

6. Acknowledgments

This work is supported by the Academy of Finland, project no. 287150, 2015-2019.
Fig. 8. Varying amplitude and phase modelled as Cameraman and Lena, respectively: visualization of the phase reconstruction for the comparatively low noise level, $\chi = 0.01$. Traces of Cameraman are not seen in these images.

References

Fig. 9. Lena phase image for 25% subsampled observations: RMSE for phase and amplitude reconstructions versus the parameter $\chi$. Comparison of the GS and SPAR algorithms.


Fig. 11. The absolute (unwrapped) phase for the Gaussian phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true phase.

Fig. 12. The wrapped phase for the Gaussian phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true wrapped phase.
Fig. 13. The absolute (unwrapped) phase for the truncated Gaussian phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true phase.

Fig. 14. The wrapped phase for the truncated Gaussian phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true wrapped phase.
Fig. 15. The absolute (unwrapped) phase for the share plane phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true phase.

Fig. 16. The wrapped phase for the share plane phase object from noisy Poissonian observations. From left to right: TWF, GS, SPAR reconstructions and true wrapped phase.