

# Optimal phase retrieval from multiple observations with Gaussian noise: augmented Lagrangian algorithm for phase objects

Artem Migukin<sup>\*</sup>, Vladimir Katkovnik, Jaakko Astola  
Department of Signal Processing, Tampere University of Technology,  
Korkeakoulunkatu 10, P.O. Box 527, 33101, Tampere, Finland

## ABSTRACT

A novel iterative phase-retrieval algorithm is developed for reconstruction of phase objects. We propose a constrained variational formulation of the phase-retrieval problem with the forward wave field propagation from the object to the measurement planes as constraints. It is assumed that noisy intensity-only observations are given at measurement planes parallel to the object plane, and the additive noise in the observations is zero-mean Gaussian. This algorithm is derived from the maximum likelihood approach what enables an optimal solution for the phase reconstruction. The advanced performance of the algorithm is demonstrated by numerical simulations.

**Keywords:** phase retrieval, prior information, maximum likelihood, augmented Lagrangian, constrained optimization.

## 1. INTRODUCTION

Phase retrieval is a problem of phase reconstruction from multiple intensity observations of a wave field, made at different measurement planes, and *a priori* information. The phase recovering is of important and exploited in many technical and scientific areas such as microscopy, crystallography, astronomy, deformation detection, etc. Moreover, experimental arrangements for phase-retrieval methods are often simpler and cheaper than for interferometric ones, which require a reference beam. An important advantage of the phase-retrieval techniques is their high robustness to disturbances (e.g. vibration), which degrade the accuracy in interferometry.

Let  $u_0(x)$  and  $u_r(x)$ ,  $r=1, \dots, K$ ,  $x \in \mathbb{R}^2$  be complex-valued wave field distributions at the object and at the  $r$ -th measurement (sensor) planes, respectively. The index  $r$  corresponds to a distance  $z_r = z_1 + (r-1) \cdot \Delta_z$  between the parallel object and the  $r$ -th observation plane, where  $\Delta_z$  is a distance between the observation planes, and  $K$  is a number of these planes. The link between these wave field distributions can be defined according to the scalar diffraction theory as

$$u_r(x) = \{g_r \otimes u_0\}(x) = \iint g_r(x-\eta) \cdot u_0(\eta) d\eta \quad (1)$$

Here  $\otimes$  denotes the convolution operation,  $g_r(x)$  is the diffraction kernel<sup>1</sup>. The phase  $\varphi_r$  of the wave field at the  $r$ -th sensor plane ( $u_r = |u_r| \cdot \exp(j \cdot \varphi_r)$ ) can not be measured directly, thus the integral (1) can not be inverted. The phase is recovered, mainly iteratively, from a number of intensity measurements made at the observation planes.

In 1972 Gerchberg and Saxton<sup>2</sup> proposed a simple and efficient iterative algorithm for phase-retrieval (GS), initially for a single observation plane. In <sup>2</sup> *prior* information on the object distribution is used for the wave field reconstruction. Fienup systematized the earlier works and in 1982 introduced some classes of iterative phase retrieval algorithms<sup>3</sup>: error-reduction, gradient search and input-output methods.

The common idea of Gerchberg-Saxton-Fienup algorithms is to replace the estimated magnitudes at the sensor planes by ones obtained from the intensity observations. The main difference between these phase-retrieval algorithms<sup>3</sup> consists in methods of the object wave field reconstruction. Contrary to <sup>3</sup>, Misell's variation of GS<sup>4</sup> suggests the wave field reconstruction at the sensor planes only. In this case, there is no connection with the object plane and no *prior* information on the object distribution is used.

---

<sup>\*</sup> artem.migukin@tut.fi; phone: +358401981310; fax: +358331154989; <http://www.cs.tut.fi/~lasip/DDT>

In our work we consider the problem of the optimal wave field reconstruction from the intensity observations. In order to achieve this goal we use a variational constrained maximum likelihood formulation. The phase-retrieval algorithm is obtained as a solution of this optimization problem, and concentrated on the phase at the object plane.

## 1.1 Observation model

Let us introduce the considered multi-plane wave field reconstruction scenario and our basic notations. It is assumed that the wave field distributions at the object and sensor planes are pixel-wise invariant. In such a discrete-to-discrete model continuous argument  $x=(x_1, x_2)$  is replaced by digital one  $k=(k_1, k_2)$  with the corresponding replacement of continuous distributions by their discrete counterparts as  $u_0(x) \rightarrow u_r[k]$ ,  $u_r(x) \rightarrow u_r[k]$ .

In this work we use a conventional vector-matrix notation for complex-valued distributions of the wave fields at the object and sensor planes as  $\mathbb{C}^{n \times 1}$  vectors. For 2D discrete distributions (matrices) of the size  $N \times M$  the complex-valued vector variables of the length  $n=N \cdot M$  are constructed by concatenating columns of the matrices. Bold lower case characters are used for these vectors. Let the wave fields at the object and sensor planes be of the same size. Then, the forward wave field propagation from the object to the  $r$ -th sensor plane can be presented as follows:

$$\mathbf{u}_r = \mathbf{A}_r \cdot \mathbf{u}_0, \quad r=1, \dots, K, \quad (2)$$

where  $\mathbf{A}_r \in \mathbb{C}^{n \times n}$  is a discrete forward propagation operator corresponding to the diffraction kernel  $g_r$  in (1). We consider a coherent light scenario with the paraxial approximation of the wave field propagation based on the Rayleigh-Sommerfield integral<sup>1</sup>. The operator  $\mathbf{A}_r$  in (2) is specified by the used discretization model of this integral, and depending on it one can use different discrete forward propagation models. It can be: discrete convolution of the single or double size<sup>5</sup>, angular spectrum decomposition (ASD)<sup>1</sup> or e.g. the recent discrete diffraction transform (DDT) given in the matrix (M-DDT)<sup>6</sup> or frequency domain (F-DDT) forms<sup>7</sup>. These DDT models are obtained for the Fresnel approximation of the Rayleigh-Sommerfield integral and enable an accurate pixel-to-pixel mapping of  $\mathbf{u}_0$  to  $\mathbf{u}_r$ . This assumption is natural for digital sensors and can be exploited as a pixel-wise approximation for the object wave field distribution. In our simulations we deal with the pixelated objects, and for numerical experiment use the DDT modeling for the accurate forward wave field propagation.

In the vector-matrix notation the discrete intensity observations are given in the form

$$\mathbf{o}_r = |\mathbf{u}_r|^2 + \boldsymbol{\varepsilon}_r, \quad r=1, \dots, K, \quad (3)$$

where the wave field intensity is measured with an additive random errors  $\varepsilon_r[k]$ . Here the modulus  $|\cdot|$  and square of modulus  $|\cdot|^2$  are the elementwise operations applied to the elements of the corresponding vectors, thus  $|\cdot|$  and  $|\cdot|^2$  are vectors. We assume that the error  $\varepsilon_r$  of the intensity observations is a result of various degradation factors such as the sensor noise, a nonideality of a laser beam and propagation media, etc. For simplicity and referring to the central limit theorem, we assume that the resulting noise can be taken as zero-mean Gaussian with the standard deviation  $\sigma_r$  for the  $r$ -th plane,  $\varepsilon_r[k] \sim N(0, \sigma_r^2)$ .

The problem is to reconstruct discrete complex-valued wave field distributions at the object  $\mathbf{u}_0$  and sensor planes  $\mathbf{u}_r$  from the noisy intensity data  $\mathbf{o}_r$ .

## 1.2 Successive scenario of phase retrieval

The wave field  $\mathbf{u}_r$  can be generated by the object distribution  $\mathbf{u}_0$  (according to Eq. (2)) or by the wave field from any previous sensor planes (e.g.  $\mathbf{u}_s$ ,  $Z_s < Z_r$ )<sup>4</sup>:

$$\mathbf{u}_r = \mathbf{A}_{r,s} \cdot \mathbf{u}_s, \quad (4)$$

where  $\mathbf{A}_{r,s}$  denotes a propagation operator from the  $s$ -th to the  $r$ -th sensor plane. It means that given an initial guess for the phase at any measurement plane, the phase recovering is performed cyclically by the *successive* propagation from one sensor plane to another one.

One of the powerful circular phase-retrieval algorithm exploited this idea is the so-called *single-beam multiple-intensity phase reconstruction* (SBMIR) algorithm proposed in <sup>8,9</sup>.

According to Eq. (4) the recursive SBMIR algorithm (as it is defined in <sup>10</sup>) can be written in the form:

$$\hat{\mathbf{u}}_{\tau(p+1)} = \mathbf{A}_{\tau(p+1),\tau(p)} \cdot \left[ \frac{\sqrt{\mathbf{o}_r}}{|\hat{\mathbf{u}}_{\tau(p)}|} \circ \hat{\mathbf{u}}_{\tau(p)} \right], \quad p = 1, 2, \dots \quad (5)$$

In (5) the operations in the square brackets are elementwise, i.e.  $\sqrt{\mathbf{o}_r} / |\hat{\mathbf{u}}_{\tau(p)}|$  is a vector, obtained by the elementwise division of the vector  $\sqrt{\mathbf{o}_r}$  by  $|\hat{\mathbf{u}}_{\tau(p)}|$ .  $\circ$  denotes the Hadamard elementwise product, thus the result in the brackets is also a vector. Note that the multiplication  $\hat{\mathbf{u}}_{\tau(p)}$  by  $\sqrt{\mathbf{o}_r} / |\hat{\mathbf{u}}_{\tau(p)}|$  means a replacement of the magnitude of the calculated estimate at the  $\tau(p)$ -th sensor plane by the magnitude from the observations, keeping the phase of  $\hat{\mathbf{u}}_{\tau(p)}$ . It corresponds to GS.

The subindex  $\tau(p)$  shows the serial number of the observation plane, where the estimate of the wave field  $\hat{\mathbf{u}}_{\tau(p)}$  is calculated. There is no specific rule determining the sequence of the sensor planes in a set. For instance, the subindexes for the forward-forward (FF)  $\tau^{\text{FF}}(p)$  and forward-backward (FB)  $\tau^{\text{FB}}(p)$  algorithms (see Fig. 2 in <sup>10</sup>) can be found as follows:

$$\tau^{\text{FF}}(p) = \tau_K(p), \quad \tau^{\text{FB}}(p) = \tau_{2K-2}(p) - 2 \cdot \tau_K(\tau_{2K-2}(p)) \cdot \lfloor \chi_{\chi_K(\tau_{2K-2}(p))} \rfloor \quad (6)$$

Here  $\tau_K(p) = \text{mod}(p-1, K) + 1$ ,  $\chi_{\chi_K(p)} = \lfloor (p-1)/L \rfloor$ , where  $\text{mod}(a, b)$  means  $a \bmod b$ , and  $\lfloor \cdot \rfloor$  denotes the floor operation.

In this work we use SBMIR for comparison of the algorithm performance and resulting imaging obtained by different phase-retrieval algorithms.

### 1.3 Parallel scenario of phase retrieval

In contrast to *successive* algorithms, which are concentrated on the sensor planes, there is another interpretation of the wave field propagation. In the so-called *parallel* algorithms the object reconstruction is performed by aggregation and processing of a number of estimates for the object wave field.

In <sup>11</sup> we presented a multi-plane variation of GS, derived from the least square estimation of  $\mathbf{u}_0$  assuming that complex-valued estimates at the sensor plane are available. We keep the magnitude obtained from the sensor observations, and the phase at the measurement planes is recalculated iteratively:

$$\hat{\mathbf{u}}_r^{(p)} = \mathbf{A}_r \cdot \hat{\mathbf{u}}_0^{(p)}, \quad p = 1, 2, \dots$$

$$\hat{\mathbf{u}}_0^{(p+1)} = \left( \sum_{r=1}^K \mathbf{A}_r^H \mathbf{A}_r + \mu \mathbf{I} \right)^{-1} \times \sum_{r=1}^K \mathbf{A}_r^H \cdot \left[ \frac{\sqrt{\mathbf{o}_r}}{|\hat{\mathbf{u}}_r^{(p)}|} \circ \hat{\mathbf{u}}_r^{(p)} \right] = \sum_{r=1}^K \mathbf{B}^{-1} \cdot \mathbf{A}_r^H \cdot \left[ \frac{\sqrt{\mathbf{o}_r}}{|\hat{\mathbf{u}}_r^{(p)}|} \circ \hat{\mathbf{u}}_r^{(p)} \right], \quad (7)$$

where  $(\cdot)^H = [(\cdot)^*]^T$  is the Hermitian conjugate transpose of the propagation operator and  $\mathbf{B} = \sum_{r=1}^K \mathbf{A}_r^H \mathbf{A}_r + \mu \mathbf{I}$ . We denote the replacement of magnitudes in (7) similar to Eq. (5). For invertible propagation operators (e.g. ASD)  $\mathbf{A}_r^H \mathbf{A}_r = \mathbf{I}$  for all  $r$ . Then, for the regularization parameter  $\mu=0$  the matrix  $\mathbf{B} = \mathbf{K} \cdot \mathbf{I}$ , and the algorithm (7) can be rewritten as follows:

$$\begin{aligned}\hat{\mathbf{u}}_r^{(p)} &= \mathbf{A}_r \cdot \hat{\mathbf{u}}_0^{(p)}, \quad p = 1, 2, \dots \\ \hat{\mathbf{u}}_0^{(p+1)} &= \frac{1}{K} \sum_{r=1}^K \mathbf{A}_r^H \cdot \left[ \frac{\sqrt{\mathbf{o}_r}}{|\hat{\mathbf{u}}_r^{(p)}|} \circ \hat{\mathbf{u}}_r^{(p)} \right],\end{aligned}\quad (8)$$

The estimates of the object are calculated by the backward propagation of  $[\hat{\mathbf{u}}_{\tau(p)} \circ \sqrt{\mathbf{o}_r} / |\hat{\mathbf{u}}_{\tau(p)}|]$  using its multiplication by  $\mathbf{B}^{-1} \cdot \mathbf{A}_r^H$ . Finally,  $\hat{\mathbf{u}}_0^{(p+1)}$  is found using the summation of these object estimates.

The generalization for ill-posed matrices in (7) allows applying the parallel algorithms for non-focal sensor planes for an arbitrary distance  $z_r$ . The only unknown variable here is the object distribution and  $K$  wave fields at the sensor planes are calculated in *parallel* in order to find  $\mathbf{u}_0$ .

The algorithm developed in this paper has the *parallel* structure as well. Moreover, the idea of the *parallel* processing is widely exploited for phase recovering, e.g. in the relative problem of computer-generated holograms<sup>12,13</sup>.

## 2. AUGMENTED LAGRANGIAN (AL) ALGORITHM

A typical variational setting for the chosen Gaussian noise distribution can be presented in the form

$$J = \sum_{r=1}^K \frac{1}{2\sigma_r^2} \|\mathbf{o}_r - |\mathbf{u}_r|^2\|_2^2 + \mu \cdot \text{pen}(\mathbf{u}_0), \quad (9)$$

where  $\|\cdot\|_2^2$  is the Euclidian norm. The first summand in (9) is the quadratic fidelity term obtained as the minus logarithm of the Gaussian likelihood function corresponding to the model (3), and the second term is the regularization including *prior* information on the object distribution  $\mathbf{u}_0$  to be reconstructed.

There are different criteria used as fidelity terms in variational formulations for wave field reconstruction<sup>14,15,16</sup>. Typically, they can be presented in the form

$$J_\gamma = \sum_{r=1}^K \|\mathbf{o}_r^{\gamma/2} - |\mathbf{u}_r|^\gamma\|_2^2, \quad (10)$$

where  $\gamma$  is a parameter. In our work the choice of the criterion is made in favor of  $\gamma=2$  due to the maximum likelihood approach and the assumption that the noise is i.i.d. Gaussian.

$\mu$  is a regularization parameter controlling a balance in (9) between the accuracy of the observation fitting and a prior given by  $\text{pen}(\mathbf{u}_0)$ . In this paper we use a simple but quite efficient quadratic Tikhonov's penalty<sup>17</sup> given in the form  $\text{pen}(\mathbf{u}_0) = \|\mathbf{u}_0\|_2^2$ .

Based on the criterion  $J$  and the forward propagation model for *parallel* algorithms (2) we formulate the object wave field reconstruction as the following constrained optimization

$$\hat{\mathbf{u}}_0 = \arg \min_{\mathbf{u}_0} \sum_{r=1}^K \frac{1}{2\sigma_r^2} \|\mathbf{o}_r - |\mathbf{u}_r|^2\|_2^2 + \mu \cdot \|\mathbf{u}_0\|_2^2 \quad \text{subject to } \mathbf{u}_r = \mathbf{A}_r \cdot \mathbf{u}_0, \quad r=1, \dots, K \quad (11)$$

We are looking for the solution of this constrained problem using the augmented Lagrangian method<sup>18,19</sup>. As in<sup>20</sup> the complex-valued object wave field is reconstructed by minimization of the following criterion

$$L = \sum_{r=1}^K \frac{1}{\sigma_r^2} \left[ \frac{1}{2} \|\mathbf{o}_r - |\mathbf{u}_r|^2\|_2^2 + \frac{1}{\gamma_r} \|\mathbf{u}_r - \mathbf{A}_r \cdot \mathbf{u}_0\|_2^2 + \frac{2}{\gamma_r} \text{Re}\{\Lambda_r^H (\mathbf{u}_r - \mathbf{A}_r \cdot \mathbf{u}_0)\} \right] + \mu \cdot \|\mathbf{u}_0\|_2^2, \quad (12)$$

where the vectors  $\Lambda_r \in \mathbb{C}^{n \times 1}$  are the Lagrange multipliers. Here  $\gamma_r$  are positive penalty coefficients. The variations  $\mathbf{u}_0$ ,  $\{\mathbf{u}_r\}$  are separated into several blocks according to their roles, and the minimization of  $L$  is performed separately in different blocks<sup>21</sup>. This *alternating minimization* of  $L$  on  $\mathbf{u}_0$ ,  $\{\mathbf{u}_r\}$  results in the following iterative algorithm:

$$\begin{aligned}
& \text{Repeat for } t = 0, 1, 2, \dots \\
& \mathbf{u}_{r,t+1} \in \arg \min_{\{\mathbf{u}_r\}} L(\mathbf{u}_{0,t}, \{\mathbf{u}_r\}, \{\Lambda_{r,t}\}) \\
& \Lambda_{r,t+1} = \Lambda_{r,t} + \alpha_r \cdot (\mathbf{u}_{r,t+1} - \Lambda_{r,t} \cdot \mathbf{u}_{0,t}) \\
& \mathbf{u}_{0,t+1} \in \arg \min_{\mathbf{u}_0} L(\mathbf{u}_0, \{\mathbf{u}_{r,t}\}, \{\Lambda_{r,t}\}) \\
& \text{End on } t
\end{aligned} \tag{13}$$

## 2.1 Derivation of the augmented Lagrangian algorithm

The developed phase-retrieval algorithm is derived from the minimization of the criterion  $L$  (12) on  $\mathbf{u}_0$ ,  $\{\mathbf{u}_r\}$  and  $\{\Lambda_r\}$  separately.

### Minimization on $\{\mathbf{u}_r\}$ and $\{\Lambda_r\}$ .

The Lagrangian (12) is additive with respect to the vectors  $\{\mathbf{u}_r\}$ . According to the minimum condition  $\nabla_{\mathbf{u}_r} L = 0$  we obtain

$$\mathbf{u}_r[k] = \frac{(\mathbf{A}_r \cdot \mathbf{u}_0)[k] - \Lambda_r[k]}{\gamma_r (|\mathbf{u}_r[k]|^2 - \mathbf{o}_r[k]) + 1} := \frac{\eta_r[k]}{\kappa_r[k]} \tag{14}$$

Taking the module from the left and right sides of (14) we arrive at the cubic equation with respect to  $|\mathbf{u}_r[k]|$ :

$$|\mathbf{u}_r[k]|^3 + |\mathbf{u}_r[k]| \cdot \left( \frac{1}{\gamma_r} - \mathbf{o}_r[k] \right) - \text{sgn}(\kappa_r[k]) \cdot \frac{|\eta_r[k]|}{\gamma_r} = 0 \tag{15}$$

We are looking for a nonnegative real root of (15), denoted as  $|\tilde{\mathbf{u}}_r[k]|$ . It can be shown that such a root always exists: the equation (15) has a single real root or three real roots. If the nonnegative real solution is not unique, a proper solution is selected with respect to the minimum value of the corresponding  $r$ -th summand of  $L$ . Then, the complex-valued estimate of the wave field at the sensor plane is calculated according to Eq. (14) as

$$\mathbf{u}_{r,t+1}[k] = G(\mathbf{o}_r[k], \mathbf{u}_{0,t}[k], \mathbf{u}_{r,t}[k], \Lambda_{r,t}[k]), \tag{16}$$

where  $G$  is a notation for the nonlinear algorithm which gives the solution for (15) with the found  $|\tilde{\mathbf{u}}_r[k]|$  in the form<sup>20</sup>

$$\mathbf{u}_r[k] = \frac{(\mathbf{A}_r \cdot \mathbf{u}_0)[k] - \Lambda_r[k]}{\gamma_r (|\tilde{\mathbf{u}}_r[k]|^2 - \mathbf{o}_r[k]) + 1} \tag{17}$$

The gradient  $\nabla_{\Lambda_r} L = \mathbf{u}_r - \Lambda_r \cdot \mathbf{u}_0$ , then, the iterative update of the Lagrangian multipliers in the gradient direction is of the form

$$\Lambda_{r,t+1} = \Lambda_{r,t} + \alpha_r \cdot (\mathbf{u}_{r,t+1} - \Lambda_{r,t} \cdot \mathbf{u}_{0,t}) \tag{18}$$

### Minimization on $\mathbf{u}_0$ .

Firstly, let us assume that the object is arbitrary of the form  $\mathbf{u}_0 = |\mathbf{u}_0| \cdot \exp(j \cdot \varphi_0)$ . The minimum condition  $\partial L / \partial \mathbf{u}_0^* = 0$  gives the following analytical solution<sup>20</sup>

$$\mathbf{u}_0 = \left( \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \mathbf{A}_r^H \mathbf{A}_r + \mu \mathbf{I} \right)^{-1} \times \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \mathbf{A}_r^H \cdot [\mathbf{u}_r + \mathbf{\Lambda}_r] \quad (19)$$

In <sup>20</sup> the object estimation (19) is implemented with the double size F-DDT approach<sup>7</sup>. In this work the algorithm for the reconstruction of the *arbitrary*  $\mathbf{u}_0$  is used for comparison with the proposed phase-retrieval algorithm derived for phase objects.

Secondly, let us assume the object wave field distribution be of the form  $\mathbf{u}_0 = a \cdot \exp(j \cdot \boldsymbol{\varphi}_0)$ , where  $a > 0$  is a constant (phase object). Let  $a$  be known. Then, the minimum condition on  $\boldsymbol{\varphi}_0$  for (12) can be presented in the form:

$$\nabla_{\boldsymbol{\varphi}_0} L = 2 \operatorname{Im} \left\{ \mathbf{u}_0^* \circ \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \mathbf{A}_r^H (\mathbf{A}_r \mathbf{u}_0 - \mathbf{u}_r - \mathbf{\Lambda}_r) \right\} = 0 \quad (20)$$

The iterative gradient descent algorithm for estimation of  $\boldsymbol{\varphi}_0$  has a form:

$$\boldsymbol{\varphi}_{0,t}^{(p+1)} = \boldsymbol{\varphi}_{0,t}^{(p)} - 2\xi \cdot \operatorname{Im} \left\{ [a \cdot \exp(-j\boldsymbol{\varphi}_{0,t}^{(p)})] \circ \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \mathbf{A}_r^H (\mathbf{A}_r [a \cdot \exp(j\boldsymbol{\varphi}_{0,t}^{(p)})] - \mathbf{u}_{r,t+1} - \mathbf{\Lambda}_{r,t}) \right\} \quad (21)$$

Note that  $\|\mathbf{u}_0\|^2 = a^2 \cdot [\exp(-j\boldsymbol{\varphi}_0)^T \cdot \exp(j\boldsymbol{\varphi}_0)] = a^2 \cdot n$ , thus the penalty term disappears in Eqs. (20), (21). If the magnitude of the object is known, we use its true value, denoted as  $a_0$ . If the constant amplitude  $a$  is unknown, then the amplitude estimate is calculated after the update of the phase. The analytical solution for the magnitude estimate  $\hat{a}$  can be obtained from the condition  $\nabla_a L = 0$  in the form:

$$\hat{a} = \left( \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \operatorname{Re} \{ \exp(-j\boldsymbol{\varphi}_0)^T \cdot \mathbf{A}_r^H (\mathbf{u}_r + \mathbf{\Lambda}_r) \} \right) / \left( \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \|\mathbf{A}_r \exp(j\boldsymbol{\varphi}_0)\|^2 + \mu n \right) \quad (22)$$

## 2.2 AL algorithm for phase objects

The proposed phase-retrieval algorithm for *phase objects* is obtained using Eqs. (16), (18), (21) and (22):

Repeat for  $t = 0, 1, 2, \dots$

$$\mathbf{u}_{r,t+1/2} = \mathbf{A}_r \cdot \mathbf{u}_{0,t}, \quad r = 1, \dots, K$$

$$\mathbf{u}_{r,t+1}[k] = G(\mathbf{o}_r[k], \mathbf{u}_{0,t}[k], \mathbf{u}_{r,t+1/2}[k], \mathbf{\Lambda}_{r,t}[k]), \quad r = 1, \dots, K \quad (23)$$

$$\mathbf{\Lambda}_{r,t+1} = \mathbf{\Lambda}_{r,t} + \alpha_r \cdot (\mathbf{u}_{r,t+1} - \mathbf{u}_{r,t+1/2}), \quad r = 1, \dots, K$$

$$\boldsymbol{\varphi}_{0,t}^{(p+1)} = \boldsymbol{\varphi}_{0,t}^{(p)} - \xi \cdot \nabla_{\boldsymbol{\varphi}_0} L(a, \boldsymbol{\varphi}_{0,t}^{(p)}, \mathbf{u}_{r,t+1}, \mathbf{\Lambda}_{r,t}), \quad p = 0, 1, 2, \dots, P$$

$$a_{t+1} = \left( \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \operatorname{Re} \{ \exp(-j\boldsymbol{\varphi}_{0,t}^{(p)})^T \cdot \mathbf{A}_r^H (\mathbf{u}_{r,t+1} + \mathbf{\Lambda}_{r,t}) \} \right) / \left( \sum_{r=1}^K \frac{1}{\gamma_r \sigma_r^2} \|\mathbf{A}_r \exp(j\boldsymbol{\varphi}_{0,t}^{(p)})\|^2 + \mu n \right)$$

$$\mathbf{u}_{0,t+1} = \begin{cases} a_0 \cdot \exp(j\boldsymbol{\varphi}_{0,t}^{(p)}), & a \text{ is known} \\ a_{t+1} \cdot \exp(j\boldsymbol{\varphi}_{0,t}^{(p)}), & a \text{ is unknown} \end{cases}$$

End on  $t$

We denote this algorithm as AL-Ph. This abbreviation highlights that this variation of the original AL algorithm uses *prior* information on the structure of the object distribution.

Here  $\zeta$  is the step-size for the gradient descent algorithm. Both in <sup>20</sup> and in (23) we use the initial guess (for  $t=0$ ) which concerns the object plane distribution  $\mathbf{u}_{0,0}$  (for instance,  $\mathbf{u}_{0,0}=\frac{1}{2}\cdot\mathbf{1}_{n\times 1}$ ) and Lagrange multipliers  $\Lambda_{r,0}$  (in our experiments  $\Lambda_{r,0}[k]=0$ ).

### 3. NUMERICAL SIMULATIONS

In our numerical simulations we consider the reconstruction of the object wave field of the form  $\mathbf{u}_0=a\cdot\exp(j\cdot\varphi_0)$ . If it is not specified the constant object amplitude is known and its true value  $a_0=1$ . We assume that  $\varphi_0[k]=\pi\cdot(w[k]-\frac{1}{2})$ , where  $w[k]$  is a test-image used for the phase modeling. For the simulation experiments the following square ( $N\times N$ ) test-images are exploited: binary *chessboard* ( $128\times 128$ ) and gray-scale smooth *Mexican Hat* ( $200\times 200$ )

$$w[k] = \frac{4}{3\sqrt{\pi}} \left( \left\| \frac{k}{40} \right\|^2 - 2 \right) \cdot \exp\left(-\frac{1}{2} \left\| \frac{k}{40} \right\|^2\right) \quad (24)$$

calculated on the 2D integer grid  $-100\leq(k_1,k_2)\leq 99$ . These test-images are normalized that  $0.1\leq w[k]\leq 1$ .

The pixelated models for the object and sensor planes have square pixels  $\Delta\times\Delta$ ,  $\Delta=6.7\mu m$  with 100% fill factors. The wavelength  $\lambda=532nm$  corresponding to a Nd:YAG green laser. The "in-focus" distance for the considered lensless scenario is calculated as  $z_f=N\cdot\Delta^2/\lambda$  (see <sup>22</sup>). The distance from the object plane to the first sensor plane  $z_1$  is expressed through this "in-focus" distance as  $z_1=d\cdot z_f$ , where the parameter  $d$  is varying in the interval  $[0.5, 3]$ . The object reconstruction is considered for several number of measurement planes  $K=5, 10$ . The algorithm performance is presented for the fixed distances between the sensor planes  $\Delta_z=2mm$ . If it is not specified, all results are presented for a low noise level ( $\sigma=0.05$ ) obtained with 100 iterations.

For simplicity, we assume that  $\alpha_r=\alpha$ ,  $\gamma_r=\gamma$ ,  $\sigma_r^2=\sigma^2$  for all  $r$ . According to the analysis presented in <sup>20,23</sup> the used parameters for AL based algorithm are as follows:

- The recommended penalty coefficient  $\gamma=10$ ;
- The step-size parameter for the Lagrange multipliers  $\alpha=1$ ;
- For noisy data  $\sigma=0.05$  we take the regularization parameter  $\mu=0.01$ . Note,  $\mu$  is fixed for different  $d$  and  $K$ ;
- The step-size  $\zeta=0.1$  and the number of iterations  $P=40$ .

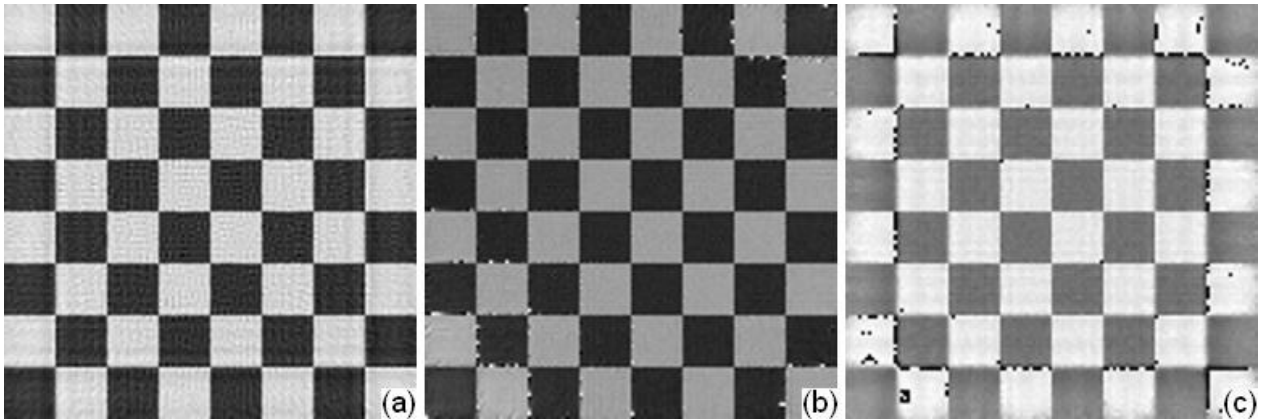


Figure 1. Comparison of the phase reconstructions obtained by (a)  $AL^{20}$ , RMSE=0.2; (b) AL-Ph, RMSE=0.28 and (c) SBMIR, RMSE=0.72. *chessboard* test-image,  $K=5$ ,  $d=1$ , noisy data  $\sigma=0.05$ .

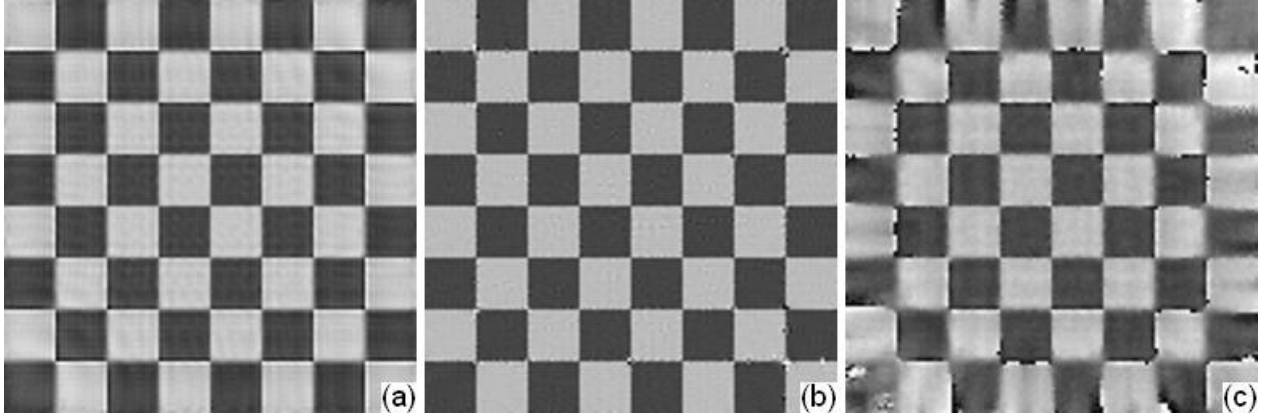


Figure 2. Comparison of the phase reconstructions obtained by (a)  $AL^{20}$ ,  $RMSE=0.26$ ; (b)  $AL-Ph$ ,  $RMSE=0.17$  and (c)  $SBMIR$ ,  $RMSE=0.66$ . *chessboard* test-image,  $K=5$ ,  $d=2$ , noisy data  $\sigma=0.05$ .

The reconstruction accuracy for the phase estimate  $\hat{\varphi}_0$  is characterized via the root-mean-square error (*RMSE*) criterion. The phase from the intensity measurements can be reconstructed up to a constant only. In order to eliminate this ambiguity, the *RMSE* values are calculated for  $\tilde{\varphi}_0 - \varphi_0$ . Here  $\tilde{\varphi}_0 = \hat{\varphi}_0 - c$ , where  $c$  is the best constant correcting the accuracy of the phase reconstruction.

Firstly, let us consider the phase reconstruction for the binary *chessboard* test-image. In Fig. 1 and Fig. 2 we compare the reconstructed phases  $\tilde{\varphi}_0$  obtained for  $d=1$  and  $d=2$ , respectively. These phase estimates are obtained by the following algorithms: (from the left to the right)  $AL$  with the double size F-DDT approach<sup>7</sup>,  $AL-Ph$  and  $SBMIR$  with the FF algorithm<sup>10</sup> (the difference in *RMSE* values comparing with  $FB$  is not valuable).

The visual advantage of the  $AL$  based algorithm is obvious: the phase reconstructions by  $AL$  and  $AL-Ph$  are quite sharp, while the  $SBMIR$  reconstruction is blurred (especially on the boarder) and significantly destroyed.  $AL-Ph$  estimates are more accurate comparing with  $AL$ : sharper boarders of the blocks of *chessboard*, smaller oscillations within blocks.

The phase estimates obtained by  $SBMIR$  can be strongly corrupted by the wrapping effects. Moreover, these phase wrapping effects may lead to an erroneous phase reconstruction, when  $SBMIR$  fails (see e.g. Table 1 in case of  $d=3$ ). In Fig. 1(c) this wrapping effect can be seen as dark spots. Moreover, the same wrapping effects may degrade the  $AL-Ph$  phase reconstruction as well, but in a less degree (see bright spots in Fig.2(b)). Therefore, despite a good imaging for  $d=1$  (with respect to  $d=2$ ) the *RMSE* values for  $AL-Ph$  and  $SBMIR$  are high. Note there is no such phase wrapping disturbances for  $AL$ . Overall the reconstruction accuracy of the  $AL$  based algorithms is significantly better (two times and more) with respect to  $SBMIR$  for different distances.

The *RMSE* values for the reconstructed  $\tilde{\varphi}_0$  (*chessboard* test-image) obtained by the considering phase-retrieval algorithms are presented for noiseless ( $\sigma=0$ ) and noisy data ( $\sigma=0.05$ ) in Table 1 and Table 2, respectively.

Table 1. The quantitative comparison of the phase reconstruction accuracy obtained by considering phase-retrieval algorithms, *chessboard* test-image,  $K=5$ , noiseless data  $\sigma=0$ .

algorithm \ $d$	0.5	0.75	1	1.25	1.5	2	3
$AL^{20}$	0.04	0.1	0.21	0.24	0.27	0.52	1.16
$AL-Ph$	0.006	0.014	0.13	0.007	0.03	0.13	0.44
$SBMIR$	0.29	3.17	0.56	0.82	2.97	0.59	2.15



Table 2. The quantitative comparison of the phase reconstruction accuracy obtained by considering phase-retrieval algorithms, *chessboard* test-image,  $K=5$ , noisy data  $\sigma=0.05$ .

algorithm \ $d$	0.5	0.75	1	1.25	1.5	2	3
AL <sup>20</sup>	0.09	0.13	<b>0.2</b>	0.23	0.24	<b>0.26</b>	0.87
AL-Ph	0.06	0.06	<b>0.28</b>	0.08	0.1	<b>0.17</b>	0.48
SBMIR	0.33	3.15	<b>0.72</b>	0.85	2.88	<b>0.66</b>	1.93

It is found that for noiseless data ( $\sigma=0$ ) the advantage of AL-Ph may be very large: more than 10 times better in RMSE values for  $d \leq 2$  comparing with AL.

It can be seen that for noisy data ( $\sigma=0.05$ ) the accuracy of AL-Ph algorithms is from 30% up to two times better comparing with AL for various  $K$  and  $d$ . Note that the reconstruction accuracy highly depends on proper algorithm parameters. Here  $\{\gamma, \mu\}$  are fixed for different  $d$ .

The convergence rates of the AL, AL-Ph and SBMIR algorithms for the phase reconstruction are compared in Fig. 3. Initially SBMIR converges very fast. After the first 30 iterations, where the accuracies of the considering algorithms are close, the convergence of SBMIR becomes slow. AL based algorithms converge more regular; but finally, after 100 iterations, show a significant advantage in the achieved RMSE values. It is found that after some initial steps AL-Ph converges faster than AL and results in a better reconstruction quality.

The above results are obtained assuming that the constant object magnitude  $a$  is known. In this case, at each iteration step the object magnitude estimate is always equal to the true value  $a_0=1$ . In Fig. 4 we show the convergence of the object magnitude estimate  $\hat{a}$  for different  $a_0=\{0.85, 1, 1.25\}$ . It is found that a larger number of observations  $K$  leads to a faster convergence, but results in significant oscillations around  $a_0$ . Moreover, these oscillations are larger for larger  $a_0$ .

Further, we consider the reconstruction of the smooth object phase distribution *Mexican Hat* (24), contrary to the discontinuous *chessboard*. In Fig. 5 we show the cross-sections of the true phase and its normalized reconstructions, obtained by the AL, AL-Ph and SBMIR algorithms.

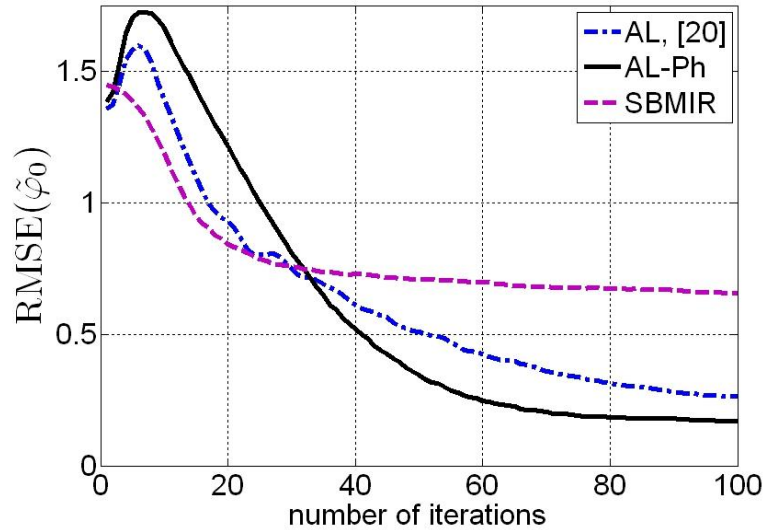


Figure 3. The convergence rates of AL, AL-Ph and SBMIR algorithms for the test presented in Fig. 2.

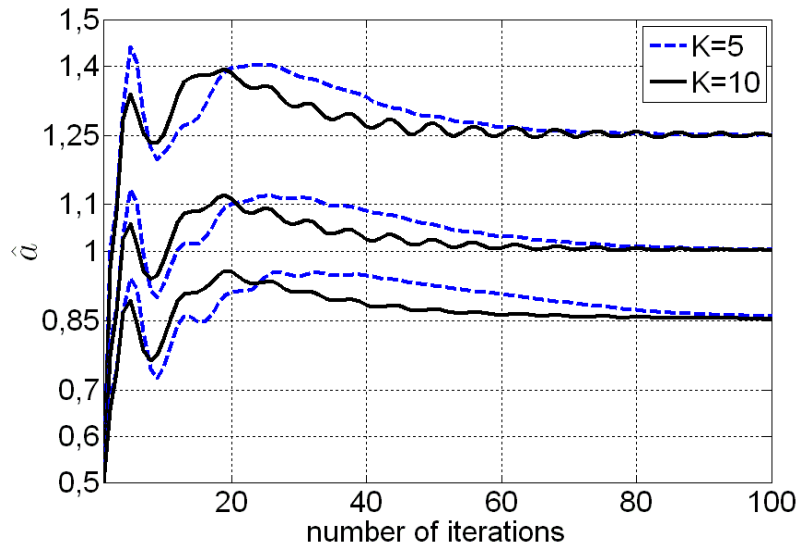


Figure 4. The convergence rates of the object magnitude estimate for the *chessboard* test-image,  $d=2$ , noisy data  $\sigma=0.05$ .

The AL-Ph algorithm demonstrates the reconstruction of the almost complete shape with the best fill to the object. The improvement of AL-Ph with respect to other algorithms is clear on the borders. It can be seen that 1000 iterations is not enough for SBMIR to fulfill the phase recovering: the hollow of *Mexican Hat* is not separated well. Moreover, the AL and SBMIR algorithms have significant errors on the borders. In addition to a better RMSE value the proposed AL-Ph algorithm converges (for smooth phase objects) much faster than SBMIR and AL.

The Matlab code used for the simulation experiments, more numerical reconstruction results and the discussions are available on the website<sup>23</sup>.

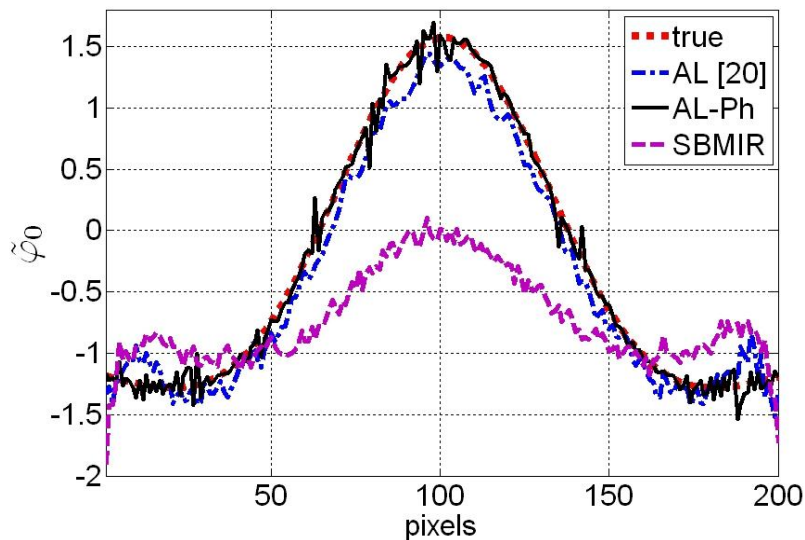


Figure 5. The cross-sections of the true phase (dotted curve) and phase reconstructions obtained for the *Mexican Hat* test-image. The best result (solid line) corresponds to AL-Ph, RMSE=0.09. The dash-dotted curve corresponds to AL<sup>20</sup>, RMSE=0.19, and dashed curve corresponds to SBMIR, RMSE=0.5. 1000 iterations,  $d=2$ ,  $K=5$ , noisy data  $\sigma=0.05$ .

## 4. CONCLUSION

In this work we present a novel variational formulation for the phase-retrieval problem based on the augmented Lagrangian technique. Being the maximum likelihood style, this setting takes into consideration the Gaussian noise distribution and *prior* information on the object: the developed phase-retrieval algorithm is treated as a particular case of AL<sup>20</sup> for *phase* objects.

The proposed algorithm demonstrates a significant improvement of the reconstruction quality, imaging and convergence rate comparing with AL both for noiseless and noisy observation data.

## REFERENCES

- [1] Goodman, J. W., [Introduction to Fourier Optics], 3rd ed., Roberts & Company, Englewood, (2005).
- [2] Gerchberg R. W. and Saxton, W. O., "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* 35, 237-246 (1972).
- [3] Fienup, J. R., "Phase retrieval algorithms: A comparison," *Appl. Opt.* 21(15), 2758-2769 (1982).
- [4] Misell, D. L., "A method for the solution of the phase problem in electron microscopy," *J. Phys. D* 6, L6--L9 (1973).
- [5] Shen, F. and Wang, A., "Fast-Fourier-transform based numerical integration method for the Rayleigh-Sommerfeld diffraction formula," *Appl. Opt.* 45(6), 1102-1110 (2006).
- [6] Katkovnik, V., Migukin, A. and Astola, J., "Backward discrete wave field propagation modeling as an inverse problem: toward perfect reconstruction of wave field distributions," *Appl. Opt.* 48(18), 3407-3423 (2009).
- [7] Katkovnik, V., Astola, J. and Egiazarian, K., "Discrete diffraction transform for propagation, reconstruction, and design of wavefield distributions," *Appl. Opt.* 47(19), 3481-3493 (2008).
- [8] Pedrini, G., Osten, W. and Zhang, Y., "Wave-front reconstruction from a sequence of interferograms recorded at different planes," *Opt. Lett.* 30(8), 833-835 (2005).
- [9] Almoró, P., Pedrini, G. and Osten, W., "Complete wavefront reconstruction using sequential intensity measurements of a volume speckle field," *Appl. Opt.* 45(34), 8596-8605 (2006).
- [10] Almoró, P., Maallo, A. M. and Hanson, S. "Fast-convergent algorithm for speckle-based phase retrieval and a design for dynamic wavefront sensing," *Appl. Opt.* 48(8), 1485-1493 (2009).
- [11] Migukin, A., Katkovnik, V. and Astola, J., "Multiple plane phase retrieval based on inverse regularized imaging and discrete diffraction transform," *AIP Conf. Proc.* 1236, ICAPMMOI, (2010).
- [12] Haist, T., Schönleber, M. and Tiziani, H. J., "Computer-generated holograms from 3D-objects written on twisted-nematic liquid crystal displays," *Opt. Commun.* 140, 299-308 (1997).
- [13] Sinclair, G., Leach, J., Jordan, P., Gibson, G., Yao, E., Laczik, Z., Padgett, M. and Courtial, J., "Interactive application in holographic optical tweezers of a multi-plane Gerchberg-Saxton algorithm for three-dimensional light shaping," *Opt. Express* 12(8), 1665-1670 (2004)
- [14] Guizar-Sicairos, M. and Fienup, J. R., "Phase retrieval with transverse translation diversity: a nonlinear optimization approach," *Opt. Express* 16(10), 7264-7278 (2008).
- [15] Xiaojun-Hu, Shengyi-Li, and Yulie-Wu, "Resolution-enhanced subpixel phase retrieval method," *Appl. Opt.* 47(32), 6079-6087 (2008).
- [16] Brady, G. R., Guizar-Sicairos, M. and Fienup, J. R., "Optical wavefront measurement using phase retrieval with transverse translation diversity," *Opt. Express* 17(2), 624-639 (2009).
- [17] Tikhonov, A.N. and Arsenin, V.Y., [Solution of Ill-Posed Problems], Wiley, (1977).
- [18] Hestenes, M. R. "Multiplier and Gradient Methods," *JOTA.* 4, 303-320 (1969).
- [19] Powell, M. J. D., "A method for nonlinear constraints in minimization problems," in R. Fletcher, ed. [Optimization], Academic Press, 283-298 (1969).
- [20] Migukin, A., Katkovnik, V. and Astola, J., "Wave field reconstruction from multiple plane intensity-only data: Augmented Lagrangian algorithm", *JOSA A*, in press.
- [21] Bertsekas, D. P. and Tsitsiklis, J. N., [Parallel and distributed computation: numerical methods], Prentice-Hall, Upper Saddle River (1989).
- [22] Kreis, Th., [Handbook of Holographic Interferometry, Optical and Digital Methods], Wiley-VCH, (2005).
- [23] [www.cs.tut.fi/~lasip/DDT/](http://www.cs.tut.fi/~lasip/DDT/)