Supplemental Document

Optics Letters

Projected refractive index framework for multi-wavelength phase retrieval: supplement

YUNHUI GAO AND LIANGCAI CAO*

State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instruments, Tsinghua University, Beijing 100084, China
Corresponding author: clc@tsinghua.edu.cn

Supplement DOI: https://doi.org/10.6084/m9.figshare.21424275

Parent Article DOI: https://doi.org/10.1364/OL.476707

Projected refractive index framework for multi-wavelength phase retrieval: supplemental document

1. SOLUTION TO THE INVERSE PROBLEM

A. Proximal gradient method

The inverse problem of Eq. (10) can be equivalently expressed as

$$\min_{\boldsymbol{u}} \underbrace{\frac{1}{2K}\sum_{i=1}^{K}\left\|\left|\boldsymbol{A}_i \exp\left[\mathrm{j}\frac{\lambda_1}{\lambda_i}\mathrm{Re}(\boldsymbol{u})-\mathrm{Im}(\boldsymbol{u})\right]\right|-\boldsymbol{y}_i\right\|_2^2}_{F(\boldsymbol{u})}+\underbrace{\tau\|\boldsymbol{D}\boldsymbol{u}\|_1+I_C(\boldsymbol{u})}_{R(\boldsymbol{u})}, \tag{S1}$$

where $\tau > 0$ is a regularization parameter, $\boldsymbol{D} \in \mathbb{R}^{2n\times n}$ denotes the finite difference operator, and I_C denotes the indicator function of a physical constraint set C. The term $\|\boldsymbol{D}\boldsymbol{u}\|_1$ can be more explicitly defined as

$$\|\boldsymbol{D}\boldsymbol{u}\|_1 = \mathrm{TV}(\boldsymbol{U}) = \sum_{i=1}^{n_\xi-1}\sum_{j=1}^{n_v}|U_{i+1,j}-U_{i,j}|+\sum_{i=1}^{n_\xi}\sum_{j=1}^{n_v-1}|U_{i,j+1}-U_{i,j}|, \tag{S2}$$

where $\boldsymbol{U} \in \mathbb{C}^{n_\xi\times n_v}$ with $n_\xi n_v = n$ denotes the non-vectorized two-dimensional image corresponding to $\boldsymbol{u}$. $F(\boldsymbol{u})$ is the data-fidelity function and $R(\boldsymbol{u})$ is the regularization function. The data-fidelity function penalizes any deviation from the forward model, whereas the regularization function is introduced to incorporate prior knowledge of the object and thus push the solution towards the desired direction.

The problem of Eq. (S1) is a composite optimization problem consisting of a smooth data-fidelity function $F(\boldsymbol{u})$ and a non-smooth regularization function $R(\boldsymbol{u})$. Therefore, we adopt an accelerated variant of the proximal gradient method to solve this problem [1, 2]. The algorithm proceeds as follows:

$$\boldsymbol{u}^{(t)} = \mathrm{prox}_{\gamma R}\left(\boldsymbol{v}^{(t-1)}-\gamma\nabla_{\boldsymbol{v}}F(\boldsymbol{v}^{(t-1)})\right), \tag{S3}$$

$$\boldsymbol{v}^{(t)} = \boldsymbol{u}^{(t)}+\beta_t(\boldsymbol{u}^{(t)}-\boldsymbol{u}^{(t-1)}), \tag{S4}$$

where $t = 1, 2, \ldots$ denotes the iteration number, $\gamma > 0$ is the step size, and $\boldsymbol{v}^{(0)} = \boldsymbol{u}^{(0)}$. We set the extrapolation parameter $\beta_t = t/(t+3)$ as is suggested by [2]. The gradient of the fidelity function is calculated in the next subsection as

$$\begin{aligned}\nabla_{\boldsymbol{u}}F(\boldsymbol{u}) = \frac{1}{K}\sum_{i=1}^{K}\Bigg[&-\frac{\mathrm{j}}{4}\left(\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\bar{\boldsymbol{x}}_i)\boldsymbol{A}_i^{\mathsf{H}}\mathrm{diag}\left(\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)\\ &-\frac{\mathrm{j}}{4}\left(-\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\boldsymbol{x}_i)\boldsymbol{A}_i^{\mathsf{T}}\mathrm{diag}\left(\overline{\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)\Bigg],\end{aligned} \tag{S5}$$

where $\boldsymbol{x}_i = \exp\left[(\mathrm{j}\lambda_1/\lambda_i)\mathrm{Re}(\boldsymbol{u})-\mathrm{Im}(\boldsymbol{u})\right]$, $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$ denote the transpose and Hermitian of a matrix, respectively. The non-smooth regularization function $R(\boldsymbol{u})$ is minimized via its proximity operator, which is defined as

$$\mathrm{prox}_{\gamma R}(\boldsymbol{u}) = \operatorname*{argmin}_{\boldsymbol{w}}\frac{1}{2}\|\boldsymbol{w}-\boldsymbol{u}\|_2^2+\gamma R(\boldsymbol{w}). \tag{S6}$$

In most cases, the proximity operator has no closed-form solutions, and an iterative solver should be invoked. When $R(\boldsymbol{u})$ takes the particular form of Eq. (S1), an efficient algorithm is readily available [3].

Because the optimization problem of Eq. (S1) is essentially non-convex, a naive implementation of the accelerated proximal gradient method does not ensure convergence. In light of this, we enforce a monotone decrease of the objective function. Whenever monotonicity is violated, a basic proximal gradient step with $\beta_t = 0$ is applied and the step size is reduced.

We experimentally observe that when the maximum OPL of the object is relatively large, the algorithm tends to get stuck in a false local minima. To address this problem, we introduce an automatic tuning scheme for the regularization parameter. Although an optimal tuning scheme still remains to be studied in the future, we found in this work that the following rule generally works well for retrieving most objects:

$$\tau^{(t)} = \frac{\tau_1 - \tau_2}{1 + \exp\left[\alpha(t/T - 1/2)\right]} + \tau_2, \tag{S7}$$

where τ_1 and τ_2 denote the regularization parameters at the beginning stage and the final stage, respectively. T denotes the total iteration number, and α is a parameter controlling the changing rate of τ. τ_1 controls the accuracy of the initial estimate, whereas τ_2 controls the consistency with the forward model. The specific choice of τ_1 depends on the phase shift induced by the object. A larger phase shift requires a larger τ_1 to locate the vicinity of the global minima. The choice of τ_2 is determined according to the noise level of the measurement. A larger noise level requires a larger τ_2 to remove the artifacts.

B. Gradient calculation

To deal with complex-valued gradient calculation, we adopt the CR-calculus for convenience. Readers may refer to [4] for a detailed introduction. The i-th term of the fidelity function is

$$F_i = \frac{1}{2}\left\| |\boldsymbol{A}_i \boldsymbol{x}_i| - \boldsymbol{y}_i \right\|_2^2 = \frac{1}{2}\left\| |\boldsymbol{A}_i \exp\left[(\mathrm{j}\lambda_1/\lambda_i)\mathrm{Re}(\boldsymbol{u}) - \mathrm{Im}(\boldsymbol{u})\right]| - \boldsymbol{y}_i \right\|_2^2. \tag{S8}$$

According to the chain rule of Wirtinger derivatives, we have

$$\frac{\partial F_i}{\partial \boldsymbol{u}} = \frac{\partial F_i}{\partial \boldsymbol{x}_i}\frac{\partial \boldsymbol{x}_i}{\partial \boldsymbol{u}} + \frac{\partial F_i}{\partial \bar{\boldsymbol{x}}_i}\frac{\partial \bar{\boldsymbol{x}}_i}{\partial \boldsymbol{u}}, \tag{S9}$$

where [5]

$$\frac{\partial F_i}{\partial \boldsymbol{x}_i} = \frac{1}{2}(|\boldsymbol{A}_i\boldsymbol{x}_i| - \boldsymbol{y}_i)^{\mathsf{H}}\mathrm{diag}\left(\frac{\overline{\boldsymbol{A}_i\boldsymbol{x}_i}}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)\boldsymbol{A}_i, \tag{S10}$$

$$\frac{\partial F_i}{\partial \bar{\boldsymbol{x}}_i} = \frac{1}{2}(|\boldsymbol{A}_i\boldsymbol{x}_i| - \boldsymbol{y}_i)^{\mathsf{T}}\mathrm{diag}\left(\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)\bar{\boldsymbol{A}}_i. \tag{S11}$$

Also notice that

$$\mathrm{Re}(\boldsymbol{u}) = \frac{\boldsymbol{u} + \bar{\boldsymbol{u}}}{2}, \quad \mathrm{Im}(\boldsymbol{u}) = \frac{\boldsymbol{u} - \bar{\boldsymbol{u}}}{2\mathrm{j}}. \tag{S12}$$

Thus, we have

$$\begin{aligned}
\frac{\partial \boldsymbol{x}_i}{\partial \boldsymbol{u}} &= \frac{\partial}{\partial \boldsymbol{u}}\exp\left[\mathrm{j}\frac{\lambda_1}{\lambda_i}\mathrm{Re}(\boldsymbol{u}) - \mathrm{Im}(\boldsymbol{u})\right] \\
&= \mathrm{diag}(\boldsymbol{x}_i)\frac{\partial}{\partial \boldsymbol{u}}\left[\mathrm{j}\frac{\lambda_1}{\lambda_i}\mathrm{Re}(\boldsymbol{u}) - \mathrm{Im}(\boldsymbol{u})\right] \\
&= \mathrm{diag}(\boldsymbol{x}_i)\frac{\partial}{\partial \boldsymbol{u}}\left(\mathrm{j}\frac{\lambda_1}{\lambda_i}\frac{\boldsymbol{u}+\bar{\boldsymbol{u}}}{2} - \frac{\boldsymbol{u}-\bar{\boldsymbol{u}}}{2\mathrm{j}}\right) \\
&= \mathrm{diag}(\boldsymbol{x}_i)\frac{\partial}{\partial \boldsymbol{u}}\left[\frac{\mathrm{j}}{2}\left(\frac{\lambda_1}{\lambda_i}+1\right)\boldsymbol{u} + \frac{\mathrm{j}}{2}\left(\frac{\lambda_1}{\lambda_i}-1\right)\bar{\boldsymbol{u}}\right] \\
&= \frac{\mathrm{j}}{2}\left(\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\boldsymbol{x}_i), \qquad \text{(S13)} \\
\frac{\partial \bar{\boldsymbol{x}}_i}{\partial \boldsymbol{u}} &= \frac{\partial}{\partial \boldsymbol{u}}\exp\left[-\mathrm{j}\frac{\lambda_1}{\lambda_i}\mathrm{Re}(\boldsymbol{u}) - \mathrm{Im}(\boldsymbol{u})\right] \\
&= \frac{\mathrm{j}}{2}\left(-\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\bar{\boldsymbol{x}}_i). \qquad \text{(S14)}
\end{aligned}$$

Substituting Eqs. (S10)-(S14) into Eq. (S9), we obtain

$$\begin{aligned}\frac{\partial F_i}{\partial \boldsymbol{u}} =& \frac{\mathrm{j}}{4}\left(\frac{\lambda_1}{\lambda_i}+1\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)^{\mathsf{H}}\mathrm{diag}\left(\frac{\overline{\boldsymbol{A}_i\boldsymbol{x}_i}}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)\boldsymbol{A}_i\mathrm{diag}(\boldsymbol{x}_i) \\ &+\frac{\mathrm{j}}{4}\left(-\frac{\lambda_1}{\lambda_i}+1\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)^{\mathsf{T}}\mathrm{diag}\left(\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)\bar{\boldsymbol{A}}_i\mathrm{diag}(\bar{\boldsymbol{x}}_i),\end{aligned} \tag{S15}$$

where $\boldsymbol{x}_i = \exp\left[(\mathrm{j}\lambda_1/\lambda_i)\mathrm{Re}(\boldsymbol{u}) - \mathrm{Im}(\boldsymbol{u})\right]$. Therefore, the Wirtinger gradient is

$$\begin{aligned}\nabla_{\boldsymbol{u}} F_i &= \left(\frac{\partial F_i}{\partial \boldsymbol{u}}\right)^{\mathsf{H}} \\ &= -\frac{\mathrm{j}}{4}\left(\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\bar{\boldsymbol{x}}_i)\boldsymbol{A}_i^{\mathsf{H}}\mathrm{diag}\left(\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i) \\ &\quad -\frac{\mathrm{j}}{4}\left(-\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\boldsymbol{x}_i)\boldsymbol{A}_i^{\mathsf{T}}\mathrm{diag}\left(\frac{\overline{\boldsymbol{A}_i\boldsymbol{x}_i}}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i).\end{aligned} \tag{S16}$$

By summing over all wavelengths, we arrive at

$$\begin{aligned}\nabla_{\boldsymbol{u}} F(\boldsymbol{u}) =& \frac{1}{K}\sum_{i=1}^{K}\nabla_{\boldsymbol{u}} F_i(\boldsymbol{u}) \\ =& \frac{1}{K}\sum_{i=1}^{K}\left[-\frac{\mathrm{j}}{4}\left(\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\bar{\boldsymbol{x}}_i)\boldsymbol{A}_i^{\mathsf{H}}\mathrm{diag}\left(\frac{\boldsymbol{A}_i\boldsymbol{x}_i}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)\right. \\ &\left.-\frac{\mathrm{j}}{4}\left(-\frac{\lambda_1}{\lambda_i}+1\right)\mathrm{diag}(\boldsymbol{x}_i)\boldsymbol{A}_i^{\mathsf{T}}\mathrm{diag}\left(\frac{\overline{\boldsymbol{A}_i\boldsymbol{x}_i}}{|\boldsymbol{A}_i\boldsymbol{x}_i|}\right)(|\boldsymbol{A}_i\boldsymbol{x}_i|-\boldsymbol{y}_i)\right].\end{aligned}$$

2. IMPLEMENTATION OF THE CONVENTIONAL MULTI-WAVELENGTH PHASE RETRIEVAL ALGORITHM

To ensure a fair comparison, we modified the conventional multi-wavelength phase retrieval algorithm [6] so that it can be applied to the specific forward model as described in the main text. The flowchart of the algorithm is shown in Fig. S1.

Free-space propagation is numerically calculated using circular convolution with the Fresnel propagation kernel, which can be implemented efficiently via Fast Fourier Transforms. To avoid boundary artifacts, the object field distribution is assumed to be confined within the sensor area, as is highlighted by the red dashed box in Fig. S1. In practice, this support constraint can be physically implemented with a square aperture. The algorithm proceeds as follows:

(1) Obtain an initial guess of the amplitude and phase of the object.

(2) Forward propagate the wavefield to the image plane at wavelength λ_i.

(3) Replace the amplitude within the sensor area by square root of the measured intensity $\boldsymbol{y}_i$, the amplitude distribution outside the sensor area and the phase distribution remain unchanged.

(4) Backward propagate the wavefield to the object plane at wavelength λ_i.

(5) Enforce object support and absorption constraints to the amplitude. The phase is converted to the next wavelength via $\boldsymbol{\phi}_{i+1} = (\lambda_i/\lambda_{i+1})\boldsymbol{\phi}_i$. When the object OPL is larger than the wavelength, phase unwrapping is performed.

(6) Repeat steps (2) - (5) for $i = 1, 2, \ldots, K$, which completes one iteration. The entire reconstruction procedure typically involves tens or hundreds of iterations.

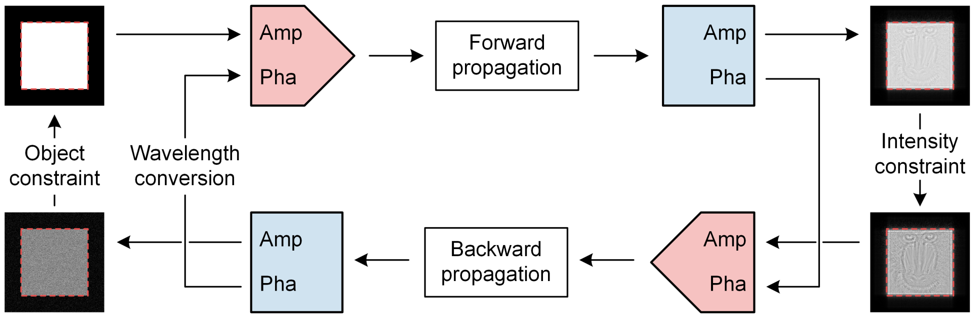

Fig. S1. Flowchart of the conventional multi-wavelength phase retrieval algorithm.

REFERENCES

1. N. Parikh and S. Boyd, "Proximal algorithms," Foundations Trends Optim. **1**, 127–239 (2014).
2. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM J. on Imaging Sci. **2**, 183–202 (2009).
3. Y. Gao and L. Cao, "A complex constrained total variation image denoising algorithm with application to phase retrieval," arXiv preprint arXiv:2109.05496 (2021).
4. K. Kreutz-Delgado, "The complex gradient operator and the CR-calculus," arXiv preprint arXiv:0906.4835 (2009).
5. Y. Gao and L. Cao, "Generalized optimization framework for pixel super-resolution imaging in digital holography," Opt. Express **29**, 28805–28823 (2021).
6. P. Bao, F. Zhang, G. Pedrini, and W. Osten, "Phase retrieval using multiple illumination wavelengths," Opt. Lett. **33**, 309–311 (2008).