Adaptive Regularization for Three-dimensional Optical Diffraction Tomography

Thanh-An Pham, Emmanuel Soubies, Ahmed Ayoub, Demetri Psaltis, Michael Unser

ADAPTIVE REGULARIZATION
FOR THREE-DIMENSIONAL OPTICAL DIFFRACTION TOMOGRAPHY

Thanh-an Pham [1], *Emmanuel Soubies* [2], *Ahmed Ayoub* [3], *Demetri Psaltis* [3], *and Michael Unser* [1]

[1] Biomedical Imaging Group, EPFL, Lausanne, Switzerland.
[2] IRIT, Université de Toulouse, CNRS, France.
[3] Optics Laboratory, EPFL, Lausanne, Switzerland.

ABSTRACT

Optical diffraction tomography (ODT) allows one to quantitatively measure the distribution of the refractive index of the sample. It relies on the resolution of an inverse scattering problem. Due to the limited range of views as well as optical aberrations and speckle noise, the quality of ODT reconstructions is usually better in lateral planes than in the axial direction. In this work, we propose an adaptive regularization to mitigate this issue. We first learn a dictionary from the lateral planes of an initial reconstruction that is obtained with a total-variation regularization. This dictionary is then used to enhance both the lateral and axial planes within a final reconstruction step. The proposed pipeline is validated on real data using an accurate nonlinear forward model. Comparisons with standard reconstructions are provided to show the benefit of the proposed framework.

Index Terms— plug-and-play, nonlinear inverse problems, dictionary learning, computational imaging.

1. INTRODUCTION

Optical diffraction tomography (ODT) is a microscopic technique that provides a three-dimensional map of the refractive index (RI) of the sample [1]. It proceeds by measuring the complex fields that are produced when the sample is illuminated with different tilted plane waves. These measurements are related to the RI of the sample through the wave equation that governs the light-scattering phenomenon. This allows for the deployment of numerical methods to recover the RI.

The reconstruction of the RI consists in solving an inverse scattering problem which is particularly challenging for several reasons. First, wave scattering is a nonlinear phenomenon and the most faithful forward model for ODT is iterative [2–5]. Second, the wavevector of the incident wave is limited to a small cone around the optical axis (Figure 1). This leads to the well-known missing-cone problem [6,7] whose effect is an elongation of the reconstructed object along the optical axis. Put simply, the ODT lateral resolution is better than its axial resolution. Finally, the measured fields contain optical aberrations or speckle noise [8,9] which are usually unknown and constitute a source of model mismatch. To mitigate these problems, former works added prior knowledge in the reconstruction process using regularization [4,7,10,11]. These include, for instance, non-negativity constraints and total-variation (TV) regularization.

In this work, we propose to go one step further and take into account the anisotropic resolution in ODT while improving the quality
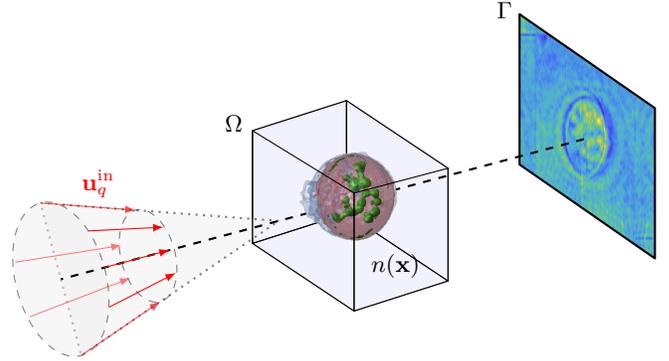
Fig. 1. Principle of optical diffraction tomography. The wave vectors $\{\mathbf{k}_q^{\text{in}}\}_{q=1}^Q \in \mathbb{R}^3$ of the Q incident plane waves $\{u_q^{\text{in}}\}_{q=1}^Q$ are limited to a cone around the optical axis.

of the reconstructions. Our motivation is to learn highly resolved features from lateral planes and use them to enhance the quality in the axial direction. Hence, inspired by the strategy proposed by Soulez [12] for the deconvolution of fluorescent microscopic images, we deploy a dictionary-based regularizer that is learnt from the lateral planes of an initial reconstructed volume.

2. PHYSICAL MODEL

We denote the RI of the sample by $n : \Omega \mapsto \mathbb{R}$ and we assume that its support is included in the region of interest $\Omega \subset \mathbb{R}^3$ (Figure 1). The RI of the surrounding medium is denoted by n_{b}. Let the sample be illuminated with the monochromatic incident field $u^{\text{in}} : \Omega \mapsto \mathbb{C}$ of wavelength λ. The resulting total field $u = u^{\text{in}} + u^{\text{sc}}$, where $u^{\text{sc}} : \Omega \mapsto \mathbb{C}$ is the scattered field, follows the Lippmann-Schwinger (LiSc) model

$$u(\mathbf{x}) = u^{\text{in}}(\mathbf{x}) + \int_\Omega g(\mathbf{x} - \mathbf{z}) f(\mathbf{z}) u(\mathbf{z}) \, \mathrm{d}\mathbf{z}, \qquad (1)$$

where $f(\mathbf{x}) = k_{\text{b}}^2 \left(n(\mathbf{x})^2 / n_{\text{b}}^2 - 1 \right)$ is the scattering potential and $k_{\text{b}} = 2\pi n_{\text{b}}/\lambda$ is the wavenumber in the surrounding medium. Finally, under Sommerfeld' radiation condition, the free-space Green function $g : \Omega \to \mathbb{C}$ is defined by [13]

$$g(\mathbf{x}) = \frac{\exp\left(j k_{\text{b}} \|\mathbf{x}\|\right)}{4\pi \|\mathbf{x}\|}. \qquad (2)$$

Following [2, 4], we consider a two-step ODT forward model. We first discretize Ω into $N = n^3$ voxels. Then, the scattered field $\mathbf{y}^{\mathrm{sc}} \in \mathbb{C}^M$ at the camera plane Γ is computed as

$$\mathbf{u} = (\mathbf{I} - \mathbf{G}\,\mathrm{diag}(\mathbf{f}))^{-1}\,\mathbf{u}^{\mathrm{in}} \tag{3}$$

$$\mathbf{y}^{\mathrm{sc}} = \tilde{\mathbf{G}}\,\mathrm{diag}(\mathbf{f})\mathbf{u}, \tag{4}$$

where $\mathbf{f} \in \mathbb{R}^N$, $\mathbf{u}^{\mathrm{in}} \in \mathbb{C}^N$, and $\mathbf{u} \in \mathbb{C}^N$ are sampled version of f, u^{in}, and u within Ω, respectively. The discrete counterpart of the convolution with the Green function in (1) is denoted by $\mathbf{G} : \mathbb{C}^N \to \mathbb{C}^N$. Finally, given $\mathbf{u}$ within Ω, $\tilde{\mathbf{G}} : \mathbb{C}^N \to \mathbb{C}^M$ provides the scattered field $\mathbf{y}^{\mathrm{sc}}$. We refer the reader to [5] for details concerning the implementation of $\mathbf{G}$ and $\tilde{\mathbf{G}}$.

3. THREE-STEP RECONSTRUCTION

In the spirit of [12], we designed a three-step reconstruction scheme.

1. TV-regularized reconstruction.
2. Dictionary learning (DL) from lateral planes.
3. Final reconstruction using the learnt dictionary.

3.1. TV-Regularized Reconstruction

The first step consists in solving the nonnegative TV-regularized problem

$$\mathbf{f}^{\mathrm{TV}} \in \left\{ \arg\min_{\mathbf{f} \in \mathbb{R}_{\geq 0}^N} \left(\sum_{q=1}^Q \|\mathbf{H}_q(\mathbf{f}) - \mathbf{y}_q^{\mathrm{sc}}\|_2^2 + \tau_{\mathrm{TV}}\|\mathbf{f}\|_{\mathrm{TV}} \right) \right\}, \tag{5}$$

where $\mathbf{H}_q : \mathbb{C}^N \mapsto \mathbb{C}^M$ denotes the two-step forward model described by (3) and (4) for the qth incident wave $\mathbf{u}_q^{\mathrm{in}}$, $\|\cdot\|_{\mathrm{TV}} \triangleq \|\boldsymbol{\nabla}\cdot\|_{2,1}$ is the TV seminorm, and $\tau_{\mathrm{TV}} > 0$ is the regularization parameter which balances between the data-fidelity term and the TV term. As in [4], we deploy a forward-backward splitting (FBS), implemented using the GlobalBioIm library[1] [14], to obtain $\mathbf{f}^{\mathrm{TV}}$.

3.2. DL from Lateral Planes

Given the nonnegative TV solution $\mathbf{f}^{\mathrm{TV}}$, our goal is to learn a dictionary $\mathbf{D} \in \mathbb{R}^{n \times K}$ formed out of K atoms of size $n < N$ such that $\mathbf{f}^{\mathrm{TV}}$ can be represented as

$$\mathbf{f}^{\mathrm{TV}} = \sum_{p=1}^{P_{\mathrm{XY}}} (\mathbf{R}_p^{\mathrm{XY}})^T \mathbf{D}\boldsymbol{\alpha}_p, \tag{6}$$

where $\mathbf{R}_p^{\mathrm{XY}} : \mathbb{R}^N \to \mathbb{R}^n$ is an operator that extracts a 2D patch of size n centered on the pth element of the input vector (its adjoint inserts the patch at the pth position), P_{XY} denotes the number of patches, and $\{\boldsymbol{\alpha}_p \in \mathbb{R}^K\}_{p=1}^{P_{\mathrm{XY}}}$ are sparse vectors. The superscript XY in $\mathbf{R}^{\mathrm{XY}}$ denotes the fact that the operation only extracts 2D patches from lateral planes.

We formulate the DL problem as

$$\{\widehat{\mathbf{D}}, \widehat{\boldsymbol{\alpha}_p}\} \in \left\{ \arg\min_{\substack{\mathbf{D} \in \mathbb{R}^{n \times K} \\ \boldsymbol{\alpha}_p \in \mathbb{R}^K}} \left(\left\| \sum_{p=1}^{P_{\mathrm{XY}}} (\mathbf{R}_p^{\mathrm{XY}})^T \mathbf{D}\boldsymbol{\alpha}_p - \mathbf{f}^{\mathrm{TV}} \right\|_2^2 \right. \right.$$
$$\left. \left. + \tau_{\mathrm{DL}} \sum_{p=1}^{P_{\mathrm{XY}}} \|\boldsymbol{\alpha}_p\|_1 \right) \right\}, \tag{7}$$

where $\tau_{\mathrm{DL}} > 0$ controls the sparsity level.

Our formulation is fundamentally different from the pioneering approaches [15, 16] where the solution was such that each extracted patch had a sparse representation in $\mathbf{D}$. Our representation $\mathbf{f} = \sum_{p=1}^{P_{\mathrm{XY}}} (\mathbf{R}_p^{\mathrm{XY}})^T \mathbf{D}\boldsymbol{\alpha}_p$ is related to convolutional dictionary learning (CDL) [17–19], as shown by Papyan *et al.* [20, 21]. As opposed to traditional DL, CDL accounts for global information in the image such as shift invariance. Hence, (8) not only enjoys the global sparse representation of CDL but also benefits from the local (patch-based) processing of DL [20, 21].

By introducing the auxiliary variable $\mathbf{s}_p = \mathbf{D}\boldsymbol{\alpha}_p$ in (7), we can deploy the alternating-direction method of multipliers (ADMM) to minimize the augmented-Lagrangian functional

$$\mathcal{L}(\widehat{\mathbf{D}}, \widehat{\boldsymbol{\alpha}_p}, \widehat{\mathbf{s}_p}, \widehat{\mathbf{w}_p}) = \left\| \sum_{p=1}^{P_{\mathrm{XY}}} (\mathbf{R}_p^{\mathrm{XY}})^T \mathbf{s}_p - \mathbf{f}^{\mathrm{TV}} \right\|_2^2$$
$$+ \sum_{p=1}^{P_{\mathrm{XY}}} \frac{\rho}{2} \left\| \mathbf{s}_p - \mathbf{D}\boldsymbol{\alpha}_p + \frac{\mathbf{w}_p}{\rho} \right\|_2^2 + \tau_{\mathrm{DL}}\|\boldsymbol{\alpha}_p\|_1, \tag{8}$$

where $\{\mathbf{w}_p \in \mathbb{R}^n\}_{p=1}^{P_{\mathrm{XY}}}$ are the dual variables and $\rho > 0$ is the Lagrangian multiplier. Using the CDL terminology [21], the auxiliary variable $\mathbf{s}_p$ is referred to as the pth *slice*. The ADMM is implemented using the SPAM toolbox[2] [22].

3.3. Final Reconstruction Using the Learnt Dictionary

Equipped with the dictionary $\widehat{\mathbf{D}} \in \mathbb{R}^{n \times K}$ learnt from lateral planes in Section 3.2, we now consider the optimization problem

$$\begin{cases} \mathbf{f}^\star = \mathbf{R}^T \mathbf{s}^*, \\ \mathbf{s}^* \in \left\{ \arg\min_{\mathbf{s} \in \mathbb{R}^{nP}} \left(\sum_{q=1}^Q \|\mathbf{H}_q(\mathbf{R}^T\mathbf{s}) - \mathbf{y}_q^{\mathrm{sc}}\|_2^2 \right. \right. \\ \left. \left. \qquad\qquad + \mathcal{R}_{\mathrm{pos}}(\mathbf{R}^T\mathbf{s}) + \mathcal{R}_{\widehat{\mathbf{D}}}(\mathbf{s}) \right) \right\}, \end{cases} \tag{9}$$

where $\mathbf{s} = [\mathbf{s}_1^T \cdots \mathbf{s}_P^T]^T \in \mathbb{R}^{nP}$ is the concatenation of all the slices, $P = P_{\mathrm{XY}} + P_{\mathrm{XZ}} + P_{\mathrm{YZ}}$ is the total number of slices, and $\mathbf{R} = [\mathbf{R}_{\mathrm{XY}}^T \; \mathbf{R}_{\mathrm{XZ}}^T \; \mathbf{R}_{\mathrm{YZ}}^T]^T \in \mathbb{R}^{nP \times N}$ with $\mathbf{R}_{\mathrm{XZ}}$ ($\mathbf{R}_{\mathrm{YZ}}$) the counterpart of $\mathbf{R}_{\mathrm{XY}}$ for the XZ (YZ, respectively) sections of the volume. We use the differentiable functional $\mathcal{R}_{\mathrm{pos}} : \mathbb{R}^N \to \mathbb{R}$

$$\mathcal{R}_{\mathrm{pos}}(\mathbf{f}) = \sum_{n=1}^N \frac{\lambda_{\mathrm{pos}}}{\mu} \log\left(\exp(-\mu f_n) + 1\right) \tag{10}$$

to favor nonnegative solutions. Here, $\lambda_{\mathrm{pos}} > 0$ is a weighting factor and $\mu > 0$ shapes the tolerance to negative values. Finally, the functional $\mathcal{R} : \mathbb{R}^{nP} \to \mathbb{R}$ in (9) is a regularization term designed to enforce the slices $\mathbf{s}_p$ to have a sparse representation in $\widehat{\mathbf{D}}$. Denoting $\mathcal{F} = \sum_{q=1}^Q \|\mathbf{H}_q(\mathbf{R}^T \cdot) - \mathbf{y}_q^{\mathrm{sc}}\|_2^2 + \mathcal{R}_{\mathrm{pos}}(\mathbf{R}^T \cdot)$, we can deploy the FBS algorithm whose iterates are given by

$$\mathbf{s}^{k+1} = \mathrm{prox}_{\gamma\mathcal{R}_{\widehat{\mathbf{D}}}} \left(\mathbf{s}^k - \gamma \boldsymbol{\nabla}\mathcal{F}(\mathbf{s}^k) \right), \tag{11}$$

where $\gamma > 0$ is a descent parameter and $\mathrm{prox}_{\gamma\mathcal{R}_{\widehat{\mathbf{D}}}}$ denotes the proximity operator of the functional $\mathcal{R}_{\widehat{\mathbf{D}}}$. Here, we follow the plug-and-play prior philosophy [23–25] and replace $\mathrm{prox}_{\gamma\mathcal{R}_{\widehat{\mathbf{D}}}}$ in (11) by the

---

[1] http://bigwww.epfl.ch/algorithms/globalbioim/

[2] http://spams-devel.gforge.inria.fr/

Algorithm 1 Proposed algorithm to solve (9)

Require: $\mathbf{y} \in \mathbb{C}^{MQ}$, $\mathbf{x}^0 \in \mathbb{C}^N$, $\widehat{\mathbf{D}} \in \mathbb{R}^{n \times K}$, $\gamma > 0, \tau_{sc} > 0$
1: Define $\mathcal{F} = \sum_{q=1}^{Q} \|\mathbf{H}_q(\mathbf{R}^T \cdot) - \mathbf{y}_q^{sc}\|_2^2 + \mathcal{R}_{pos}(\mathbf{R}^T \cdot)$
2: $\mathbf{s}^0 = \frac{1}{n^2}\mathbf{R}\mathbf{x}^0$
3: $k = 1$
4: **while** (not converged) **do**
5: $\quad \mathbf{z}^k = \mathbf{s}^k - \gamma \boldsymbol{\nabla}\mathcal{F}(\mathbf{s}^k)$
6: $\quad \mathbf{s}_p^{k+1} = \mathcal{C}_{\tau_{sc},\widehat{\mathbf{D}}}(\mathbf{z}_p^k), \quad \forall p \in \{1, \ldots, P\}$
7: $\quad k \leftarrow k + 1$
8: **end while**
9: **return** $\mathbf{x}^\star = \mathbf{R}^T \mathbf{s}^{k-1}$

Fig. 3. Dictionary $\widehat{\mathbf{D}}$ learnt from lateral planes of the TV regularized solution ($\mathbf{f}^{TV}$, dotted rectangle in Figure 2).

Rytov model [28]. We used diverse regularization parameters τ_{TV} for TV (Figure 2). When the regularization is weak, artifacts due to model mismatch are hindering the quality of reconstruction. On the contrary, over-regularization results in cartoon-like solutions.

4.2. Learning the Dictionary

We learned the dictionary by minimizing (8). We used patches of size (8×8) ($n = 64$) and $K = 64$ atoms. We set $\tau_{DL} = 1/\sqrt{8}$ and $\rho = 0.5 \max(\mathbf{f}^{TV})n^2$. The learnt atoms of the dictionary $\widehat{\mathbf{D}}$ are shown in Figure 3.

4.3. Final Reconstruction

We solved the optimization problem (9) and encouraged the nonnegativity of the solution with $\lambda_{pos} = 1/96^3$ and $\mu = 5$ in (10). The initial guess was $\mathbf{f}^{TV}$. The denoising operator (12)–(13) was used with regularization parameter $\tau_{sc} = 10^{-4}$.

We observe that the Rytov-based solution suffers from the missing-cone problem whereas the regularized solutions (*i.e.,* TV and the proposed one) mitigate its effect. In addition, some features are enhanced with the proposed solution in comparison to the TV solution (Figure 4). Finally, the proposed method is able to recover features in deeper axial position whereas the TV-regularized solution is over-regularized (Figure 4, right column).
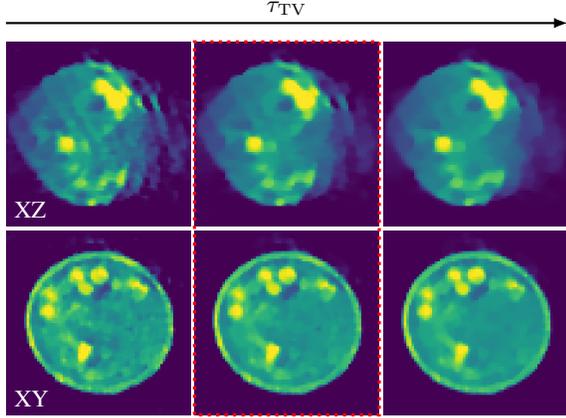

Fig. 2. TV-regularized solutions obtained with regularization parameters $\tau_{TV} = \{0.15, 0.3, 0.5\}/96^3$ (from left to right, respectively).

"denoising" operator

$$\mathcal{C}_{\tau_{sc},\widehat{\mathbf{D}}} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$$
$$\mathbf{s} \longmapsto \widehat{\mathbf{D}}\boldsymbol{\alpha}^\star, \qquad (12)$$

where $\tau_{sc} > 0$, and

$$\boldsymbol{\alpha}^\star \in \left\{ \arg\min_{\boldsymbol{\alpha} \in \mathbb{R}^K} \left(\frac{1}{2}\|\mathbf{s} - \widehat{\mathbf{D}}\boldsymbol{\alpha}\|_2^2 + \tau_{sc}\|\boldsymbol{\alpha}\|_1 \right) \right\}. \qquad (13)$$

Numerous solvers exist to solve (13) [16, 26]. Again, we use the GlobalBioIm library together with the SPAMS toolbox for this step. Finally, we summarize the complete reconstruction scheme of this section in Algorithm 1.

4. RESULTS

We validated our method on real data acquired with the ODT setup described in [27]. We acquired 61 views with illumination of wavelength $\lambda = 532$nm within a cone of $35°$. Each view has (300×300) (complex) measurements focused at the center of the sample. The sample is a yeast cell (Figure 1) immersed in water ($n_b = 1.3388$). The size of the reconstructed volume is $(96 \times 96 \times 96)$ with a sampling step of 99.3nm (9.53μm in each dimension).

4.1. TV-Regularized Reconstruction

We first reconstructed the sample using the method described in [4] by minimizing (5). The initial guess was the solution provided by the

5. CONCLUSION

We designed an adaptive regularization that allowed us to improve the quality of optical diffraction tomography (ODT) reconstructions in the axial direction using features learned in lateral planes. The proposed regularization relies on a dictionary that is learnt from the lateral planes of an initial total-variation reconstruction. This dictionary is then used in a final step to enhance the quality of the reconstruction in all XY, XZ, and YZ sections. We applied this strategy to the reconstruction of real ODT measurements. Our results show the superior performance of the proposed pipeline over conventional regularizations.

6. REFERENCES

[1] E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Optics Communications*, vol. 1, no. 4, pp. 153–156, 1969.

[2] H.-Y. Liu, D. Liu, H. Mansour, P. T. Boufounos, L. Waller, and U. S. Kamilov, "Seagle: Sparsity-driven image reconstruction under multiple scattering," *IEEE Transactions on Computational Imaging*, vol. 4, no. 1, pp. 73–86, 2017.

[3] T. Zhang, C. Godavarthi, P. C. Chaumet, G. Maire, H. Giovannini, A. Talneau, M. Allain, K. Belkebir, and A. Sentenac,
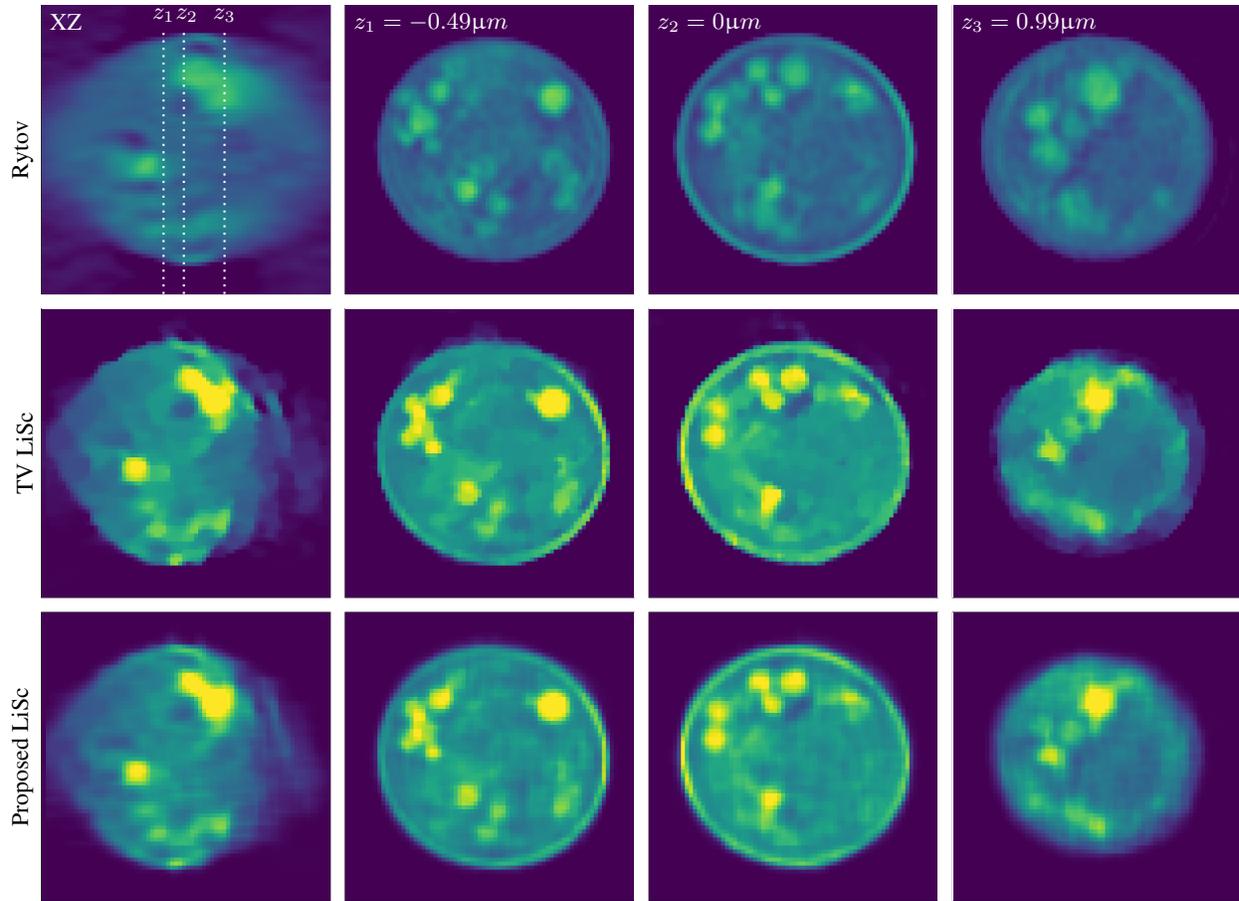
Fig. 4. Yeast reconstructions. Views of the reconstructed sample at the XZ plane and three different XY positions. The Rytov-based solution (Rytov), the TV-regularized solution (TV LiSc), and the proposed regularized solution (Proposed LiSc) are displayed in the top, middle, and bottom rows, respectively. The TV-regularized solution was obtained with $\tau_{\mathrm{TV}} = 0.3/96^3$.

"Far-field diffraction microscopy at $\lambda/10$ resolution," *Optica*, vol. 3, no. 6, pp. 609–612, 2016.

[4] E. Soubies, T.-A. Pham, and M. Unser, "Efficient inversion of multiple-scattering model for optical diffraction tomography," *Optics Express*, vol. 25, no. 18, pp. 21786–21800, 2017.

[5] T.-A. Pham, E. Soubies, A. Ayoub, J. Lim, D. Psaltis, and M. Unser, "Three-dimensional optical diffraction tomography with Lippmann-Schwinger model," *IEEE Transactions on Computational Imaging*, to be published.

[6] V. Lauer, "New approach to optical diffraction tomography yielding a vector equation of diffraction tomography and a novel tomographic microscope," *Journal of Microscopy*, vol. 205, no. 2, pp. 165–176, 2002.

[7] J. Lim, K. Lee, K. H. Jin, S. Shin, S. Lee, Y. Park, and J. C. Ye, "Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography," *Optics Express*, vol. 23, no. 13, pp. 16933–16948, 2015.

[8] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*, Roberts & Company, Englewood, Colorado, 2007.

[9] L. Che, W. Xiao, F. Pan, B. Dong, and Z. Zhong, "Reduction of speckle noise in digital holography by combination of averaging several reconstructed images and modified nonlocal means filtering," *Optics Communications*, vol. 426, pp. 9 – 15, November 1, 2018.

[10] Y. Sung, W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, and M. S. Feld, "Optical diffraction tomography for high resolution live cell imaging," *Optics Express*, vol. 17, no. 1, pp. 266–277, 2009.

[11] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Optical tomographic image reconstruction based on beam propagation and sparse regularization," *IEEE Transactions on Computational Imaging*, vol. 2, no. 1, pp. 59–70, 2016.

[12] F. Soulez, "A learn 2D, apply 3D method for 3D deconvolution microscopy," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, Beijing, People's Republic of China, April 29-May 2 2014, pp. 1075–1078.

[13] J. A. Schmalz, G. Schmalz, T. E. Gureyev, and K. M. Pavlov, "On the derivation of the Green's function for the Helmholtz equation using generalized functions," *American Journal of Physics*, vol. 78, no. 2, pp. 181–186, January 2010.

[14] E. Soubies, F. Soulez, M. McCann, T.-A. Pham, L. Donati, T. Debarre, D. Sage, and M. Unser, "Pocket guide to solve inverse problems with GlobalBioIm," *Inverse Problems*, vol. 35, no. 10, pp. 1–20, 2019, paper no. 104006.

[15] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image processing*, vol. 15, no. 12, pp. 3736–3745, 2006.

[16] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proceedings of the 26th Annual International Conference on Machine Learning*, Montreal, QC, Canada, June 14 - 18 2009, ACM, pp. 689–696.

[17] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional networks," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, June 13-18 2010, pp. 2528–2535.

[18] M. Mørup and M. N. Schmidt, "Transformation invariant sparse coding," in *IEEE International Workshop on Machine Learning for Signal Processing*, Beijing, People's Republic of China, September 18-21 2011, pp. 1–6.

[19] C. Garcia-Cardona and B. Wohlberg, "Convolutional dictionary learning: A comparative review and new algorithms," *IEEE Transactions on Computational Imaging*, vol. 4, no. 3, pp. 366–381, Sep. 2018.

[20] V. Papyan, Y. Romano, M. Elad, and J. Sulam, "Convolutional dictionary learning via local processing," in *International Conference on Computer Vision*, Venice, Italy, October 22-29 2017, pp. 5306–5314.

[21] V. Papyan, J. Sulam, and M. Elad, "Working locally thinking globally: Theoretical guarantees for convolutional sparse coding," *IEEE Transactions on Signal Processing*, vol. 65, no. 21, pp. 5687–5701, 2017.

[22] J. Mairal, F. Bach, and J. Ponce, "Sparse modeling for image and vision processing," *Foundations and Trends in Computer Graphics and Vision*, vol. 8, no. 2-3, pp. 85–283, 2014.

[23] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, Austin, Texas, U.S.A., December 3-5 2013, pp. 945–948.

[24] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Transactions on Computational Imaging*, vol. 2, no. 4, pp. 408–423, 2016.

[25] U. S. Kamilov, H. Mansour, and B. Wohlberg, "A plug-and-play priors approach for solving nonlinear imaging inverse problems," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1872–1876, 2017.

[26] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.

[27] A. B. Ayoub, T.-A. Pham, J. Lim, M. Unser, and D. Psaltis, "A method for assessing the fidelity of optical diffraction tomography reconstruction methods using structured illumination," *Optics Communications*, vol. 454, no. 124486, pp. 1–6, January 1 2020.

[28] A. Devaney, "Inverse-scattering theory within the Rytov approximation," *Optics Letters*, vol. 6, no. 8, pp. 374–376, 1981.