Computational Super-Sectioning for Single-Slice Structured-Illumination Microscopy

Emmanuel Soubies and Michael Unser, *Fellow, IEEE*

***Abstract*—While structured-illumination microscopy (SIM) is inherently a three-dimensional (3-D) technique, many biological questions can be addressed from the acquisition of a single focal plane with high lateral resolution. Unfortunately, the single-slice reconstruction of thick samples suffers from defocusing. In this paper, however, we take advantage of a 3-D model of the acquisition system to derive a reconstruction method out of a single two-dimensional (2-D) SIM measurement. It enables the estimation of the out-of-focus signal and improves the quality of the reconstruction, without the need of acquiring additional slices. The proposed algorithm relies on a specific formulation of the optimization problem together with the derivation of computationally efficient proximal operators. These developments allow us to deploy an efficient inner-loop-free alternating-direction method of multipliers (ADMM), with guaranteed convergence.**

***Index Terms*—Structured-illumination microscopy, super-resolution, reconstruction algorithms, inner-loop-free ADMM, proximal operators.**

I. Introduction

STRUCTURED-ILLUMINATION microscopy (SIM) offers an excellent tradeoff between spatial and temporal resolution for fluorescence microscopy. In its conventional form, sinusoidal illuminations—formed out of two (2D-SIM [1], [2]), sometimes three (3D-SIM [3]–[5]) interfering laser beams—are used to excite fluorescent probes. This procedure shifts the high-frequency components of the imaged structure to the bandpass of the microscope. These components can then be numerically recovered and shifted back to their correct location to produce an image with twice the resolution of conventional systems. This twofold resolution enhancement can also be obtained using speckle illuminations [6] and, theoretically, unlimited resolution can even be achieved from nonlinear SIM [7], such as the saturated-SIM that was originally introduced in [8], [9].

A. Related Work on SIM Reconstruction

With conventional 2D sinusoidal illumination, the spectrum of the acquired image is a linear combination of three shifted versions of the imaged structure spectrum. Moreover, varying the phase of the illumination grating allows to change the coefficients of this linear combination. Hence, pioneering works [1], [2] were extracting these unknown components through the resolution of a linear system formed out of three acquisitions with different phases. Then, a generalized Wiener filter was used to assemble these components, yielding a single super-resolved image with extended frequency information. We refer the reader to [10] for a detailed description of this reconstruction approach. Because of its simplicity and speed, this direct method is currently the most used in practice, especially on commercial systems. Non-direct, iterative approaches have the potential to significantly improve the reconstruction quality and robustness of direct approaches. In particular, the use of sparsity-promoting regularization can produce high-quality reconstructions from a limited amount of measurements. To the best of our knowledge, Orieux *et al.* [11] were the first to report a variational approach for SIM microscopy. Several other works have then promoted these methods [10], [12]–[18].

Recently, various open-source softwares have been developed for SIM reconstruction. Those include FairSIM [19], OpenSIM [20], and Simtoolbox [21], [22]. They all implement the classical direct method together with a maximum *a posteriori* (MAP) estimation for Simtoolbox. While they are limited to 2D-SIM reconstruction, they can still perform slice-by-slice reconstructions with computational sectioning[1] improvement. For instance, FairSIM handles out-of-focus signal through attenuation of the optical-transfer function (OTF). Simtoolbox merges reconstructions obtained by homodyme detection [23], [24] for computational sectioning and MAP estimation to enhance the lateral resolution, as in [25]. Finally, an alternative method proposed by Jost *et al.* [26] considers a few additional planes during the reconstruction process in order to collect out-of-focus light. This method is called thick-slice SIM reconstruction and bridges 2D and 3D-SIM reconstruction. For more details about SIM literature, we refer the reader to the review paper by Heintzmann and Huser [27].

B. Contributions and Road Map

This paper builds upon the prior work [26] and its thick-slice reconstruction algorithm. The proposed method extends this work with a new reconstruction algorithm that improves speed

Manuscript received July 2, 2018; revised October 9, 2018 and November 30, 2018; accepted December 10, 2018. Date of publication December 17, 2018; date of current version May 7, 2019. This work was supported in part by the Research-IDEAS initiative of ZEISS and EPFL and in part by the European Research Council under the European Union's Horizon 2020 research and innovation programme, Grant Agreement 692726 GlobalBioIm: Global integrative framework for computational bio-imaging. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Laura Waller. *(Corresponding author: Emmanuel Soubies.)*

The authors are with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland (e-mail: emmanuel.soubies@epfl.ch; michael.unser@epfl.ch).

Digital Object Identifier 10.1109/TCI.2018.2887136

[1]Optical sectioning capitalizes on optical means to obtain better sections of the sample. Computational sectioning deploys computational methods to reach the same goal, typically by rejecting out-of-focus light algorithmically.

and that offers new regularization opportunities (e.g., sparsity-based). Our main contribution is an efficient inner-loop-free alternating-direction method of multipliers (ADMM) [28]. It derives from a specific formulation of the optimization problem (Section III-B), together with a closed-form proximal operator provided in Theorem III.1. The convergence of the algorithm is guaranteed in Proposition III.3. Moreover, we show that the proposed algorithm is significantly faster than competing algorithms (Section IV-A). Finally, we validate the proposed framework (thick-slice model and ADMM) on simulations and real data, and compare it to the open-source FairSIM software (Sections IV-B and IV-C). In particular, we emphasize the ability of our method to reject out-of-focus light by virtue of the additional reconstruction planes. We also show that our computational super-sectioning SIM reconstruction of single-slice data compares favorably to reconstructions that capitalize on the availability of full 3D-SIM acquisitions.

C. Notations

Scalar and continuously defined functions are denoted by italic letters (e.g., $x \in \mathbb{R}$, $f \in L_2(\mathbb{R})$). Vectors are denoted by bold lowercase letters (e.g., $\mathbf{f}$) and matrices (i.e., discrete linear operators) by bold uppercase letters (e.g., $\mathbf{H}$). Given a column vector $\mathbf{x} = [x_1 \ \cdots \ x_N]^T \in \mathbb{R}^N$, its p-norm is defined as $\|\mathbf{x}\|_p = (\sum_{n=1}^N |x_n|^p)^{\frac{1}{p}}$. For an operator $\mathbf{H} \in \mathbb{R}^{M\times N}$, $\|\mathbf{H}\| = \sigma_{\max}(\mathbf{H})$ denotes its spectral norm, which is equal to its largest singular value. The adjoint of an operator $\mathbf{H} \in \mathbb{R}^{M\times N}$ is denoted by $\mathbf{H}^*$ and verifies $\langle \mathbf{Hx}, \mathbf{y}\rangle_{\mathbb{R}^M} = \langle \mathbf{x}, \mathbf{H}^*\mathbf{y}\rangle_{\mathbb{R}^N}$, where $\langle\cdot,\cdot\rangle_{\mathbb{R}^M}$ ($\langle\cdot,\cdot\rangle_{\mathbb{R}^N}$, respectively) is the usual scalar product in $\mathbb{R}^M$ ($\mathbb{R}^N$, respectively). The set of nonnegative vectors of $\mathbb{R}^N$ is denoted by $\mathbb{R}^N_{\geq 0} = \{\mathbf{x} \in \mathbb{R}^N : x_n \geq 0, \forall n = 1, \ldots, N\}$.

For a vector $\mathbf{v} \in \mathbb{R}^N$, $\mathbf{diag}(\mathbf{v}) \in \mathbb{R}^{N\times N}$ defines the diagonal operator whose diagonal entries are given by the elements of $\mathbf{v}$. The notation $\mathbf{1}_N = [1 \ \cdots \ 1]^T \in \mathbb{R}^N$ ($\mathbf{0}_N = [0 \ \cdots \ 0] \in \mathbb{R}^N$, respectively) stands for a vector of ones (zeros, respectively). Then, $\mathbf{I}_N = \mathbf{diag}(\mathbf{1}_N) \in \mathbb{R}^{N\times N}$ is the identity operator of size N. We write the N-point unitary discrete Fourier transform (DFT) as $\mathbf{F}_N \in \mathbb{C}^{N\times N}$. It is defined by[2] $[\mathbf{F}_N\mathbf{x}]_k = \frac{1}{\sqrt{N}}\sum_{n=1}^N x_n \mathrm{e}^{-\frac{2\mathrm{i}\pi}{N}kn}$ and verifies $\mathbf{F}_N^*\mathbf{F}_N = \mathbf{F}_N\mathbf{F}_N^* = \mathbf{I}_N$. Finally, $\otimes$ denotes the Kronecker product, $\odot$ the Hadamard product, and, for $x \in \mathbb{R}$, $\lfloor x \rfloor$ is the greatest integer that does not exceed x.

II. Variational Formulation of the Inverse Problem

In order to solve an inverse problem, one has to (i) model the acquisition process; (ii) formulate the reconstruction problem using an adequate regularization; (iii) deploy an efficient optimization strategy. In this section, we describe the first two steps for structured-illumination microscopy.

[2]The extension to higher dimensions is straightforward by separability of the Fourier transform.

A. Image Formation Model in 2D-SIM

A 2D SIM acquisition $\mathbf{y} \in \mathbb{R}^M$ (with $M = M_1M_2$) can be described by

$$y_m = (h * (wf))(\mathbf{x}_m, z_{\mathrm{fp}}) + n_m, \tag{1}$$

for all $m \in \{1, \ldots, M\}$. Here $f \in L_2(\mathbb{R}^3)$ is the (3D) biological sample (fluorophores density map), $w \in L_2(\mathbb{R}^3)$ is the illumination grating, $h \in L_2(\mathbb{R}^3)$ is the point-spread function (PSF) of the optical system, $\{\mathbf{x}_m \in \mathbb{R}^2\}_{m=1}^M$ is the list of camera sampling points, z_{fp} denotes the focal plane position, and $\mathbf{n} \in \mathbb{R}^M$ is a random disturbance (noise vector).

Model (1) maps the 3D continuously defined object f to the 2D discrete measurements $\mathbf{y}$. Then, to numerically solve the SIM reconstruction problem, one has to discretize the object f. The standard practice [10]–[12] is to define a 2D discrete version of f at focal plane, i.e., $\mathbf{f} \in \mathbb{R}^N$ with $N = N_1N_2$ such that $f_n = f(\mathbf{x}_n, z_{\mathrm{fp}})$ and $\{\mathbf{x}_n \in \mathbb{R}^2\}_{n=1}^N$. In contrast, we consider in this work P lateral sections (2D) of f with positions $\{z_p\}_{p=1}^P$, so that $\mathbf{f} = [\mathbf{f}_1^T \ \cdots \ \mathbf{f}_P^T]^T \in \mathbb{R}^{NP}$ and $[\mathbf{f}_p]_n = f(\mathbf{x}_n, z_p)$. Accordingly, the discrete version of (1) reads

$$\mathbf{y} = \sum_{p=1}^P \mathbf{SH}_p\mathbf{diag}(\mathbf{w}_p)\,\mathbf{f}_p + \mathbf{n}, \tag{2}$$

where $\mathbf{H}_p \in \mathbb{R}^{N\times N}$ is a discrete convolution operator whose kernel $\mathbf{h}_p \in \mathbb{R}^N$ is the sampled version of $h(\cdot, z_p)$ and $\mathbf{w}_p \in \mathbb{R}^N$ is the sampled version of $w(\cdot, z_p)$. Finally, $\mathbf{S} \in \mathbb{R}^{M\times N}$ is a decimation operator with a downsampling factor of two in each dimension (i.e., $2M_1 = N_1$ and $2M_2 = N_2$). This is required because we aim at doubling the lateral resolution. Hence, for data acquired at Nyquist rate, the reconstruction grid needs to be twice finer than the acquisition grid.

To simplify the notations, we introduce the vector $\mathbf{w} = [\mathbf{w}_1^T \ \cdots \ \mathbf{w}_P^T]^T$, as well as the operators

$$\mathbf{H} = [\mathbf{H}_1 \ \cdots \ \mathbf{H}_P], \ \mathbf{W} = \mathbf{diag}(\mathbf{w}). \tag{3}$$

Model (2) can then be rewritten in the compact form

$$\mathbf{y} = \mathbf{SHWf} + \mathbf{n}. \tag{4}$$

An illustration of this model is presented in Fig. 1.

It is noteworthy that this model is generic, in the sense that it can cope with any illumination pattern (e.g., purely sinusoidal [1], [2], [5], speckle [29], grid of lines [30], and nonlinear [7], [8]). Because several acquisitions are needed to reconstruct the sample, we denote by $\{\mathbf{W}_l\}_{l=1}^L$ the operators associated to the L illumination patterns that lead to the acquisitions $\{\mathbf{y}_l\}_{l=1}^L$.

B. A Large-Scale Optimization Problem

Following the standard practice, we consider the optimization problem

$$\hat{\mathbf{f}} \in \left\{ \arg\min_{\mathbf{f}\in\mathbb{R}^{NP}} \left(\sum_{l=1}^L \mathcal{D}_l(\mathbf{f}, \mathbf{y}_l) + \mu\mathcal{R}(\mathbf{Lf}) + i_{\geqslant 0}(\mathbf{f}) \right) \right\}, \tag{5}$$

where the objective lets the tradeoff between fidelity to data and regularization be controlled by the parameter $\mu > 0$. In this

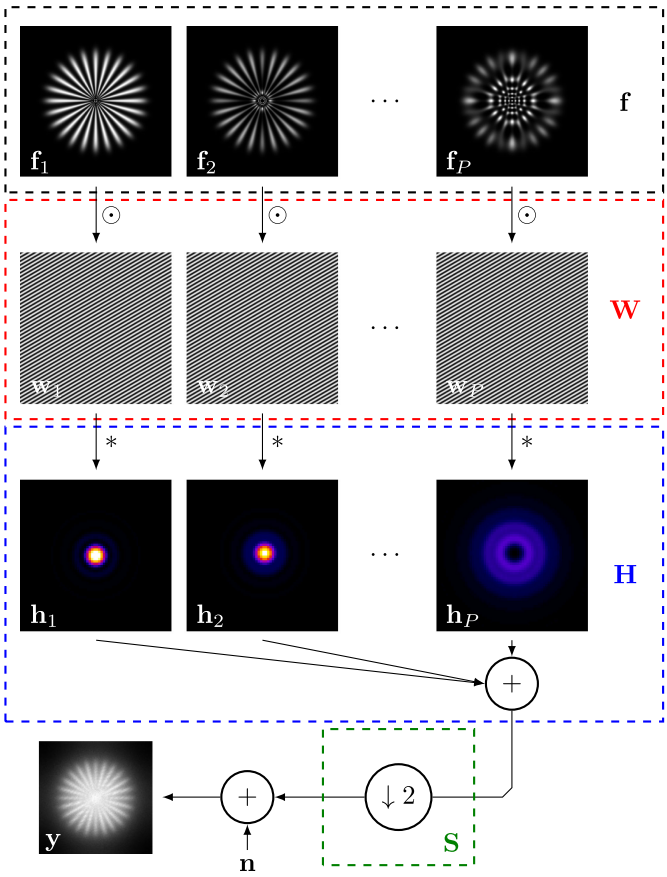

Fig. 1. Illustration of the forward model (4). The notation $\downarrow 2$ denotes a downsampling by a factor of two. In this example, two-dimensional patterns are used (i.e., w in (1) is constant along z). Hence, all the $\{\mathbf{w}_p\}_{p=1}^P$ are identical. However, Model (4) is more general and can be used with patterns that vary along z.

work, we consider the least-squares data-fidelity term

$$\mathcal{D}_l(\mathbf{f}, \mathbf{y}_l) = \frac{1}{2} \|\mathbf{SHW}_l \mathbf{f} - \mathbf{y}_l\|_2^2 . \tag{6}$$

In (5), the problem is regularized using the combination of the nonnegativity constraint

$$i_{\geqslant 0}(\mathbf{f}) = \begin{cases} 0, & \text{if } \mathbf{f} \in \mathbb{R}^{NP}_{\geqslant 0} \\ +\infty, & \text{otherwise} \end{cases} \tag{7}$$

with the sparsity-promoting convex functional $\mathcal{R} : \mathbb{R}^K \to \mathbb{R}$ composed with the linear operator $\mathbf{L} \in \mathbb{R}^{K \times N}$. A popular regularizer is the total-variation (TV) norm [31] obtained by composing the (ℓ_2, ℓ_1)-mixed norm $\mathcal{R} = \|\cdot\|_{2,1}$ with the gradient operator $\mathbf{L} = \boldsymbol{\nabla}$ (see Appendix B-A). In the present paper, we use instead the Schatten-norm (of order 1) of the Hessian operator ($\mathcal{R} = \|\cdot\|_{\mathcal{S}_1}$ and $\mathbf{L} = \mathbf{H}_\mathbf{e}$, see Appendix B-B) which avoids the staircasing effect of TV [32] and is thus better suited to biological samples.

III. Inner-Loop-Free ADMM

The field of convex optimization has experienced an important development during the past two decades. This offers several possibilities for solving (5). Because the objective function is the sum of three terms that involve non-smooth functionals such as $i_{\geqslant 0}$ and possibly $\mathcal{R}$, a possible approach is to use splitting-based algorithms. These include ADMM [28], [33] or primal-dual proximal algorithms [34]. In this work we consider ADMM.

Algorithm 1: ADMM [28] for Minimizing (8).

Require: $\mathbf{f}^0 \in \mathbb{R}^N$, $(\rho_q)_{q\in\{1,\dots,Q\}} \in \mathbb{R}^Q_{\geqslant 0}$

1: $\mathbf{u}_q^0 = \mathbf{A}\mathbf{f}^0,\ \forall q \in \{1,\dots,Q\}$
2: $\mathbf{v}_q^0 = \mathbf{u}_q^0,\ \forall q \in \{1,\dots,Q\}$
3: $k = 0$
4: **while** (not converged) **do**
5: $\quad \mathbf{u}_q^{k+1} = \text{prox}_{\frac{1}{\rho_q}\mathcal{F}_q}\left(\mathbf{A}_q\mathbf{f}^k - \mathbf{v}_q^k\right),\ \forall q \in \{1,\dots,Q\}$
6: $\quad \mathbf{f}^{k+1} = \left(\sum_{q=1}^Q \rho_q \mathbf{A}_q^*\mathbf{A}_q\right)^{-1}\left(\sum_{q=1}^Q \rho_q \mathbf{A}_q^* (\mathbf{u}_q^{k+1} + \mathbf{v}_q^k)\right)$
7: $\quad \mathbf{v}_q^{k+1} = \mathbf{v}_q^k - (\mathbf{A}_q\mathbf{f}^{k+1} - \mathbf{u}_q^{k+1}),\ \forall q \in \{1,\dots,Q\}$
8: $\quad k = k + 1$
9: **end while**

A. ADMM Principle

ADMM is designed to minimize cost functions of the form

$$\mathcal{J}(\mathbf{f}) = \sum_{q=1}^Q \mathcal{F}_q(\mathbf{A}_q\mathbf{f}), \tag{8}$$

where $\{\mathbf{A}_q \in \mathbb{R}^{N_q \times N}\}_{q=1}^Q$ are linear operators and $\{\mathcal{F}_q : \mathbb{R}^{N_q} \to \mathbb{R}\}_{q=1}^Q$ are convex functions for which one can efficiently evaluate the proximal operator [35]

$$\text{prox}_{\mathcal{F}_q}(\mathbf{z}) = \arg\min_{\mathbf{f}\in\mathbb{R}^N} \left(\frac{1}{2}\|\mathbf{f} - \mathbf{z}\|_2^2 + \mathcal{F}_q(\mathbf{f})\right). \tag{9}$$

Introducing the auxilliary variables $\{\mathbf{u}_q = \mathbf{A}_q\mathbf{f}\}_{q=1}^Q$, we obtain a constrained optimization problem which admits the augmented Lagrangian formulation

$$\mathcal{L}(\mathbf{f}, \mathbf{u}, \mathbf{v}) = \sum_{q=1}^Q \Big(\mathcal{F}_q(\mathbf{u}_q) + \langle \mathbf{v}_q, \mathbf{A}_q\mathbf{f} - \mathbf{u}_q\rangle + \frac{\rho_q}{2}\|\mathbf{A}_q\mathbf{f} - \mathbf{u}_q\|_2^2\Big). \tag{10}$$

ADMM alternates between a minimization of $\mathcal{L}$ with respect to $\mathbf{f}$, a minimization of $\mathcal{L}$ with respect to $\{\mathbf{u}_q\}_{q=1}^Q$, and an update of the dual variables $\{\mathbf{v}_q\}_{q=1}^Q$. The iterations are summarized in Algorithm 1. The parameters $\{\rho_q\}_{q=1}^Q$ are the Lagrangian mutipliers. When $Q > 2$ (i.e., the objective has more than two terms), this algorithm is termed as simultaneous-direction method of multipliers (SDMM) [36], [37].

The computational burden of Steps 5 and 6 in Algorithm 1 is directly related to the splitting strategy. Hence, to reduce the cost within each ADMM iteration, the optimal splitting is the one for which Steps 5 and 6 admit a closed-form solution.

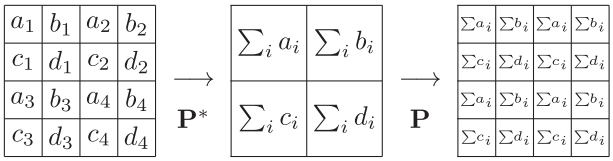

Fig. 2. Effect of the periodization operator $\mathbf{P}$ and its adjoint $\mathbf{P}^*$.

B. Problem Reformulation and Proposed Splitting Strategy

There exists different ways of splitting Problem (5) to deploy ADMM for its minimization. The simplest solution is to set

$$\mathcal{F}_l = \frac{1}{2}\|\cdot - \mathbf{y}_l\|_2^2,\ \mathbf{A}_l = \mathbf{SHW}_l,\ \forall l \in \{1,\dots,L\}, \quad (11)$$

$$\mathcal{F}_{L+1} = \mu\mathcal{R},\ \mathbf{A}_{L+1} = \mathbf{L}, \quad (12)$$

$$\mathcal{F}_{L+2} = i_{\geqslant 0},\ \mathbf{A}_{L+2} = \mathbf{I}, \quad (13)$$

which we call full-splitting (FS). With this choice, the proximal operator required at Line 5 of Algorithm 1 admits a closed-form expression. However, the main limitation of this splitting lies in Step 6 of Algorithm 1. Indeed, it requires to invert the matrix

$$\mathbf{B}_{\mathrm{FS}} = \sum_{l=1}^{L} \rho_l \mathbf{W}_l^*\mathbf{H}^*\mathbf{S}^*\mathbf{SHW}_l + \rho_{L+1}\mathbf{L}^*\mathbf{L} + \rho_{L+2}\mathbf{I}, \quad (14)$$

which cannot be done in a direct way. Hence, one would need to use an inner iterative procedure. This drawback was one of the motivations of the authors in [12] for using the primal-dual proximal method [34], rather than ADMM.

However, we now show that it is possible to derive a more efficient inner-loop-free ADMM for the resolution of (5).

Let us first introduce the following alternative splitting

$$\mathcal{F}_l = \frac{1}{2}\|\mathbf{SH}\cdot - \mathbf{y}_l\|_2^2,\ \mathbf{A}_l = \mathbf{W}_l,\ \forall l \in \{1,\dots,L\}, \quad (15)$$

$$\mathcal{F}_{L+1} = \mu\mathcal{R},\ \mathbf{A}_{L+1} = \mathbf{L}, \quad (16)$$

$$\mathcal{F}_{L+2} = i_{\geqslant 0},\ \mathbf{A}_{L+2} = \mathbf{I}. \quad (17)$$

The difference with the aforementioned full-splitting lies in equation (15). From Line 5 of Algorithm 1, one can see that this modification entails the need to evaluate the proximal operator of $\frac{\gamma}{2}\|\mathbf{SH}\cdot - \mathbf{y}\|_2^2$. We provide a closed-form expression for the latter in Theorem III.1.

Theorem III.1: Let $\mathbf{S} \in \mathbb{R}^{M\times N}$ be a uniform downsampling operator and $\mathbf{H}$ be the convolution operator defined by (3). Then, the proximal operator of $g = \frac{\gamma}{2}\|\mathbf{SH}\cdot - \mathbf{y}\|_2^2$ (for $\gamma > 0$) admits the closed-form expression

$$\mathrm{prox}_g(\mathbf{z}) = \mathbf{F}^*\left(\mathbf{I} - \frac{\gamma}{d}\mathbf{\Lambda}^*\mathbf{P}\mathbf{D}^{-1}\mathbf{P}^*\mathbf{\Lambda}\right)\mathbf{Fr}. \quad (18)$$

where $d = N/M$ and

- $\mathbf{r} = \mathbf{z} + \gamma\mathbf{H}^*\mathbf{S}^*\mathbf{y}$,
- $\mathbf{F} = \mathbf{I}_P \otimes \mathbf{F}_N$,
- $\mathbf{D} = \mathbf{I} + \frac{\gamma}{d}\mathbf{diag}\left(\mathbf{P}^*\mathbf{\Lambda}\mathbf{\Lambda}^*\mathbf{1}_N\right)$,
- $\mathbf{\Lambda} = [\mathbf{\Lambda}_1\ \cdots\ \mathbf{\Lambda}_P]$ is such that $\mathbf{H} = \mathbf{F}_N^*\mathbf{\Lambda}\mathbf{F}$,
- $\mathbf{P}^*$ is the adjoint of the periodization operator $\mathbf{P} = \mathbf{1}_d \otimes \mathbf{I}_M \in \mathbb{R}^{N\times M}$ (see Fig. 2).

The proof is provided in Appendix A-B.

Remark III.1: For $P = 1$, Theorem III.1 retrieves a result of [38]. However, while the proof in [38] relies on the study of the operator $\mathbf{F}_N\mathbf{S}^*\mathbf{S}\mathbf{F}_N^*$, we use here a result concerning the operator $\mathbf{SHH}^*\mathbf{S}^*$ (see Lemma A.3), which yields a shorter proof.

Then, with the proposed splitting strategy, the matrix to invert at Step 6 of Algorithm 1 becomes

$$\mathbf{B} = \sum_{l=1}^{L} \rho_l \mathbf{W}_l^*\mathbf{W}_l + \rho_{L+1}\mathbf{L}^*\mathbf{L} + \rho_{L+2}\mathbf{I}. \quad (19)$$

Although this matrix allows for faster matrix-vector products than $\mathbf{B}_{\mathrm{FS}}$ in (14), its inversion still requires an inner iterative procedure. To sidestep this bottleneck, we propose hereafter an alternative and equivalent formulation of (5) that makes $\mathbf{B}$ directly invertible in the Fourier domain.

Definition III.2: We define $\mathcal{N}_{\geqslant 0}$ as the set of linear operators $\mathbf{T} \in \mathbb{R}^{N\times N}$ that preserve the set $\mathbb{R}^N_{\geqslant 0}$ and its complement in $\mathbb{R}^N$. In other words, $\mathbf{T} \in \mathcal{N}_{\geqslant 0}$ if and only if

$$\mathbf{Tf} \in \mathbb{R}^N_{\geqslant 0}, \forall \mathbf{f} \in \mathbb{R}^N_{\geqslant 0} \quad (20)$$

$$\mathbf{Tf} \in \mathbb{R}^N \backslash \mathbb{R}^N_{\geqslant 0}, \forall \mathbf{f} \in \mathbb{R}^N \backslash \mathbb{R}^N_{\geqslant 0}. \quad (21)$$

Proposition III.2: For any functional $\mathcal{J} : \mathbb{R}^N \to \mathbb{R}$ and $\mathbf{T} \in \mathcal{N}_{\geqslant 0}$, we have that

$$\arg\min_{\mathbf{f}\in\mathbb{R}^N} \mathcal{J}(\mathbf{f}) + i_{\geqslant 0}(\mathbf{f}) = \arg\min_{\mathbf{f}\in\mathbb{R}^N} \mathcal{J}(\mathbf{f}) + i_{\geqslant 0}(\mathbf{Tf}). \quad (22)$$

Proof: The proof is straightforward since, for all $\mathbf{T} \in \mathcal{N}_{\geqslant 0}$, $i_{\geqslant 0}(\cdot) = i_{\geqslant 0}(\mathbf{T}\cdot)$. ∎

Following Proposition III.2, we reformulate Problem (5) as

$$\hat{\mathbf{f}} \in \left\{ \arg\min_{\mathbf{f}\in\mathbb{R}^{NP}} \left(\sum_{l=1}^{L} \mathcal{D}_l(\mathbf{f}, \mathbf{y}_l) + \mu\mathcal{R}(\mathbf{Lf}) + i_{\geqslant 0}(\mathbf{Tf}) \right) \right\}, \quad (23)$$

where

$$\mathbf{T} = \left(\alpha\mathbf{I}_N - \sum_{l=1}^{L} \mathbf{W}_l^*\mathbf{W}_l \right)^{\frac{1}{2}}, \quad (24)$$

with $\alpha > \|\sum_{l=1}^{L} \mathbf{W}_l^*\mathbf{W}_l\|$ to ensure that $\mathbf{T} \in \mathcal{N}_{\geqslant 0}$. In practice, we normalize the illumination patterns such that $\|\sum_{l=1}^{L} \mathbf{W}_l^*\mathbf{W}_l\| = 1$ and we set $\alpha = 2$.

Then, we replace the third splitting in (17) by

$$\mathcal{F}_{L+2} = i_{\geqslant 0},\ \mathbf{A}_{L+2} = \mathbf{T}. \quad (25)$$

Setting $\rho_l = \rho_{L+2} = \rho_{\mathcal{D}} > 0,\ \forall l \in \{1,\dots,L\}$, and $\rho_{L+1} = \rho_{\mathcal{R}} > 0$, the linear step of ADMM now amounts to inverting the matrix

$$\tilde{\mathbf{B}} = \sum_{l=1}^{L} \rho_{\mathcal{D}} \mathbf{W}_l^*\mathbf{W}_l + \rho_{\mathcal{R}}\mathbf{L}^*\mathbf{L} + \rho_{\mathcal{D}}\mathbf{T}^*\mathbf{T}$$

$$\overset{(24)}{=} \rho_{\mathcal{R}}\mathbf{L}^*\mathbf{L} + \rho_{\mathcal{D}}\alpha\mathbf{I}. \quad (26)$$

When $\mathbf{L}$ is the gradient or the Hessian operator (with periodic boundary conditions), $\mathbf{L}^*\mathbf{L}$ is a convolution operator (see Appendix B) and (26) is inverted easily at the cost of one FFT

Algorithm 2: Proposed Inner-Loop-Free ADMM for the Minimization of (23).

Require: $\mathbf{f}^0 \in \mathbb{R}^N$, $\rho_{\mathcal{D}} > 0$, $\rho_{\mathcal{R}} > 0$, $\alpha > \|\sum_{l=1}^{L} \mathbf{W}_l^* \mathbf{W}_l\|$

1: $\mathbf{u}_l^0 = \mathbf{W}_l \mathbf{f}^0,\ \forall l \in \{1, \dots, L\}$
2: $\mathbf{u}_{L+1}^0 = \mathbf{L}\mathbf{f}^0$
3: $\mathbf{u}_{L+2}^0 = \mathbf{f}^0$
4: $\mathbf{v}_l^0 = \mathbf{u}_l^0,\ \forall l \in \{1, \dots, L+2\}$
5: $k = 0$
6: **while** (not converged) **do**
7: **for** $l = 1 \dots L$ **do**
8: $\mathbf{u}_l^{k+1} = \operatorname{prox}_{\frac{1}{2\rho_{\mathcal{D}}}\|\mathbf{SH}\cdot - \mathbf{y}_l\|_2^2} \left(\mathbf{W}_l \mathbf{f}^k - \mathbf{v}_l^k\right)$
9: **end for**
10: $\mathbf{u}_{L+1}^{k+1} = \operatorname{prox}_{\frac{\mu}{\rho_{\mathcal{R}}}\mathcal{R}} \left(\mathbf{L}\mathbf{f}^k - \mathbf{v}_{L+1}^k\right)$
11: $\mathbf{u}_{L+2}^{k+1} = \operatorname{prox}_{\frac{1}{\rho_{\mathcal{D}}} i_{\geqslant 0}} \left(\mathbf{T}\mathbf{f}^k - \mathbf{v}_{L+2}^k\right)$
12: $\mathbf{b} = \rho_{\mathcal{R}} \mathbf{L}^* \left(\mathbf{u}_{L+1}^{k+1} + \mathbf{v}_{L+1}^k\right) + \rho_{\mathcal{D}} \mathbf{T}^*$
13: $\left(\mathbf{u}_{L+2}^{k+1} + \mathbf{v}_{L+2}^k\right) + \sum_{l=1}^{L} \rho_{\mathcal{D}} \mathbf{W}_l^* \left(\mathbf{u}_l^{k+1} + \mathbf{v}_l^k\right)$
14: $\mathbf{f}^{k+1} = \left(\rho_{\mathcal{R}} \mathbf{L}^* \mathbf{L} + \rho_{\mathcal{D}} \alpha \mathbf{I}\right)^{-1} \mathbf{b}$
15: $\mathbf{v}_l^{k+1} = \mathbf{v}_l^k - (\mathbf{W}_l \mathbf{f}^{k+1} - \mathbf{u}_l^{k+1}),\ \forall l \in \{1, \dots, L\}$
16: $\mathbf{v}_{L+1}^{k+1} = \mathbf{v}_{L+1}^k - (\mathbf{L}\mathbf{f}^{k+1} - \mathbf{u}_{L+1}^{k+1})$
17: $\mathbf{v}_{L+2}^{k+1} = \mathbf{v}_{L+2}^k - (\mathbf{T}\mathbf{f}^{k+1} - \mathbf{u}_{L+2}^{k+1})$
18: $k = k + 1$
19: **end while**

and one iFFT. Hence, we obtain an inner-loop-free ADMM, which is summarized in Algorithm 2.

C. Algorithm Complexity

We briefly discuss the complexity of one iteration of our modified ADMM. Line 8 in Algorithm 2 requires L evaluations of the proximal operator given in Theorem III.1. It corresponds to applying LP Fourier transforms of size N and as many inverse Fourier transforms, yielding a complexity of $\mathcal{O}(2LPN\log(N))$. Then, the complexity of steps 10 and 11 is generally linear (with classical regularizers $\mathcal{R}(\mathbf{L}\cdot)$ such as TV or Hessian-Schatten). It is thus negligible compared to step 8. Finally, the linear step in Line 14 is computed in the Fourier domain at the cost of one Fourier transform and one inverse Fourier transform of size NP, which corresponds to a complexity of $\mathcal{O}(2NP\log(NP))$. The overall complexity of one iteration is thus $\mathcal{O}(2NP(L\log(N) + \log(NP)))$.

D. Convergence Analysis

We prove the convergence of our inner-loop-free algorithm in Proposition III.3.

Proposition III.3: Assume that $(\bigcap_{l=1}^{L} \ker(\mathbf{SHW}_l)) \cap \ker(\mathbf{L}) = \{\mathbf{0}_N\}$ and that $\mathcal{R}$ is coercive. Then, Algorithm 2 converges to a solution of (23) and, thus, to a solution of (5).

Proof: The functions $\mathcal{F}_l$, $l \in \{1, \dots, L+2\}$, (defined by (15), (16), and (25)) are proper, closed, and convex (by definition for $\mathcal{R}$). This implies the same properties for the objective function in (23). Moreover, the latter is also coercive because $\mathcal{R}$ is coercive and $(\bigcap_{l=1}^{L} \ker(\mathbf{SHW}_l)) \cap \ker(\mathbf{L}) = \{\mathbf{0}_N\}$. Hence its set of minimizers is nonempty [39, Theorem 2.5.1 (ii)]. Then, because the identity matrix $\mathbf{I}_N$ has a full rank, $\mathbf{G} = [\mathbf{W}_l \cdots \mathbf{W}_L\ \mathbf{L}\ \mathbf{I}_N]^*$ has also a full rank. This property, combined with the fact that all steps in Algorithm 2 are solved exactly by our new formulation of the problem, guarantees the convergence of the algorithm to a solution of (23) using [40, Proposition 1]. ■

The assumptions in Proposition III.3 are generally satisfied. For instance, considering the TV regularizer, $\mathcal{R}$ is coercive and the first assumption reduces to $\mathbf{1}_N \notin (\bigcap_{l=1}^{L} \ker(\mathbf{SHW}_l))$. This is always satisfied because the patterns are nonzero and $\mathbf{H}$ is a lowpass filter in SIM.

IV. Validation of the Proposed Method

In this section, we present a series of numerical experiments that are dedicated to the evaluation of several aspects of our reconstruction method. First, we study the efficiency of Algorithm 2 in terms of convergence speed, and compare it to other algorithms that minimize the same objective. Then, using simulated data, we evaluate the ability of the proposed method to reject the out-of-focus signal (within the additional planes that are considered in the model). Finally, we present reconstruction results on real data and compare them to those obtained by the open-source FairSIM software, as well as to the corresponding plane of a full 3D SIM-reconstructed volume.

A. Efficiency of the Proposed Inner-Loop-Free ADMM

Fig. 3 depicts the empirical convergence of the objective function in (23) for the proposed inner-loop-free ADMM. We also provide the corresponding curves for ADMM with full-splitting (FS) (i.e., equations (11)–(13)) and for the primal-dual algorithm [34] with the splitting strategy proposed in [12]. We refer the reader to [12] for details. All methods are implemented with the GlobalBioIm library [41]. Algorithm parameters (i.e., Lagrangian multipliers ρ for ADMM and τ for the primal-dual method[3]) have been tuned empirically in order to get the fastest convergence. For ADMM with FS, inner conjugate-gradient (CG) [42] iterations are required to solve the linear step. One can observe that performing more CG iterations does not improve the convergence in terms of iterations (Fig. 3, top) but consequently increase the execution time of the algorithm (Fig. 3, bottom) for this large-scale dataset. Concerning the primal-dual algorithm, it requires more iterations to converge but is much less computationally demanding than ADMM with FS. Finally, the proposed inner-loop-free ADMM is significantly more efficient. Not only does it enjoy a better convergence with respect to the number of iterations (Fig. 3, top), but its cost per iteration is also lower, leading to much faster computations (Fig. 3, bottom).

[3]There is also two others parameters which are chosen as suggested in [12].

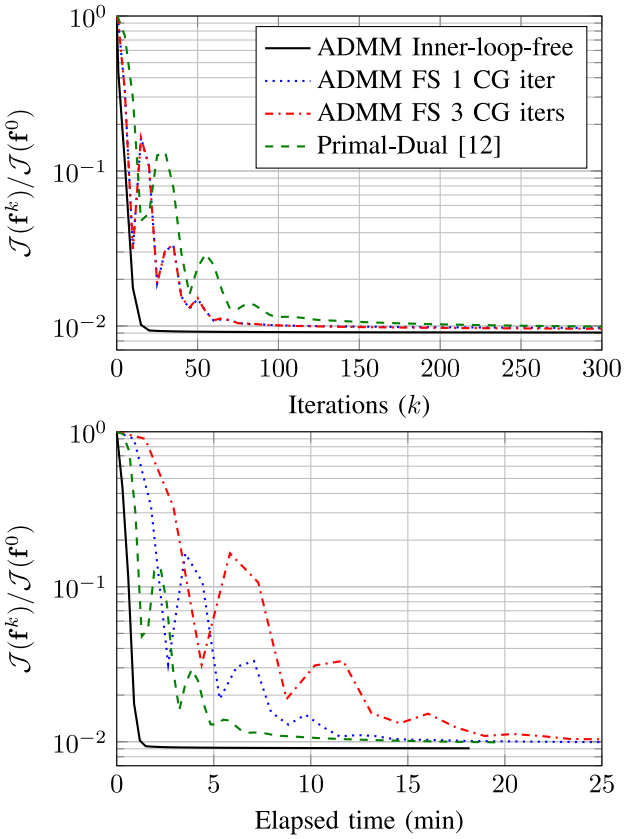

Fig. 3. Empirical convergence of the objective function in (23) for the proposed inner-loop-free ADMM on a $(1024 \times 1024 \times 4)$ size problem (i.e., $N_1 = N_2 = 1024$, $P = 4$). For comparison, we provide the convergence curves for ADMM with full-splitting (FS), using different numbers of inner CG iterations, as well as the curve for the primal-dual method proposed in [12]. Top: convergence curves with respect to iterations. Bottom: convergence curves with respect to elapsed time.

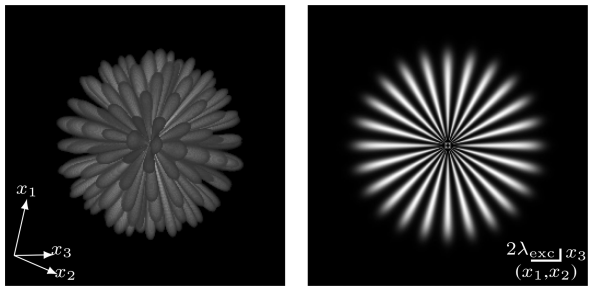

Fig. 4. Three-dimensional sample used in the experiments. Left: 3D rendering. Right: slice $x_1 = 0$. For this sample, the three planes ($x_1 = 0$, $x_2 = 0$, and $x_3 = 0$) are identical.

B. Rejection of Out-of-Focus Light

1) Simulation Setting: We consider the three-dimensional sample depicted in Fig. 4. Its size is $(256 \times 256 \times 256)$. Illumination patterns are generated according to the two-beam model

$$w(\mathbf{x}, z) \propto a_0 + a_1 \cos\left(2\left(k_1 x_1 + k_2 x_2 + \varphi\right)\right), \qquad (27)$$

where $a_0 > 0$, $a_1 > 0$ are weight parameters, $\mathbf{k} = [k_1\, k_2]$ denotes the wave vector, and φ corresponds to a phase shift. One can note that w corresponds to a two-dimensional

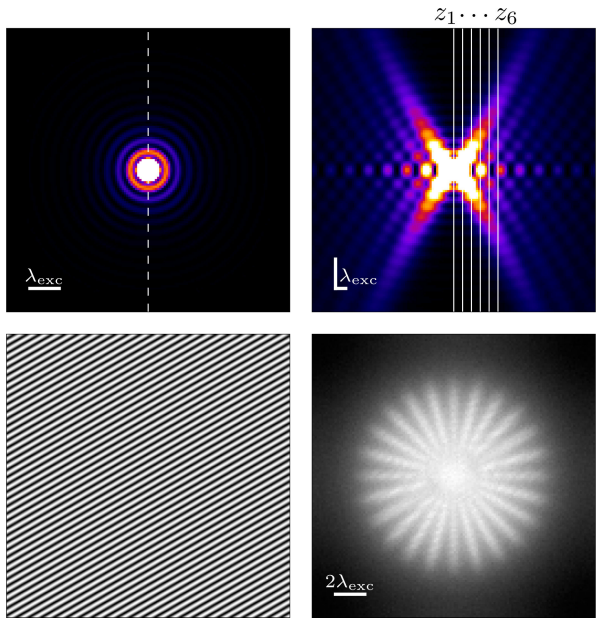

Fig. 5. Synthesized data. Top row: Born-and-Wolf PSF used for data simulation. Focal plane (left) and axial ($x_1 = 0$) plane (right). Vertical lines represent the different planes used in the reconstruction $\{z_p\}_{p=1}^{P}$. Bottom row: illumination pattern (left) and corresponding simulated acquisition at focal plane (right).

pattern that does not depends on the axial dimension. Nine patterns (i.e., $L = 9$) are generated by using three lateral orientations $\{0, \pi/3, 2\pi/3\}$ of the wave vector $\mathbf{k}$ and three lateral phase shifts φ $\{0, \pi/3, 2\pi/3\}$. The angle between each beam and the optical axis is $\beta = \arcsin(\mathrm{NA}/n_{\text{sam}})$, where the objective numerical aperture is $\mathrm{NA} = 1.4$ and the refractive index of the sample is $n_{\text{sam}} = 1.333$. The excitation wavelength is set to $\lambda_{\text{exc}} = 561$ nm and the objective is immersed in oil ($n_i = 1.518$). Lateral and axial resolutions are set to 40nm and 100nm, respectively. We use the shift-invariant Born-and-Wolf PSF model (Fig. 5, top) and generate it using the PSF generator from [43]. Only the central plane of the volume is kept (but simulations are 3D on the whole $256 \times 256 \times 256$ volume) and downsized by a factor of two (using averaging). This results in nine two-dimensional acquisitions $\{\tilde{\mathbf{y}}_l\}_{l=1}^{9}$ of size $(M_1 = M_2 = 128)$. These noiseless measurements are normalized such that the average number of photons in the sum of all images is 10^4, so that

$$\frac{1}{M}\left\|\sum_{l=1}^{9} \tilde{\mathbf{y}}_l\right\|_1 = 10^4. \qquad (28)$$

Finally, the noisy data are obtained according to $\mathbf{y}_l = \mathcal{P}(\tilde{\mathbf{y}}_l)$, where $\mathcal{P}$ denotes the Poisson distribution. A simulated acquisition for one structured illumination is presented in Fig. 5 (bottom). It is noteworthy that the simulated sample is very thick, which leads to a strong out-of-focus signal.

2) Reconstruction Results and Discussion: We applied to our simulated data the proposed inner-loop-free ADMM. We considered the Hessian-Schatten-norm regularizer [32], independently on each lateral slice of the reconstructed volume. With this regularizer and periodic boundary conditions, the

TABLE I
ACQUISITION PARAMETERS FOR THE THREE REAL DATASETS OF FIG. 7

	Dataset 1	Dataset 2	Dataset 3
Size (M_1, M_2)	256×256	1024×1024	1024×1024
Lateral resolution	64.5 nm	64.5 nm	64.5 nm
Wavelength λ_{exc}	488 nm	561 nm	488 nm
Numerical aperture	1.4	1.4	1.4
# orientations	5	5	5
# phase shifts	5	5	5

operator in (26) is an invertible convolution operator (see Appendix B-B). Because we used a symmetric PSF, the considered planes $\{z_p\}_{p=1}^{P}$ are selected from the same side of the PSF (Fig. 5, top-right), as proposed in [26]. This allows us to further increase the computational efficiency. For each reconstruction, the hyper-parameter μ is tuned so as to maximize the reconstruction signal-to-noise ratio (RSNR) of the reconstructed image. It is defined by $\text{RSNR} = 20\log(\|\mathbf{f}^\star\|/\|\hat{\mathbf{f}} - \mathbf{f}^\star\|)$ where $\hat{\mathbf{f}}$ ($\mathbf{f}^\star$, respectively) is the reconstructed (ground truth, respectively) image. More precisely, we selected the optimal μ among 15 values logarithmically equally spaced between 10^{-5} and 10^{-2}.

Reconstruction results for different number of planes $P \in \{1, 2, 4, 6\}$, always spaced by 400nm, are presented in Fig. 6. The quality clearly improves when additional planes are considered. A significant part of the strong out-of-focus light is rejected. Moreover, only few additional planes (e.g., up to three) are sufficient to reveals details that are occluded on the classical 2D reconstruction (i.e., no additional planes, $P = 1$). One can also appreciate the substantial improvement with respect to the FairSIM reconstruction [19], even when the OTF attenuation option is activated (see [19] for details). Finally, as expected, we observe that SIM reconstructions always have a better resolution than a more conventional thick-slice 2D deconvolution, as realized by the direct adaptation of the proposed method to 2D-deconvolution. (It corresponds to Model (4) without the diagonal operator $\mathbf{W}$.) There are structural details that are absent from the deconvolved profile but clearly distinguishable on the reconstructed SIM profiles (bottom graph in Fig. 6).

C. Reconstructions of Real Data

We validated our method on several datasets acquired with the Zeiss Elyra microscope.[4] Acquisition parameters are summarized in Table I. Reconstructions have been performed with the proposed method and with FairSIM [19] for comparison. FairSIM implements the classic two-dimensional multichannel Wiener reconstruction. For each dataset, the Wiener parameter was set to 0.1 and an OTF attenuation (default setting) was used to improve axial sectioning. For both methods, the PSF was approximated by a theoretical expression (Born-and-Wolf, Fig. 5, top row) using the acquisition parameters provided in Table I. For the proposed computational super-sectioning reconstruction, operators $\mathbf{W}_l$, $l \in \{1, \ldots, L\}$, were derived from

[4]Courtesy of Carl Zeiss Research Department.

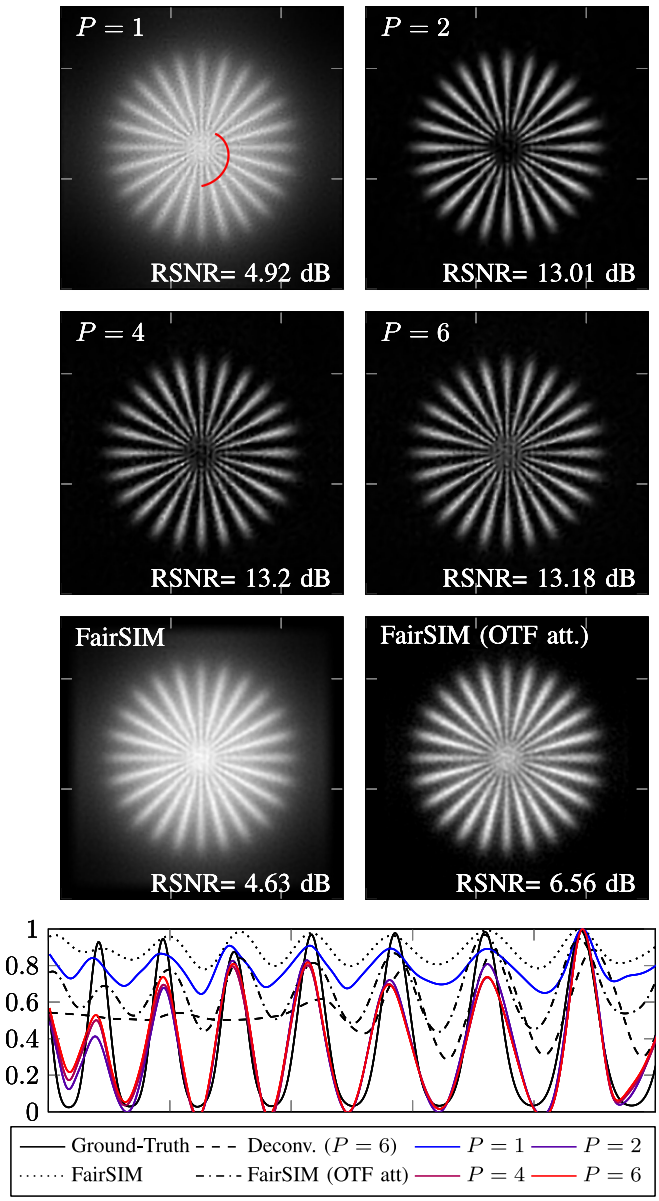

Fig. 6. Effect of the number $P \in \{1, 2, 4, 6\}$ of planes considered for reconstruction. The Hessian-Schatten-norm is used as regularizer. Some results obtained with the FairSIM software [19] (Wiener parameter 0.5), with and without OTF attenuation, are provided for comparison. Bottom graph: one-dimensional (1-D) profiles that correspond to the spiral portion shown in the top-left image. For comparison, we provide the profile of the ground truth as well as of a deconvolved image with $P = 6$.

the pattern parameters (wave vector $\mathbf{k}$ and phase shift φ) estimated by FairSIM. Hence, the comparison of both methods is fair. As regularizer, we used the Hessian-Schatten-norm (see Appendix B-B) with $\mu = 10^{-5}$. Because data were acquired at Nyquist rate and because we aim at doubling the resolution, reconstructions were performed on a twice-finer grid. According to the results in Fig. 6, we used one out-of-focus plane ($P = 2$, tradeoff between problem size and reconstruction quality), leading to a reconstruction problem of size ($N_1 = N_2 = 512$, $P = 2$) for dataset 1 and ($N_1 = N_2 = 2048$, $P = 2$) for the two other datasets. Finally, the inner-loop-free ADMM was stopped either after 100 iterations or when the relative difference of the

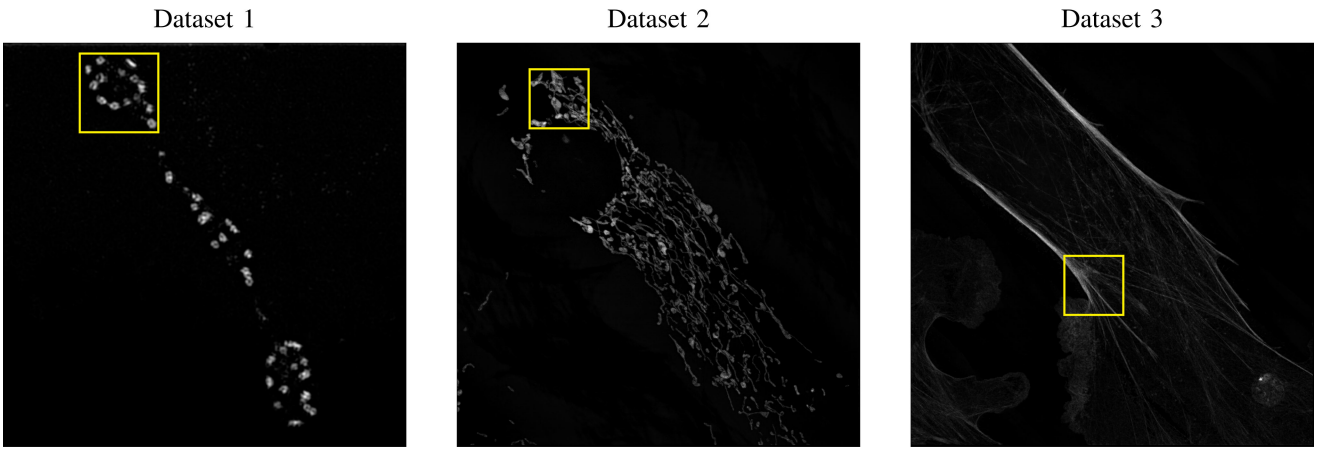

Fig. 7. Reconstruction results for the three datasets in Table I. Squares indicate the regions presented in Fig. 8.

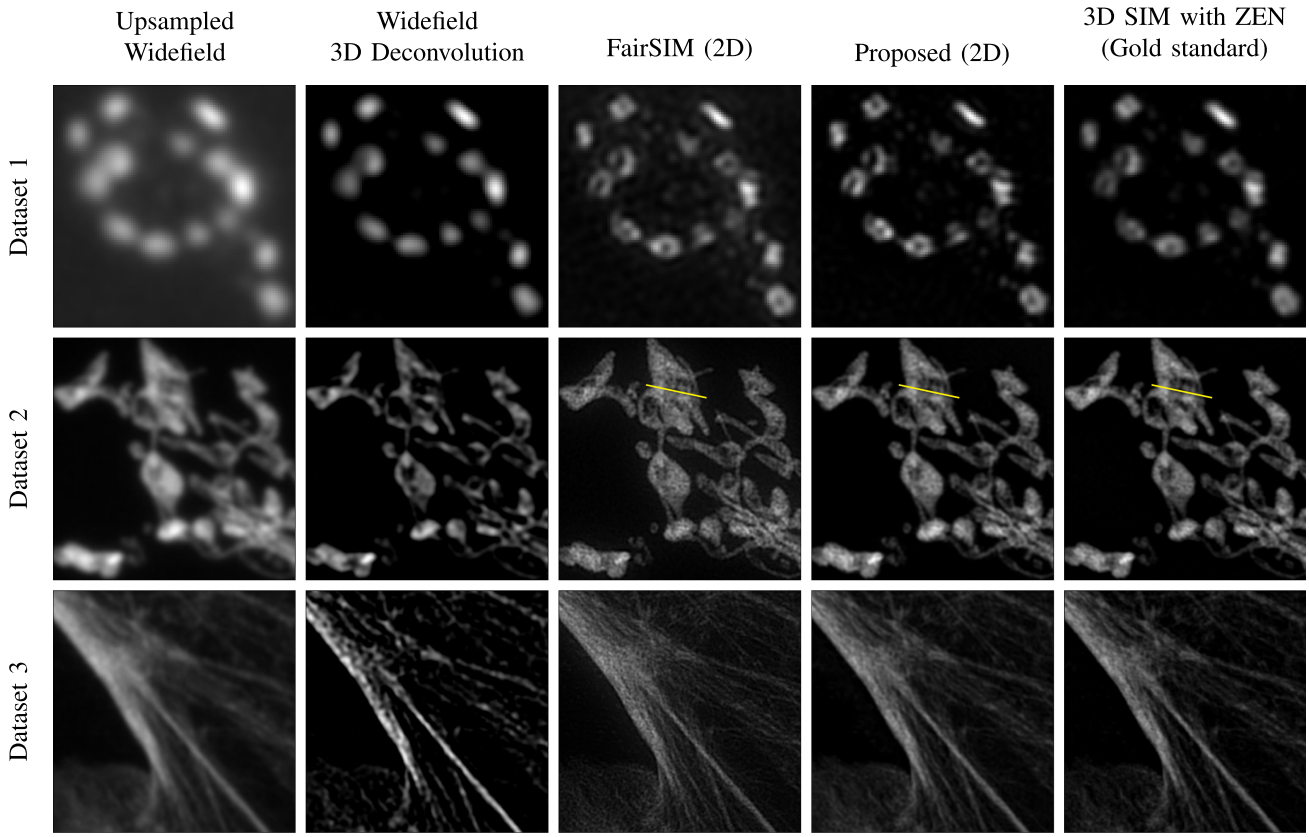

Fig. 8. Cutouts of Fig. 7. Profile plots corresponding to the yellow lines are presented in Fig. 9.

cost function between two successive iterates was below 10^{-5}, whichever happened first.

Reconstructions at focal plane z_1 obtained with the proposed method are presented in Fig. 7, with cutouts of specific regions in Fig. 8. The corresponding plane of a 3D deconvolved stack, as well as the one of a 3D SIM-reconstructed volume using the ZEN software developed by ZEISS, is also provided for reference.

First, the resolution enhancement is clearly visible by comparing the SIM results with the deconvolved images. Second, the proposed method provides reconstructed images which are really close to those obtained by ZEN (full 3D SIM reconstruction). Although only a single data slice per pattern is used, the algorithm is able to properly reject the out-of-focus signal within the additional planes so as to reach an axial sectioning comparable to the one obtained while taking into account the full 3D dataset. This is not the case for FairSIM, for which the reconstructed structures are less sharp. These observations are further exemplified with the line plots depicted in Fig. 9. We observe spurious peaks (top arrows) in the FairSIM reconstruction, which are absent from the proposed and 3D-SIM (ZEN) reconstructions. However, the corresponding line profile for a plane

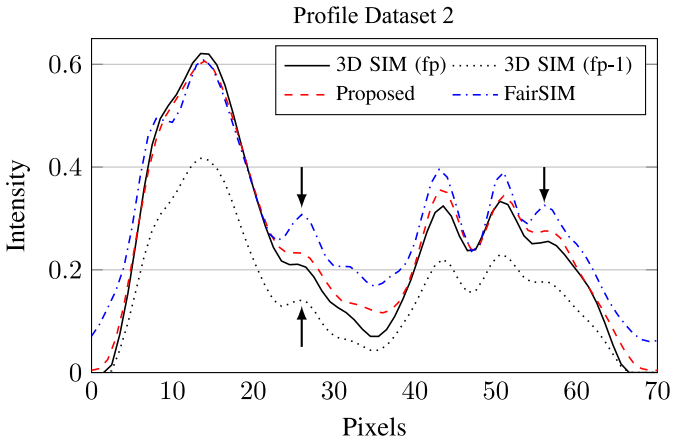

Fig. 9. Line plots of Fig. 8. The abbreviation fp stands for "focal plane".

in front of the focal one in the 3D reconstruction (dotted line) reveals that the spurious details in the FairSIM reconstruction correspond in fact to structures that live in a different plane of the 3D reconstruction (bottom arrow). By contrast, the proposed method was able to reject this signal from the plane of interest.

V. Conclusion

It is well documented that a full 3D reconstruction can yield a superior quality than its slice-by-slice counterpart. However, we demonstrated in this paper that the latter can be improved by computational sectioning, that is, by considering additional planes in the model. By doing so, the 3D nature of the sample is considered, allowing for the rejection of most of the out-of-focus light that affects the reconstructions obtained from 2D acquisitions. On several real datasets, we showed that the proposed method provides super-resolved images with an axial sectioning comparable to the performance of a full 3D reconstruction. The significance of the approach is furthered by the fact that many biological studies only require 2D super-resolved images, which simplifies sample preparation and acquisition. To tackle the challenging large-scale inverse problem, we proposed an efficient inner-loop-free alternating-direction method of multipliers that relies on a suitable formulation of the optimization problem together with closed-form expressions of proximal operator. It offers a new fast solution to solve the problem of reconstructing structured-illumination microscopy slices. Moreover, it is noteworthy to mention that, although only sinusoidal illuminations were used in this paper, the proposed approach can work with any other pattern (e.g., speckle, nonlinear). An interesting future direction to improve the proposed method would be to adapt the inner-loop-free algorithm to data-fidelity terms that account for Poisson noise.

Appendix A
Closed-Form Proximal Operator

A. Preliminary Results

This section gathers some results that will be used to prove Theorem III.1.

Lemma A.1 (Woodbury matrix identity [44]): Let $\mathbf{A}$, $\mathbf{U}$, $\mathbf{C}$, and $\mathbf{V}$ be matrices with appropriate sizes. Let $\mathbf{A}$ and $\mathbf{C}$ be invertible. Then,

$$(\mathbf{A}+\mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{U}\left(\mathbf{C}^{-1}+\mathbf{VA}^{-1}\mathbf{U}\right)^{-1}\mathbf{VA}^{-1}. \quad (29)$$

Lemma A.2 (Stretch property of the DFT): Let $\mathbf{F}_N$ ($\mathbf{F}_M$, respectively) be the N-point (M-point, respectively) unitary DFT and $\mathbf{S}\in\mathbb{R}^{M\times N}$ be a d-decimation operator (i.e., $dM=N$). Then, $\mathbf{S}^*$ is a stretch operator, so that, for $\mathbf{x}\in\mathbb{R}^M$,

$$[\mathbf{S}^*\mathbf{x}]_n = \begin{cases} x_{n/d} & \text{if } n/d \text{ integer},\\ 0 & \text{otherwise}, \end{cases} \quad (30)$$

and we have the equality

$$\mathbf{PF}_M = \sqrt{d}\mathbf{F}_N\mathbf{S}^*, \quad (31)$$

where $\mathbf{P}=\mathbf{1}_d\otimes\mathbf{I}_M\in\mathbb{R}^{N\times M}$ is a periodization operator (see Fig. 2).

Proof: Let $\mathbf{x}\in\mathbb{R}^M$, then

$$\begin{aligned}
[\mathbf{F}_N\mathbf{S}^*\mathbf{x}]_k &= \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}[\mathbf{S}^*\mathbf{x}]_n \mathrm{e}^{-\frac{2\mathrm{i}\pi}{N}kn}\\
&\overset{(30)}{=} \frac{1}{\sqrt{N}}\sum_{m=0}^{M-1} x_m \mathrm{e}^{-\frac{2\mathrm{i}\pi}{N}kmd}\\
&\overset{N=dM}{=} \frac{1}{\sqrt{dM}}\sum_{m=0}^{M-1} x_m \mathrm{e}^{-\frac{2\mathrm{i}\pi}{M}(k-\lfloor k/M\rfloor M)m}\\
&= \frac{1}{\sqrt{d}}\left[\mathbf{F}_M\mathbf{x}\right]_{k-\lfloor k/M\rfloor M}\\
&= \frac{1}{\sqrt{d}}\left[\mathbf{PF}_M\mathbf{x}\right]_k, \quad (32)
\end{aligned}$$

which completes the proof. ■

Remark A.1: The extension of Lemma A.2 to a multi-dimensional unitary DFT is straightforward. The same result holds with $d=N/M=\prod_k d_k$, where $d_k\in\mathbb{R}$ are the downsampling factors in each dimension.

Lemma A.3: Let $\mathbf{H}\in\mathbb{R}^{N\times N}$ be a convolution operator with kernel $\mathbf{h}\in\mathbb{R}^N$ and $\mathbf{S}\in\mathbb{R}^{M\times N}$ be a d-decimation operator (i.e., $dM=N$). Then, $\mathbf{I}+\gamma\mathbf{SHS}^*$ (for $\gamma>0$) is a convolution operator. More precisely, we have that

$$\mathbf{I}+\gamma\mathbf{SHS}^* = \mathbf{F}_M^*\left(\mathbf{I}+\frac{\gamma}{d}\mathbf{diag}\left(\mathbf{P}^*\mathbf{\Lambda}\mathbf{1}_N\right)\right)\mathbf{F}_M, \quad (33)$$

where $\mathbf{\Lambda}=\mathbf{diag}\left(\mathbf{F}_N\mathbf{h}\right)$ and $\mathbf{P}^*$ is the adjoint of the periodization operator $\mathbf{P}=\mathbf{1}_d\otimes\mathbf{I}_M\in\mathbb{R}^{N\times M}$ (see Fig. 2).

Proof: Let $\mathbf{S}$ and $\mathbf{\Lambda}$ be defined as in the statement of Lemma A.3. Then, we have that

$$\begin{aligned}
\mathbf{I}+\gamma\mathbf{SHS}^* &= \mathbf{I}+\gamma\mathbf{SF}_N^*\mathbf{\Lambda}\mathbf{F}_N\mathbf{S}^*\\
&\overset{\text{Lem.A.2}}{=} \mathbf{I}+\frac{\gamma}{d}\mathbf{F}_M^*\mathbf{P}^*\mathbf{\Lambda}\mathbf{P}\mathbf{F}_M\\
&= \mathbf{I}+\frac{\gamma}{d}\mathbf{F}_M^*\mathbf{diag}\left(\mathbf{P}^*\mathbf{\Lambda}\mathbf{1}_N\right)\mathbf{F}_M. \quad (34)
\end{aligned}$$

The last equality comes from the fact that $\mathbf{P}^*\mathbf{\Lambda}\mathbf{P}\mathbf{x}=\sum_{k=1}^d \mathbf{\Lambda}_k\odot\mathbf{x}=\mathbf{diag}(\sum_{k=1}^d\mathbf{\Lambda}_k)\mathbf{x}=\mathbf{diag}(\mathbf{P}^*\mathbf{\Lambda}\mathbf{1}_N)\mathbf{x}$, where $\mathbf{\Lambda}_k\in\mathbb{R}^M$ are blocks of the diagonal of $\mathbf{\Lambda}$. ■

B. Proof of Theorem III.1

We first recall the notation $\mathbf{H} = [\mathbf{H}_1 \ \cdots \ \mathbf{H}_P]$ introduced in Section II-A which can also be expressed as $\mathbf{H} = \mathbf{F}_N^* \mathbf{\Lambda} \mathbf{F}$ with $\mathbf{F} = \mathbf{I}_P \otimes \mathbf{F}_N$, $\mathbf{\Lambda} = [\mathbf{\Lambda}_1 \ \cdots \ \mathbf{\Lambda}_P]$, and $\mathbf{\Lambda}_p = \mathbf{diag}\left(\mathbf{F}_N \mathbf{h}_p\right)$, $p \in \{1, \dots, P\}$. Then, from Definition (9) of the proximal operator and letting $g = \frac{\gamma}{2}\|\mathbf{SH} \cdot - \mathbf{y}\|_2^2$, we have that

$$\mathrm{prox}_g(\mathbf{z}) = \left(\mathbf{I} + \gamma \mathbf{H}^* \mathbf{S}^* \mathbf{S} \mathbf{H}\right)^{-1} \left(\mathbf{z} + \gamma \mathbf{H}^* \mathbf{S}^* \mathbf{y}\right). \quad (35)$$

In the sequel we set $\mathbf{r} = \mathbf{z} + \gamma \mathbf{H}^* \mathbf{S}^* \mathbf{y}$. Then, from the Woodbury matrix identity in Lemma A.1 with $\mathbf{A} = \mathbf{C} = \mathbf{I}$, $\mathbf{U} = \gamma \mathbf{H}^* \mathbf{S}^*$, and $\mathbf{V} = \mathbf{SH}$, we get

$$\mathrm{prox}_g(\mathbf{z}) = \left(\mathbf{I} - \gamma \mathbf{H}^* \mathbf{S}^* (\mathbf{I} + \gamma \mathbf{S} \mathbf{H} \mathbf{H}^* \mathbf{S}^*)^{-1} \mathbf{S} \mathbf{H}\right) \mathbf{r}. \quad (36)$$

Noticing that $\mathbf{H}\mathbf{H}^* = \mathbf{F}_N^* \mathbf{\Lambda} \mathbf{\Lambda}^* \mathbf{F}_N$, Lemma A.3 leads to

$$\mathbf{I} + \gamma \mathbf{S} \mathbf{H} \mathbf{H}^* \mathbf{S}^* = \mathbf{F}_M^* \left(\mathbf{I} + \frac{\gamma}{d} \mathbf{diag}\left(\mathbf{P}^* \mathbf{\Lambda} \mathbf{\Lambda}^* \mathbf{1}_N\right)\right) \mathbf{F}_M. \quad (37)$$

Then, letting $\mathbf{D} = \mathbf{I} + \frac{\gamma}{d} \mathbf{diag}\left(\mathbf{P}^* \mathbf{\Lambda} \mathbf{\Lambda}^* \mathbf{1}_N\right)$, and injecting (37) into (36), we obtain

$$\begin{aligned} \mathrm{prox}_g(\mathbf{z}) &= \left(\mathbf{I} - \gamma \mathbf{F}^* \mathbf{\Lambda}^* \mathbf{F}_N \mathbf{S}^* \mathbf{F}_M^* \mathbf{D}^{-1} \mathbf{F}_M \mathbf{S} \mathbf{F}_N^* \mathbf{\Lambda} \mathbf{F}\right) \mathbf{r} \\ &\overset{\text{Lem. A.2}}{=} \left(\mathbf{I} - \frac{\gamma}{d} \mathbf{F}^* \mathbf{\Lambda}^* \mathbf{P} \mathbf{F}_M \mathbf{F}_M^* \mathbf{D}^{-1} \mathbf{F}_M \mathbf{F}_M^* \mathbf{P}^* \mathbf{\Lambda} \mathbf{F}\right) \mathbf{r} \\ &= \mathbf{F}^* \left(\mathbf{I} - \frac{\gamma}{d} \mathbf{\Lambda}^* \mathbf{P} \mathbf{D}^{-1} \mathbf{P}^* \mathbf{\Lambda}\right) \mathbf{F} \mathbf{r}. \end{aligned} \quad (38)$$

which completes the proof.

Appendix B
Total-Variation and Hessian-Schatten Norm Regularizers

A. Total Variation

In its isotropic form, the total-variation seminorm penalizes the ℓ_1-norm of the gradient magnitudes of $\mathbf{f}$. It can be expressed as the composition of the gradient operator $\mathbf{\nabla} : \mathbb{R}^N \to \mathbb{R}^{N \times D}$ ($D > 0$, the number of dimensions) with the (ℓ_2, ℓ_1)-mixed norm

$$\forall \mathbf{x} \in \mathbb{R}^{N \times D}, \ \|\mathbf{x}\|_{2,1} = \sum_{n=1}^{N} \|x_{n,\cdot}\|_2, \quad (39)$$

where $x_{n,\cdot} \in \mathbb{R}^D$. It leads to $\|\mathbf{x}\|_{\mathrm{TV}} = \|\mathbf{\nabla} \mathbf{x}\|_{2,1}$. In order to use the proposed ADMM with the isotropic TV, the proximal operator of $\|\cdot\|_{2,1}$ is required. It admits the following closed-form expression [45], $\forall \mathbf{x} \in \mathbb{R}^{N \times D}$,

$$\left[\mathrm{prox}_{\gamma \|\cdot\|_{2,1}}(\mathbf{x})\right]_{n,\cdot} = x_{n,\cdot} \left(1 - \frac{\gamma}{\|x_{n,\cdot}\|_2}, 0\right)_+, \quad (40)$$

where $(\cdot)_+ = \max(\cdot, 0)$. Then, $\mathbf{\nabla}$ is involved in ADMM through the operator $\mathbf{\nabla}^* \mathbf{\nabla}$ which constitutes one term of the operator to invert at Line 14 of Algorithm 2. Considering periodic boundary conditions (i.e., $\mathbf{x} \in \mathbb{R}^N$ is extended according to $\forall n \in \mathbb{N}$, $x_n = x_{n - \lfloor n/N \rfloor N}$), $\mathbf{\nabla}^* \mathbf{\nabla}$ is the convolution operator

$$\mathbf{\nabla}^* \mathbf{\nabla} = \mathbf{F}_N^* \mathbf{\Lambda}_{\mathbf{\nabla}} \mathbf{F}_N, \quad (41)$$

with a properly defined diagonal operator $\mathbf{\Lambda}_{\mathbf{\nabla}}$ that one can easily derive (Laplacian filter).

B. Hessian-Schatten Norm

This regularization has been proposed in [32]. It is defined as the composition of the Hessian operator $\mathbf{H_e} : \mathbb{R}^N \to \mathbb{R}^{N \times D \times D}$ ($D > 0$, the number of dimensions)

$$[\mathbf{H_e} \mathbf{x}]_{n,\cdot,\cdot} = \left\{[\mathbf{D}_{ij} \mathbf{x}]_n\right\}_{1 \le i,j \le D}, \quad (42)$$

where $x_{n,\cdot,\cdot} \in \mathbb{R}^{D \times D}$, and $\mathbf{D}_{ij}$ is the second-order derivative operator along dimensions i and j, with the $(\mathcal{S}_p, \ell_1)$-mixed norm ($p \ge 1$)

$$\forall \mathbf{x} \in \mathbb{R}^{N \times D \times D}, \ \|\mathbf{x}\|_{\mathcal{S}_p,1} = \sum_{n=1}^{N} \|x_{n,\cdot,\cdot}\|_{\mathcal{S}_p}, \quad (43)$$

where $\|\cdot\|_{\mathcal{S}_p}$ denotes the Schatten norm of order p. It is defined as the ℓ_p-norm of the vector of singular values of its argument $x_{n,\cdot,\cdot} \in \mathbb{R}^{D \times D}$. As for the TV regularizer, the proximal operator of the $(\mathcal{S}_p, \ell_1)$-mixed norm has the closed-form expression [46], [47], $\forall \mathbf{x} \in \mathbb{R}^{N \times D \times D}$,

$$\left[\mathrm{prox}_{\gamma \|\cdot\|_{\mathcal{S}_p,1}}(\mathbf{x})\right]_{n,\cdot,\cdot} = \mathbf{U}_n \left(\mathrm{prox}_{\gamma \|\cdot\|_p}(\mathbf{\Sigma}_n)\right) \mathbf{V}_n^T, \quad (44)$$

where $\mathbf{U}_n$, $\mathbf{V}_n$, and $\mathbf{\Sigma}_n$ are obtained from a singular-value decomposition of $x_{n,\cdot,\cdot} = \mathbf{U}_n \mathbf{\Sigma}_n \mathbf{V}_n$. Moreover, considering periodic boundary conditions, $\mathbf{H_e}^* \mathbf{H_e}$ is also a convolution operator.

References

[1] R. Heintzmann and C. G. Cremer, "Laterally modulated excitation microscopy: Improvement of resolution by using a diffraction grating," *Proc. SPIE*, vol. 3568, pp. 185–196, Jan. 1999.

[2] M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microscopy*, vol. 198, no. 2, pp. 82–87, Dec. 2001.

[3] J. T. Frohn, H. F. Knapp, and A. Stemmer, "Three-dimensional resolution enhancement in fluorescence microscopy by harmonic excitation," *Opt. Lett.*, vol. 26, no. 11, pp. 828–830, Jun. 2001.

[4] M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "Doubling the lateral resolution of wide-field fluorescence microscopy using structured illumination," *Proc. SPIE*, vol. 3919, pp. 141–150, May 2000.

[5] M. G. Gustafsson *et al.*, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys. J.*, vol. 94, no. 12, pp. 4957–4970, Jun. 2008.

[6] J. Idier, S. Labouesse, M. Allain, P. Liu, S. Bourguignon, and A. Sentenac, "On the superresolution capacity of imagers using unknown speckle illuminations," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 87–98, Mar. 2018.

[7] M. G. L. Gustafsson, "Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution," *Proc. Nat. Acad. Sci. United States Amer.*, vol. 102, no. 37, pp. 13081–13086, Sep. 2005.

[8] R. Heintzmann, T. M. Jovin, and C. Cremer, "Saturated patterned excitation microscopy — A concept for optical resolution improvement," *JOSA A*, vol. 19, no. 8, pp. 1599–1609, Aug. 2002.

[9] R. Heintzmann, "Saturated patterned excitation microscopy with two-dimensional excitation patterns," *Micron*, vol. 34, no. 6, pp. 283–291, Oct. 2003.

[10] T. Lukeš, G. M. Hagen, P. Křížek, Z. Švindrych, K. Fliegel, and M. Klíma, "Comparison of image reconstruction methods for structured illumination microscopy," *Proc. SPIE*, vol. 9129, May 2014, Art. no. 91293J.

[11] F. Orieux, E. Sepulveda, V. Loriette, B. Dubertret, and J.-C. Olivo-Marin, "Bayesian estimation for optimized structured illumination microscopy," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 601–614, Feb. 2012.

[12] J. Boulanger, N. Pustelnik, L. Condat, L. Sengmanivong, and T. Piolot, "Nonsmooth convex optimization for structured illumination microscopy image reconstruction," *Inverse Probl.*, vol. 34, no. 9, 2018, Art. no. 095004.

[13] K. Chu *et al.*, "Image reconstruction for structured-illumination microscopy with low signal level," *Opt. Express*, vol. 22, no. 7, pp. 8687–8702, Apr. 2014.
[14] N. Chakrova, B. Rieger, and S. Stallinga, "Deconvolution methods for structured illumination microscopy," *JOSA A*, vol. 33, no. 7, pp. B12–B20, May 2016.
[15] E. Mudry *et al.*, "Structured illumination microscopy using unknown speckle patterns," *Nature Photon.*, vol. 6, no. 5, pp. 312–315, Apr. 2012.
[16] R. Ayuk *et al.*, "Structured illumination fluorescence microscopy with distorted excitations using a filtered blind-SIM algorithm," *Opt. Lett.*, vol. 38, no. 22, pp. 4723–4726, Nov. 2013.
[17] S. Labouesse *et al.*, "Joint reconstruction strategy for structured illumination microscopy with unknown illuminations," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2480–2493, May 2017.
[18] L.-H. Yeh, L. Tian, and L. Waller, "Structured illumination microscopy with unknown patterns and a statistical prior," *Biomed. Opt. Express*, vol. 8, no. 2, pp. 695–711, Jan. 2017.
[19] M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, and T. Huser, "Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ," *Nature Commun.*, vol. 7, Mar. 2016, Art. no. 10980.
[20] A. Lal, C. Shan, and P. Xi, "Structured illumination microscopy image reconstruction algorithm," *IEEE J. Sel. Topics Quantum Electron.*, vol. 22, no. 4, pp. 50–63, Jul./Aug. 2016.
[21] P. Křížek, T. Lukeš, M. Ovesný, K. Fliegel, and G. M. Hagen, "SIMToolbox: A MATLAB toolbox for structured illumination fluorescence microscopy," *Bioinformatics*, vol. 32, no. 2, pp. 318–320, Jan. 2016.
[22] T. Lukeš *et al.*, "Three-dimensional super-resolution structured illumination microscopy with maximum a posteriori probability image estimation," *Opt. Express*, vol. 22, no. 24, pp. 29805–29817, Nov. 2014.
[23] M. A. A. Neil, R. Ju Kškaitis, and T. Wilson, "Method of obtaining optical sectioning by using structured light in a conventional microscope," *Opt. Lett*, vol. 22, no. 24, pp. 1905–1907, Dec. 1997.
[24] R. Heintzmann, "Structured illumination methods," in *Handbook of Biological Confocal Microscopy*, J. B. Pawley, Ed. New York, NY, USA: Springer, 2006, pp. 265–279.
[25] K. O'Holleran and M. Shaw, "Optimized approaches for optical sectioning and resolution enhancement in 2D structured illumination microscopy," *Biomed. Opt. Express*, vol. 5, no. 8, pp. 2580–2590, Jul. 2014.
[26] A. Jost, E. Tolstik, P. Feldmann, K. Wicker, A. Sentenac, and R. Heintzmann, "Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice blind-SIM reconstruction," *PLOS ONE*, vol. 10, no. 7, Jul. 2015, Art. no. e0132174.
[27] R. Heintzmann and T. Huser, "Super-resolution structured illumination microscopy," *Chem. Rev.*, vol. 117, no. 23, Nov. 2017, Art. no. 13890.
[28] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2011.
[29] A. Negash *et al.*, "Improving the axial and lateral resolution of three-dimensional fluorescence microscopy using random speckle illuminations," *JOSA A*, vol. 33, no. 6, pp. 1089–1094, May 2016.
[30] P. Křížek, I. Raška, and G. M. Hagen, "Flexible structured illumination microscope with a programmable illumination array," *Opt. Express*, vol. 20, no. 22, pp. 24 585–24 599, Oct. 2012.
[31] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenom.*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992.
[32] S. Lefkimmiatis, J. P. Ward, and M. Unser, "Hessian Schatten-norm regularization for linear inverse problems," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1873–1888, May 2013.
[33] M. Fortin and R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*, vol. 15. Amsterdam, The Netherlands: Elsevier, 2000.
[34] L. Condat, "A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, Dec. 2013.
[35] J.-J. Moreau, "Fonctions convexes duales et points proximaux dans un espace hilbertien," *Comptes Rendus de l'Académie des Sciences Série A Mathématiques*, vol. 255, pp. 2897–2899, 1962.
[36] S. Setzer, G. Steidl, and T. Teuber, "Deblurring Poissonian images by split Bregman techniques," *J. Vis. Commun. Image Representation*, vol. 21, no. 3, pp. 193–199, Apr. 2010.
[37] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. New York, NY, USA: Springer, May 2011, pp. 185–212.
[38] N. Zhao, Q. Wei, A. Basarab, N. Dobigeon, D. Kouamé, and J.-Y. Tourneret, "Fast single image super-resolution using a new analytical solution for ℓ_2-ℓ_2 problems," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3683–3697, Aug. 2016.
[39] C. Zalinescu, *Convex Analysis in General Vector Spaces*. Singapore: World scientific, 2002.
[40] M. S. Almeida and M. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3074–3086, Aug. 2013.
[41] M. Unser, E. Soubies, F. Soulez, M. McCann, and L. Donati, "GlobalBioIm: A unifying computational framework for solving inverse problems," in *Proc. OSA Imag. Appl. Opt. Congr. Comput. Opt. Sens. Imag.*, San Francisco CA, USA, Jun. 2017, paper CTu1B.
[42] M. R. Hestenes and E. Stiefel, *Methods of Conjugate Gradients for Solving Linear Systems*, vol. 49, no. 6. Washington, DC, USA: NBS, Dec. 1952.
[43] H. Kirshner, F. Aguet, D. Sage, and M. Unser, "3-D PSF fitting for fluorescence microscopy: Implementation and localization application," *J. Microscopy*, vol. 249, no. 1, pp. 13–25, Nov. 2013.
[44] W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.*, vol. 31, no. 2, pp. 221–239, 1989.
[45] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse Probl.*, vol. 24, no. 6, Nov. 2008, Art. no. 065014.
[46] S. Lefkimmiatis and M. Unser, "Poisson image reconstruction with Hessian Schatten-norm regularization," *IEEE Trans. Image Process.*, vol. 22, no. 11, pp. 4314–4327, Nov. 2013.
[47] G. Chierchia, N. Pustelnik, B. Pesquet-Popescu, and J.-C. Pesquet, "A nonlocal structure tensor-based approach for multicomponent image recovery problems," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5531–5544, Dec. 2014.

Emmanuel Soubies received the graduate degree from the Institut National des Sciences Appliquées de Toulouse, Toulouse, France, the M.Sc. degree in operational research from the University of Toulouse, Toulouse, France, in 2013, and the Ph.D. degree from the University of Nice Sophia Antipolis, Nice, France, in 2016. Since 2016, he has been a Postdoctoral Fellow with the Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His main research interests include inverse problems for imaging and sparse optimization.

Michael Unser (M'89–SM'94–F'99) is currently a Professor and the Director of Biomedical Imaging Group, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. His primary area of investigation is biomedical image processing. From 1985 to 1997, he was with the Biomedical Engineering and Instrumentation Program, National Institutes of Health, Bethesda, MD, USA, conducting research on bioimaging. He is the author with P. Tafti of the book *An Introduction to Sparse Stochastic Processes* (Cambridge University Press, 2014). He has authored/coauthored more than 300 journal papers on topics, which include sampling theory, wavelets, the use of splines for image processing, stochastic processes, and computational bioimaging.

Dr. Unser has served on the editorial board of most of the primary journals in his field including the IEEE TRANSACTIONS ON MEDICAL IMAGING (Associate Editor-in-Chief 2003–2005), IEEE TRANSACTIONS IMAGE PROCESSING, PROCEEDINGS OF IEEE, and *SIAM Journal of Imaging Sciences*. He is the Founding Chair of the technical committee on Bio Imaging and Signal Processing of the IEEE Signal Processing Society. He is a EURASIP Fellow (2009), and a member of the Swiss Academy of Engineering Sciences. He is the recipient of several international prizes including four IEEE-SPS Best Paper awards and two Technical Achievement awards from the IEEE (2008 SPS and EMBS 2010).