Efficient compressed sensing reconstruction for 3D fluorescence microscopy using OptoMechanical Modulation Tomography (OMMT) with a 1+2D regularization

FRANÇOIS MARELLI[1,2,*], MICHAEL LIEBLING[1,3]

[1] *Computational Bioimaging, Idiap Research Institute, Martigny, CH-1920, Switzerland*
[2] *Electrical Engineering, École Polytechnique Fédérale de Lausanne, Lausanne, CH-1015, Switzerland*
[3] *Electrical & Computer Engineering, University of California, Santa Barbara, CA 93106, USA*
[*] *francois.marelli@idiap.ch*

Abstract: OptoMechanical Modulation Tomography (OMMT) exploits compressed sensing to reconstruct high resolution microscopy volumes from fewer measurement images compared to exhaustive section sampling in conventional light sheet microscopy. Nevertheless, the volumetric reconstruction process is computationally expensive, making it impractically slow to use on large-size images, and prone to generating visual artefacts. Here, we propose a reconstruction approach that uses a 1+2D Total Variation (TV_{1+2}) regularization that does not generate such artefacts and is amenable to efficient implementation using parallel computing. We evaluate our method for accuracy and scaleability on simulated and experimental data. Using a high quality, but computationally expensive, Plug-and-Play (PnP) method that uses the BM4D denoiser as a benchmark, we observe that our approach offers an advantageous trade-off between speed and accuracy.

1. Introduction

Compressed sensing techniques exploit prior knowledge of the imaged object to reconstruct a high quality image [1, 2] from fewer measurements than traditional sampling would require. In 3D medical imaging, reconstruction algorithms exploit the specificities of the imaging procedure to yield improved performance and faster convergence speed [3].

In fluorescence microscopy, compressed sensing offers the prospect of high quality imaging while also reducing light exposure, which can lead to photobleaching or induce toxicity during in vivo imaging. Furthermore it could increase acquisition speed while taking fewer images [4]. Compressive sensing approaches have been proposed for structured illumination microscopy [5] and widefield imaging [6].

For Selective Plane Illumination Microscopy (SPIM) [7], compressive sensing has been implemented by collecting axial projections after illuminating the sample with a spatial modulation pattern, either by shaping the illumination with digital micromirror devices [8] or by applying optomechanical modulation [9]. The latter approach mechanically moves the sample along the axial (focal) direction while simultaneously modulating the overall light sheet intensity in time, all while keeping the camera shutter open. Measurements are repeated for multiple scans, each with a different modulation pattern. The volume is reconstructed computationally. We will use an approach similar to [9] as it is a particularly attractive way to achieve spatial light modulation. First, it can be implemented with minimal hardware modifications of simple light sheet microscopes such as, e.g., the OpenSPIM microscope [10]. Second, the spatial illumination pattern is created by modulating the intensity in time, which leaves great freedom for designing a suitable modulation pattern. We will refer to this technique as OptoMechanical Modulation Tomography (OMMT).

Despite the potential benefits of compressive sensing approaches, the computational cost of reconstructing 3D volumes is particularly high. Naive extensions of 2D reconstruction schemes

to 3D are impractical for the large images that can be acquired with modern microscopes. Furthermore, overly simplified 1D or 2D regularizing terms may not capture the objects' 3D nature, a prior that is key to ensure quality reconstructions.

Recently, in the field of 2D image reconstruction, the introduction of Plug-and-Play (PnP) methods has allowed replacing explicit regularization priors with denoising algorithms for image reconstruction [11, 12]. This led to using high-quality denoising algorithms such as Block-Matching and 3D filtering (BM3D) [13] to improve image reconstruction [14–16]. More recently, the use of pretrained deep learning based denoisers has further pushed the performance of PnP methods [17, 18].

Here, we propose an efficient reconstruction method for OMMT that exploits the 3D nature of the data, thereby reducing the reconstruction artefacts from which this techniques suffered [9]. To that end, we introduce a hybrid 1+2D regularization function that takes into account the anisotropy of the problem, while allowing efficient 3D computations thanks to parallel computing. We also implement a high quality, but computationally expensive, Plug-and-Play (PnP) regularization function to serve as a benchmark against which we compare our proposed approach. We evaluate our method for accuracy and scaleability on simulated and experimental data.

2. Method

We briefly recall the method by Woringer et al. [9], who combine temporal illumination modulation combined with focal plane scanning to generate patterned illumination along the depth axis and implement 3D compressed sensing. In this approach, the mechanized focus stage moves at a constant speed v during a single camera exposure time Δ_E, enabling the acquisition of projections by optical integration. Simultaneously, the intensity of the light sheet illumination varies according to a temporal modulation function, which creates spatial light patterns along the depth, as illustrated in Fig. 1 (a-b).
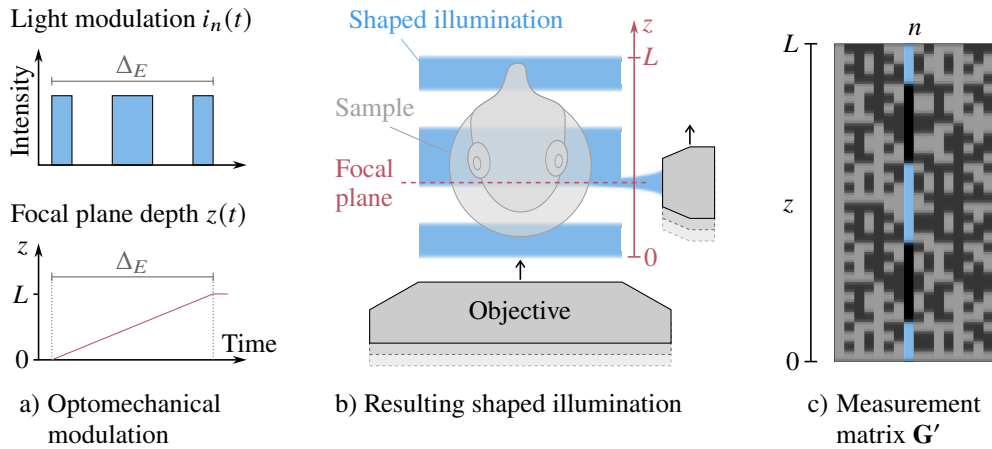

Fig. 1. Overview of OMMT imaging. (a) Optomechanical modulation combines focus sweeping synchronized with temporal light modulation. (b) The modulation creates a shaped illumination pattern in the depth axis. (c) This is repeated to acquire N projections with different modulation functions, which are stacked into the measurement matrix $\mathbf{G}'$.

2.1. Forward model

Let $f(x, y, z)$ be a three-dimensional object, with x, y, z the horizontal, vertical, and focus axes respectively. We write the Point Spread Function (PSF) of the system as $h(x, y, z)$, and the position in depth of the focal plane over time t as $z(t) = vt$. We acquire N projections $p_n(x, y)$ $(n = 1, \ldots, N)$ by modulating the light intensity over time using a function $i_n(t)$ defined over the interval $[0, \Delta_E]$, resulting in the following imaging model:

$$p_n(x, y) = \int_0^{\Delta_E} [f * h] (x, y, vt) \, i_n(t) \, dt \, + \, \epsilon_n(x, y), \tag{1}$$

where $*$ is a 3D convolution operator and $\epsilon_n(x, y)$ is an additive measurement noise.

The optomechanical modulation is particularly flexible, as the illumination functions $i_n(t)$ can be arbitrarily chosen, with the only constraint that all their values are positive. We choose to use an incomplete Hadamard basis for our experiments, as it can easily be generated by switching the illumination on and off over time, and has been used successfully in computational imaging and compressed sensing applications [4, 8, 19–23].

We start by constructing a Hadamard matrix $\mathbf{H}_M \in \mathbb{R}^{M \times M}$ of order M using Sylvester's construction, setting its negative values to 0 to conform to the positivity constraint. We then build an incomplete Hadamard matrix $\mathbf{H}_N \in \mathbb{R}^{N \times M}$ of size $N \times M$ by selecting the first row (full of ones) of $\mathbf{H}_M$ and stacking it with $N - 1$ other rows of $\mathbf{H}_M$ that are randomly sampled without replacement with uniform probability. By using $N \leq M$, we can reduce the number of measurements while keeping the high frequency sampling offered by a higher order Hadamard matrix. Finally, we define our illumination modulation functions as:

$$i_n(t) = H_N[n, j] \qquad \text{s.t.} \quad j - 1 \leq \frac{Mt}{\Delta_E} < j, \tag{2}$$

with $n \in \{1, \ldots, N\}$, and $j \in \{1, \ldots, M\}$.

By applying a variable substitution $z = vt$, we can express Eq. (1) as an integral over depth rather than time:

$$p_n(x, y) = \int_0^L [f * h] (x, y, z) \, g_n(z) \, dz \, + \, \epsilon_n(x, y), \tag{3}$$

where $L = v\Delta_E$ is the scanning depth, and $g_n(z) = \frac{1}{v} i_n(\frac{z}{v})$, defined over $[0, L]$, is the spatial illumination pattern created by the temporal modulation i_n combined with focus sweeping, illustrated in Fig. 1 (b).

Reconstructing the sample f from the imaging model in Eq. (3) is a three-dimensional problem whose computational complexity increases with the size of the imaged volume. It makes this model unusable to reconstruct volumetric data in reasonable time when working with the typical image sizes acquired by modern microscopes.

To address this issue, Woringer et al. [9] approach the problem as a stack of 2D images, which they split into smaller regions that they reconstruct independently in a parallel computing cluster. For these 2D reconstructions, they first perform empirical measurements of the 3D PSF that they crop to a single xz plane. Because this approach solves 2D reconstructions independently, it cannot exploit the 3D nature of the sample to improve its reconstruction consistency, leading to the adjacent sections showing visual incoherences because they are computed separately. Moreover, the computational cost is still high and the method requires access to a computational cluster to run in reasonable time, as reconstructing a $512 \times 512 \times 100$ volume on a single machine takes longer than a full day to complete.

In order to reduce the computational complexity of this problem, we simplify Eq. (3) by ignoring the lateral extent of the PSF and modelling it with a one-dimensional function $h'(z) = h(0, 0, z)$.

This is motivated by the fact that light sheet devices have a thin truncated PSF, which vanishes much faster in the lateral directions than in the axial one [24]. Using this simplified PSF, we write an approximated projection p'_n:

$$p'_n(x, y) = \int_0^L \left(\int_{-\infty}^{\infty} f(x, y, u) \, h'(u - z) \, \mathrm{d}u \right) g_n(z) \, \mathrm{d}z \; + \; \epsilon_n(x, y) \tag{4}$$

$$p'_n(x, y) = \int_{-\infty}^{\infty} f(x, y, u) \left(\int_0^L h'(u - z) \, g_n(z) \, \mathrm{d}z \right) \mathrm{d}u \; + \; \epsilon_n(x, y) \tag{5}$$

$$p'_n(x, y) = \int_{-\infty}^{\infty} f(x, y, z) \, g'_n(z) \, \mathrm{d}z \; + \; \epsilon_n(x, y), \tag{6}$$

where $g'_n(z) = g_n(z) \circledast h'(-z)$ is the modulation function blurred by the simplified PSF by the 1D convolution $\circledast$. At each location (x, y) in a projection, the imaging model reduces to a one-dimensional expression that is much simpler and faster to compute than Eq. (3).

When discretized, we can write Eq. (6) as the following linear expression:

$$\mathbf{P}[j, k, :] = \mathbf{G}' \mathbf{F}[j, k, :] \; + \; \mathbf{E}[j, k, :], \tag{7}$$

where $\mathbf{F} \in \mathbb{R}^{W \times H \times D}$ is a discrete sampling of f with width W, height H, and depth D, the latter corresponding to the focus axis z. $\mathbf{G}' \in \mathbb{R}^{N \times D}$ is the matrix operator corresponding to the multiplication with the modulation function g'_n and the axial integration, and constitutes the compressed sensing measurement matrix. The values in $\mathbf{P} \in \mathbb{R}^{W \times H \times N}$ correspond to our discrete measurements, and $\mathbf{E} \in \mathbb{R}^{W \times H \times N}$ is the measurement noise. We use : as a shorthand notation to describe all elements in the corresponding dimension of the array [25]. For each pair of coordinates $j = 1, \ldots, W$ and $k = 1, \ldots, H$, the model is a simple matrix-vector multiplication, which can be computed very efficiently. It can therefore realistically be applied on big experimental data to compute volumetric reconstructions in reasonable time.

To derive the measurement matrix, we estimate $h'(z)$ based on a simulated PSF. Since the 1D model is an approximation, we do not require accurate experimental measurements of the 3D PSF which can be tedious to obtain. We compute the rows of $\mathbf{G}'$ by discretizing the blurred light patterns g'_n according to the axial resolution of $\mathbf{F}$, as illustrated in Fig. 1 (c).

2.2. Inverse problem

Reconstructing $\mathbf{F}$ from the measurements requires solving the linear inverse problem corresponding to Eq. (7). Since the reconstructed axial resolution is higher than the number of acquired projections $D > N$, this inverse problem is ill-posed and requires additional constraints to be solved uniquely. A common approach in compressed sensing is to formulate the reconstruction as a minimization problem [4, 9, 26]:

$$\hat{\mathbf{F}} = \underset{\mathbf{F} \in \mathbb{R}^{W \times H \times D}}{\arg\min} \left(\sum_{j=1}^{W} \sum_{k=1}^{H} C \left(\mathbf{p}[j, k, :], \mathbf{GF}[j, k, :] \right) + \lambda \, \mathcal{R} \left(\mathbf{F} \right) \right), \tag{8}$$

where $C : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}_+$ is a loss functional that ensures consistency of the solution with the measurements, $\mathcal{R} : \mathbb{R}^{W \times H \times D} \to \mathbb{R}_+$ is a regularization term that imposes prior constraints and the parameter $\lambda \in \mathbb{R}_+$ allows adjusting the strength of the regularization. More specifically, we use a squared ℓ_2 loss for data consistency:

$$C \left(\mathbf{p}[j, k, :], \mathbf{G}\hat{\mathbf{F}}[j, k, :] \right) = ||\mathbf{p}[j, k, :] - \mathbf{G}\hat{\mathbf{F}}[j, k, :]||_2^2. \tag{9}$$

We solve Eq. (8) using an iterative optimization algorithm, the alternating direction method of multipliers (ADMM [27–29]), which alternates between minimizing C and $\mathcal{R}$. Choosing a good

regularization $\mathcal{R}$ is crucial as it impacts the quality of the reconstruction but also the computational complexity of the minimization. We consider three different regularization candidates that we compare below.

2.2.1. ℓ_1 sparsity constraint

In the original method, Woringer et al. [9] use a spatial sparsity constraint expressed as:

$$\mathcal{R}(\mathbf{F}) = ||\mathbf{F}||_1. \tag{10}$$

This regularization favours reconstructions that are sparse [30], and is common in compressed sensing applications [2,4]. It has a low computational cost, but it does not enforce any continuity between adjacent values in the volume and therefore cannot take advantage of the additional information provided by neighbouring voxels in the reconstruction.

2.2.2. TV_{1+2} sparsity constraint

Another common regularization is Total Variation (TV), which aims at generating smoother solutions while preserving sharp edges [31,32]. We introduce a 1+2D formulation that penalizes TV separately along the z axis, which we call TV_{1+2}:

$$\mathcal{R}(\mathbf{F}) = \rho\, TV_{1D}(\mathbf{F}) \;+\; TV_{2D}(\mathbf{F}) \tag{11}$$

where TV_{1D} is the 1D TV along the depth axis and TV_{2D} is the isotropic 2D TV on xy sections:

$$TV_{1D}(\mathbf{F}) = \sum_{i=1}^{W} \sum_{j=1}^{H} \sum_{k=1}^{D} |\mathbf{F}[i, j, k+1] - \mathbf{F}[i, j, k]| \tag{12}$$

$$TV_{2D}(\mathbf{F}) = \sum_{i=1}^{W} \sum_{j=1}^{H} \sum_{k=1}^{D} \sqrt{|\mathbf{F}[i+1, j, k] - \mathbf{F}[i, j, k]|^2 \;+\; |\mathbf{F}[i, j+1, k] - \mathbf{F}[i, j, k]|^2}. \tag{13}$$

The factor ρ allows us to tune the regularization strength independently in the compression axis. We use that additional degree of freedom to factor in the anisotropic properties of the reconstructed volume, which stem from the compressed imaging and the shape of the PSF.

This regularization enforces visual coherence across space in the solution, thus benefitting from the additional information contained in neighbouring voxels. Although this comes at the cost of a higher computational complexity, we can massively speed it up using parallel computing. Indeed, all the minimization steps are simple operations that can be computed independently: both the data consistency term and TV_{1D} are 1D operations, and TV_{2D} is a simple 2D computation, as illustrated on Fig. 2. Moreover, the ADMM algorithm allows computing these terms simultaneously at each iteration, making the minimization fall into the 'embarrassingly parallel' category. We can therefore expect fast computations on machines supporting multithreaded execution, or with Graphical Processing Unit (GPU) acceleration.

2.2.3. PnP BM4D constraint

ADMM does not compute the gradient of $\mathcal{R}$ but can instead rely on its proximal function to carry out the minimization. This makes it suitable for implementing PnP methods [11] by using a denoiser as the proximal function to an implicit regularization prior. This is not possible with the SPIRAL solver [33] used by Woringer et al. [9], as it requires an explicit regularization function.

Although deep learning denoisers provide state-of-the-art performance for PnP methods in 2D, publicly available pretrained universal 3D denoisers based on deep learning are lacking, and the massive computational cost of building a volumetric dataset and training our own model prevents us from using one in this work. Instead, we use the BM4D denoiser [34], an extension
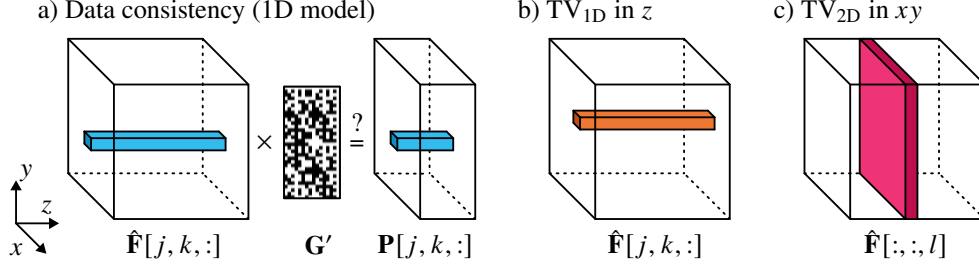
a) Data consistency (1D model) b) TV$_{1D}$ in z c) TV$_{2D}$ in xy

$\hat{\mathbf{F}}[j,k,:]$ $\mathbf{G}'$ $\mathbf{P}[j,k,:]$ $\hat{\mathbf{F}}[j,k,:]$ $\hat{\mathbf{F}}[:,:,l]$

Fig. 2. Regularization with TV$_{1+2}$ implements 3D prior constraints while keeping efficient computations using a parallel implementation. (a) Data consistency relies on the 1D forward model. (b) TV$_{1D}$ regularization along the compression axis is tuned separately to account for anisotropy in the volume. (c) Isotropic TV$_{2D}$ on xy sections allows information sharing between the many 1D problems.

of the well-known BM3D algorithm to volumetric data, as a PnP prior for our problem. By enforcing a 3D spatial coherence in the reconstruction, this prior also benefits from information sharing between adjacent voxels. The regularization strength λ in Eq. (8) is replaced by a new parameter σ specific to BM4D, which is the strength of the denoising applied at each iteration of the ADMM optimization.

2.3. Reliable choice of the hyperparameters

The performance of the reconstruction algorithm is impacted by the selected hyperparameter values for the optimization problem (depending on the regularization, these can be λ, ρ or σ). We have found the minimization of the data consistency cost C to be a good proxy for the performance of the reconstruction. Although the minimum cost does not always correspond to the best reconstruction accuracy, we have measured that it consistently gives a very similar performance to the actual best on our experiments.

Choosing hyperparameters that minimize the final data consistency cost is therefore a reliable way to guarantee good performance using a reproducible method. This does not require knowing the ground truth, and can be applied regardless of the regularization constraints used. That is particularly helpful in the case of the BM4D PnP prior, as the lack of an explicit regularization function prevents from using common hyperparameter selection methods such as the L-curve [35].

3. Experiments and results

In order to take advantage of the highly parallel nature of our algorithm, we implemented the reconstruction using the scientific imaging library SCICO [36]. This framework provides an efficient implementation of ADMM based on JAX [37, 38], a high-performance Python library that uses just-in-time compilation to provide parallel computing and hardware acceleration (such as GPU). We used a Python implementation of an updated version of BM4D provided by its authors [39, 40].

We first characterize our method on simulated data, then validate it on experimental images.

3.1. Characterization on simulated data

We used our previously introduced simulation framework [41] to generate synthetic volumes of size $128 \times 128 \times 128$ with physical properties matching those of data acquired with the OpenSPIM platform [10]. We simulated the imaging 3D PSF using an implementation of the Born & Wolf model [42, 43]. After the imaging simulation, we corrupted the data with shot noise generated using a Poisson distribution. We set the strength of the noise by rescaling the data to a chosen

maximum photon count before sampling the Poisson distribution, with higher photon counts corresponding to weaker overall noise. Finally, we quantized the data to a 12 bit representation to emulate acquisition with a digital camera.

In practice, the PSF simulation model used for reconstruction in our method does not exactly match the actual imaging PSF. In order to account for this discrepancy in our experiments, we used a different simplified Gaussian beam model [44] to compute the 1D PSF when building the measurement matrix.

To select the hyperparameters, we ran loose grid searches (in logarithmic space for λ and ρ, and in linear space for σ) and selected the solution with minimal consistency cost as described in Section 2.3. We use $N = 16$ and $M = 32$ to build our measurement matrices. We measured that this ratio of 2 for undersampling of the Hadamard matrix yields the best performance for a fixed number of projections.

In order to evaluate how the use of different regularizations can affect the artefacts present in the reconstruction, we simulated an object with a moon-shaped cross-section and imaged it with low noise (10^4 photons). The 3D reconstructions obtained using our method with each regularization prior are shown in Fig. 3, alongside the central xy and xz sections.
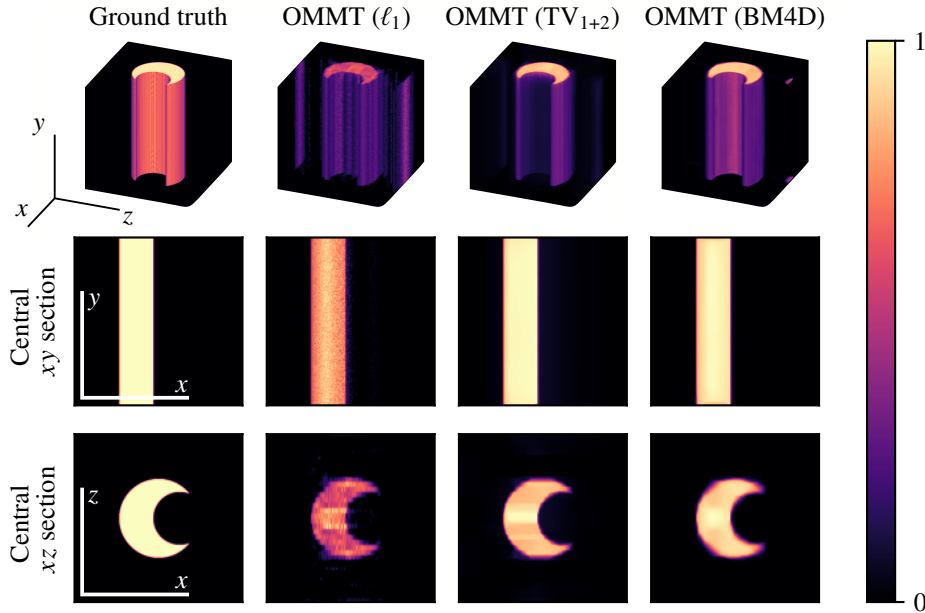

Fig. 3. Reconstruction of a simulated sample using the different regularization priors. The ℓ_1 result contains more noise and has fuzzy edges in the xz section, while TV_{1+2} and BM4D give a much smoother result with better defined edges. 3D views use attenuated mean intensity projection. Scale bar axes are 50 µm.

Reconstructions performed with the ℓ_1 regularization contain visual noise. While the xy section appears good visually, the quality of the xz section is visibly worse, with the strong noise leading to a very uneven intensity in the object and fuzzy edges. Similar artefacts also appeared in the results obtained by Woringer et al. [9]. It is also visible on the 3D views that the ℓ_1 reconstruction contains non-zero values in empty areas around the object.

Both our proposed TV_{1+2} and BM4D priors give a much cleaner result and better defined edges, with an even intensity in the xy section and only small variations in the xz section. They also contain only few non-zero values outside the object. This reduced number of artefacts shows the benefit of using a 3D-aware regularization prior that better enforces spatial continuity. The

information sharing between neighbouring voxels leads to a visually more coherent volumetric reconstruction and reduced noise.

3.1.1. Comparison with plane-by-plane acquisition for sparse object detection

When introducing compressed sensing based on optomechanical modulation, Woringer et al. [9] showed that one of its main benefits is the reduction of phototoxicity during acquisition. They demonstrated how the method outperforms plane-by-plane acquisition when matching the light dose by reducing the exposure of the SPIM acquisition, which increases the level of noise in the stack. Here we want to measure if these advantages are maintained when comparing to a plane-by-plane acquisition with fewer imaged sections, which gives a comparable level of phototoxicity without adding noise. To that end, we consider the scenario where we image a number of uniformly spaced sections equal to the number of projections $N = 16$ acquired with OMMT. By using the same average light intensity per frame and exposure time for the plane-by-plane as for OMMT, the sample receives the same light dose and the noise level in images is comparable. We then resample the SPIM stack along the z axis with cubic interpolation to achieve the same final spatial sampling as the OMMT reconstructions.

We simulate a sample composed of thin vertical cylinders uniformly spaced along the diagonal of the volume. We choose spacing along the depth axis to not be a multiple of the distance between acquired sections with the plane-by-plane method, and position the central cylinder to be exactly aligned with one of the imaged planes. That way, the cylinders are gradually less aligned with the plane-by-plane sampling as they get further away from the centre.

Figure 4 shows the SPIM axial sampling and the obtained volumes with both methods. While the central cylinders are correctly visible in the plane-by-plane image, misaligned objects highlighted in the figure are barely visible when they fall exactly between two sampled sections. All cylinders are clearly visible in the OMMT reconstructions regardless of their position, thanks to the fact that all regions of the sample are illuminated at least once during acquisition due to the choice of illumination functions. This makes OMMT more reliable than plane-by-plane acquisition to detect spatially sparse objects, at equal phototoxicity and noise levels.
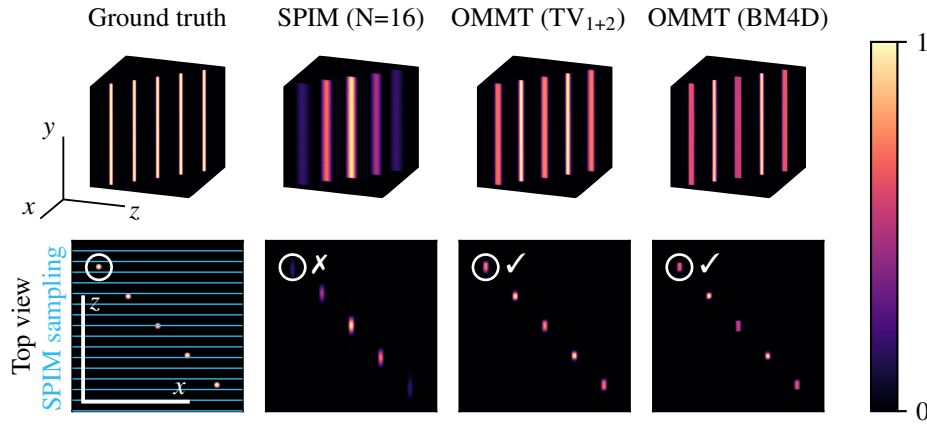

Fig. 4. Comparison of undersampled plane-by-plane SPIM and OMMT with equal phototoxicity. SPIM fails to detect objects that are misaligned with its axial sampling, while OMMT reconstructions contain all the elements. All views use mean intensity projection, and axes are 50 µm scale bars.

3.1.2. Quantitative performance comparison and impact of noise

In order to quantify the performance of our algorithm, we simulated a sample containing multiple thin objects crossing the volume at various angles and with different cross-sections and intensities. To measure the impact of noise on the reconstruction, we generated images for a virtually noise-free scenario with a maximum photon count per pixel of 10^4, and a noisy scenario with a maximum photon count of 400.

Our comparison criterion is the Peak Signal-to-Noise Ratio (PSNR), defined as:

$$\text{PSNR} = 10 \log_{10} \left(\max(\mathbf{F})^2 / \text{MSE} \right), \tag{14}$$

where MSE is the Mean Squared Error between our reconstructions and the original sample $\mathbf{F}$. We ran the reconstructions on a multithreaded machine (AMD EPYC 7742 2.25 GHz, limited to 8 cores), and measured the total running time of the ADMM minimization to compare the computational complexity of the different regularization priors.

We repeated each experiment 5 times with different random seeds for the noise and the choice of the illumination functions, and averaged the obtained metrics. The measured deviation between repetitions was negligible, guaranteeing good confidence in the presented numbers.

Table 1 summarizes the results obtained with the different regularization priors for both clean and noisy scenarios. As a reference, we also simulated a fully sampled SPIM stack with 128 sections and an undersampled SPIM stack with 16 sections corresponding to equivalent phototoxicity, as described in Section 3.1.1.

	Max photons: 10^4		Max photons: 400	
	PSNR	Time	PSNR	Time
SPIM (N=128)	36.08 dB	/	34.63 dB	/
SPIM (N=16)	28.64 dB	/	28.49 dB	/
OMMT (ℓ_1)	32.77 dB	**4 s**	32.22 dB	**4 s**
OMMT (TV$_{1+2}$)	34.22 dB	26 s	34.05 dB	21 s
OMMT (BM4D)	**35.26 dB**	72 min	**35.21 dB**	68 min

Table 1. Performance characterization on simulated sample for different regularization priors and with varying noise strengths, with full and undersampled SPIM stacks as reference. BM4D performs best, but is significantly slower than using other priors. OMMT reconstruction consistently outperforms undersampled SPIM, and beats the full SPIM stack in the noisy scenario with BM4D prior.

The results clearly show that the PnP BM4D prior consistently outperforms all other regularizations in both scenarios. However, it is very slow to compute, with the ℓ_1 prior being 1000× faster to reach convergence. Our proposed TV$_{1+2}$ regularization offers a trade-off between speed and accuracy, being only 5× slower than ℓ_1, while improving over its performance by 1.5 dB on average.

OMMT outperforms the undersampled SPIM regardless of the photon count. Adding noise does not significantly impact its reconstruction performance, and with the BM4D prior OMMT beats the full SPIM stack in the noisy scenario. This confirms the observations of Woringer et al. [9] that compressed sensing methods match the performance of classically sampling noisy images as the noise level increases.
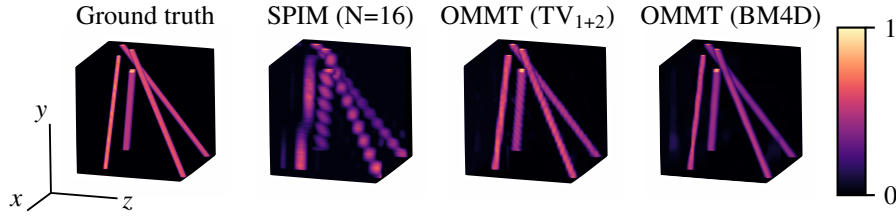
Fig. 5. Volumetric reconstructions on simulated objects in a noisy scenario compared to an undersampled SPIM with equal phototoxicity. OMMT reconstructions are smoother, with a minor improvement when using BM4D over TV_{1+2}. Attenuated mean intensity projection is shown, and axes are $50\,\mu m$ scale bars.

We show the reconstructions obtained with TV_{1+2} and BM4D priors for the noisy scenario in Fig. 5 alongside the undersampled SPIM volume. OMMT reconstructions are visibly smoother in regions where the SPIM stack appears to have gaps in the objects. Using the BM4D PnP prior yields a smoother reconstruction but the visual improvement over TV_{1+2} is not significant.

3.2.   Analysis of the computational cost

To efficiently explore many scenarios and methods efficiently, we performed our simulations on undersized data. In order to apply our method in practice, it must be able to compute a solution on real-scale images with a size of at least 512×512 in reasonable time. To assess the scaleability of the different regularization priors, we measured the time to run a single iteration of the ADMM optimization loop for increasing data sizes. This gives us a computational efficiency measurement that is not influenced by the number of iterations to converge. The obtained single iteration run time is sufficient to estimate the relative total run time of the algorithm with various regularizations, as we observed that the different priors lead to similar total number of iterations to convergence. Any arbitrary data can be used for this speed assessment as we run a fixed number of iterations that do not need to converge. We used random uniform input data and averaged the time over 5 runs of 2 iterations each, after running 2 warm-up iterations.

We measured the performance of our implementation on a multithreaded machine (AMD EPYC 7742 2.25 GHz, limited to 8 cores). In the case of ℓ_1 and TV_{1+2}, we also measured the performance on a machine with GPU (Nvidia GeForce RTX 3090). We did not measure it for BM4D as its implementation does not support GPU acceleration.

Figure 6 shows our measurements for output volumes with equal depth, width, and height. We used a number of projections N equal to half the depth of the reconstructed volume. As observed before, the BM4D PnP prior is much slower than the other regularizers. It becomes unusable for volumes bigger than $256 \times 256 \times 256$ as its computation time become unreasonably high. Even if ℓ_1 is consistently the fastest, TV_{1+2} is only slightly slower for bigger images. Their computation time grows less quickly with data size, making them realistically useable to reconstruct big volumes. Moreover, using GPU acceleration brings a massive speed-up of over 30× to the algorithm, which shows that our method is suitable for reconstructing real-scale experimental data very efficiently without requiring access to a computational cluster.

3.3.   Validation on experimental data

To confirm the real-case applicability of our method, we mounted two fluorescent textile fibres with a diameter of $25\,\mu m$ in $2\,\%$ low melting agarose solution inside a fluorinated ethylene propylene tube. We imaged the sample using an implementation of the OpenSPIM platform [10] with a 561 nm Vortran Stradus laser, a UMPLFLN 20XW semi-apochromat water dipping
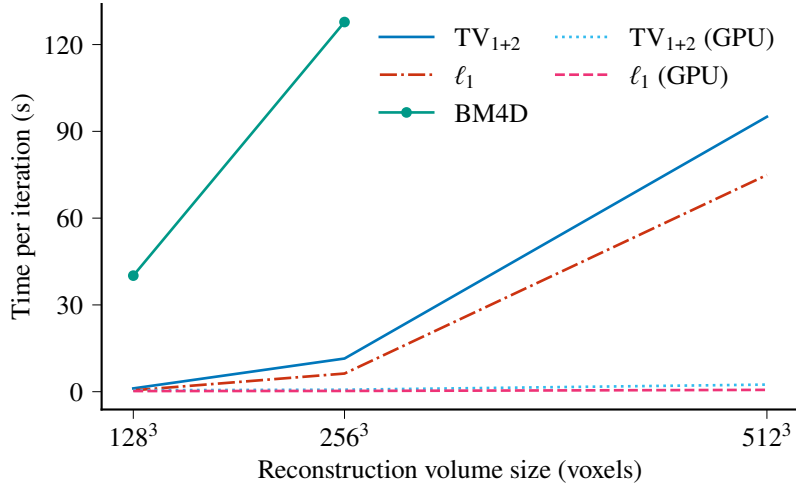
Fig. 6. Average time per iteration of our ADMM optimization loop for different regularization priors. BM4D is not realistically useable to reconstruct big images, but both TV_{1+2} and ℓ_1 are fast enough to be applied to real size experimental data. This is particularly true with the 30× speed-up when running with GPU acceleration.

objective lens and an Andor Zyla 4.2 sCMOS camera mounted on a U-TV0.5XC-3 adapter. We used a custom modulation controller based on Arduino that we introduced in previous works [45] to control the intensity of the laser directly at emission, which is a cost-efficient alternative to the more expensive acousto-optic tunable filter used in the original method [9].

We acquired $N = 16$ projections with $M = 32$, sweeping the focal plane over a total depth of $600\,\mu m$ and reconstructed a volume of size $640 \times 640 \times 128$ using the TV_{1+2} prior. As a reference, we acquired a full SPIM stack with 128 sections. We also imaged an undersampled SPIM containing 16 sections, which yields equivalent phototoxicity to OMMT.

Figure 7 shows the obtained volumes. For an equal light exposure, OMMT does not feature the jagged artefacts visible in the undersampled SPIM and does not contain gaps in the reconstructed fibres. This confirms our results obtained on simulations. The solution matches the geometry of the reference volume even with a very low number of acquired images.
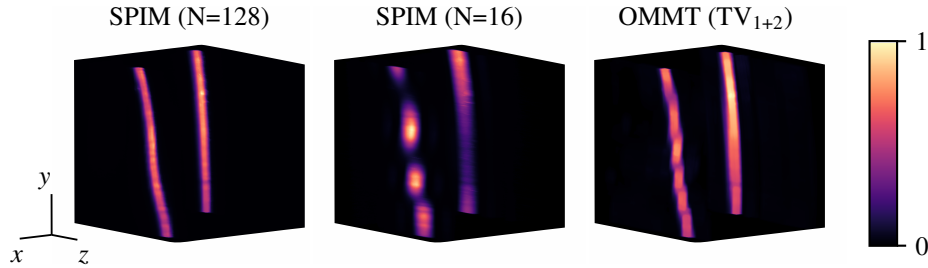

Fig. 7. OMMT reconstruction of fluorescent textile fibres and comparison to undersampled SPIM with equal light dose. The reconstructed image with TV_{1+2} prior is smoother, and looks similar to the reference SPIM volume. Views use mean intensity projection, and axes are $100\,\mu m$ scale bars.

4. Discussion

Our reconstruction method is able to use regularization priors that enforce a 3D coherence in the solution while being efficient to compute given its parallel implementation. Since our proposed TV_{1+2} formulation is specifically tailored to the geometry of the problem, we have been able to improve over the existing method with ℓ_1 prior and reduce reconstruction artefacts while having a minor additional computational cost. This better performance likely stems from the fact that the method takes into account the volumetric nature of the problem where neighbouring voxels are considered along all dimension to improve the overall accuracy. Using the same acquisition procedure as in [9]), our updated method retains the advantages of the original technique over plane-by-plane imaging (lower phototoxicity, faster acquisition) while the efficient implementation of our reconstruction method eliminates the need for a computational cluster to obtain solutions in reasonable time.

Using a universal 3D denoiser as PnP prior for reconstruction outperforms the TV_{1+2} regularization but is not realistically applicable in practice due to its slower computation (up to a factor 1000) on bigger data. However, this result seems promising when considering the use of a 3D deep learning-based denoiser instead of BM4D, as neural networks are usually running efficiently with GPU acceleration. This would alleviate the speed problem by shifting part of the computational load from reconstruction time to the training of the model. Moreover, recent works in image reconstruction are trying to reduce the memory and computational requirements of PnP methods by using subsets of the data at each iteration [16, 46]. These methods could further improve computation speed if they can be adapted to our 3D problem.

5. Conclusion

We have implemented an ADMM algorithm for OMMT reconstruction that uses parallel computing and shown that it can efficiently solve the inverse problem on a single machine in reasonable time. Our efficient 1+2D formulation of the TV regularization prior allows introducing volumetric smoothness constraints while taking into account the specific geometry of the problem while keeping the computational overhead in check.

We used simulated data to characterize the performance of our method, showing that using fully 3D-aware priors reduces visual artefacts over ℓ_1 and that a PnP method with BM4D denoising gives the best accuracy. We further established the advantages of the OMMT compressed sensing acquisition scheme over plane-by-plane imaging by demonstrating its superior ability to detect sparse objects at comparable phototoxicity.

We have measured the scaleability of our implementation on big volumes, proving its applicability to real-scale images. We validated these results by successfully applying our method to an experimental dataset.

Although we showed that using a PnP prior for reconstruction is too computationally demanding with the BM4D denoiser, we expect that efficient implementations of 3D deep learning denoising algorithms could further improve the reconstruction accuracy with a lower computational cost.

Disclosures.
The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are publicly available in the OMMT-Fibre dataset [47].

We provide the implementation of the simulation and reconstruction framework in the Idiap CBI Toolbox library: https://github.com/idiap/cbi_toolbox.

References

1. D. Donoho, "Compressed sensing," IEEE Trans. on Inf. Theory **52**, 1289–1306 (2006).
2. E. J. Candes and M. B. Wakin, "An introduction to compressive sampling," IEEE Signal Process. Mag. **25**, 21–30 (2008).
3. L. B. Montefusco, D. Lazzaro, S. Papi, and C. Guerrini, "A fast compressed sensing approach to 3D MR image reconstruction," IEEE Trans. on Med. Imaging **30**, 1064–1075 (2011).
4. G. Calisesi, M. Ghezzi, D. Ancora, C. D'Andrea, G. Valentini, A. Farina, and A. Bassi, "Compressed sensing in fluorescence microscopy," Prog. Biophys. Mol. Biol. **168**, 66–80 (2022).
5. W. Meiniel, P. Spinicelli, E. D. Angelini, A. Fragola, V. Loriette, F. Orieux, E. Sepulveda, and J.-C. Olivo-Marin, "Reducing data acquisition for fast structured illumination microscopy using compressed sensing," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017),* (2017), pp. 32–35.
6. Q. Guo, H. Chen, Y. Wang, Y. Guo, P. Liu, X. Zhu, Z. Cheng, Z. Yu, S. Yang, M. Chen, and S. Xie, "High-speed compressive microscopy of flowing cells using sinusoidal illumination patterns," IEEE Photonics J. **9**, 1–11 (2017).
7. J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, and E. H. Stelzer, "Optical sectioning deep inside live embryos by selective plane illumination microscopy," Science **305**, 1007–1009 (2004).
8. G. Calisesi, M. Castriotta, A. Candeo, A. Pistocchi, C. D'Andrea, G. Valentini, A. Farina, and A. Bassi, "Spatially modulated illumination allows for light sheet fluorescence microscopy with an incoherent source and compressive sensing," Biomed. optics express **10**, 5776–5788 (2019).
9. M. Woringer, X. Darzacq, C. Zimmer, and M. Mir, "Faster and less phototoxic 3D fluorescence microscopy using a versatile compressed sensing scheme," Opt. Express **25**, 13668–13683 (2017).
10. P. G. Pitrone, J. Schindelin, L. Stuyvenberg, S. Preibisch, M. Weber, K. W. Eliceiri, J. Huisken, and P. Tomancak, "OpenSPIM: an open-access light-sheet microscopy platform," Nat. methods **10**, 598–599 (2013).
11. 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,* (2013), pp. 945–948.
12. U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications," IEEE Signal Process. Mag. **40**, 85–97 (2023).
13. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," in *Image Processing: Algorithms and Systems VI,* vol. 6812 J. T. Astola, K. O. Egiazarian, and E. R. Dougherty, eds., International Society for Optics and Photonics (SPIE, 2008), p. 681207.
14. S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-Play ADMM for image restoration: Fixed-point convergence and applications," IEEE Trans. on Comput. Imaging **3**, 84–98 (2017).
15. C. J. Pellizzari, R. Trahan, H. Zhou, S. Williams, S. E. Williams, B. Nemati, M. Shao, and C. A. Bouman, "Optically coherent image formation and denoising using a plug and play inversion framework," Appl. Opt. **56**, 4735–4744 (2017).
16. Y. Sun, B. Wohlberg, and U. S. Kamilov, "An online Plug-and-Play algorithm for regularized image reconstruction," IEEE Trans. on Comput. Imaging **5**, 395–408 (2019).
17. E. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin, "Plug-and-Play methods provably converge with properly trained denoisers," in *Proceedings of the 36th International Conference on Machine Learning,* vol. 97 of *Proceedings of Machine Learning Research* K. Chaudhuri and R. Salakhutdinov, eds. (PMLR, 2019), pp. 5546–5557.
18. K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-Play image restoration with deep denoiser prior," IEEE Trans. on Pattern Anal. Mach. Intell. **44**, 6360–6376 (2022).
19. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," IEEE Signal Process. Mag. **25**, 83–91 (2008).
20. F. Soldevila, P. Clemente, E. Tajahuerce, N. Uribe-Patarroyo, P. Andrés, and J. Lancis, "Computational imaging with a balanced detector," Sci. Reports **6**, 29181 (2016).
21. E. Salvador-Balaguer, P. Latorre-Carmona, C. Chabert, F. Pla, J. Lancis, and E. Tajahuerce, "Low-cost single-pixel 3D imaging by using an LED array," Opt. express **26**, 15623–15631 (2018).
22. A. Zunino, F. Garzella, A. Trianni, P. Saggau, P. Bianchini, A. Diaspro, and M. Duocastella, "Multiplane encoded light-sheet microscopy for enhanced 3D imaging," ACS Photonics **8**, 3385–3393 (2021).
23. S. Crombez, P. Leclerc, C. Ray, and N. Ducros, "Computational hyperspectral light-sheet microscopy," Opt. Express **30**, 4856–4866 (2022).
24. O. E. Olarte, J. Andilla, E. J. Gualda, and P. Loza-Alvarez, "Light-sheet microscopy: a tutorial," Adv. Opt. Photon. **10**, 111–179 (2018).
25. G. H. Golub and C. F. Van Loan, *Matrix computations* (JHU press, 2013).
26. 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 Trans. on Comput. Imaging **2**, 59–70 (2016).
27. R. Glowinski and A. Marroco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de dirichlet non linéaires," Revue française d'automatique, informatique, recherche opérationnelle. Anal. numérique **9**, 41–76 (1975).
28. D. Gabay and B. Mercier, "A dual algorithm for the solution of nonlinear variational problems via finite element approximation," Comput. & Math. with Appl. **2**, 17–40 (1976).

29. 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. **3**, 1–122 (2011).
30. M. Elad and A. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," IEEE Trans. on Inf. Theory **48**, 2558–2567 (2002).
31. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," Phys. D: Nonlinear Phenom. **60**, 259–268 (1992).
32. A. Chambolle, "An algorithm for total variation minimization and applications," J. Math. imaging vision **20**, 89–97 (2004).
33. Z. T. Harmany, R. F. Marcia, and R. M. Willett, "Sparse poisson intensity reconstruction algorithms," in *2009 IEEE/SP 15th Workshop on Statistical Signal Processing,* (2009), pp. 634–637.
34. M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," IEEE Trans. on Image Process. **22**, 119–133 (2013).
35. P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," SIAM Rev. **34**, 561–580 (1992).
36. T. Balke, F. Davis, C. Garcia-Cardona, S. Majee, M. McCann, L. Pfister, and B. Wohlberg, "Scientific computational imaging code (SCICO)," J. Open Source Softw. **7**, 4722 (2022).
37. J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, "JAX: composable transformations of Python+NumPy programs," (2018). http://github.com/google/jax, Version 0.4.2.
38. R. Frostig, M. J. Johnson, and C. Leary, "Compiling machine learning programs via high-level tracing," Syst. for Mach. Learn. **4** (2018).
39. Y. Mäkinen, L. Azzari, and A. Foi, "Collaborative filtering of correlated noise: Exact transform-domain variance for improved shrinkage and patch matching," IEEE Trans. on Image Process. **29**, 8339–8354 (2020).
40. Y. Mäkinen, S. Marchesini, and A. Foi, "Ring artifact and Poisson noise attenuation via volumetric multiscale nonlocal collaborative filtering of spatially correlated noise," J. Synchrotron Radiat. **29**, 829–842 (2022).
41. F. Marelli and M. Liebling, "Optics versus computation: Influence of illumination and reconstruction model accuracy in focal-plane-scanning optical projection tomography," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI),* (2021), pp. 567–570.
42. M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light* (Elsevier, 2013).
43. H. Kirshner, F. Aguet, D. Sage, and M. Unser, "3-D PSF fitting for fluorescence microscopy: implementation and localization application," J. Microsc. **249**, 13–25 (2013).
44. A. K. Trull, J. van der Horst, W. J. Palenstijn, L. J. van Vliet, T. van Leeuwen, and J. Kalkman, "Point spread function based image reconstruction in optical projection tomography," Phys. Med. & Biol. **62**, 7784 (2017).
45. F. Marelli, A. Ernst, N. Mercader, and M. Liebling, "PAAQ: Paired Alternating AcQuisitions for virtual high frame rate multichannel cardiac fluorescence microscopy," (2023). Submitted for publication.
46. Y. Sun, Z. Wu, X. Xu, B. Wohlberg, and U. S. Kamilov, "Scalable Plug-and-Play ADMM with convergence guarantees," IEEE Trans. on Comput. Imaging **7**, 849–863 (2021).
47. F. Marelli and M. Liebling, "OMMT-Fibre: textile fibres acquired using OptoMechanical Modulation Tomography (OMMT) [Data set]," (2023). https://doi.org/10.34777/6m4a-0c80.