

# NON-SMOOTH CONVEX OPTIMIZATION FOR AN EFFICIENT RECONSTRUCTION IN STRUCTURED ILLUMINATION MICROSCOPY

Jérôme Boulanger<sup>1\*</sup>, Nelly Pustelnik<sup>2</sup>, Laurent Condat<sup>3</sup>

<sup>1</sup>UMR144 CNRS/Institut Curie, 12 rue Lhomond, F-75005 Paris, France.

<sup>2</sup>Physics Laboratory, CNRS and ENS Lyon, F-69007 Lyon, France.

<sup>3</sup>GIPSA-lab, CNRS and University of Grenoble, F-38402 St Martin d'Hères, France.

Contact: jerome.boulanger@curie.fr

## ABSTRACT

This work aims at proposing a new reconstruction procedure for structured illumination microscopy. The proposed method is based on some recent development in non-smooth convex optimization that allows to deal with Poisson negative log-likelihood as data fidelity term and with regularization terms allowing to extract sharp features. The performances of the proposed method are compared to the state-of-the-art of SIM reconstruction techniques.

**Index Terms**— Structured illumination microscopy, image restoration, deconvolution, Poisson noise, proximal algorithms.

## 1. INTRODUCTION

Improving the resolution of optical microscopy beyond the limits defined by Abbe remains an important challenge. To achieve this goal, several techniques have been developed to explore living samples at higher resolutions [1]. Among these techniques, *structured illumination microscopy* (SIM) plays a specific role, with its ability to double the spatial resolution with no specific constraints on tagging. If recent advances in deconvolution show that deconvolution microscopy might offer resolution close to SIM for a lower complexity [2], one may combine deconvolution techniques and SIM, as suggested in [3]. This paper deals with such a combination.

The direct model involved in SIM generates a set of noisy low-resolution images, denoted by  $z = (z_\ell)_{1 \leq \ell \leq L}$  and belonging to  $\mathbb{R}^N$ , from a high-resolution image  $x \in \mathbb{R}^N$  (the 2D images are represented as vectors). The linear operator involved during the acquisition process can be modeled by the combination of amplitude modulation, expressed with the diagonal matrix  $M_\ell \in \mathbb{R}^{N \times N}$ , and a convolution by the point spread function (PSF) of the optical system, which is denoted  $A \in \mathbb{R}^{N \times N}$  in the following. The photon-counting process involved in the image formation and the thermal noise in the electronics of the sensor can be modeled by a combination of Poisson and Gaussian noise. However, for the sake of simplicity, we will consider that the gain of the sensor is 1 and that the dark current noise can be factorized as a single Poisson random variable corresponding to the number of generated photo-electrons. Finally, the degradation model can be summarized as follows:

$$z_\ell = \mathcal{P}(AM_\ell \bar{x} + b), \quad (1)$$

where  $b \in \mathbb{R}_+$  denotes a dark current level that is supposed to be known in this study and  $\mathcal{P}$  models for the Poisson noise. Note that

\*This work was supported by the grant PEPS PROMIS funded by the French agencies CNRS, INSERM and INRIA. The authors greatly acknowledge the Nikon Imaging Centre at Institut Curie-CNRS.

when  $L = 1$  and  $M_1$  denotes the identity matrix, the direct model (1) reduces to the one involved in usual *wide-field microscopy* (WF).

The reconstruction procedure aims at estimating a high resolution image  $\hat{x}$  close to the unknown original data  $\bar{x}$  from the observations  $(z_\ell)_{1 \leq \ell \leq L}$ . It can be noticed that very few reconstruction techniques for SIM have been proposed in the literature so far. The standard reconstruction method performs a sequence of demodulation and recombination, as originally proposed in [3]. More recently, another approach based on a variational procedure has been investigated in [4]. This method aims at restoring the acquired data by minimizing a smooth criterion involving a quadratic data fidelity term and a Tikhonov regularization. Moreover, a Bayesian procedure has been proposed, in order to estimate the involved regularization parameter. Dealing with a quadratic criterion such as in [4] allows the authors to use efficient iterative procedures, like gradient based algorithms. However, this kind of criterion does not allow to model accurately the degradation model, e.g. Poisson noise degradation, or to extract some specific features (Tikhonov regularization being adapted for smooth features rather than sharp ones).

Based on recent advances in non-smooth convex optimization [5], the present work aims at proposing an algorithm allowing to deal with a large class of data fidelity terms and regularization terms, such as Poisson negative log-likelihood or regularization promoting sparsity. In the experimental part, we evaluate on simulated data the impact of dealing with such a criterion compare to usual quadratic data fidelity term and regularization term.

Section 2 is dedicated to the description of the criterion we propose to minimize. Section 3 aims at describing the proposed primal-dual algorithm. In Section 4 we describe the numerical experiments and the impact of a criterion involving *Kullback-Leibler* (KL) divergence and *total variation* (TV) for the SIM reconstruction.

**Notations** The  $n$ -th coordinate of a vector  $x \in \mathbb{R}^N$  is denoted by  $x^{(n)}$ . The notation  $\cdot^{[i]}$  stands for the  $i$ -th iteration in the iterative procedure. The notation  $\cdot_\ell$  denotes the  $\ell$ -th component involved in the SIM process. Let  $C$  denote a non-empty closed convex subset of  $\mathbb{R}^N$ , the indicator function  $\iota_C$  is such that  $\iota_C(x) = 0$  if  $x \in C$  and  $+\infty$  otherwise. Let  $f$  be a convex function from  $\mathbb{R}^N$  to  $]-\infty, +\infty]$ .  $\text{Argmin } f = \{x \in \mathbb{R}^N | (\forall y \in \mathbb{R}^N) f(x) \leq f(y)\}$  denotes a set of minimizers while  $\arg \min$  is employed when the set of minimizers is reduced to a singleton.

## 2. INVERSE PROBLEM FORMULATION

We aim at finding a high resolution image  $\hat{x}$  fitting the observations  $(z_\ell)_{1 \leq \ell \leq L}$  through the degradation model described in (1). This im-

plies to look for the solution that minimizes a distance/divergence between the two quantities  $x$  and  $z$ . This can be modelled through the function  $\sum_{\ell=1}^L \sum_{n=1}^N \psi_n((AM_\ell x)^{(n)} + b, z_\ell^{(n)})$  where, for every modulation  $\ell \in \{1, \dots, L\}$  and every pixel  $n \in \{1, \dots, N\}$ ,  $\psi_n(\cdot + b, z_\ell^{(n)})$  denotes a convex, lower semi-continuous, and proper function from  $\mathbb{R}$  to  $]-\infty, +\infty]$ . These assumptions are usual in the convex optimization literature and they allow to deal with quadratic data fidelity term as well as Kullback-Leibler divergence (KL) i.e. for every  $u \in \mathbb{R}$ ,

$$\psi_n(u, z_\ell^{(n)}) = \begin{cases} -z_\ell^{(n)} \log u + u & \text{if } z_\ell^{(n)} > 0 \text{ and } u > 0 \\ u & \text{if } z_\ell^{(n)} = 0 \text{ and } u \geq 0 \\ +\infty & \text{otherwise.} \end{cases} \quad (2)$$

The Kullback-Leibler divergence models the Poisson negative log-likelihood [6]. Consequently, according to a Maximum A Posteriori interpretation, the Kullback-Leibler divergence should be an appropriate data fidelity term for SIM reconstruction.

Moreover, in order to impose some regularity on the solution, we consider a penalization term  $\phi(Hx)$  where  $H \in \mathbb{R}^{K \times N}$  and  $\phi$  denotes a convex, lower semi-continuous, proper function from  $\mathbb{R}^K$  to  $]-\infty, +\infty]$ . For instance, this model allows to constrain the solution to have a minimal total variation considering

$$\phi(H \cdot) = \sum_{n=1}^N \|(H \cdot)^{(n)}\|_{2,1} = \sum_{n=1}^N \sqrt{|(H_1 \cdot)^{(n)}|^2 + |(H_2 \cdot)^{(n)}|^2} \quad (3)$$

where the matrices  $H_1$  and  $H_2$  belongs to  $\mathbb{R}^{N \times N}$  and model the horizontal and vertical finite differences. In this formulation  $H = [H_1^*, H_2^*]^*$  and  $K = 2N$ .

In order to constrain the data dynamic of the data, a convex constraint  $C \subset \mathbb{R}^N$  is usually used for image reconstruction. It may model a positivity constraint as well as a dynamic range constraint. To summarize, the criterion we have to deal with is

$$\hat{x} \in \underset{x \in C}{\text{Argmin}} \sum_{\ell=1}^L \sum_{n=1}^N \psi_n((AM_\ell x)^{(n)} + b, z_\ell^{(n)}) + \lambda \phi(Hx) \quad (4)$$

where  $\lambda > 0$  denotes the regularization parameter. This criterion is convex but non-necessarily differentiable (due to the regularization term and the constraint  $C$ ) and the data fidelity term may not be differentiable with a Lipschitz gradient (e.g. KL). For these reasons, we cannot deal with usual gradient based algorithms and we have to deal with recent proximal techniques [5].

### 3. PROPOSED ALGORITHM

Over the last decade, following the increasing interest for sparsity, several optimization procedures with convergence guarantees have been proposed in order to minimize non-smooth convex criteria involving linear operators and non-finite functions (see [5] and the references therein), i.e., problems formulated as follows

$$\hat{x} \in \underset{x \in \mathbb{R}^N}{\text{Argmin}} \sum_{q=1}^Q f_q(T_q x) \quad (5)$$

where, for every  $q \in \{1, \dots, Q\}$ ,  $T_q \in \mathbb{R}^{N_q \times N}$  and  $f_q$  denotes a convex, lower semi-continuous, and proper function from  $\mathbb{R}^{N_q}$  to  $]-\infty, +\infty]$ .

In order to deal with non-smoothness, the usual gradient descent is replaced with subgradient based iterations involving the computation of the proximity operators of  $(f_q)_{1 \leq q \leq Q}$ ; that is, for every  $q \in \{1, \dots, Q\}$ ,

$$(\forall y \in \mathbb{R}^{N_q}) \quad \text{prox}_{f_q} y = \arg \min_{x \in \mathbb{R}^{N_q}} \|x - y\|^2 + f_q(x).$$

Problem (4) is a particular case of (5) with  $Q = 3$ ,  $N_1 = NL$ ,  $N_2 = K$ ,  $N_3 = N$ ,  $T_1 = [M_1^* A^*, \dots, M_L^* A^*]^*$ ,  $T_2 = H$ ,  $T_3 = \text{Id}$ ,  $f_2 = \lambda \phi$ ,  $f_3 = \iota_C$ , and

$$f_1: u = (u_\ell)_{1 \leq \ell \leq L} \in \mathbb{R}^{NL} \mapsto \sum_{\ell=1}^L \sum_{n=1}^N \psi_n(u_\ell^{(n)} + b, z_\ell^{(n)}).$$

Consequently, we have to compute the proximity operator associated to  $\iota_C$ ,  $\psi_n$ , and  $\phi$ . The proximity operator of  $\iota_C$  reduces to the projection onto the convex set  $C$ , denoted  $P_C$ . The computation of  $\text{prox}_{\psi_n}$  and  $\text{prox}_{\lambda \phi}$  have closed form expressions [5] when  $\psi_n$  and  $\phi$  are expressed by (2) and (3). More precisely, for every  $u = (u_1, u_2) \in \mathbb{R}^N \times \mathbb{R}^N$ ,

$$\text{prox}_{\lambda \phi} u = \left( \text{prox}_{\lambda \|\cdot\|_{2,1}} (u_1^{(n)}, u_2^{(n)}) \right)_{1 \leq n \leq N}$$

where, for every  $(\eta_1, \eta_2) \in \mathbb{R} \times \mathbb{R}$ ,

$$\text{prox}_{\lambda \|\cdot\|_{2,1}}(\eta_1, \eta_2) = \begin{cases} 1 - \lambda \frac{(\eta_1, \eta_2)}{\sqrt{\eta_1^2 + \eta_2^2}} & \text{if } \sqrt{\eta_1^2 + \eta_2^2} \leq \lambda \\ (0, 0) & \text{otherwise} \end{cases}$$

and, for every  $\eta \in \mathbb{R}$ ,

$$\text{prox}_{\psi_n(\cdot + b, z_\ell^{(n)})} \eta = \frac{1}{2} \left( \eta + b - 1 + \sqrt{|\eta + b - 1|^2 + \frac{4}{\sigma} z_\ell^{(n)}} \right) - b.$$

In order to efficiently deal with the linear operators  $(T_q)_{1 \leq q \leq Q}$ , the large panel of algorithms designed for non-smooth convex optimization can be split in two classes: primal algorithms [7] that require to invert large size linear operator (i.e., compute  $(\sum_{q=1}^Q T_q^* T_q)^{-1}$ ) and primal-dual algorithms [8, 9] for which the inversion stands in for the computation of the operator (e.g.  $T_q$ ) and its adjoint (e.g.  $T_q^*$ ). In order to avoid large size matrix inversion that does not have a closed form in SIM, we derive iterations from a revisited and extended version of Chambolle-Pock algorithm [10] proposed in [9].<sup>1</sup> The iterations resulting from [9] and allowing to find a minimizer of (4) are presented in Figure 1. Conditionally to the existence of  $\hat{x}$ , the sequence  $(x^{[i]})_{i \in \mathbb{N}}$  generated by Algorithm 1 converges to  $\hat{x}$ , the solution of (4).

### 4. RESULTS

We consider a synthetic  $512 \times 512$  2-D test image (see Fig. 3(a)). This image contains different types of structures typically observed in life cells such as filaments (actin and microtubules), spots (single molecules or vesicles), and a complex background produced by a fractional Brownian motion providing a wide range of frequencies. The operator  $A$  is approximated by a Gaussian function with a limited frequency support corresponding to the Abbe limit. The modulation operator  $M$  is built so that the modulation frequency is near the cut-off frequency of  $A$  and contains  $L = 15$  modulations corresponding to 3 angles and 5 phases. In order to simulate photon noise typical of optical imaging, we simulate a Poisson random variable whose intensity is given by the simulated image. Finally, in order to be able to fairly compare the WF deconvolution approach with SIM reconstruction, we consider that the intensity of WF simulated test image is 15 times higher corresponding to a 15 times longer exposure time (see Fig. 3 (c) and (d)).

<sup>1</sup>Note that we also derived the iterations resulting from algorithms such as M+LFBF [8] or even the iterations for a primal algorithm PPXA+ [7]. However, for the SIM purpose the extended version of Chambolle-Pock algorithm appears to converge faster than the other methods. For this reason, we have decided to only describe the iterations resulting from the extended version of Chambolle-Pock algorithm proposed in [9].

### STEP 0 – Initialization

Set  $\beta = \|\sum_{\ell=1}^L M_\ell^* A^* A M_\ell + H^* H\|$   
Initialize  $x^{[0]} \in \mathbb{R}^N$   
Initialize  $(\bar{v}_{1,\ell}^{[0]})_{1 \leq \ell \leq L} \in \mathbb{R}^{NL}$  and  $v_2^{[0]} \in \mathbb{R}^K$   
Set  $\tau > 0$  and  $\sigma = \frac{1}{\tau\beta}$   
Set  $i = 0$

### STEP 1 – Steps involving the linear operators

Set  $\rho_i \in [0, 2]$   
 $y^{[i+1]} = PC(x^{[i]} - \tau \sum_{\ell=1}^L M_\ell^* A^* v_{1,\ell}^{[i]} - \tau H^* v_2^{[i]})$   
 $x^{[i+1]} = \rho_i y^{[i+1]} + (1 - \rho_i) x^{[i]}$   
For  $\ell = \{1, \dots, L\}$

$$\begin{cases} u_{1,\ell}^{[i]} = v_{1,\ell}^{[i]} + \sigma A M_\ell (2y^{[i+1]} - x^{[i]}) \\ u_{2,\ell}^{[i]} = v_2^{[i]} + \sigma H (2y^{[i+1]} - x^{[i]}) \end{cases}$$

### STEP 2 – Proximal operator computation

For  $\ell = \{1, \dots, L\}$

$$\begin{cases} p_{1,\ell}^{[i+1]} = u_{1,\ell}^{[i]} - \sigma \left( \text{prox}_{\frac{1}{\sigma} \psi_n(\cdot + b, z_\ell^{(n)})} \left( \frac{u_{1,\ell}^{[i]}}{\sigma} \right) \right)_{1 \leq n \leq N} \\ p_{2,\ell}^{[i+1]} = u_{2,\ell}^{[i]} - \sigma \text{prox}_{\frac{1}{\sigma} \phi} \left( \frac{u_{2,\ell}^{[i]}}{\sigma} \right) \end{cases}$$

### STEP 3 – Update

For  $\ell = \{1, \dots, L\}$

$$\begin{cases} v_{1,\ell}^{[i+1]} = \rho_i p_{1,\ell}^{[i+1]} + (1 - \rho_i) v_{1,\ell}^{[i]} \\ v_2^{[i+1]} = \rho_i p_{2,\ell}^{[i+1]} + (1 - \rho_i) v_2^{[i]} \end{cases}$$

**STEP 4 – Increment  $i = i + 1$  and return to STEP 1 until convergence.**

**Fig. 1.** Proposed algorithm

Figure 4 presents the results corresponding to the WF Microscopy deconvolution and Figure 5 is dedicated to the SIM reconstruction. We compare the reconstruction results in term of SNR and SSIM [11] (the closest from 1 the value is the best is the reconstruction) for two data terms and two regularization terms corresponding respectively to  $\psi_n$  and  $\phi(H \cdot)$ . “KL” stands for the Kullback-Leibler data fidelity term expressed in (2). “Quadratic” models the quadratic fidelity term that is  $\psi_n(\cdot, z_\ell^{(n)}) = (\cdot - z_\ell^{(n)})^2$ . “Tikhonov” regularization stands for  $h = \|\cdot\|^2$ ,  $H$  modelling a Laplacian operator, and  $K = N$ . “TV” corresponds to the regularization term described in (3). One can note that the SIM results associated with “L2 + Tikhonov” leads to the results obtained by Orieux et al. [4], which appears to be the state-of-the-art in term of SIM reconstruction. For this specific smooth criterion the proposed algorithm would probably under-perform gradient-based algorithms which incorporate the Lipschitz differentiability assumption of the criterion.

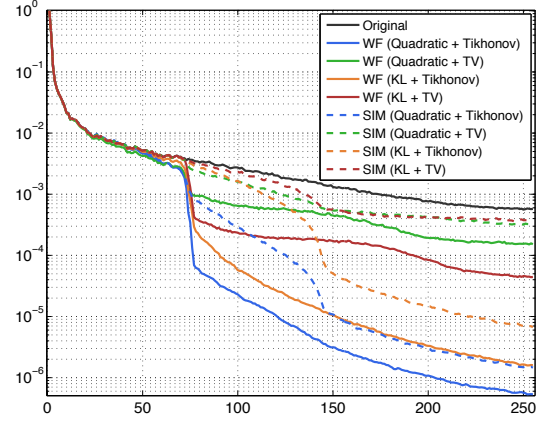
Algorithm 1 can be fitted for dealing with WF deconvolution. It requires to set  $L = 1$ ,  $M_1 = \text{Id}$ , and  $z_1 = \sum_{\ell=1}^L z_\ell$ .

For each test, 2000 iterations are used,  $\tau = 10$ ,  $\rho_i \equiv 1$ , and the regularization parameter  $\lambda^*$  is chosen as the one maximizing the SNR in a range of  $[10^{-5}, 10^5]$ . The first row in Figure 4 (resp. 5) presents the evolution of the SNR w.r.t  $\lambda$ . For each optimal estimate, we computed a radially averaged power spectrum as show on Fig. 2.

In Figs. 4 and 5, the comparison in terms of SNR and SSIM between WF deconvolution and SIM reconstruction leads to the conclusion that, indeed, SIM provides a much higher resolution. In Fig. 2 we can clearly notice that high frequencies are better preserved using a SIM formalism and an adapted criterion (cf. dashed red and green plots in Fig. 2). On the other hand, Figure 5 (g) illustrates the performances of a SIM reconstruction based on a KL-TV criterion. Such a criterion outperforms usual “L2 + Tikhonov” criterion.

## 5. CONCLUSION

In this paper, we proposed a proximal algorithm in order to test several criteria for the restoration of SIM images. It clearly appears that

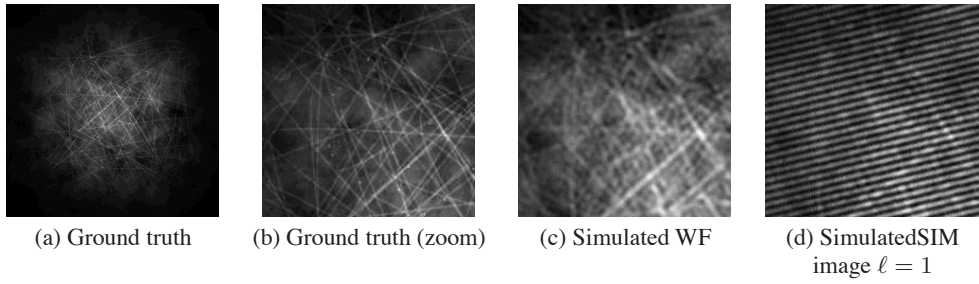


**Fig. 2.** Radially averaged power spectrum of the estimates obtained for the described methods.

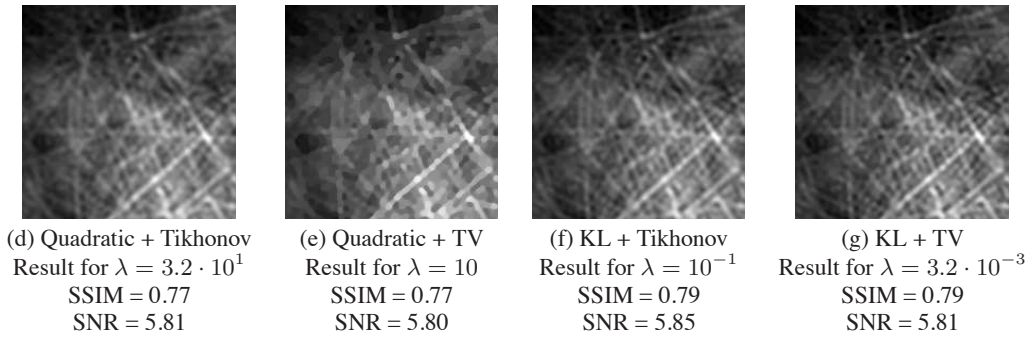
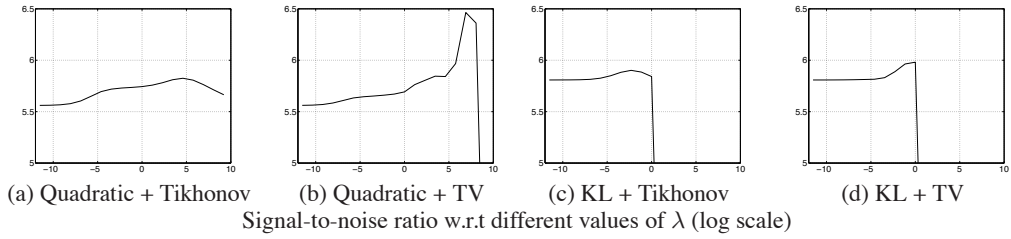
considering suited functionals (such as TV that models sharp features or the Kullback-Leibler divergence, which aims to better deal with Poisson noise) improves the SIM restoration results. In the future, we plan to extend this approach to 3D microscopy and to real images. Finally, other type of functionals could also be explored in this context.

## 6. REFERENCES

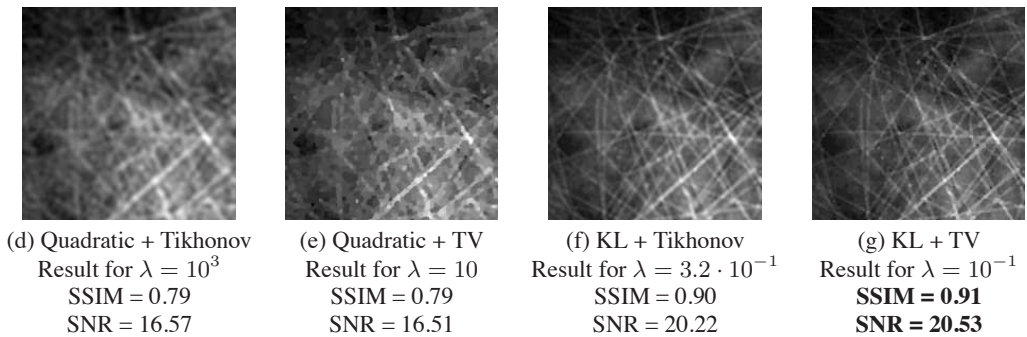
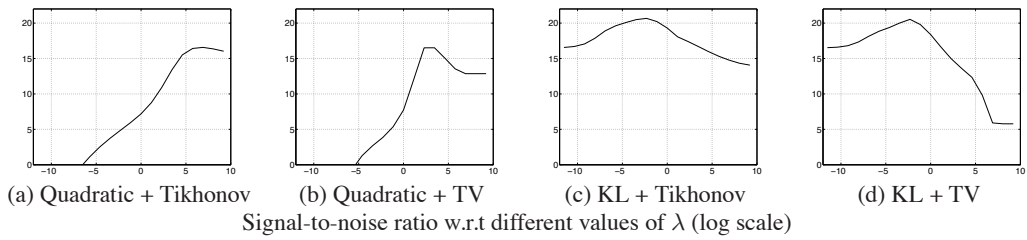
- [1] L. Schermelleh, R. Heintzmann, and H. Leonhardt, “A guide to super-resolution fluorescence microscopy,” *The Journal of Cell Biology*, vol. 190, no. 2, pp. 165–175, July 2010, PMID: 20643879.
- [2] F. Soulez, L. Denis, Y. Tournier, and E. Thiebaud, “Blind deconvolution of 3D data in wide field fluorescence microscopy,” in *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, 2012, pp. 1735–1738.
- [3] M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *Journal of microscopy*, vol. 198, no. Pt 2, pp. 82–87, May 2000, PMID: 10810003.
- [4] F. Orieux, E. Sepulveda, V. Lorient, 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.
- [5] P. L. Combettes and J.-C. Pesquet, “Proximal splitting methods in signal processing,” in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds., pp. 185–212. Springer-Verlag, New York, 2010.
- [6] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J.C. Olivo-Marin, and J. Zerubia, “Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution,” *Microscopy Research Technique*, vol. 69, pp. 260–266, 2006.
- [7] J.-C. Pesquet and N. Pustelnik, “A parallel inertial proximal optimization method,” *Pac. J. Optim.*, vol. 8, no. 2, pp. 273–305, Apr. 2012.
- [8] P. L. Combettes and J.-C. Pesquet, “Primal-dual splitting algorithm for solving inclusions with mixtures of composite, Lipschitzian, and parallel-sum type monotone operators,” *Set-Valued Var. Anal.*, 2011.
- [9] 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, 2013.
- [10] A. Chambolle and T. Pock, “A first-order primal-dual algorithm for convex problems with applications to imaging,” *J. Math. Imag. Vis.*, vol. 40, no. 1, pp. 120–145, 2011.
- [11] Z. Wang and A. C. Bovik, “Mean squared error: love it or leave it?,” *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.



**Fig. 3.** Data



**Fig. 4.** Wild Field results



**Fig. 5.** SIM results