A Non-convex Nonseparable Approach to Single-Molecule Localization Microscopy

Raymond H. Chan[1], Damiana Lazzaro[2], Serena Morigi[2(✉)], and Fiorella Sgallari[2]

[1] Department of Mathematics, City University of Hong Kong, Kowloon Tong, Hong Kong
rchan.sci@cityu.edu.hk

[2] Department of Mathematics, University of Bologna, Bologna, Italy
{damiana.lazzaro,serena.morigi,fiorella.sgallari}@unibo.it

Abstract. We present a method for high-density super-resolution microscopy which integrates a sparsity-promoting penalty and a blur kernel correction into a nonsmooth, non-convex, nonseparable variational formulation. An efficient majorization minimization strategy is applied to reduce the challenging optimization problem to the solution of a series of easier convex problems.

1 Introduction

Single-molecule localization microscopy (SMLM) is a powerful microscopical technique that is used to detect with high precision molecule localization by sequentially activating and imaging only a random sparse subset of fluorescent molecules in the sample at the same time, localizing these few emitters very precisely, deactivating them and activating another subset. Repeating the process several thousand times ensures that all fluorophores can go through the bright state and are recorded sequentially in frames. A high density map of fluorophore positions is then reconstructed by a sequential imaging process of sparse subsets of fluorophores distributed over thousands of frames. Even when theoretical characteristics on the blur kernel involved in the formation of the images are given, the acquisition process is so complicated that also the slightest difference to the theoretical ideal conditions, results in distortions which affect Point Spread Function (PSF), and, consequently, the image recovering process [12]. Several algorithms have been developed for point source localization in the context of the SMLM challenge. In [5] the variational model is equipped with a sparsity-promoting CEL0 penalty and solved by iterative reweighting. In [10] the blur kernel inaccuracy is addressed with a Taylor approximation of the PSF. For a detailed list of the software proposed to solve the SMLM challenge, and on the physical background of SMLM, we refer the reader to [12]. We formulate the localization problem as a variational sparse image reconstruction problem which integrates a nonseparable structure-preserving penalty. To overcome the problem of inaccurate blur kernel which can cause severe distorsions on the solution,

J. Lellmann et al. (Eds.): SSVM 2019, LNCS 11603, pp. 498–509, 2019.
https://doi.org/10.1007/978-3-030-22368-7_39

we combine our sparsity-promoting formulation with an effective blur kernel model correction. We perform a nonsmooth nonconvex optimization algorithm for the minimization task, based on a majorization minimization strategy. The proposed algorithm is validated on both simulated and experimental datasets and compared with other challenging high density localization softwares.

2 Image Formation Modelling

Let $u \in \mathbb{R}^{N\times N}$ be the unknown high resolution image to be reconstructed, and $g \in \mathbb{R}^{n\times n}$ the acquisition following the molecules activation, with $n = \frac{N}{d}$, and d is the downsampling factor. The linear acquisition process can be formulated as

$$g = \mathcal{P}(M_d(\mathcal{B}_t * u)) + \eta, \tag{1}$$

where $\mathcal{P}$ models the degradation with Poisson noise, η is the zero-mean Gaussian image noise, $\mathcal{B}_t$ is the convolution blurring operator with Gaussian kernel, and $M_d : \mathbb{R}^{N\times N} \to \mathbb{R}^{n\times n}$ is the downsampling operator which averages pixels by patches of size $d \times d$ in order to map the high resolution image to the coarser one. In the image formation model (1) we assumed that the given blur kernel model of the optical system (microscope) is free of error. The challenge in [12] provided parameters to model a Gaussian PSF model for each experiment data set. However, as assessed in [12] a simple Gaussian PSF model can be sufficiently accurate for low-density data, whereas the quality of high-density imaging depends strongly on the model of the PSF and the PSF model will have an even more significant role in 3D SMLM applications. When an inaccurate blur kernel is used as the input, significant distortions can appear in the recovered image. In this work, we assume an inaccurate blur kernel $\mathcal{B}$, with unknown model error $\delta_{\mathcal{B}}$, that is the true blur kernel $\mathcal{B}_t = \mathcal{B} - \delta_{\mathcal{B}}$ and we neglect the Poisson shot noise contribution, thus (1) becomes

$$g = M_d((\mathcal{B} - \delta_{\mathcal{B}}) * u) + \eta = M_d((\mathcal{B} * u) - (\delta_{\mathcal{B}} * u)) + \eta. \tag{2}$$

In Fig. 1 the high pass nature of the model error $\delta_{\mathcal{B}}$ is shown. Since $\delta_{\mathcal{B}}$ is the difference between two low pass filter, the input blur kernel $\mathcal{B}$ and the unknown true blur kernel $\mathcal{B}_t$, the corrector term $\delta_{\mathcal{B}} * u$ has an enhancing effect of the edges in the image, see [7].

Let us introduce a matrix-vector notation that will be useful in the algorithmic description. In particular, let $B \in \mathbb{R}^{N\times N}$ be the blurring matrix corresponding to the operator $\mathcal{B}$, then

$$\mathcal{B} * u = BuB^T = (B \otimes B)vec(u) = \bar{B}vec(u), \tag{3}$$

where $\otimes$ is the Kronecker product, and $vec(u)$ denotes the vectorization of u. Let $M \in \mathbb{R}^{n\times N}$ be the downsampling matrix such that

$$M_d(u) = MuM^T = (M \otimes M)vec(u) = \bar{M}vec(u). \tag{4}$$

To accurately estimate u in (2) we only need to know the residual term $\delta_{\mathcal{B}} * u$ instead of the perturbation operator $\delta_{\mathcal{B}}$ itself which is hard to estimate due to the lacking of information of the blurring process.

3 Penalty Function

In this section we introduce the sparsity-inducing function used in our variational model, and we highlight some of its properties: be non-convex, parameterized with μ so that we can tune its non-convex behaviour, structure preserving, as required by the high density molecule localization problem. In light of the previous requirements, the proposed penalty function is defined on the local data set consisting of a neighborhood of the pixel (i, j). In particular, we consider a square window centered at u_{ij} containing all the $(2\ell+1)^2$ neighbors, $\ell \geq 1$, and we denote by $I_{ij} = \{(i+r, j+s) : r, s = -\ell, \ldots, \ell\}$ the neighborhood index set of size ℓ, and by $t := u|_{I_{ij}},\ t \in \mathbb{R}_+^{(2\ell+1)\times(2\ell+1)}$ the restriction of u to the window I_{ij}. Following [9], to fulfill our goals, we define the non-convex nonseparable penalty function $\psi : \mathbb{R}_+^{(2\ell+1)\times(2\ell+1)} \to \mathbb{R}$ as follows:

$$\psi(t;\mu) = \frac{1}{\log(2)} \log\left(\frac{2}{1 + exp(-\|vec(t)\|_1/\mu)}\right), \tag{5}$$

where $\mu > 0$ represents a parameter which controls the degree of non-convexity of the penalty function and $vec(t) \in \mathbb{R}_+^{(2\ell+1)^2}$. The partial derivatives of $\psi(t;\mu)$ in (5), $\forall (r,s) \in I_{ij}$ are given by

$$\frac{\partial \psi}{\partial |u_{r,s}|}(t;\mu) = \frac{1}{\mu \log(2)} \frac{1}{1 + exp(\|vec(u|_{I_{rs}})\|_1/\mu)}. \tag{6}$$

Simple investigations of the first and second order partial derivatives lead to the following properties for $\psi(t;\mu)$, which characterize a sparsity-promoting function:

- $\psi(t;\mu)$ is concave and non-decreasing;
- $\psi(t;\mu)$ has continuous bounded partial derivatives for $t \neq 0$, and $\psi(0;\mu) = 0$;
- for μ values approaching to zero, $\psi(t;\mu)$ tends to the ℓ_0 quasi-norm.

4 Optimization Model NCNS for SMLM Problem

In the SMLM problem the aim is to recover sparse images with non-zero pixels clustered into elongated structures, whose number, dimension and position are unknown. The problem can also be classified as a blind cluster structured sparse image recovery problem [9]. For its solution we propose to minimize the following nonconvex cost function involving the non-convex nonseparable (NCNS) penalty function introduced in Sect. 3. Let $h := \delta_{\mathcal{B}} * u$ be the correction term. Then we will denote by NCNS model the following optimization problem

$$\min_{u,h \in \mathbb{R}^{N\times N}} \{J(u,h;\lambda_1,\lambda_2) := F(u,h) + \lambda_1 R(u;\mu) + \lambda_2 H(h)\} \tag{7}$$

where $\lambda_1, \lambda_2 > 0$ are regularization parameters,

$$F(u,h) = \frac{1}{2}\|M_d((\mathcal{B} * u) - h) - g\|_2^2, \tag{8}$$

is the fidelity term, the penalty function $R(u;\mu)$ reads as

$$R(u;\mu) = \sum_{i=1}^{N}\sum_{j=1}^{N} \psi_{i,j}(u;\mu), \tag{9}$$

where $\psi_{i,j}(u;\mu) = \psi(u|_{I_{ij}};\mu)$ is defined in (5), and $H(h) = \|h\|_p^p, p = \{1,2\}$. The aforementioned properties of $\psi_{i,j}$ induce similar properties in the sparsity-promoting function $R(\cdot;\mu)$, which turns out to be both non-convex and non-separable. From (9) we can define the partial derivative with respect to a pixel $(p,q) \in I_{ij}$ as

$$\frac{\partial R(u;\mu)}{\partial |u_{p,q}|} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial \psi_{i,j}}{\partial |u_{p,q}|}(u;\mu) = \sum_{(r,s)\in I_{ij}} \frac{\partial \psi_{r,s}}{\partial |u_{p,q}|}(u;\mu). \tag{10}$$

Formula (10) is obtained taking into account that, due to the local support of ψ, the partial derivatives $\frac{\partial \psi_{i,j}}{\partial |u_{p,q}|}$ that are non-zero are those defined on the $(2\ell+1)^2$ windows containing the pixel $u_{p,q}$ itself.

The effect of the nonseparable penalty $R(u;\mu)$ on a pixel $u_{p,q}$ depends on its neighbors defined in I_{pq}. In particular, the pixel $u_{p,q}$ is considered as belonging to a structure and thus preserved if the ℓ_1 norm of the vector of the pixels in its surrounding window is greater than μ, otherwise, it is forced to be zero, because it could be an isolated artifact. This fulfills the requirements of the SMLM data, where the fluorescent molecules are in general aggregated forming elongated thin structures.

Proposition 1. *For any couple of positive parameters (λ_1, λ_2) the functional $J(u,h;\lambda_1,\lambda_2) : \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} \to \mathbb{R}$, defined in (7) is non-convex, proper, continuous, bounded from below by zero but not coercive in u, hence the existence of global minimizers for J is not guaranteed.*

The lack of coercivity not only stems from $R(u;\mu)$, but also from the down-sampling operator $\mathcal{M}_d$ which has a nontrivial kernel, and the non-convexity is due to $R(u;\mu)$. The problem (7) is in general a challenging non-convex nonseparable optimization problem. A minimizer for J in (7) is carried out by applying the Majorization-Minimization (MM) strategy which iteratively minimizes a convexification of J obtained by replacing R with its linearization $\tilde{R}$ around the previous iterate, [8].

In the kth **majorization step**, we generate a tangent majorant of the function (surrogate functional) $J(u,h;\lambda_1,\lambda_2)$ defined as

$$\tilde{J}(u,h;\lambda_1,\lambda_2,u^{(k)},\mu^{(k)}) = F(u,h) + \lambda_1 \tilde{R}(u;u^{(k)},\mu^{(k)}) + \lambda_2 H(h), \tag{11}$$

where the linear tangent majorant of $R(u;\mu^{(k)})$ at $u^{(k)}$ is

$$\tilde{R}(u;u^{(k)},\mu^{(k)}) = R(u^{(k)};\mu^{(k)}) + \sum_{i=1}^{N}\sum_{j=1}^{N} \left(\frac{\partial R(u;\mu^{(k)})}{\partial |u_{i,j}|}\Big|_{u=u^{(k)}} (|u_{i,j}| - |u_{i,j}^{(k)}|) \right). \tag{12}$$

A suitable reduction of the parameter $\mu^{(k)}$ is carried out at each iteration k, namely $\mu^{(k+1)} = c_\mu \mu^{(k)}$, with $0 < c_\mu < 1$, in such a way that, as the number of iterations increases, the sparsity inducing function gets closer to its limit ℓ_0 quasi-norm.

In the **minimization step**, the following convex nonsmooth minimization problem is solved

$$\{u^{(k+1)}, h^{(k+1)}\} = \arg\min_{u,h} \left\{ \tilde{J}(u, h; \lambda_1, \lambda_2, u^{(k)}, \mu^{(k)}) \right\}. \tag{13}$$

By neglecting the constant terms, problem (13) can be simplified to:

$$\{u^{(k+1)}, h^{(k+1)}\} = \arg\min_{u,h} \{F(u, h) + \lambda_1 \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j}^{(k)} |u_{i,j}| + \lambda_2 H(h)\} \tag{14}$$

where, using (10) and (6), the positive weights are defined as

$$w_{i,j}^{(k)} = \frac{\partial R(u; \mu^{(k)})}{\partial |u_{i,j}|}\Big|_{u=u^{(k)}} = \sum_{(r,s)\in I_{ij}} \frac{1}{\mu^{(k)} \log(2)} \frac{1}{1 + exp(\|vec(u|_{I_{ij}})\|_1/\mu^{(k)})}. \tag{15}$$

Equation (14) can be rewritten in vectorized form as

$$\{u^{(k+1)}, h^{(k+1)}\} = \arg\min_{u,h} \{F(u, h) + \lambda_1 \underbrace{\|W^{(k)} u\|_1}_{G(u)} + \lambda_2 H(h)\}, \tag{16}$$

where $W^{(k)} \in \mathbb{R}^{N^2 \times N^2}$ is a diagonal matrix of weights $w_{ij}^{(k)}$, which assume high values for isolated pixel (i, j) and small values for pixels representing structures. For the sake of simplicity, from now on we will represent the image variables in vectorized form.

4.1 Solving the Minimization Step

In this section we determine an approximate solution of the minimization step (16), which can be rewritten in the form

$$\{u^*, h^*\} = \arg\min_{u,h} \{F(u, h) + \lambda_1 G(u) + \lambda_2 H(h)\}, \tag{17}$$

where we neglected the iteration index (k).

A standard approach for solving (17) is thus to adopt an alternating minimization strategy. However, its convergence is only guaranteed under restrictive assumptions. Therefore, alternative strategies based on proximal tools have been proposed [4]. In particular, in this work, following [1], we propose to adopt the alternating accelerated Forward Backward algorithm which alternates the minimization on the two variable blocks (u, h).

Assuming that $F(u, h)$ is a C^1 coupling function which is required to have only partial Lipschitz continuous gradients $\nabla_u(F(u, h))$ and $\nabla_h(F(u, h))$, and

that each of the regularizers $G(u)$ and $H(h)$ is proper, lower semicontinuous with an efficiently computable proximal mapping. In particular, $G(u)$ is convex and nonsmooth, while $H(h)$ is convex and, eventually, nonsmooth. We cannot claim the same for the optimization problem (7).

Proposition 2. *For any fixed h the function $u \to F(u,h)$ has partial Lipschitz continuous gradient with moduli $L_1 = \rho(A^TA)$, with ρ denoting the spectral radius, that is*

$$\|\nabla_u F(x,h) - \nabla_u F(y,h)\| \le L_1\|x-y\|, \quad \forall x,y \in \mathbb{R}^{N^2},$$

where

$$\nabla_u(F(u,h)) = A^T(Au - g + \bar{M}h), \tag{18}$$

with $A = \bar{M}\bar{B}$, $\bar{M}$ defined in (4) and $\bar{B}$ in (3). For any fixed u the function $h \to F(u,h)$ has partial Lipschitz continuous gradient $\nabla_h F(u,h)$ with moduli $L_2 = \rho(M^TM)$ that is

$$\|\nabla_h F(u,x) - \nabla_h F(u,y)\| \le L_2\|x-y\|, \quad \forall x,y \in \mathbb{R}^{N^2},$$

where

$$\nabla_h(F(u,h)) = \bar{M}^T(\bar{M}h - Au + g). \tag{19}$$

Formulas (18) and (19) can be derived from (8), which is rewritten as

$$\begin{aligned} F(u,h) &= \frac{1}{2}\|\bar{M}(\bar{B}u - h) - g\|_2^2, \\ &= \frac{1}{2}(u^TA^TAu + h^T\bar{M}^T\bar{M}h - 2u^TA^T\bar{M}h + 2g^T\bar{M}h - 2u^TA^Tg + g^Tg). \end{aligned}$$

Let $0 < \beta_1 < \frac{1}{L_1}$ and $0 < \beta_2 < \frac{1}{L_2}$, the approximate solution of the optimization problem (17) is obtained by the iterative procedure sketched below.

- Initialization: start with $u_0 = \tilde{u}_0 = u^{(k)}, h_0 = \tilde{h}_0 = h^{(k)}$, $\lambda_1, \lambda_2 > 0$,
- For each $\ell \ge 1$ generate the sequence (u_ℓ, h_ℓ) by iterating
  - Accelerated FB for u

$$v_\ell = u_{\ell-1} - \beta_1\nabla_u(F(u_{\ell-1}, h_{\ell-1})) \tag{20}$$

$$\tilde{u}_\ell = \arg\min_u\{\frac{1}{2\beta_1}\|u - v_\ell\|_2^2 + \lambda_1 G(u)\} \tag{21}$$

$$u_\ell = \tilde{u}_\ell + \tau_\ell(\tilde{u}_\ell - \tilde{u}_{\ell-1}) \tag{22}$$

  - Accelerated FB for h

$$s_\ell = h_{\ell-1} - \beta_2\nabla_h(F(u_\ell, h_{\ell-1})) \tag{23}$$

$$\tilde{h}_\ell = \arg\min_h\{\frac{1}{2\beta_2}\|h - s_\ell\|_2^2 + \lambda_2 H(h)\} \tag{24}$$

$$h_\ell = \tilde{h}_\ell + \tau_\ell(\tilde{h}_\ell - \tilde{h}_{\ell-1}). \tag{25}$$

The FB procedure is stopped when the functional (11), evaluated in the current u_ℓ, h_ℓ solutions, drops below 10^{-8}.

The weights τ_ℓ in (22) and (25) used for convergence acceleration are computed as in [2]. The optimization subproblem (21) for u reduces to a weighted soft thresholding with an explicitly given closed-form solution

$$\tilde{u}_\ell = S_{\lambda_1 \beta_1 diag(W)}(v_\ell),$$

where $S_t(\nu)$ is a point-wise soft-thresholding function which, for given vectors t and ν, applies soft thresholding with parameter t_i to the element ν_i of ν, namely $[S_t(\nu)]_i = \text{sign}(\nu_i)\max(0, |\nu_i| - t_i), \quad \forall i.$

The minimization of (24) is easily obtained as follows

$$\begin{array}{l} \tilde{h}_\ell = \frac{1}{1+\lambda_2\beta_2} s_\ell \text{ for } H(h) = \|h\|_2^2, \\ \tilde{h}_\ell = S_{\lambda_2\beta_2}(s_\ell) \text{ for } H(h) = \|h\|_1. \end{array} \tag{26}$$

At each Majorization step the parameter λ_1 is decreased following the well-known continuation framework [6], that significantly reduces the number of iterations required. In particular, we adopt the following reduction:

$$\lambda_1^{(k+1)} = c_\lambda \cdot \tilde{J}(u^{(k+1)}, h^{(k+1)}; \lambda_1^{(k)}, \lambda_2^{(k)}, u^{(k)}), \qquad 0 < c_\lambda < 1. \tag{27}$$

The algorithm starts with an initial over-regularized problem and then, at each subsequent majorization step, it reduces the value of parameter λ_1 proportionally to the decreasing of the functional.

For what concerns the parameter choice for λ_2 we consider an a priori fixed value which can be estimated regarding the accuracy of the PSF.

Finally, each nonzero entry $u^*_{i,j}$ in the minimizer of (7) is selected as a fluorescent molecule with localization (X_i, Y_j).

5 Numerical Experiments

We compared the proposed NCNS algorithm, applied with $\ell = 1$ and $\mu^0 = 1$ in (5), with the methods FALCON [10], ThunderSTORM [11], IRL1-CEL0 [5], which are super-resolution localization algorithms currently among the best state-of-the-art methods for high-density molecules estimation according to the 2013/2016 IEEE ISBI Single-Molecule Localisation Microscopy (SMLM) challenge [12]. The algorithms have been provided by the authors. In the experimental results, the methods FALCON and ThunderSTORM are equipped with a post-processing phase, while the proposed NCNS method does not exploit any post-processing. Further improvements will be integrated for removing false positive using a centroid method as suggested in [3].

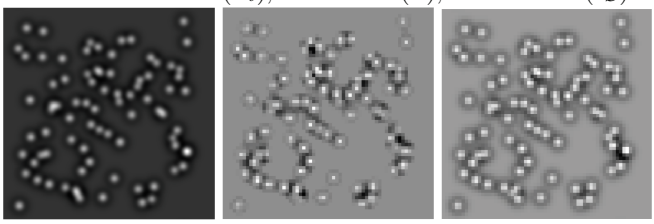
For all the examples, the reconstructed images $N \times N$ are obtained from the acquired images $n \times n$ where $N = n \times d$, $d = 4$ and $n = 64$, that is $N = 256$.

In the simulated data delivered the xy-Gaussian PSF $\mathcal{B}$ is applied to very high resolution images ($n \times 20$) and is characterized by a Full Width at Half

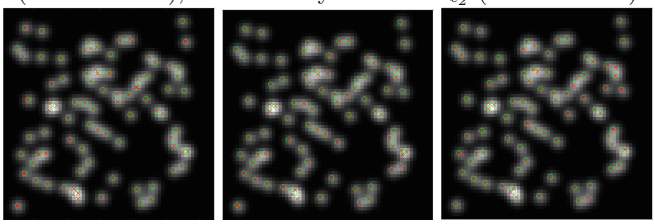
Maximum (FWHM) parameter provided with the dataset which is related to the standard deviation σ by the relation $\sigma = FWHM/2.355$ nm. The Gaussian PSF $\mathcal{B}$ applied in the reconstruction algorithms which process high resolution images of size N, is characterized by a standard deviation σ obtained from the relation

$$\sigma = \left(\frac{N}{(n \times 20)} FWHM \right) /2.355.$$

PSF kernels: true ($\mathcal{B}_t$), inaccurate ($\mathcal{B}$), model error ($\delta_{\mathcal{B}}$)



Correction term h: ground-truth ($h = \delta_{\mathcal{B}} * u_{GT}$), estimated by the NCNS_{ℓ_1} (RMSE=10^{-3}), estimated by the NCNS_{ℓ_2} (RMSE=10^{-4})



Molecule localizations (JAC(%)): NCNS-0 (86.6), NCNS_{ℓ_1} (89.6), NCNS_{ℓ_2} (90.7).

Fig. 1. Comparisons among different regularization terms for h in (7).

5.1 Performance Evaluation

The performances are evaluated in terms of molecule localizations measured by the **detection rate** via the Jaccard index (JAC), and the **localization accuracy**, measured by the root-mean-square-error (RMSE). The evaluation of both the metrics are performed by the tool in http://bigwww.epfl.ch/smlm/challenge2016/. In particular, let R and T be the two sets of reference (ground truth) molecules and test molecules respectively, the localized molecules successfully paired with some test molecules are classified as true positives (TP), while the remaining localized molecules unpaired are categorized as false positives (FP), and the ground truth molecules not associated with any localized

molecules are categorized as false negatives (FN) and related by $FN = |R| - TP$ and $FP = |T| - TP$. A test molecule is paired with a reference one only if the distance between them is lower than a tolerance TOL which should be less than the FWHM of the PSF. The Jaccard index defined by

$$JAC(\%) := \frac{TP}{TP + FP + FN} \times 100 = \frac{|R \cap T|}{|R| + |T| - |R \cap T|}. \tag{28}$$

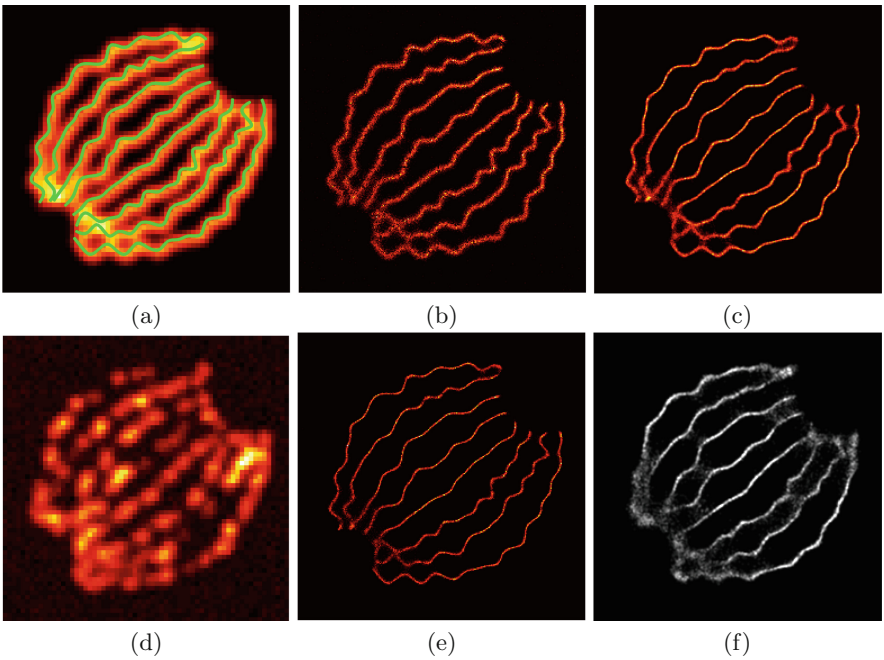


Fig. 2. Example 2: Averaged image (a) acquisition; (b) NCNS_{ℓ_2}; (c) FALCON; (d) single image (frame 58); (e) IRL1-CEL0; (f) ThunderSTORM.

Example 1: Performance of the Blur Correction Term. We first illustrate the benefits introduced by the proposed blur kernel correction in NCNS algorithm applied to a simple synthetic image "Toy" provided by the authors of [5], for which also the ground-truth u_{GT} is given. In particular, we compare the results obtained by the proposed NCNS algorithm without the h regularization term in (7) by optimizing only over u (NCNS−0), with NCNS and $H(h) = \|h\|_2^2$ (NCNS_{ℓ_2}), and with NCNS and $H(h) = \|h\|_1$ (NCNS_{ℓ_1}).

The test image "Toy" of dimension 256×256 has been blurred by the true blur Gaussian kernel $\mathcal{B}_t$ with unknown standard deviation illustrated in Fig. 1

Table 1. Example 2: JAC (and RMSE) for different JAC TOLs.

Method - TOL (nm)	100	150	200	250
NCNS_{ℓ_2}	55.95 (52.12)	64.55 (60.70)	66.07 (64.01)	66.96 (65.75)
IRL1-CEL0	46.79 (**43.15**)	49.33 (**47.91**)	50.06 (**50.59**)	50.49 (**53.15**)
FALCON	**61.92** (49.75)	**72.58** (59.80)	**76.34** (65.66)	**78.09** (69.76)
ThunderSTORM	14.17 (51.70)	17.43 (68.41)	18.61 (77.41)	18.83 (80.10)

Table 2. Example 3 - Dataset MT0.N1.HD and MT0.N2.HD: JAC and RMSE values for different reconstruction methods using a fixed $TOL = 250$.

	MT0.N1.HD		MT0.N2.HD	
Method	JAC	RMSE	JAC	RMSE
NCNS_{ℓ_2}	**59.18**	69.20	**49.10**	72.4
NCNS-0	56.75	69.70	48.65	72.4
IRL1-CEL0	37.90	73.04	34.33	71.4
FALCON	44.78	**56.23**	44.55	87.0
ThunderSTORM	52.92	59.61	46.06	**61.7**

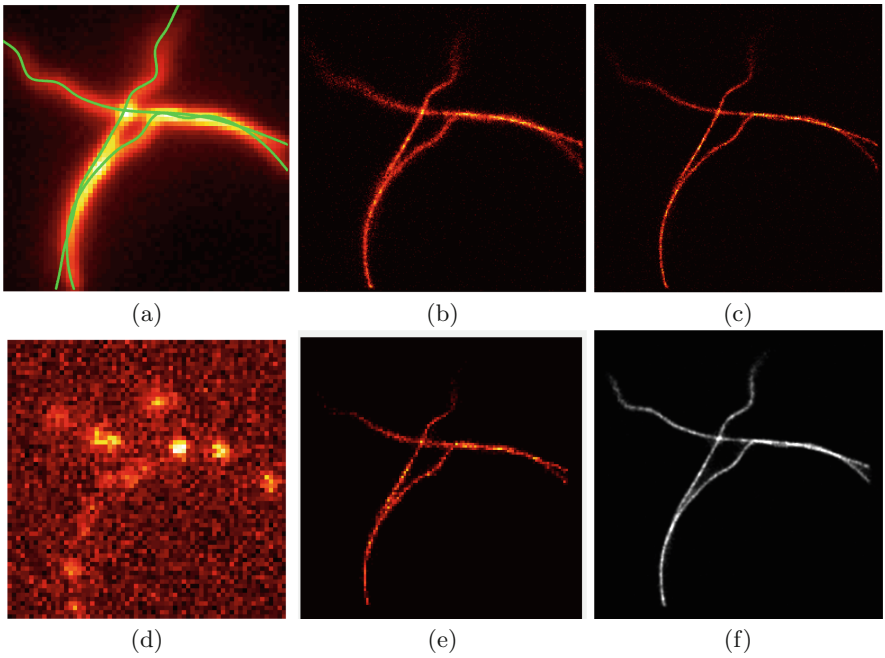


Fig. 3. Example 3 - `MT0.N1.HD` dataset (a)–(f): Averaged image (a) acquisition; (b) NCNS_{ℓ_2}; (c) IRL1-CEL0; (d) single image (frame 700); (e)FALCON; (f) ThunderSTORM.

(first row, left), and subsampled according to the M_d operator with $d = 4$. All the algorithms have been initialized by an inaccurate Gaussian blur kernel $\mathcal{B}$ with $\sigma = 10^{-4}$, shown in Fig. 1 (first row, center). The model error $\delta_{\mathcal{B}}$ obtained by the difference between $\mathcal{B}-\mathcal{B}_t$ is shown in Fig. 1 (first row, right). The algorithms minimizing the functional (7) produce the approximate solutions (u^*, h^*). Figure 1 (second row) reports from left to right: the true h computed by $h = \delta_{\mathcal{B}} * u_{GT}$, the solution h^* of NCNC_{ℓ_1} and of NCNC_{ℓ_2}. In the third row of Fig. 1 we illustrate the acquired blurred image $g \in \mathbb{R}^{64\times 64}$ with overimposed the ground-truth molecule locations by green circles, and the estimated molecule locations by red crosses. For each method we also reported the Jaccard index obtained and the RMSE computed on h^* results with respect to the ground truth. For what concerns the h reconstructions, qualitative and quantitative results confirm that the use of ℓ_2 norm regularization term in (7) instead of the ℓ_1 norm, provides a more accurate and smooth reconstruction, avoiding the well-known staircase effects. The Jaccard indices highlight the noticeable advantages of the presence of the model error $\delta_{\mathcal{B}}$ to correct the inaccuracy of the guessed blur kernel during the reconstruction process. More pronounced is the error of the initial blur kernel $\mathcal{B}$ compared to the one that has really corrupted the data $\mathcal{B}_t$, and more significant is the contribution of the correction term.

Example 2: Challenge 2013 Bundled Tubes HD. The Bundled Tubes HD SMLM challenge is part of the Challenge 2013 which represents a set of high density simulated acquisitions of a bundle of 8 simulated tubes of 30 nm diameter. For this simulation, the camera resolution is 64×64 pixels of PixelSize 100 nm, the PSF is modelled by a Gaussian function whose $FWHM = 258.21$ nm, and the stack simulates 81049 emitters activated on 361 different frames. Figure 2 shows the averaged acquisition image with the ground truth in green (Fig. 2(a)), a single image extracted from the stack (Fig. 2(c)), together with the averaged reconstructions of the whole stack, given by the average of the reconstructions of the 361 frames obtained by the compared methods. In Table 1 the Jaccard index results are reported for different tolerances TOL; the best results are shown in bold.

Example 3: Challenge 2016 MT0.N1.HD and MT0.N2.HD. The datasets MT0.N1.HD and MT0.N2.HD in the Challenge 2016 represent three microtubules in the field of view of $6.4 \times 6.4 \times 1.5\,\mu$m. The resolution of the camera is 64 pixels, the pixelsize is 100 nm, the stack simulates 31612 emitters activated on 2500 different frames, the PSF is modelled by a Gaussian function whose $FWHM = 270.21$ nm. The two datasets MT0.N1.HD and MT0.N2.HD differ in the noise corruption, and in the molecule density which are respectively of 2.0 and 0.2. Figure 3 shows the reconstructions of the whole stack MT0.N2.HD, given by the average of the reconstructions of the 2500 frames, processed by the several methods. In Table 2 the Jaccard index (JAC) and RMSE values are reported for the different reconstruction methods for the two different cases.

Results shown in Tables 1, 2 and illustrated in Figs. 2, 3, highlight the good performance of the proposed NCNS algorithm, further improved in Example 3 where the data sparsity is more pronounced with respect to data in Example 2.

6 Conclusion and Future Work

In this paper, we have proposed a non-convex nonseparable optimization algorithm for the 2D molecule localization in high-density super-resolution microscopy which combines a sparsity-promoting formulation with an accurate estimate of the inaccurate blur kernel. The performance results confirm the efficacy of the proposed variational model in the SMLM context.

References

1. Bolte, J., Sabach, S., Teboulle, M.: Proximal alternating linearized minimization for nonconvex and nonsmooth problems. Math. Program. **146**, 459–494 (2014)
2. Chambolle, A., Dossal, C.: On the convergence of the iterates of the "fast iterative shrinkage/thresholding algorithm". J. Optim. Theory Appl. **166**(3), 968–982 (2015)
3. Chan, R., Wang, C., Nikolova, M., Plemmons, R., Prasad, S.: Non-convex optimization for 3D point source localization using a rotating point spread function. SIAM J. Imaging Sci. (2018, in press)
4. Chouzenoux, E., Pesquet, J.C., Repetti, A.: A block coordinate variable metric forward-backward algorithm. J. Glob. Optim. **66**(3), 457–485 (2016)
5. Gazagnes, S., Soubies, E., Blanc-Féraud, L.: High density molecule localization for super-resolution microscopy using CEL0 based sparse approximation. In: IEEE 14th International Symposium on Biomedical Imaging, Melbourne, VIC, pp. 28–31 (2017)
6. Hale, E.T., Yin, W., Zhang, Y.: Fixed-point continuation for ℓ_1-minimization: methodology and convergence. SIAM J. Optim. **19**(3), 1107–1130 (2008)
7. Ji, H., Wang, K.: Robust image deblurring with an inaccurate blur kernel. IEEE Trans. Image Process. **21**(4), 1624–1634 (2012)
8. Lanza, A., Morigi, S., Selesnick, I., Sgallari, F.: Nonconvex nonsmooth optimization via convex-nonconvex majorization-minimization. Numer. Math. **136**(2), 343–381 (2017)
9. Lazzaro, D., Montefusco, L.B., Papi, S.: Blind cluster structured sparse signal recovery: a nonconvex approach. Signal Process. **109**, 212–225 (2015)
10. Min, J., Holden, S.J., Carlini, L., Unser, M., Manley, S., Ye, J.C.: 3D high-density localization microscopy using hybrid astigmatic/ biplane imaging and sparse image reconstruction. Biomed. Opt. Express **5**(11), 3935–3948 (2014)
11. Ovesny, M., Křížek, P., Borkovec, J., Svindrych, Z., Hagen, G.: ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. Bioinformatics **30**(16), 2389–2390 (2014)
12. Sage, D., et al.: Quantitative evaluation of software packages for single molecule localization microscopy. Nature Methods **12**(8), 717 (2015)