Edge-preserving nonnegative deconvolution of hyperspectral fluorescence microscopy images

Simon Henrot, Charles Soussen, Said Moussaoui, David Brie

Edge-preserving nonnegative deconvolution of hyperspectral fluorescence microscopy images

Simon Henrot, Charles Soussen, *Member, IEEE*, Saïd Moussaoui, and David Brie, *Member, IEEE*

***Abstract*—In many hyperspectral image restoration problems, some prior information is available, such as the spatial and/or spectral regularity of the solution. Additionally, a nonnegativity constraint must often be imposed to provide a physically meaningful estimate. The restored image is then obtained as the constrained minimizer of a penalized convex criterion. In this paper, we propose a fast algorithm for edge-preserving hyperspectral image restoration.**

***Index Terms*—Deconvolution, Hyperspectral imaging, Constrained convex optimization, Half-quadratic criterion**

I. Introduction

In many imaging situations, the data are blurred during the acquisition process; for instance, microscopy images are corrupted by diffraction-limited blur arising from the finite aperture of the objective. Under the common assumption of linear space-invariant blur, the direct problem models the observed image as the two-dimensional (2D) convolution of the true underlying light distribution and the microscope *point-spread function*. Solving the inverse problem of *deconvolution* leads to restored images with higher resolution and signal-to-noise ratio, a crucial processing step in fields as diverse as biomedical imaging [1], astronomy [2] or geophysics [3]. In this paper, we focus on a specific fluorescence microscopy problem, although ideas developed within are readily applicable to other applications. The problem of interest is the deconvolution of hyperspectral images, which capture a spatial scene at many overlapping spectral bands (*channels*) and thus requires an efficient restoration strategy. It is worth noting here that in addition to enhancing interpretation of the data, deconvolution of hyperspectral images has recently been shown to improve hyperspectral unmixing [4], [5].

From a physical standpoint, the intensity of each pixel is obtained by counting individual photons incident on a detector, *e.g.* a charged coupled device cell. Hence, the image is known to be nonnegative, which adds another constraint to the restoration process. In a deterministic framework, the image estimate is usually obtained as the minimizer of a penalized convex objective criterion, computed as the weighted sum of: *1)* a quadratic *data-fidelity* term, measuring the goodness of fit between the observed data and the blurred candidate; *2)* a convex *regularization* term which incorporates prior knowledge on the solution, *e.g.* spatial and spectral regularity. In [6], we proposed to use quadratic (ℓ_2) regularization functions to reconstruct spatially and spectrally smooth images, leading to a closed-form minimizer of the criterion. Denoting by N the number of pixels and L the number of spectral bands in the hyperspectral data, straightforward computation of the solution would require a costly $LN \times LN$ matrix inversion. The proposed method instead exploits the properties of 2D convolution operations to operate in the Fourier domain, requiring only $L \times L$ matrix inversions.

Unfortunately, this approach is not suitable to preserve sharp structures in the restored images, such as edges. Edge-preserving restoration may be achieved by selecting a convex, non-quadratic function instead of the ℓ_2 regularizer. Popular alternatives include both non-differentiable (*e.g.*, ℓ_1) and differentiable (*e.g.*, a piecewise $\ell_2 - \ell_1$ penalty) functions. The resulting non-quadratic criterion no longer has a closed-form minimizer, and one needs to resort to numerical iterative methods to carry out the optimization. The problem arises in many applications and minimization of such objective functions received particular attention in the last decade [7]–[9].

Many iterative schemes (*e.g.* conjuguate gradient methods) are based on line-search procedures, which involve a stepsize parameter selection problem at each iteration. In this paper, we aim instead at recovering closed-form expressions at each iteration of the algorithm, and exploiting the fast Fourier domain structure of [6] to compute these solutions. To the best of our knowledge, no hyperspectral deconvolution algorithm accounts for both spatial and spectral regularity prior information and nonnegativity in a fast manner. Thus, we choose instead to compare various implementations of the proposed approach, using state-of-the art non-quadratic minimization techniques. Among the relevant techniques in the literature, we focus on

- *majorization-minimization* strategies, with quadratic majorizing approximations also known as *Half-Quadratic* (HQ) methods in the field of image processing [10], [11];
- the *Alternative Direction Method of Multipliers* (ADMM) [12], which falls in the general framework of *proximal* methods [13].

We note that in addition to lending themselves to the Fourier domain approach, these methods have been successfully used in image restoration [14]–[18].

We mention here that this work is an extension of ideas presented in [19]. Specifically, this paper proposes two main contributions in addition to those presented in [19] : 1) the use of proximal methods to account for non-differentiable objective functions; 2) experiments on real fluorescence data. The following is organized as follows. In section II, the

This work was supported by the French Agence Nationale de la Recherche under Grant ANR-09-BLAN-0336-04. Simon Henrot is with Thales Alenia Space, 5 Allée des Gabians, 06150 Cannes - France (e-mail: simon.henrot@thalesaleniaspace.com). Charles Soussen and David Brie are with CRAN, UMR 7039, Université de Lorraine, CNRS, Faculté des Sciences, B.P. 70239, F-54506 Vandoeuvre-lès-Nancy, France (e-mail: firstname.lastname@univ-lorraine.fr). Saïd Moussaoui is with IRCCyN, UMR CNRS 6597, Ecole Centrale Nantes, France (e-mail: said.moussaoui@irccyn.ec-nantes.fr).

imaging model is introduced and the problem is formulated in the framework of constrained optimization. Section III deals with the restoration of hyperspectral data using HQ and proximal methods. Experimental results and empirical discussions on the convergence of the algorithms are presented in section IV. Finally, we conclude in section V.

II. Problem statement

A. Imaging model

Let us consider a hyperspectral image of N pixels, acquired in L spectral bands using the emission filter of a microscope. In each band ℓ, the observed image $\mathbf{y}_\ell$ is obtained from the true image $\mathbf{x}_\ell^0$ according to

$$\mathbf{y}_\ell = \mathbf{H}_\ell \mathbf{x}_\ell^0 + \mathbf{n}_\ell \tag{1}$$

where column vectors $\mathbf{y}_\ell$ and $\mathbf{x}_\ell^0$ are respectively the observed and true image after lexicographical ordering, $\mathbf{H}_\ell$ is the degradation matrix *e.g.* accounting for diffraction by the aperture of the objective in microscopy, and $\mathbf{n}_\ell$ is an additive noise term accounting for measurement and model errors, assumed to be white and independent and identically distributed (i.i.d.) Gaussian. $\mathbf{H}_\ell$ corresponds to the 2D convolution matrix for the channel point spread function (PSF), which is assumed to be experimentally measured *e.g.* using sub-resolution beads in fluorescence microscopy, or theoretically modeled [20], [21]. Model (1) actually combines various noise sources, such as the random nature of fluorescence emission (following a Poisson distribution), background radiation, dark current, calibration errors and thermal noise [22]. The use of a Gaussian noise model is a standard approximation when the signal-to-noise ratio (SNR) is high, justified by the central limit theorem [23], which is crucial to the optimization procedure discussed below. For the deconvolution of Poisson noise-corrupted images, see *e.g.* [24]–[26]. Denoting the entire observed hyperspectral image cube by $\mathbf{y} = [\mathbf{y}_1^t \dots \mathbf{y}_L^t]^t$ and the hyperspectral degradation matrix by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{H}_L \end{bmatrix}, \tag{2}$$

(since spectral bands do not interfere with each other), the hyperspectral degradation model simply reads:

$$\mathbf{y} = \mathbf{H}\mathbf{x}^0 + \mathbf{n} \tag{3}$$

The hyperspectral deconvolution problem consists in estimating the true image $\mathbf{x}^0$ given $\mathbf{y}$ and $\mathbf{H}$. This problem will be reduced to a nonnegative optimization problem since the sought image should present nonnegative pixel values.

B. Problem formulation

The nonnegative restoration problem is formulated as the constrained minimization problem:

$$\min_{\mathbf{x}} \mathcal{J}(\mathbf{x}) \text{ } s.t. \text{ } \mathbf{x} \geq \mathbf{0} \tag{4}$$

where $\mathbf{x}$ is a candidate solution. To tackle the ill-posed nature of the inverse problem [27], we propose to write criterion $\mathcal{J}(\mathbf{x})$ as the weighted sum of a data-fidelity term and two regularization terms penalizing large variations in the spatial and spectral dimensions of the image :

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \mu_s \Phi_s(\mathbf{x}) + \mu_\lambda \Phi_\lambda(\mathbf{x}) \tag{5}$$

where $\|\,.\,\|_2$ is the ℓ_2 norm, (μ_s, μ_λ) and (ϕ_s, ϕ_λ) are the regularization parameters and operators in the spatial and spectral dimensions, respectively denoted by the s and λ indices. Note that depending on the application, one may choose to regularize the solution only in a specific dimension, by setting the other regularization parameter to zero. However, setting μ_λ to 0 leads to the resolution of several equivalent small-size problems in the spatial domain. The generic form of both operators is given by $\Phi(\mathbf{x}) = \sum_{i\sim j} d_{i,j}\phi(x_i - x_j)$ where ϕ is usually called a *potential function* and $\sum_{i\sim j} d_{i,j}$ denotes a weighted summation over i and j that are in the same clique according to a spatial or spectral neighborhood system. For simplicity, we rewrite

$$\Phi(\mathbf{x}) = \sum_i \phi(\{\mathbf{D}\mathbf{x}\}_i)$$

where $\{.\}_i$ denotes the i-th entry of a vector, and $\mathbf{D}$ is the matrix for the appropriate finite-difference operator. Without loss of generality, we set $\mathbf{D}_s$ as the 2D convolution matrix of a Laplacian filter (analogous to a second-order derivative) and $\mathbf{D}_\lambda$ as the matrix corresponding to a first-order difference in the spectral dimension.

The choice of potential functions ϕ_s and ϕ_λ significantly impacts the texture of the reconstructed image : for instance, retaining a quadratic function (*Tikohonov regularization*) typically yields smooth estimates. Because hyperspectral bands are contiguous and overlapping, this prior is appropriate for ϕ_λ. However, it is not suitable for regularizing the spatial dimension when the imaged object contains sharp structures. To prevent an over-smoothing of the estimate, one may instead choose a convex, non-quadratic potential ϕ_s such as the absolute value, similarly to *total variation* approaches [8], [10]. Operator Φ_s is then simply the ℓ_1 norm of $\mathbf{D}\mathbf{x}$. Recalling the sparsity-enforcing property of the ℓ_1 norm (see *e.g.* [28]), this strategy favors sparse entries of a filtered version of $\mathbf{x}$. When the filter $\mathbf{D}$ is analogous to a first-order derivative, the resulting estimate is piecewise-constant, where the spatial edges (boundaries) between the regions of the image are preserved. Unfortunately, selecting ϕ_s as the modulus means that the objective function is no longer differentiable. An alternative is to use so-called $\ell_2 - \ell_1$ potentials such as *Huber's function* whose piecewise definition is

$$\phi_{\text{Huber}}(t;\eta) = \begin{cases} t^2 & \text{for } |t| < \eta \\ \eta(2|t| - \eta) & \text{otherwise} \end{cases} \tag{6}$$

where parameter η controls the shape of the function: it is quadratic near the origin and linear towards infinity. This relaxed approach is sometimes thought of as more suitable for natural images, and leads to 'piecewise-smooth' image estimates. In the following, we will investigate the restoration

properties of both ℓ_1 and $\ell_2 - \ell_1$ potential functions for spatial regularization.

The final objective function writes

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \frac{\mu_s}{2}\sum_i \phi_s(\{\mathbf{D}_s\mathbf{x}\}_i) + \frac{\mu_\lambda}{2}\|\mathbf{D}_\lambda\mathbf{x}\|_2^2. \tag{7}$$

III. Optimization

In the original approach of [6], Φ_s is the ℓ_2 norm; since all terms in (7) are then quadratic, the minimizer has a closed-form expression involving matrices $\mathbf{H}$, $\mathbf{D}_s$ and $\mathbf{D}_\lambda$. The efficient resolution is carried out by taking advantage of the *circulant-block-circulant* (CBC) structure of 2D convolution matrices, *e.g.* for a given wavelength ℓ matrix $\mathbf{H}_\ell$ is block-circulant, with each block being a circulant matrix.[1] A well-known property of CBC matrices is that they can be diagonalized by a 2D Discrete Fourier transform (2D-DFT). The proposed method applies 2D-DFTs to each channel of the hyperspectral data and PSF, and computations in the (spatial) Fourier domain boil down to $N^2/2$ inversions of $L \times L$ matrices.

Hence, adapting this approach to the problem at hand (non-quadratic regularization) requires somehow rewriting (7) as a quadratic function. We will first discuss handling the nonnegativity constraint, then lay out two optimization strategies to deal with the non-quadratic term $\Phi_s(\mathbf{x})$.

A. Nonnegativity constraint

Constrained optimization methods are generally classified as interior point, active set and exterior penalty methods [29], referring to the location of the current estimate w.r.t. the feasible domain. As in [6], we choose the *quadratic penalty method* which belongs to the latter class, for its low computational complexity and its quadratic nature. We introduce a set of auxiliary variables $\mathbf{p}$ and an augmented criterion

$$\mathcal{K}(\mathbf{x}, \mathbf{p}; \xi) = \mathcal{J}(\mathbf{x}) + \frac{\xi}{2}\|\mathbf{x} - \mathbf{p}\|_2^2 \tag{8}$$

in such a way that the constrained minimization of $\mathcal{J}$ is replaced by the minimization of $\mathcal{K}$ with respect to $(\mathbf{x}, \mathbf{p})$. Here, the non-negativity constraint has been transferred to $\mathbf{p}$. Minimization of $\mathcal{K}$ is carried out in an alternate fashion:

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \mathcal{K}(\mathbf{x}|\mathbf{p}^k; \xi^k) \tag{9}$$

$$\mathbf{p}^{k+1} = \arg\min_{\mathbf{p}} \mathcal{K}(\mathbf{p}|\mathbf{x}^{k+1}; \xi^k) \text{ s.t. } \mathbf{p} \geq \mathbf{0} \tag{10}$$

where $.^k$ denotes the iteration index. Parameter ξ is gradually increased to force the solution towards the constrained region. Several update schemes ensure convergence of the method, among which the simplest one is a linear rule of the form

$$\xi^{k+1} = \alpha\xi^k \tag{11}$$

[1] In the usual case of direct convolution, these matrices are actually block-Tœplitz with Tœplitz blocks; however, the CBC structure can be retained using zero-padding: see [6] for more details.

Data: Images $\mathbf{y}_\ell$, PSFs $\mathbf{H}_\ell$ for $\ell = 1 \dots L$
Result: Restored images $\mathbf{x}_\ell$ for $\ell = 1 \dots L$
Set $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{D}_s$, $\mathbf{D}_\lambda$, ϕ_s, $\mu_s > 0$, $\mu_\lambda > 0$, $\tau < 0$, $\alpha > 1$, $\epsilon > 0$;

repeat
  while $(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2 / \|\mathbf{x}^k\|_2 > \epsilon)$ **do**
    Compute $\mathbf{x}$ (9) based on [6];
    Update $\mathbf{p}$ (12);
  end
  Update ξ (11);
until $(\min(\mathbf{x}) > \tau)$;

Algorithm 1: Quadratic penalty method

TABLE I
General algorithmic framework

for some constant α greater than one. Within this algorithmic framework, $\mathbf{x}$ is guaranteed to converge towards the constrained minimum [29, Theorem 17.1, page 494]. It is worth mentioning that the quadratic penalty method can be generalized by the *augmented lagrangian technique*, which includes explicit estimates of the Lagrange multipliers in objective function (8). Although both schemes are suitable for our approach, we found that both converged within a few iterations, and the explicit computation of Lagrange multipliers did not yield substantial improvements.

Because $\mathcal{K}(\mathbf{x}, \mathbf{p}; \xi)$ is separable w.r.t. to each variable p_i, it is easy to see that (10) rewrites

$$\mathbf{p}^{k+1} = (\mathbf{x}^{k+1})_+ \tag{12}$$

where the notation $(.)_+$ means that negative values are thresholded to zero, *i.e.* $\mathbf{p}$ is simply obtained by projecting $\mathbf{x}$ onto the feasible domain. However, because of the non-quadratic term $\Phi_s(\mathbf{x})$, minimizing w.r.t. $\mathbf{x}$ is a much more computationally expensive problem. (9) thus must typically be carried out in an iterative fashion. To prevent computations based on an out-of-date version of $\mathbf{p}$ (*i.e.* an out-of-date thresholded version of $\mathbf{x}$), we choose to update $\mathbf{p}$ each time a new estimate $\mathbf{x}$ is computed. Hence, the method can be thought of as alternating between only one iteration of (9) and computing $\mathbf{p}$ according to (12).

In order to reduce the computation time in this alternating scheme in large scale problems, recent research on HQ minimization study the influence of computing an approximate minimizer of $\mathcal{K}$ w.r.t. $\mathbf{x}$ given $\mathbf{b}$ (e.g., by using truncated conjugate gradient techniques) on the convergence of the relaxation algorithm [30]. Our approach is different: here, a fast computation of the exact solution of this subproblem is performed in the Fourier domain [6]. We now turn our attention to iterative schemes for solving (9).

B. Half-quadratic method

The half-quadratic methodology for image restoration was originally introduced by Geman and Reynolds (GR) [14] and Geman and Yang (GY) [15], with respective numerical

implementations ARTUR and LEGEND [31], [32]; see also e.g. [10], [11], [33]. The key idea is to construct a quadratic majorizing approximation of $\mathcal{K}$ at the current estimate $\mathbf{x}$. The quadratic function is then minimized to get the new estimate, and the procedure is carried out alternatively. The GR and GY approaches differ in their construction of the quadratic majorizing approximation and we adopt the GY construction in the following[2]. First, assume ϕ_s is Huber's function defined by (6) and define function ψ_s through a *convex duality* relation with potential ϕ_s [10], [34]:

$$\psi_s(b) = \sup_t \left(\phi_s(t) - \frac{1}{2}(t-b)^2 \right). \tag{13}$$

Although ψ_s does not have a closed-form expression in general, its explicit computation will not be needed. By definition, $\phi_s(t)$ is upper bounded by the function $t \mapsto \frac{1}{2}(t-b)^2 + \psi_s(b)$ for a fixed scalar b. Gathering auxiliary (dual) variables in vector $\mathbf{b}$, the quantity

$$\frac{1}{2}\|\mathbf{D}_s\mathbf{x} - \mathbf{b}\|_2^2 + \sum_i \psi_s(b_i)$$

is thus a majorizing approximation of the spatial regularization term $\Phi_s(\mathbf{x}) = \sum_i \phi_s(\{\mathbf{D}_s\mathbf{x}\}_i)$. Notice that this upper bound is a quadratic function of $\mathbf{x}$, but not of $\mathbf{b}$, hence the name of the HQ approach. Substituting $\Phi_s(\mathbf{x})$ with its majorizing approximation yields an augmented criterion

$$\begin{aligned}\mathcal{Q}(\mathbf{x},\mathbf{b},\mathbf{p};\xi) = &\frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \frac{\mu_\lambda}{2}\|\mathbf{D}_\lambda\mathbf{x}\|_2^2 + \frac{\xi}{2}\|\mathbf{x}-\mathbf{p}\|_2^2 \\ &+ \frac{\mu_s}{2}\left(\frac{1}{2}\|\mathbf{D}_s\mathbf{x} - \mathbf{b}\|_2^2 + \sum_i \psi_s(b_i)\right)\end{aligned} \tag{14}$$

which depends on both sets of variables. In this framework, minimizing $\mathcal{K}(\mathbf{x};\mathbf{p},\xi)$ w.r.t. to $\mathbf{x}$ is equivalent to jointly minimizing $\mathcal{Q}(\mathbf{x},\mathbf{b};\mathbf{p},\xi)$ w.r.t. to $\mathbf{x}$ and $\mathbf{b}$, which is carried out in an alternate fashion.

1) Minimization of $\mathcal{Q}(\mathbf{x}|\mathbf{b}^k,\mathbf{p}^k;\xi^k)$: this quadratic problem has a closed-form solution

$$\begin{aligned}\mathbf{x}^{k+1} = &\left(\mathbf{H}^T\mathbf{H} + \frac{\mu_s}{2}\mathbf{D}_s^T\mathbf{D}_s + \mu_\lambda\mathbf{D}_\lambda^T\mathbf{D}_\lambda + \xi^k\mathbf{I}_{NL}\right)^{-1} \\ &\left(\mathbf{H}^T\mathbf{y} + \frac{\mu_s}{2}\mathbf{D}_s^T\mathbf{b}^k + \xi^k\mathbf{p}^k\right)\end{aligned} \tag{15}$$

which can be efficiently computed in the Fourier domain using the aforementioned approach of [6]. Note that the normal matrix to be inverted does not depend on $\mathbf{b}$ and thus its eigenvalues in the Fourier domain can be stored at each iteration of the quadratic penalty method.

2) Minimization of $\mathcal{Q}(\mathbf{b}|\mathbf{x}^{k+1},\mathbf{p}^k;\xi^k)$: the objective function is separable w.r.t. to each b_i and can be shown to possess a closed-form minimizer [10]:

$$b_i^{k+1} = \{\mathbf{D}_s\mathbf{x}^{k+1}\}_i - \phi_s'(\{\mathbf{D}_s\mathbf{x}^{k+1}\}_i) \tag{16}$$

where ϕ_s' is the derivative of potential ϕ_s. Intuitively, this step amounts to computing the next quadratic majorizing approximation of $\mathbf{K}$.

C. Proximal method

The half-quadratic approach only allows to perform ℓ_1 penalization asymptotically, by driving Huber parameter η to zero. We now propose another method that accounts for both ℓ_1 and $\ell_2 - \ell_1$ regularizations; in the following, potential ϕ_s indifferently denotes the modulus or Huber's function.

Proximal methods have been the subject of particular attention in the literature of the past decade [7], [8], [13], [17], notably to solve non-differentiable problems. In this paper, we choose the framework of the *Alternated Direction Method of Multipliers* (ADMM) algorithm [12], [35], [36], which combines dual decomposition of the objective function with the Lagrange multipliers method. The structure of the resulting technique will turn out to be close to the HQ method proposed in the previous section. The general idea consists in performing *variable splitting* of the objective function, so that the spatial regularization term becomes separable. First, let us rewrite problem (4) as

$$\min_{\mathbf{x},\mathbf{p},\mathbf{z}} f(\mathbf{x},\mathbf{p};\xi) + g(\mathbf{z}) \text{ s.t. } \mathbf{D}_s\mathbf{x} - \mathbf{z} = \mathbf{0} \tag{17}$$

where $\mathbf{p}$ is again a set of auxiliary variables introduced to deal with the nonnegativity constraint, and the decomposed objective functions f and g are given by

$$f(\mathbf{x},\mathbf{p};\xi) = \frac{1}{2}\|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2 + \frac{\mu_\lambda}{2}\|\mathbf{D}_\lambda\mathbf{x}\|_2^2 + \frac{\xi}{2}\|\mathbf{x}-\mathbf{p}\|_2^2 \tag{18}$$

$$g(\mathbf{z}) = \frac{\mu_s}{2}\sum_i \phi_s(z_i) \tag{19}$$

that is, f gathers all quadratic terms of $\mathcal{K}$ and g is the non-quadratic spatial regularization term. Note that the substitution $\mathbf{z} = \mathbf{D}_s\mathbf{x}$ takes the form of an equality constraint in (17). The augmented Lagrangian of this constrained problem writes

$$\begin{aligned}\mathcal{L}_\rho(\mathbf{x},\mathbf{p},\mathbf{z},\mathbf{u};\xi,\rho) = &f(\mathbf{x},\mathbf{p};\xi) + g(\mathbf{z}) \\ &+ \frac{\rho}{2}\|\mathbf{D}_s\mathbf{x} - \mathbf{z} + \mathbf{u}\|_2^2 - \frac{\rho}{2}\|\mathbf{u}\|_2^2\end{aligned} \tag{20}$$

where ρ is called the *barrier parameter* and $\mathbf{u}$ are the so-called *normalized Lagrange multipliers*, which are scaled by a factor of ρ so that the augmented Lagrangian only incorporates quadratic terms [12]. The procedure then consists in alternatively minimizing $\mathcal{L}_\rho(\mathbf{x},\mathbf{p},\mathbf{z},\mathbf{u};\xi,\rho)$ w.r.t. $\mathbf{x},\mathbf{p},\mathbf{z}$, and updating the normalized Lagrange multipliers $\mathbf{u}$ (*dual update*) and possibly ρ.

1) Minimization of $\mathcal{L}_\rho(\mathbf{x}|\mathbf{p}^k,\mathbf{z}^k,\mathbf{u}^k;\xi^k,\rho^k)$: this is a quadratic problem, whose solution is given by

$$\begin{aligned}\mathbf{x}^{k+1} = &(\mathbf{H}^T\mathbf{H} + \rho^k\mathbf{D}_s^T\mathbf{D}_s + \mu_\lambda\mathbf{D}_\lambda^T\mathbf{D}_\lambda + \xi^k\mathbf{I}_{NL})^{-1} \\ &(\mathbf{H}^T\mathbf{y} + \rho\mathbf{D}_s^T(\mathbf{z}^k - \mathbf{u}^k) + \xi^k\mathbf{p}^k).\end{aligned} \tag{21}$$

and is efficiently computed in the Fourier domain [6], as the similar step of (15) .

[2] In Geman and Reynolds's approach, the curvature of the quadratic majorizing function changes with each estimate $\mathbf{x}$, yielding a better approximation and a more efficient algorithm. However, it does not lend itself to Fourier domain computations

2) Minimization of $\mathcal{L}_\rho(\mathbf{z}|\mathbf{x}^{k+1},\mathbf{p}^k,\mathbf{u}^k;\xi^k,\rho^k)$: the problem is separable w.r.t. each variable z_i:

$$z_i^{k+1} = \arg\min_{z_i}\left\{\mu_s 2\phi(z_i) + \rho 2(z_i - \{\mathbf{D}_s\mathbf{x}^{k+1}\}_i - u_i^k)^2\right\}. \tag{22}$$

While this problem is nonconvex, a closed-form solution can be found by exploiting its particular structure. The *proximal operator* of a given function h with penalty ρ is defined as

$$\text{prox}_{h,\rho}(t) \triangleq \arg\min_z\{h(z) + (\rho/2)(z-t)^2\}. \tag{23}$$

For some functions h, the corresponding proximal operators have known closed-form expressions. For instance, the proximal operator of Huber's function is given by [13]

$$\text{prox}_{\phi_\eta,\rho}(t) = \begin{cases} \frac{t}{1/(\eta\rho)+1} & \text{for } |t| < \frac{1}{\rho} + \eta \\ t - \eta \text{ sgn }(t) & \text{otherwise.} \end{cases} \tag{24}$$

The proximal operator of the modulus can be obtained by driving parameter η to zero, resulting in the function $t \mapsto \mathcal{S}_{1/\rho}(t)$ where $\mathcal{S}$ denotes the well-known *soft thresholding* operator [12], defined by:

$$\mathcal{S}_a(t) \triangleq (t-a)_+ - (-t-a)_+. \tag{25}$$

Putting it all together, the closed-form solution of sub-problem (22) is thus given by

$$z_i^{k+1} = \text{prox}_{\phi_\eta, 2\rho/\mu_s}(\{\mathbf{D}_s\mathbf{x}^{k+1}\}_i + u_i^k) \tag{26}$$

and does not involve costly operations.

3) Dual update: this step corresponds to the maximization of the dual problem w.r.t. dual variables $\mathbf{u}$, and also possesses a closed-form solution [12]:

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{D}_s\mathbf{x}^{k+1} - \mathbf{z}^{k+1}. \tag{27}$$

4) Barrier parameter ρ update: finding optimal values for ρ is still an open problem, some authors advocating for a fixed value [37] while others recommending that it should be adjusted at each iteration [38]. The experimental section displays good results on confocal fluorescence images achieved with a fixed value of this parameter. Our perspectives include an assessment on the value of varying parameter ρ at each iteration. Of course, an advantage of the previously discussed HQ method is that it does not need adjusting this parameter.

D. Proposed algorithms

The proposed algorithms are detailed in table II. Stopping criteria are discussed in section IV-B.

IV. Experimental results

A. Edge preserving deconvolution of biosensor hyperspectral fluorescence images

In this section, we evaluate the performance of the proposed algorithms on real hyperspectral images acquired on a fluorescence confocal microscope[3]. The observed scene comprises *Pseudomonas* biosensors, *i.e.* a strain of bacteria which was genetically modified by introduction of two reporter genes encoding fluorescent proteins. In our setup, the genetic modifications cause the bacteria to produce *green fluorescent proteins (GFP)* when they are grown in iron-rich media and to produce *E2-orange* proteins in iron-poor (lean) media. The images are acquired on 512×512 pixels of size 49 nm and 16 channels covering the spectral range of $455 - 605$ nm.

The confocal microscope PSF is estimated using the method proposed in [21]: at each spectral band, the central lobe is approximated by a 2D gaussian function of size 8×8, while the standard deviation is computed from physical acquisition parameters (pinhole radius, refraction index of the medium and excitation and emission wavelengths). Regularization parameters are empirically fixed to $\mu_s = 15$ and $\mu_\lambda = 1$. We set additional parameters to $\eta = 10$, $\alpha = 10$, $\rho = 1$, $\tau = -1e^{-6}$, $\epsilon = 1e^{-3}$ and $\rho = 1$.

Figure 1 displays the results obtained on an image patch, at three different spectral channels : $\{495, 525, 575\}$ nm, resp. corresponding to the modes of autofluorescence, GFP and E2-Orange emission spectra. The autofluorescence process acts as a reference providing information about the 'normal' behavior of the cell. The first row corresponds to the observed data, while the second and third row display restored images, resp. obtained using $\ell_2 - \ell_1$ regularization applied by Huber's function with parameter $\eta = 10$, and ℓ_1 regularization. In the former case, both algorithms converge to the restored image, while ℓ_1 regularization can only be achieved by the proximal method. Respective computational times are 142 and 155 seconds, for a MATLAB implementation on a 2,4 Ghz Intel Core 2 Duo processor with 4Go RAM. Note how deconvolved images are significantly denoised and bacteria edges are restored for both regularizations. The denoising role of the deconvolution process is especially important for this data set which suffers from low signal-to-noise ratio. As expected, Huber's potential preserves both edges and small variations (details), while ℓ_1 penalty favors piecewise-constant images (hence better denoising). Since the spectral mode corresponding to E2-Orange is more intense than the GFP one, we can conclude that the observed region is deprived of iron.

B. Empirical remarks on the convergence of both methods

In both the HQ and proximal approaches, the global structure is of an outer loop corresponding to the quadratic penalty method, and an inner loop performing non-quadratic minimization w.r.t. $\mathbf{x}$. An interesting point, crucial for the convergence properties of both methods, is the number of iterations necessary to achieve convergence in the inner loop. Experimental results indicate that the number of inner iterations required to converge is a decreasing function of Huber's parameter η. Intuitvely, problem (9) is akin to a ℓ_2 minimization problem when η is very large and a much harder

[3]The authors would like to thank Christian Mustin, DR CNRS at LIEC (UMR 7360) for his confocal microscopy work, as well as his precious hindsights on experimental rests.

Data: Images $\mathbf{y}_\ell$, PSFs $\mathbf{H}_\ell$ for $\ell = 1 \dots L$
Result: Restored images $\mathbf{x}_\ell$ for $\ell = 1 \dots L$
Set $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{D}_s$, $\mathbf{D}_\lambda$, ϕ_s, $\mu_s > 0$, $\mu_\lambda > 0$, $\tau < 0$, $\alpha > 1$, $\epsilon > 0$;

repeat
  while $(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2 / \|\mathbf{x}^k\|_2 > \epsilon)$ **do**
    Compute $\mathbf{x}$ (15) based on [6];
    Compute $\mathbf{b}$ (16);
    Update $\mathbf{p}$ (12);
  end
  Update ξ (11);
until $(\min(\mathbf{x}) > \tau)$;

Algorithm 2: Half-quadratic method

repeat
  while $(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|_2 / \|\mathbf{x}^k\|_2 > \epsilon)$ **do**
    Compute $\mathbf{x}$ (21) based on [6];
    Compute $\mathbf{z}$ (26);
    Update $\mathbf{u}$ (27), $\mathbf{p}$ (12);
  end
  Update ξ (11);
until $(\min(\mathbf{x}) > \tau)$;

Algorithm 3: Proximal method

TABLE II
PROPOSED ALGORITHMS

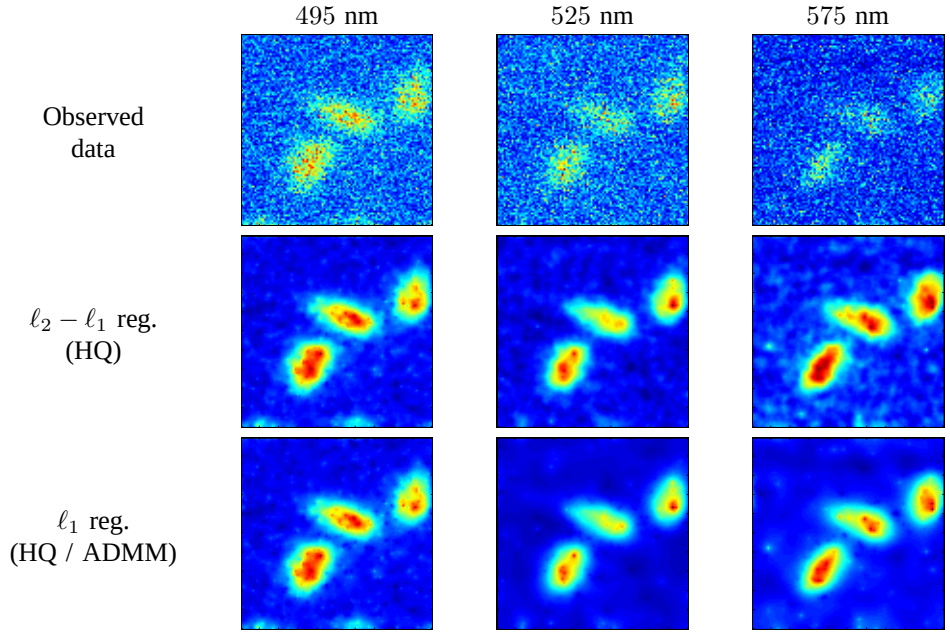
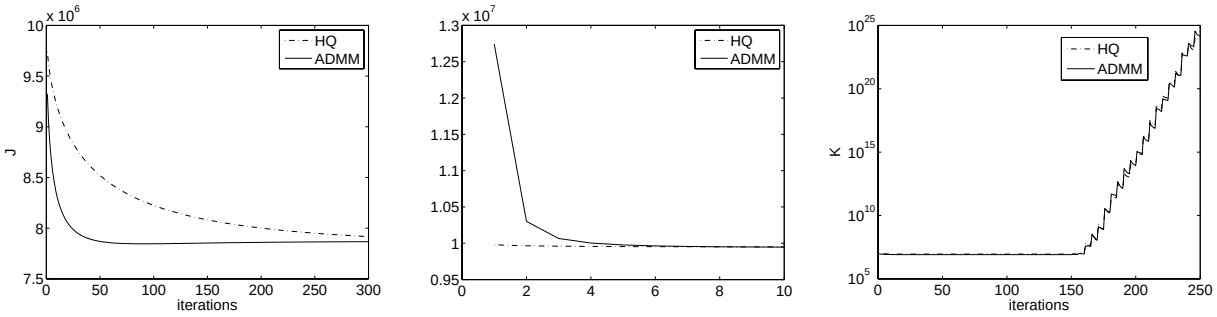


Fig. 1. Real bacterial biosensor images acquired on a confocal fluorescence microscope. First row: observed data; second row; restored data using Huber's function as the spatial potential ($\ell_2 - \ell_1$ regularization); third row: restored data using the absolute value as the spatial potential (ℓ_1 regularization). First column: spectral mode of the autofluorescence process; second column: spectral mode of the emission spectrum of the GFP (activated by the presence of iron in the medium); third column: spectral mode of the emission spectrum of E2-Orange (activated by the lack of iron in the medium).



(a) **(b)** **(c)**

Fig. 2. Criteria evolution for the half-quadratic and proximal algorithms. (a) Unconstrained case: evolution of $\mathcal{J}$ with Huber parameter $\eta = 1\bar{e}^1$. (b) Unconstrained case: evolution of $\mathcal{J}$ with Huber parameter $\eta = 10$. (c) Constrained case: evolution of $\mathcal{K}$ displayed with a log-scale.

ℓ_1 minimization problem when η is very small. Majorizing schemes with variable curvature such as GR's approach will be more appropriate in this case.These differences are illustrated in figure 2 (a) and (b) for $\eta = 1e^{-1}$ and $\eta = 10$. When considering both algorithms, the proximal approach has been more efficient than the half-quadratic one for small values of η, and both are equivalent for large values. This fact probably owes to the fixed radius of curvature inherent to the Geman and Yang's approach.

As for the outer loop, we empirically remarked that the quadratic penalty method typically only needs a few iterations to converge. From a theoretical standpoint, nonnegativity constraints are only fulfilled asymptotically. However, it is sufficient in many practical cases to impose that each pixel must only be greater than a small negative value, say 1% of the dynamic range. The stopping criterion is designed based on this observation. This point is illustrated in figure 2 (c), which explores the effects of truncating the resolution of (9) by setting a fixed number of iterations in the inner loop. The figure represents the evolution of crterion $\mathcal{K}(\mathbf{x}, \mathbf{s}; \xi)$ when the inner loops is truncated to 5 iterations. After 150 outer iterations, the value of weight ξ has considerably increased and all terms in (8) are dominated by the term $\xi\|\mathbf{x}-\mathbf{s}\|_2^2$. $\mathcal{K}(\mathbf{x}, \mathbf{s}; \xi)$ thus displays abrupt rises corresponding to an increase of ξ, followed by reductions achived by half-quadratic or proximal minimizations. By only requiring a few outer iterations, the stopping criterion at hand allows to prevent this behavior.

We also make an additional observation regarding the updating rule for weight $\xi^{k+1} = \beta\xi^k$. The value of β is linked to the computational cost of sub-problem (9) : if the minimizer can be found cheaply, one may choose an ambitious value (*e.g.* $\beta = 10$). On the other hand, if the problem is costly, a lesser value (*e.g.* $\beta = 1.5$) allow better initialization of $\mathbf{x}$ [29, p. 493]. This observation is important to account for if one decides to truncate the resolution of (9) at each iteration of the outer loop.

V. Conclusion

In this paper, we proposed two algorithms to restore blurred hyperspectral data while preserving spatial edges and accounting for spectral smoothness and pixel non-negativity. While these approaches has been derived in the context of fluorescence microscopy, the ideas within can be extended to many hyperspectral deconvolution problems. The methods are based on the fast inversion algorithm derived in [6] for smooth image restoration. We have shown that it can be adapted to non-quadratic regularizations at the cost of performing repeated fast inversions. Both algorithms present similar performances in terms of restoration quality and computational cost. Perspectives include studying the effect of adjusting the barrier parameter in the proximal / ADMM algorithm, incorporating the nonnegativity constraint in the augmented Lagrangian framework and accounting for non separable regularization terms [39].

References

[1] M. Bertero, P. Boccacci, G. Desiderà, and G. Vicidomini, "Image deblurring with Poisson data: from cells to galaxies," *Inverse Problems*, vol. 12, pp. 1–26, Dec. 2009.

[2] J.-F. Giovannelli and A. Coulais, "Positive deconvolution for superimposed extended source and point sources," *Astronomy and Astrophysics*, vol. 439, pp. 401–412, 2005.

[3] AJ Berkhout, "Least-squares inverse filtering and wavelet deconvolution," *Geophysics*, vol. 42, no. 7, pp. 1369–1383, 1977.

[4] N. Gillis and R.J. Plemmons, "Dimensionality reduction, classification, and spectral mixture analysis using non-negative underapproximation," *Optical Engineering*, vol. 50, no. 2, pp. 027001–027001, 2011.

[5] S. Henrot, C. Soussen, M. Dossot, and D. Brie, "Does deblurring improve geometrical hyperspectral unmixing?," *IEEE Transactions on Image Processing*, vol. 23, no. 3, pp. 1169–1180, March 2014.

[6] S. Henrot, C. Soussen, and D. Brie, "Fast positive deconvolution of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 828–833, Feb. 2013.

[7] P. Combettes and J.-C. Pesquet, "Proximal thresholding algorithm for minimization over orthonormal bases," *SIAM Journal on Optimization*, vol. 18, no. 4, pp. 1351–1376, 2007.

[8] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[9] M. Zibulevsky and M. Elad, "$\ell_1 - \ell_2$ optimization in signal and image processing," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 76–88, 2010.

[10] J. Idier, "Convex half-quadratic criteria and iteracting auxiliary variables for image restoration," *IEEE Trans. Image Processing*, vol. 10, no. 7, pp. 1001–1009, Jul. 2001.

[11] M. Allain, J. Idier, and Y. Goussard, "On global and local convergence of half-quadratic algorithms," *IEEE Trans. Image Process*, vol. 15, no. 5, pp. 1130–1142, May 2006.

[12] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[13] C. Chaux, P. Combettes, J.-C. Pesquet, and V. Wajs, "A variational formulation for frame-based inverse problems," *Inverse Problems*, vol. 23, no. 4, pp. 1495, 2007.

[14] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367 –383, Mar. 1992.

[15] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Trans. Image Processing*, vol. 4, no. 7, pp. 932 –946, Jul. 1995.

[16] M.A.T. Figueiredo and J.M. Bioucas-Dias, "Restoration of poissonian images using alternating direction optimization," *IEEE Trans. Image Processing*, vol. 19, no. 12, pp. 3133–3145, 2010.

[17] N. Pustelnik, *Méthodes proximales pour la résolution de problèmes inverses. Application à la Tomographie par Emission de Positrons*, thèse de doctorat, Université de Paris-Est, Marne-la-Vallée, Dec. 2010.

[18] P. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pp. 185–212. Springer, 2011.

[19] S. Henrot, S. Moussaoui, C. Soussen, and D. Brie, "Edge-preserving nonnegative hyperspectral image restoration," in *Proc. IEEE ICASSP*, May 2013.

[20] P.A. Stokseth, "Properties of a defocused optical system," *J. Opt. Soc. Amer. A.*, vol. 59, pp. 1314–1321, 1969.

[21] B. Zhang, J. Zerubia, and J.-C. Olivo-Marin, "Gaussian approximations of fluorescence microscope point-spread function models," *Applied Optics*, vol. 46, pp. 1819–1829, Mar. 2007.

[22] C. Vonesch, S. Ramani, and M. Unser, "Recursive risk estimation for non-linear image deconvolution with a wavelet-domain sparsity constraint," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE, 2008, pp. 665–668.

[23] P. Sarder and A. Nehorai, "Deconvolution methods for 3D fluorescence microscopy images," *IEEE Sig. Proc. Mag*, vol. 23, pp. 32–45, 2006.

[24] F.-X. Dupé, J. M. Fadili, and J.-L. Starck, "A proximal iteration for deconvolving Poisson noisy images using sparse representations," *IEEE Trans. Image Processing*, vol. 18, no. 2, pp. 310–321, Feb. 2009.

[25] A Jezierska, H. Talbot, C. Chaux, J. Pesquet, and G. Engler, "Poisson-gaussian noise parameter estimation in fluorescence microscopy imaging," in *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, May 2012, pp. 1663–1666.

[26] S. Lefkimmiatis and M. Unser, "Poisson image reconstruction with Hessian Schatten-norm regularization," *IEEE Trans. Image Processing*, vol. 22, no. 11, pp. 4314–4327, nov 2013.

[27] G. Demoment, "Image reconstruction and restoration: Overview of common estimation structures and problems," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 12, pp. 2024–2036, 1989.
[28] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
[29] J. Nocedal and S. Wright, *Numerical Optimization*, Springer, 2006.
[30] E. Chouzenoux, S. Moussaoui, J. Idier, et al., "Algorithme primal-dual de points intérieurs pour l'estimation pénalisée des cartes d'abondances en imagerie hyperspectrale," *Actes de GRETSI 2011*, 2011.
[31] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Two deterministic half-quadratic regularization algorithms for computed imaging," in *IEEE International Conference on Image Processing*, 1994, vol. 2, pp. 168–172.
[32] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 298–311, Feb. 1997.
[33] N. Nikolova and M. K. Ng, "Analysis of half-quadratic minimization methods for signal and image recovery," *SIAM J. Sci. Comput.*, vol. 27, no. 3, pp. 937–966, Dec. 2005.
[34] R.T. Rockafellar, *Convex analysis*, Princeton university press, 1997.
[35] A. Matakos, S. Ramani, and J. Fessler, "Accelerated edge-preserving image restoration without boundary artifacts," *IEEE Trans. Image Process*, vol. 22, no. 5, may 2013.
[36] M.S.C. Almeida and M.A.T. Figueiredo, "Deconvolving images with unknown boundaries using the alternating direction method of multipliers," *IEEE Trans. Image Process*, vol. 22, no. 8, pp. 3074–3086, 2013.
[37] M.V. Afonso, J.M. Bioucas-Dias, and M.A.T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2345–2356, 2010.
[38] E. Thiébaut, F. Soulez, and L. Denis, "Exploiting spatial sparsity for multiwavelength imaging in optical interferometry," *JOSA A*, vol. 30, no. 2, pp. 160–170, 2013.
[39] F. Soulez and E. Thiébaut, "Joint deconvolution and demosaicing," in *2009 16th IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 145–148.