Sparsity and Compressed Sensing

Deconvolution under Poisson noise

Author:
Jules Scholler

Professor:
G. Peyre

January 7, 2017

Contents

Abstract

In this report, we propose a method [5] for solving deconvolution problems when the observations are corrupted by Poisson noise. Poisson noises (also called shot noises) are encountered when counting photons in various field such as photography, medical imaging and telecommunications. As it is related to the quantum nature of photons, this noise can't be avoided during a photometric measurement. In all imaging systems the "scene" is not directly measured. Indeed, the measurement is a convolution between the scene and the impulse response (also called Point Spread Function) of the optical system. One can be interested in recovering the original scene and to do that, one need to measure the impulse response of the imaging system and perform a deconvolution. The problem is if the measurement is noisy, the "raw deconvolution" will perform poorly and the result will be far from satisfying. In order to attenuate the Poisson noise, we study in this report a post-processing method that consider the Poisson statistics of the noise for deconvolution. Towards this goal, the log-likelihood is introduced as a data-fidelity term to reflect the Poisson statistics of the noise. As a prior, the images to restore are assumed to be positive (counted photons can't be negative) and sparsely represented in a dictionary of waveforms. We will show that the deconvolution problem boils down to the minimization of non-smooth convex functionals and establish the well-posedness of each optimization problem. Proximal splitting algorithms are used to so solve this optimization problem. We present experimental results on real images[1] and compare them to the variance stabilization method and Sobolev deconvolution.

[1]Thanks to Ignacio Izeddin, world specialist in super-resolution microscopy, who shared some of his last measurements for the purpose of this project.

1 Introduction

Presentation of the problem

In optics and imaging, deconvolution is used to reverse the optical distortion that takes place in an optical microscope, electron microscope, or other imaging instrument, thus creating clearer images. When the point spread function (PSF) is unknown, it may be possible to deduce it by systematically trying different possible PSFs and assessing whether the image has improved. This procedure is called blind deconvolution. Blind deconvolution is a well-established image restoration technique in fluorescence microscopy for image restoration and this is why we decided to test our algorithm on such data. In fact, deconvolution are often crucial for exploiting images and extracting scientific content. In general, the measurement can be written:

$$y = h \circledast x + \omega \tag{1}$$

where h is the PSF (play the role of a convolutional kernel), x is the original signal, y is the measured signal and ω is the noise, considered Poissonian i.e. dominated by shot noise. Our aim is to recover x with the minimum possible error. The naive approach would be to directly perform deconvolution on (1) in the Fourier domain by dividing by the impulse response $\tilde{h}$:

$$x = \mathcal{F}^{-1}\left(\frac{\tilde{y}}{\tilde{h}} + \frac{\tilde{\omega}}{\tilde{h}}\right) \tag{2}$$

which would dramatically amplify the noise. One should therefore take into account the noise ω when deconvolving in order to stabilize the process. Without handing carefully the noise, the deconvolution results is worse than the original measurement, this phenomena is presented on Fig. 1.

As explained in the abstract, photon counting is a classic Poisson process, and the number of photons N measured by a given sensor element over a time interval t is described by the discrete probability distribution:

$$P(N = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \tag{3}$$

where $\lambda > 0$ is the expected number of photons per unit time interval. In the literature, photon noise is often modeled using a Gaussian distribution whose variance depends on the expected photon count [1]:

$$N \sim \mathcal{N}(\lambda t, \lambda t) \tag{4}$$

This approximation is very accurate if the photon flux is important (the central limit theorem ensures that the Poisson distribution approaches a Gaussian). Nonetheless, for small photon counts, the central limit theorem cannot be applied. This is typically the case in fluorescence microscopy when imaging a single molecule. These measurements are widely done since it is a super-resolution technique, awarded by a Nobel Prize in 2014 (PALM-STORM). In STochastic Optical Reconstruction Microscopy (STORM), the final resolution is limited by the noise [7]. Improving the accuracy of the images would lead to a better resolution and could lead to new discovers in bio-physics.

Previous works

In presence of Poisson noise, several deconvolution methods have been proposed such as Tikhonov-Miller inverse filter [9] and Richardson-Lucy (RL) algorithm [8]. The RL has been extensively used in many applications, but as it tends to amplify the noise after a few iterations, several extension have been proposed. In the context of deconvolution with either Gaussian or Poisson noise, sparsity-promoting regularization over orthobasis or frame dictionaries has been also proposed [3]. The authors proposed to stabilize the Poisson noise using the Anscombe variance stabilizing transform to bring the problem back to deconvolution with additive white Gaussian noise, but at the price of a non-linear data fidelity term. However, stabilization has a cost, and the performance of such an approach degrades in low intensity regimes, see section 2.4.

Contributions

In this report, based on [5], we implemented an image deconvolution algorithm for data blurred and contaminated by Poisson noise using sparsity priors. In order to form the data fidelity term, we take the exact Poisson distribution likelihood. Putting together the data fidelity and the prior terms, the deconvolution problem is formulated as the minimization of a maximum a posteriori (MAP) objective functional involving three terms: the data fidelity term; a non-negativity constraint (as Poisson data are positive by definition); and a regularization term, in the form of a non-smooth penalty that enforces the sparsity of the sought after image over a dictionary of waveforms. The well-posedness of this optimization problems is established in section 2. The problem is solved by means of proximal splitting [2] algorithms originating from the field of non-smooth convex optimization theory. More precisely a generalization of Douglas-Rachford splitting is used. In order to use this proximal splitting algorithm, the proximity operator of each individual term in the objective functional is established. This method is well suited for deconvolving under Poisson noise because it takes into account the Poissonian nature of the noise directly into the data fidelity term whereas the majority of others approaches try to transform the original image in order to transform Poisson noise into Gaussian noise.

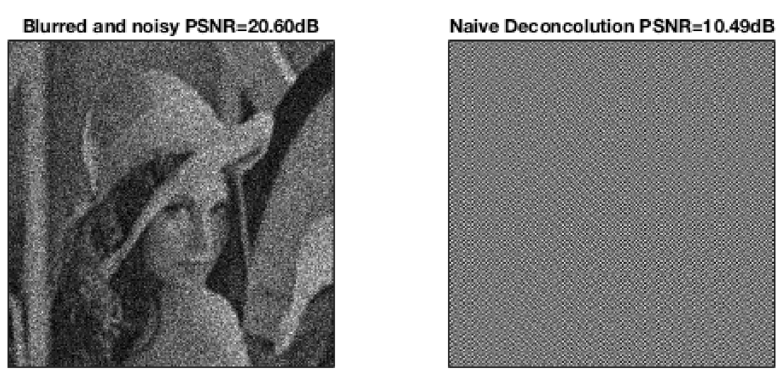

Figure 1: Naive deconvolution results

2 Deconvolution under Poisson noise

2.1 Problem statement

We consider the image acquisition process presented in the introduction where an input image of n pixels x is blurred by a PSF h and contaminated by Poisson noise. The observed image is then a discrete collection of counts $y = (y_i)_{1\leq i\leq n}$ which are bounded, i.e. $y \in \ell_\infty$. Each count y_i is a realization of an independent Poisson random variable with a mean $(h \circledast x)_i$, where $\circledast$ is the circular convolution operator:

$$y_i \sim \mathcal{P}((h \circledast x)_i) \tag{5}$$

For the sake of clarity, we denote by $\boldsymbol{H}$ a circular convolution matrix, we can then write (5) in a vector form as $y \sim \mathcal{P}(\boldsymbol{H}x)$. The deconvolution problem is to restore x from the observed count image y. The authors tackled this problem with a maximum a posteriori (MAP) bayesian framework with an appropriate likelihood function reflecting the Poisson statistics of the noise. As a prior, the image is supposed to be sparsely represented in a dictionary ϕ (we used a translation invariant wavelets tight frame) as measured by a sparsity-promoting penalty Ψ supposed to be proper, lower semi-continuous (lsc) and convex but non-smooth (we used the ℓ_1 norm). From (3), the likelihood writes:

$$p(y|x) = \prod_i \frac{((\boldsymbol{H}x)[i])^{y[i]} e^{-(\boldsymbol{H}x)[i]}}{y[i]!} \tag{6}$$

The log-likelihood function L writes:

$$L(y|x) = \sum_i (y[i]log((\boldsymbol{H}x)[i]) - (\boldsymbol{H}x)[i] - log(y[i]!)) \tag{7}$$

Taking negative log-likelihood and neglecting the constant term, we arrive at the following data fidelity term:

$$f_1 : \eta \in \mathbb{R}^n \mapsto \sum_{i=1}^{n} f_{poisson}(\eta[i]) \tag{8}$$

$$\text{if } y[i] > 0, f_{poisson}(\eta[i]) = \begin{cases} -y[i]log(\eta[i]) + \eta[i] & \text{if } \eta[i] > 0 \\ +\infty & \text{otherwise} \end{cases}$$

$$\text{if } y[i] = 0, f_{poisson}(\eta[i]) = \begin{cases} \eta[i] & \text{if } \eta[i] \in [0, +\infty) \\ +\infty & \text{otherwise} \end{cases}$$

Our aim is then to solve the following optimization problems, with a synthesis-type prior:

$$argmin_{\alpha\in\mathbb{R}^L} \quad \underbrace{f_1 \circ H \circ \phi(\alpha)}_{\text{data fidelity}} + \underbrace{\gamma\Psi(\alpha)}_{\text{sparsity penalization}} + \underbrace{\iota_c \circ \phi(\alpha)}_{\text{Positivity constraint}} \tag{9}$$

where ι_c is the indicator function of the closed convex set c (C is the positive orthant here because Poisson variables are positives), $\gamma > 0$ is a regularization parameter. In (9) we are seeking a sparse set of coefficients $\alpha_s^\star$ and the solution image is synthesized from these representation coefficients and the dictionary ϕ as $x_s^\star = \phi\alpha_s^\star$.

2.2 Well-posedeness

Using the properties of convex functions, we have the following:

i f_1 is a convex function and so are $f_1 \circ H$ and $f_1 \circ H \circ \phi$

ii f_1 is strictly convex if $\forall i, y[i] \neq 0$ and $f_1 \circ H \circ \phi$ remains stricly convex if $ker(H) = \emptyset$ and ϕ is an orthobasis.

iii (9) has at least one solution

iv (9) has a unique solution if ψ is stricly convex, or under (ii)

2.3 Minimization algorithms

2.3.1 Proximity operator of a composition with a linear operator

In order to present the proximal splitting algorithm to solve (9) we introduce the proximity operator which is a generalization of the projection onto a closed convex set:

$$prox_f(x) = argmin_{y \in c} \;\; f(y) + \frac{1}{2}||x-y||_2^2 \tag{10}$$

In our case, the difficulty is how to deal with the composition with a bounded linear operator, here the circular convolution operator and the dictionary. The proximity operator of a function f composed with an affine operator $A : \mathbb{R}^n \to \mathbb{R}^m, x \mapsto Fx - y, y \in \mathbb{R}^m$ where $F : \mathbb{R}^n \to \mathbb{R}^m$ is a bounded linear operator.

Theorem 1 A dictionary ϕ is said to be a frame with bounds c_1 and c_2, $0 < c_1 \leq c_2 < +\infty$ if $c_1||x||^2 \leq ||\phi^T x||^2 \leq c_2||x||^2$. A frame is tight when $c_1 = c_2 = c$, i.e. $\phi\phi^T = cId$.
Let F be a linear bounded operator, then $F \circ A$ is a proper lsc convex function and

i If F is a tight frame:

$$prox_{f \circ A}(x) = y + c^{-1} F^T (prox_{cf} - Id) A(x) \tag{11}$$

ii If F is a general frame, apply algorithm 1 with $\tau_t \in (0, \frac{2}{c_2})$. Then, $(u_t)_{t \in \mathbb{N}}$ converges to $\bar{u}$, and $(p_t)_{t \in \mathbb{N}}$ converges to $prox_{f \circ A}(x) = x - F^T \bar{u}$.

Algorithm 1 Forward-backward algorithm to compute $prox_{f \circ A}(x)$

Parameters: The function f, the linear bounded operator A, number of iterations N_{int} and step size $\tau_t \in (0, \frac{2}{c_2})$.
Initialization: Choose $u_0 \in \mathbb{R}^m, \; p_0 = x - F^T u_0$
for $t = 0$ **to** $N_{int} - 1$ **do**
$u_{t+1} = \tau_t (Id - prox_{f/\tau_t})(u_t/\tau_t + Ap_t)$
$p_{t+1} = x - F^T u_{t+1}$
end for
Output: The proximity operator of $f \circ A$ at $x : p_{Nint}$

Proof 1 As A is an affine operator, it preserves convexity. Thus, $f \circ A$ is convex. Moreover, A is continuous and f is lsc, and so is $f \circ A$. By the domain qualification condition, we then have $f \circ A$ a proper, lsc and convex function. By Frenchel-Rockafellar duality, we have

$$\bar{p} = prox_{prox_{f \circ A}} \iff \bar{p} = arginf_{p \in \mathbb{R}^n} \frac{1}{2}||p - x||^2 + f \circ A(p) \tag{12}$$

$$\iff \bar{u} = argmin_{u \in \mathbb{R}^m} \frac{1}{2}||x - F^T u||^2 + < u, y > + f^*(u) \tag{13}$$

and the primal solution $\bar{p}$ is recovered from the dual solution $\bar{u}$ as $\bar{p} = x - F^T \bar{u}$

i F is a tight frame with $c > 0$. Applying Moreau Identity:

$$\begin{aligned}
(Fx - y) - c\bar{u} &\in \partial f^*(\bar{u}) \\
c^{-1}(Fx - y) - \bar{u} &\in \partial(c^{-1} f^*)(\bar{u}) \\
\bar{u} &= prox_{c^{-1} f^*}(c^{-1}(Fx - y)) \\
\bar{u} &= c^{-1}(Id - prox_{cf})(Fx - y) \\
\bar{p} &= x - F^T(c^{-1}(Id - prox_{cf})A(x)
\end{aligned}$$

ii From (13), $\frac{1}{2}||x - F^T u||^2 + < ., y >$ is continuous with c_2-Lipschitz continuous gradient. Therefore, applying forward-backward splitting to (13) lead to algorithm 1. It converges provided that $0 < inf_t \ \tau_t \leq sup_t \ \tau_t < \frac{2}{||F||^2} = \frac{2}{c_2}$

2.3.2 Proximity operator of the sparsity penalty

To solve (9) we need to calculate $prox_{\gamma\Psi}$ where $\Psi = ||.||_1$ is the ℓ_1 norm. We have seen during the class that in this particular case the associated proximity operator is the soft-thresholding:

$$prox_{\gamma\Psi}(x) = max(0, 1 - \frac{\gamma}{|x|})x \tag{14}$$

2.3.3 Proximity operator of the data fidelity

Taking f_1 (Poisson anti log-likelihood) in (8) and solving (10) we have:

$$prox_{\mu f_1}(x) = \left(\frac{x[i] - \mu + \sqrt{(x[i] - \mu)^2 + 4\mu y[i]}}{2} \right)_{1 \leq i \leq n} \tag{15}$$

2.3.4 Poisson deconvolution algorithm

Now that we have the expression of the proximity operators, we can solve (9) using the splitting framework described in algorithm 2.

Algorithm 2 Image deconvolution with Poisson noise, solve (9)

Parameters: Observed image counts y, the dictionary ϕ (translation invariant orthogonal wavelets), number of main (N_{ext}) and sub (N_{int}) iterations, $\mu > 0$, the regularization parameter γ and the step θ_t.

Initialization:

$\forall i \in \{0, 1, 2\},\ p_{(0,i)} = \phi^T y$

$\alpha_0 = \phi^T y$

for $t = 0$ **to** $N_{ext} - 1$ **do**

$\xi_{(t,0)} = prox_{\mu f_1 \circ H \circ \phi/3}(p_{(t,0)})$

$\xi_{(t,1)} = prox_{\mu\gamma\Psi/3}(p_{(t,1)}) = \text{SoftThresholding}(p_{(t,1)}, \frac{\mu\gamma}{3})$

$\xi_{(t,2)} = prox_{\iota_C \circ \phi/3}(p_{(t,3)})$

$\xi_t = \frac{\xi_{(t,0)} + \xi_{(t,1)} + \xi_{(t,2)}}{3}$

for $i = 0$ **to** 2 **do**

$p_{t+1,i} = p_{t,i} + \theta_t(2\xi_t - \alpha_t - \xi_{(t,i)})$

end for

$\alpha_{t+1} = \alpha_t + \theta_t(\xi_t - \alpha_t)$

end for

Output: Deconvolded image $x^\star = \phi\alpha_{N_{Ext}}$

2.4 Numerics

We compare the presented method with the Sobolev deconvolution and Anscomb variance stabilization transform [4], inspired from the code presented on the Numerical Tours website.

Sobolev deconvolution We performed a first deconvolution using variational methods with the Sobolev prior energy. Doing that, we want to minimize over f the following problem:

$$f^\star \in argmin_f \frac{1}{2}||f_{noisy} - f||^2 + \lambda J_{Sobolev}(f) \qquad J_{Soboloev}(f) = \sum_x ||\nabla f(x)||^2 \tag{16}$$

This corresponds to penalize high frequencies ($\sum_x ||\nabla f(x)||^2 = \sum_\omega \omega^2 ||\hat{f}(\omega)||^2$) and thus should reduce the noise.

Anscomb VST A variance stabilization transform (VST) is a mapping ψ applied to a noisy image f so that the distribution of each pixel $\phi(f(x))$ is approximately Gaussian. The Anscombe VST is given by the non-linear mapping:

$$\forall i \in \{1, ..., n\},\ \phi(y[i]) = 2\sqrt{y[i] + \frac{3}{8}} \tag{17}$$

We followed the framework presented in [4], which differs from the one presented in this report in the data fidelity term (they apply Anscomb VST and consider Gaussian noise).

2.4.1 Synthetic results

In this section we artificially created blurred noisy images by applying a Gaussian filter (6x6 pixels mask) and adding Poissonian noise (images are rescaled in $[0, 30]$). In real application we can't have access to the PSF of the imaging system and it should therefore be guessed.

Figure 2: Deconvolution results on cameraman

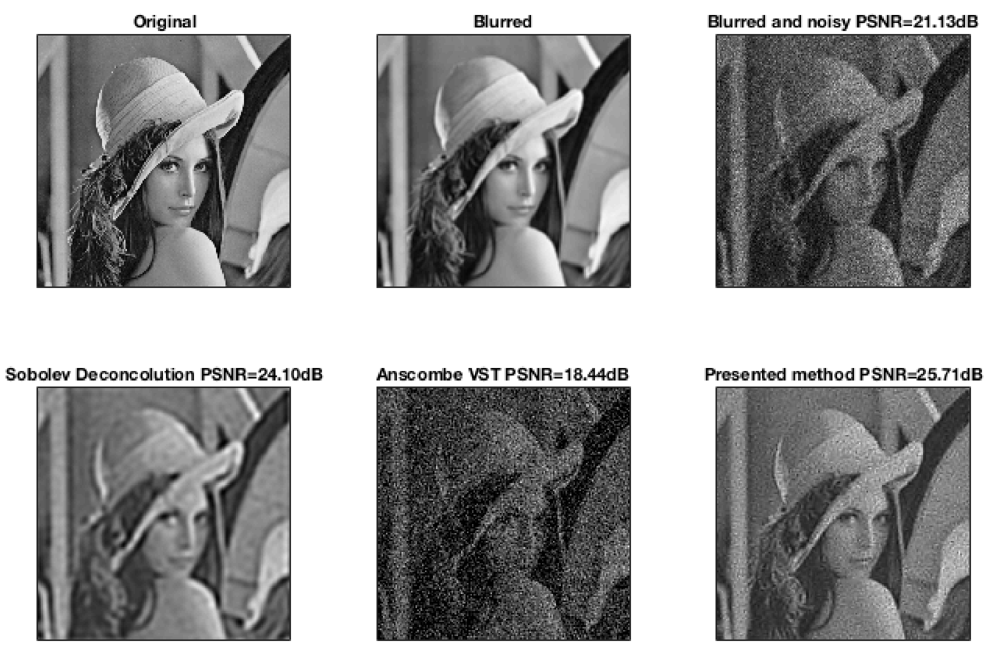

Figure 3: Deconvolution results on Lena

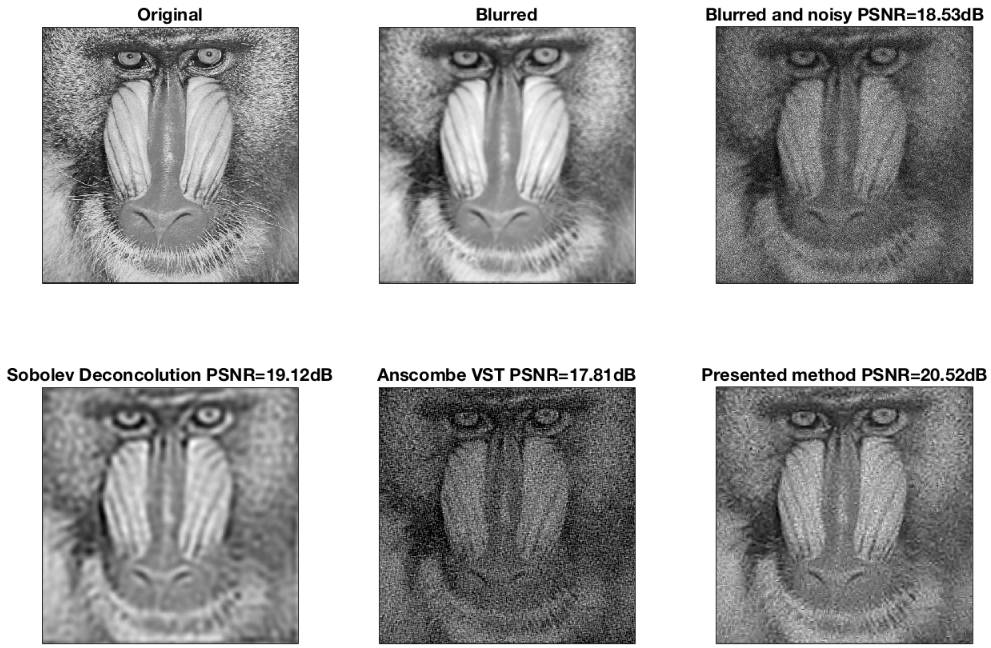

Figure 4: Deconvolution results on mandrill

We can see that the Sobolev regularization perform a denoising but also tends to blur the edges because it penalizes high frequencies. This is not satisfying. On the contrary, the proposed algorithm and the Anscomb VST succeed in deconvolving the image. Nonetheless there are artifacts on the results (more with Anscomb VST). The proposed algorithm works particularly well when there are textures in the images (Fig. 4). For the three images the results are better (PSNR and visually) with the proposed method than the others.

2.4.2 Real data results

In this section we apply the deconvolution algorithms to real data. We don't know the PSF so we needed to run the deconvolution several times to optimize the size of the PSF which is supposed to be Gaussian (see remark). The best deconvolution results were obtained by considering a PSF size of 2 pixels (which physically corresponds to 320 nm, typically the size of the airy disk in the condition of the experimental setup).

Remark Considering a Gaussian PSF makes sense because the experimental setup consisted in concave lenses and a circular aperture for the camera (EMCCD: Electron Multiplying Charge Coupled Device), whose Fourier transform is a Bessel function that is well approximated by a Gaussian.

As explained in the introduction, the noisy images of single molecules are processed to extract the center of the PSF. An example is shown of Figs. 5-6.

Figure 5: Original raw image

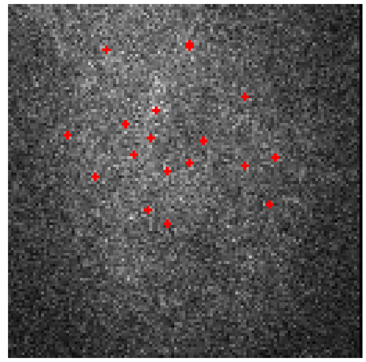

Figure 6: Marked center single molecules

Because single molecules are note very bright, the signal to noise ratio is critical in such experiments to reconstruct the microscopy image. Even if the experiment is done in complete dark with the camera sensor cooled at -80°C, one can see that the raw image is very noisy. This is typically Poisson noise and we applied the presented algorithm on these data.

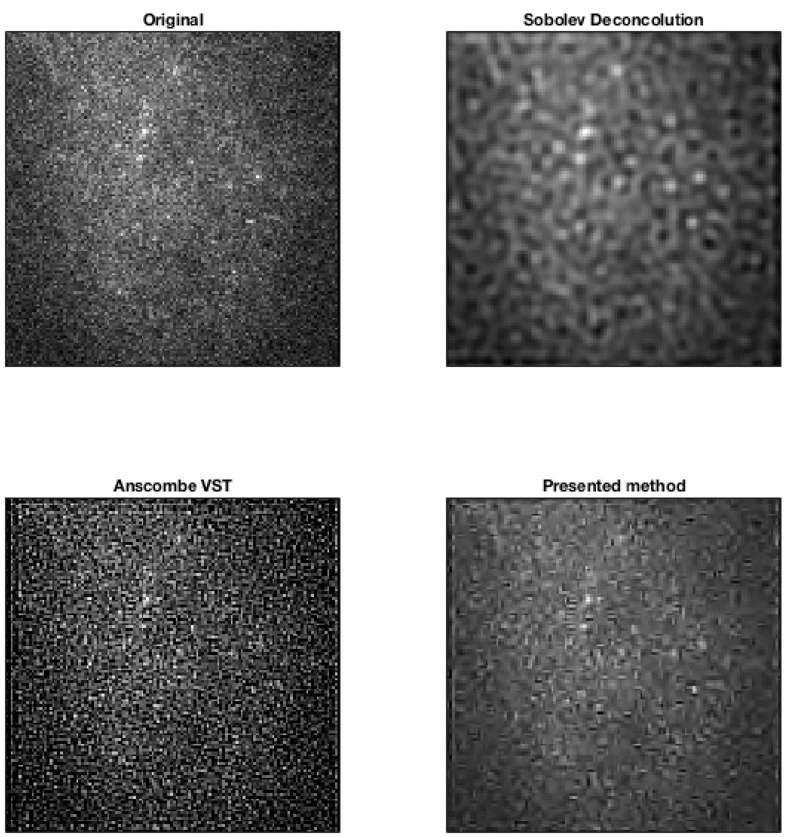

Figure 7: Deconvolution results on real data

In this case the algorithm is not very helpful for localizing single molecules. The sparsity introduces some artifacts that could be dramatic for the reconstruction because it is hard to distinguish real emitting molecules from artifacts. The presented method is better suited for classical images as shown in section 2.4.1. We tried to localize the single emitting molecules following the same procedure as the STORM specialists [6]:

1. Filter/denoise the image (in their current framework they perform a wavelet filter with β-Splines, I replaced this step with the denoising). This step is used to increase the signal to noise ratio.

2. Apply a threshold (automatically calculated in function of the image standard deviation), and keep a 8x8 images when a pixel is above the threshold (keep only local maximums).

3. Fit the PSF (Bessel function or Gaussian for more efficiency) and retrieve the center of the PSF.

The results are presented on the following figures.

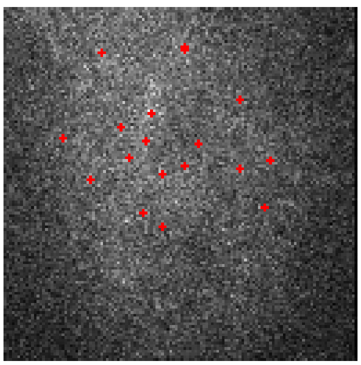

Figure 8: Ground truth for emitting molecules

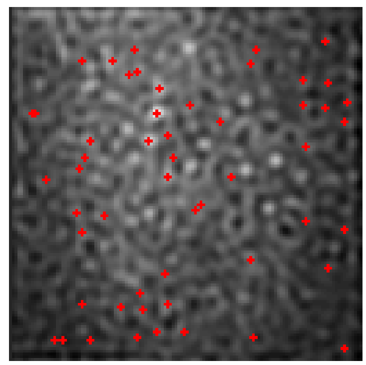

Figure 9: Sobolev deconvolution

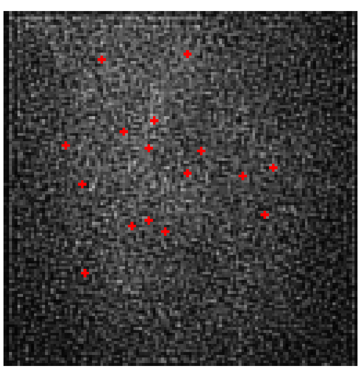

Figure 10: Anscomb VST

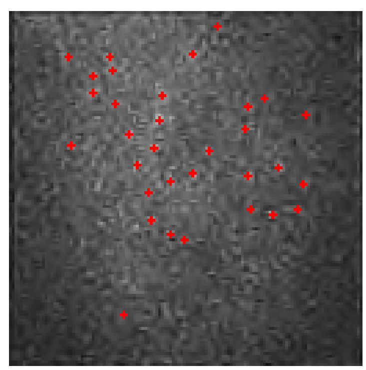

Figure 11: Proposed algorithm

We can see that for this application the best results are obtained with Anscomb VST. The Sobolev deconvolution leads to a high number of false positive.

3 Conclusion

We presented here a denoising algorithm intented to be used on images that are corrupted by Poisson noise. Indeed, the method introduced in the paper is a sparsity-based iterative thresholding deconvolution algorithm and the Poisson noise is handled properly using its associated likelihood function to construct the data fidelity term. We saw that this algorithm gives satisfactory results for traditional images. It performed better than the deconvolution using Sobolev regularization and than a method using Anscomb VST [4]. The author's approach uses a proximal algorithm which is a generalization of the Douglas-Rachford splitting applied to the primal problem, but its convergence rate is not known and is still an open problem. This means that the number of iterations before the convergence can't be calculated for computational optimization.

We tried to apply this method for denoising stochastic optical reconstruction microscopy images that consists in isolated emitting molecules. Unfortunately the results were not satisfying, i.e. were not leading to an improvement on the resolution of the reconstruction. Nonetheless, even with artifacts, the presented method was the one which gave the best denoising results.

Finally, the last limit of the presented method is the computational complexity. The denoising takes about 1 minute for an image. To be applied to microscopy image, which are taken at 200 Hz, it would require about 3 hours of computational calculation to process 1 second of measurement. A STORM acquisition takes typically more than 10 hours which would lead to 5000 days of calculation to process on an ordinary computer. These techniques are to greedy to be routinely applied to image sequences denoising.

References

[1] Sing Bing Kang-C. Lawrence Zitnick Ce Liu, Richard Szeliski and William T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 299–314.

[2] Patrick L. Combettes and Pesquet. *Proximal Splitting Methods in Signal Processing*, pages 185–212. Springer New York, 2011.

[3] Fadili M.-J. Starck-J.-L. Dupe, F.-X. A proximal iteration for deconvolving poisson noisy images using sparse representations. *IEEE Transactions on Image Processing 18 (2)*, pages 310–321.

[4] J.-L. Starck F.-X. Dupe, M. J. Fadili. A proximal iteration for deconvolving poisson noisy images using sparse representations. 2008.

[5] J.-L. Starck F.-X. Dupe, M. J. Fadili. Deconvolution under poisson noise using exact data fidelity and synthesis or analysis sparsity priors. 2011.

[6] I. Izeddin, J. Boulanger, V. Racine, C.G. Specht, A. Kechkar, D. Nair, A. Triller, D. Choquet, M. Dahan, and J.B. Sibarita. Wavelet analysis for single molecule localization microscopy. *Opt. Express*, 20(3):2081–2095, Jan 2012.

[7] Daniel R. Larson Russell E. Thompson and Watt W. Webb. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal Vol. 82*, pages 2775–2783.

[8] Nehorai A. Sarder, P. Deconvolution method for 3-d fluorescence microscopy images. *IEEE Sig. Pro. Mag. 23*, pages 32–45, 2006.

[9] Murtagh F. Starck, J.-L. Astronomical image and data analysis. *Springer*, 2006.