PROXIMAL-GRADIENT METHODS FOR POISSON IMAGE RECONSTRUCTION WITH BM3D-BASED REGULARIZATION

Willem Marais and Rebecca Willett

Department of Electrical and Computer Engineering,
Wisconsin Institute for Discovery,
Cooperative Institute for Meteorological Satellite Studies,
University of Wisconsin–Madison, Madison, WI

ABSTRACT

This paper considers estimation and inversion problems of intensity images where the observed images are corrupted by Poisson noise. Poisson noise arise in the context of counting the emission or scattering of photons. In various application domains, such as astronomy and medical imaging, the amount of photons counts are low resulting in very low signal-to-noise images. Recently, Azzari and Foi investigated using BM3D for Poisson image denoising in a course-to-fine image resolution framework. Specifically, the denoised result at a course resolution is used to improve the denoising result of the next finer resolution, resulting in state-of-the-art denoising results. This paper presents a regularized maximum likelihood formulation of the reconstruction problem, and it is explained how it can be solved using a coarse-to-fine proximal gradient optimization algorithm. The proposed methods of this paper are compared to the methods of Azzari and Foi, highlighting their strong similarities. The advantage of the proposed method of this paper is that it easily generalizes to inverse problem settings, which is demonstrated in the context of denoising a Poisson noisy image with missing pixels (i.e. image inpainting); in contrast there is no known generalization of the course-to-fine BM3D denoising method that was proposed by Azzari and Foi.

***Index Terms*—**

1. INTRODUCTION

This paper considers a Poisson intensity estimation problem in which observations are modeled as [1]

$$y \sim \text{Poisson}(z^\star),$$

where $y \in \mathbb{Z}_+^n$ is an n-dimensional vector of observed photon counts, and $z^\star \in \mathbb{R}_+^n$ is an n-dimensional vector of the ground truth image of interest (assumed to be piecewise smooth). From observations y, we are interested in recovering an estimate of $z^\star$. When the photon counts are low, this problem is referred to as *photon-limited imaging*. A related problem arises when the intensity function $z^\star$ is subjected to a linear projection operator, such as a blur, tomographic projections, missing pixels, or compressive measurements [2, 3]. In this case, we model our observations as

$$y \sim \text{Poisson}(Az^\star) \qquad A \in \mathbb{R}_+^{m \times n}$$
$$y_i \sim \text{Poisson}\left(\sum_{j=1}^{m} A_{i,j} z_j^\star\right) \qquad i = 1, \ldots, n$$

Thanks to XYZ agency for funding.

where the elements of y are independent conditioned on $z^\star$. Such problems arise frequently in astronomy, medical imaging, microscopy, atmospheric Light Detection and Ranging (lidar) and night vision.

A variety of methods have been proposed for reconstructing $z^\star$ from photon-limited measurements (*cf.* [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]). Azzari and Foi's iterative application of BM3D [14] with *variance-stabilizing transforms* (VSTs) [15] is in many ways representative of the current state of the art [16]. Despite the strong empirical performance of that approach, however, it is strictly for denoising and its application to inverse problems is unclear. Extensions to deblurring [17] rely heavily on the structure of the blur operator for determining how to best apply the VSTs, and these methods do not generalize to other inverse problem settings.

In this paper, we present a regularized maximum likelihood formulation of the reconstruction problem, and show how it can be solved using a coarse-to-fine proximal gradient optimization algorithm. This algorithm is highly reminiscent of the method in [16]. We provide a side-by-side comparison of the two methods both from an algorithmic perspective and empirically, highlighting their strong similarities. The advantage of the proposed derivation over [16] is that it easily generalizes to inverse problem settings, which we demonstrate in the context of denoising a Poisson noisy image with missing pixels.

2. BM3D AND POISSON IMAGE DENOISING

Recently, Azzari and Foi investigated using BM3D for Poisson image denoising [16]. Plugging Poisson noisy images into BM3D works very poorly. One possible approach is to use a *variance stabilizing transform* to make the Poisson noise more Gaussian-like, perform BM3D denoising, and then apply the inverse VST. Let $\mathcal{A}$ denote the Anscombe VST; then VST-based denoising can be described using Alg. 1.

Algorithm 1 Poisson VST BM3D denoising

1: $w = \mathcal{A}(y)$
2: $\hat{w} = \text{BM3D}(w)$
3: $\hat{z} = \mathcal{A}^{-1}(\hat{w})$

One challenge with this approach is that BM3D starts by breaking the noisy image into patches and then clustering those patches. If the patches are very noisy, then the clusters are uninformative. With Poisson noisy images, in practice it can often help a lot to downsample the noisy image (i.e. go from a $\sqrt{n} \times \sqrt{n}$ image to

a $\sqrt{n}/3 \times \sqrt{n}/3$ image by summing 3×3 blocks of pixels) before running VST denoising, as discussed in [4].

The recent work of [16] takes this downsampling approach a step further. Let $\mathcal{D}_h(z)$ denote downsampling the image z by a factor of h, and $\mathcal{D}_h^{-1}$ a corresponding upsampling operation (e.g. interpolation). Alg. 2 considers a decreasing series of downsampling factors h, and for each one it 1) computes a convex combination of the current estimate and the noisy data, 2) downsamples the result by a factor of h, 3) applies the Anscombe transform, 4) applies BM3D denoising, 5) applies the inverse Anscombe transform, 6) upsamples the result by a factor of h. Although this approach works quite well

Algorithm 2 Iterative Poisson VST BM3D denoising

1: Set initial estimate $\hat{f}^{(0)} = y$
2: **for** $k = 1, \ldots, K$ **do**
3: $\quad \tilde{f}^{(k)} = \lambda_k y + (1 - \lambda_k)\hat{f}^{(k-1)}$
4: $\quad z^{(k)} = \mathcal{A}[\mathcal{D}_{h_k}(\tilde{f}^{(k)})]$
5: $\quad \hat{z}^{(k)} = \text{BM3D}(z^{(k)})$
6: $\quad \hat{f}^{(k)} = \mathcal{D}_{h_k}^{-1}[\mathcal{A}^{-1}(\hat{z}^{(k)})]$
7: **end for**
8: **return** $\hat{f} = \hat{f}^{(K)}$

empirically, it has many elements that are difficult to justify, making its extension to new settings challenging.

3. ALTERNATIVE OPTIMIZATION FORMULATION

In this section, we derive an alternative to Alg. 2 based on a regularized maximum likelihood objective function and a coarse-to-fine optimization strategy with a proximal gradient algorithm at its heart.

We focus our attention on estimating the log intensity, $x^\star = \log_e z^\star$, as discussed in prior work in the context of total variation reconstruction [18, 10]. Given candidate estimate x, we can compute the likelihood of y given x as

$$p(y|x) = \prod_{i=1}^{n} \frac{e^{-e^{x_i}} e^{x_i y_i}}{y_i!}.$$

One approach to estimating $z^\star$ is to solve a regularized likelihood estimation problem; *e.g.,*

$$\begin{aligned}\hat{x} &= \operatorname*{argmin}_x -\log_e p(y|x) + \rho(x) \\ &= \operatorname*{argmin}_x \sum_{i=1}^{n} e^{x_i} - y_i x_i + \rho(x) \\ \hat{z} &= \exp(x),\end{aligned}$$

where ρ is a regularization function. If ρ is convex, then this optimization problem is also convex. One approach to the above optimization problems is via a proximal gradient method; see for example [19]. In particular, if we wish to solve

$$\min_x \ell(x) + \rho(x) \text{ where } \ell(x) = -\log_e p(y|x), \tag{1}$$

we could iteratively solve this minimization problem via

$$\begin{aligned}\tilde{x}^{(t)} &= x^{(t)} - \eta_t \nabla \ell(x^{(t)}),\ \nabla \ell(x) = e^x - y \\ x^{(t+1)} &= \text{prox}_{\eta_t \rho}(\tilde{x}^{(t)}) := \operatorname*{argmin}_x \frac{1}{2}\|x - \tilde{x}^{(t)}\|_2^2 + \eta_t \rho(x). \end{aligned} \tag{2}$$

The *proximal (prox) operator* $\text{prox}_{\eta_t \rho}$ can be interpreted as essentially performing a denoising step on the current iterate; the step size is denoted by $\eta_t > 0$. We explore replacing the prox operator with the BM3D denoising method [14], which is widely considered to be the state-of-the-art for image denoising in the presence of Gaussian noise (*i.e.,* when the data-fit term is a sum of squared errors). However, there is no closed-form expression for a function ρ such that the corresponding prox operator is BM3D denoising.

Alg. 3 is an alternative to Alg. 2 which employs a proximal gradient method to minimizes the objective function of (1) using BM3D as a proximal operator. The main idea of Alg. 3 is that for each k we run a proximal gradient algorithm where the proximal operator corresponds to downsampling the image, using BM3D, and then upsampling the result; each k corresponds to a different resolution, where small k is course resolution and large k is fine resolution. The objective function of (1) is non-convex since the implicit BM3D regularizer function ρ is non-convex; as a result, the output of a prox gradient method will depend on the initial value. Thus, we need a good initial point in order for the prox gradient method to converge to a global minimizer. Hence, the initialization of the prox gradient method is important, which is why the outer loop over k is critical. The higher the SNR of the noisy image is, the closer the noisy image itself is to a global minimum. Hence, we start of by computing the estimate of the log intensity at a course image resolution ($k = 1$) using the log intensity of the noisy image as an initial point. Then for subsequent finer image resolutions ($k > 1$) we use the previous low-resolution estimate as the initialization; this course-to-fine processing is analogous to Alg. 2. Course-to-fine processing has been previously used to reduce the processing time of a piecewise constant function estimator [20]. In this paper the aim of course-to-fine processing is to generate a sequence of good initial points for a non-convex optimization problem.

Algorithm 3 Coarse-to-Fine Poisson Denoising

1: Set initial estimate $z_i^{(0)} = \max(y_i, 10^{-6}),\ i = 1, \ldots, n$
2: **for** $k = 1, \ldots, K$ **do**
3: $\quad$ Set $t = 0$
4: $\quad x^{(t)} = \log_e z^{(k-1)}$
5: $\quad$ **repeat**
6: $\quad\quad t = t + 1$
7: $\quad\quad \tilde{x}^{(t)} = x^{(t-1)} - \eta_k \nabla \ell(x^{(t-1)})$ (gradient descent step)
8: $\quad\quad x^{(t)} = \mathcal{D}_{h_k}^{-1}\left(\text{BM3D}\left[\mathcal{D}_{h_k}\left(\tilde{x}^{(t)}\right), \eta_k \sigma_k\right]\right)$ (binned BM3D prox operation on log intensity)
9: $\quad$ **until** converged
10: $\quad z^{(k)} = \exp\left(x^{(t)}\right)$ (convert from log intensity to intensity)
11: **end for**
12: **return** $\hat{z} = z^{(K)}$

The σ_k parameter is an input to BM3D that indicates the noise standard deviation and can roughly be considered a tuning parameter in front of a corresponding regularization function ρ. The tuning parameter σ_k, for a fixed k, can be selected by finding σ_k which minimizes the difference between the true log intensity and the log intensity estimate $x^{(t)}$.

Note that the gradient descent step in line 7 of Alg. 3 amounts to

$$\tilde{x}^{(t)} = \eta_k y + (1 - \eta_k)\exp(x^{(t-1)}) + x^{(t-1)} - \exp(x^{(t-1)}).$$

Except that with Alg. 3 the log intensity (represented by $\tilde{x}^{(t)}$) is estimated, this equation is akin to the convex combination in line

3 of Alg. 2, where the parameter λ_k varies from pixel to pixel and can be based on a step size parameter η_k that we can choose using methods from optimization theory.

4. GENERALIZATION TO INVERSE PROBLEMS

In this section we derive an alternative to the deblurring method that was introduced in [17], by extending the proposed denoising algorithm Alg. 3. Similar to Alg. 3 we minimize a regularized maximum likelihood objective function using a coarse-to-fine optimization strategy with a proximal gradient algorithm. The deblurring method of [17] is an extension of Alg. 2, and it is relies heavily on the structure of the blur operator to make principled use of VTS. The inversion algorithm that we introduce in this section does not rely on the structure of the linear projection operator of an inverse problem, because it depends on the proximal gradient optimization framework rather than VSTs.

We focus our attention on estimating the log intensity $x^\star = \log_e z^\star$ from a Poisson noisy image $y \sim \text{Poisson}(Az^\star)$, where $A \in \mathbb{R}_+^{m\times n}$ represents a linear projection operator. We estimate the log intensity by solving a regularized likelihood estimate problem with a penalty function ρ. The estimation problem amounts to solving the optimization problem

$$\min_x \ell_A(x) + \rho(x), \tag{3}$$

where $\ell_A(x)$ is the Poisson negative log likelihood function which is defined by

$$\ell_A(x) = \sum_{j=1}^{m} e_j^{\mathrm{T}} A \exp(x) - y_j \log_e \left(e_j^{\mathrm{T}} A \exp(x) \right);$$

$e_j \in \mathbb{R}^m$ is a canonical vector. The gradient of the negative log likelihood function is

$$\nabla \ell_A(x) = \exp(x) \cdot \left[A^{\mathrm{T}} \left(\mathbb{1} - \frac{y}{A\exp(x)} \right) \right],$$

where $\mathbb{1} \in \mathbb{R}^m$ is a vector of ones, and $\cdot$ and $-$ denotes pointwise multiplication and division, respectively. The optimization problem (3) can be iteratively solved using a proximal gradient algorithm which relies on a prox operator $\text{prox}_{\eta_t \rho}$; see (2). As with Alg. 3, we explore replacing the prox operator with the BM3D denoising method [14].

Alg. 4 is an alternative to the deblurring method of [17], which employs a proximal gradient descent method to minimizes the objective function of (3) using BM3D as a proximal operator. As with Alg. 3, the main idea is that for each k we run a proximal gradient algorithm where the proximal operator corresponds to downsampling the image, using BM3D, and then upsampling the result. The objective function of (3) is non-convex since the implicit BM3D regularization function ρ is non-convex. Like Alg. 3, we need a good initial point in order for the prox gradient method to converge to a global minimizer. Accurate initialization is achieved by first estimating the log intensity at the coarsest image resolution ($k = 1$), and for subsequent finer resolution images ($k = 2, \ldots, K$) the prox gradient algorithm is initialized with a courser resolution estimate. The initialization for the coarsest resolution image estimate can be obtained using Tikhonov regularization.

The σ_k parameter is an input to the BM3D that indicates the noise standard deviation and can roughly be considered as a tuning parameter for the regularized function ρ.

Algorithm 4 Coarse-to-Fine Poisson Inverse

1: Use Tikhonov-regularized method to set initial estimate $z^{(0)} = \max((A^{\mathrm{T}}A + \epsilon I)^{-1}A^{\mathrm{T}}y, 10^{-6})$, e.g $\epsilon = 10^{-2}$
2: **for** $k = 1, \ldots, K$ **do**
3: Set $t = 0$
4: $x^{(t)} = \log_e z^{(k-1)}$
5: **repeat**
6: $t = t + 1$
7: $\tilde{x}^{(t)} = x^{(t-1)} - \eta_k \nabla \ell_A(x^{(t-1)})$ (gradient descent step)
8: $x^{(t)} = \mathcal{D}_{h_k}^{-1}\left(\text{BM3D}\left[\mathcal{D}_{h_k}\left(\tilde{x}^{(t)}\right), \eta_k\sigma_k\right]\right)$ (binned BM3D prox operation on log intensity)
9: **until** converged
10: $z^{(k)} = \exp\left(x^{(t)}\right)$ (convert from log intensity to intensity)
11: **end for**
12: **return** $\hat{z} = z^{(K)}$

5. EXPERIMENTS

5.1. Algorithm parameters

The current implementation of both Alg. 3 and 4 are determined by following parameters: 1) the step sizes η_k, 2) the tuning parameters σ_k, 3) the maximum bin size h_1, 4) the minimum bin size h_K, and 5) the gradient descent convergence tolerance; for Alg. 4 another parameter is the Tikhonov regularization tuning parameter ϵ. The other values of h_k are defined by $h_k = \max(h_K, h_1 - 2k + 2)$ similar to what has been done in [16]. For all our experiments the maximum h_1 and minimum h_K bin sizes were set to 5 and 1, respectively, unless it is noted otherwise. The parameters of Alg. 3 and 4 were chosen to yield strong performance for the images that were processed. We plan, in our future research, to find the optimum tuning parameters σ_k, h_1, h_K, η_k and ϵ using a Monte Carlo simulation methodology similar to what was presented in [16, §IV].

The step sizes η_k can be chosen using optimization theory. Specifically, when the proximal gradient method is used to minimize a convex objective function and if the gradient of the negative log likelihood function is Lipschitz continous, convergence is guaranteed if the step size is upper bounded by the reciprocal of the gradient Lipschitz constant [19]. The same idea can be applied to minimizing non-convex objective functions if 1) the initialization of the gradient method is near a local or global minimum, and 2) the objective function is locally convex near a local or global minimum. As it was described in sections §3 and §4, we can achieve good initialization by doing course-to-fine processing.

The gradient Lipschitz constant can be upper bounded using the Hessian matrices of the relevant negative log likelihood functions. Specifically, the gradient Lipschitz constant is upper bounded by the maximum Eigen value of the corresponds Hessian matrix which is a function of the maximum intensity of the image $z^\star$. For Alg. 3 we found that the gradient Lipschitz constant can be upper bounded by the maximum intensity of the true image $z^\star$ through 1) numerical simulations and 2) analyzing the Hessian matrices of the negative log likelihood function of the objective function (1). For Alg. 4, if the linear projection operator A is a Gaussian blur operator, we also found empirically and through analysis of the relevant Hessian matrices that the gradient Lipschitz constant can also be upper bounded by the maximum intensity of the true image $z^\star$. We leave it to future research to estimate the gradient Lipschitz upper bound from the noisy images.

5.2. Results

In this section we juxtapose the performances of the proposed denoising and deblurring algorithms (Algs. 3 and 4, respectively) against the iterBM3D [16] (Alg. 2) and iterBM3D-Deblurring [17] algorithms. The advantage of our proposed inversion algorithm (Alg. 4) is that it easily generalizes to inverse problem settings other than deblurring. We demonstrate this assertion by showing an inversion result of denoising an image with missing pixels.

In our experiments the upsampling operator $\mathcal{D}_{h_k}^{-1}$ was selected to be similar to what has been used in [16]. Specifically, our upsampling operator does spline interpolation whereas in [16] an algorithm is used that iteratively does spline interpolation with a non-negative projection operator.

Table 1 compares Alg. 3 against the iterBM3D algorithm [16] over a set of four images (Bridge, Peppers, House and USA Flag) where all of the images have a peak intensity of one. The PSNR (dB) results are averaged over 10 Poisson noisy realizations. From Table 1 we observe that the performance of Alg. 3 is comparable to that of the iterBM3D algorithm. Fig. 1 confirms that the result of Alg. 3 is comparable to the iterBM3D algorithm by juxtaposing the denoised Bridge and USA images of the two algorithms.

Table 1: Poisson denoising results expressed by PSNR (dB) values averaged over 10 Poisson noisy realizations. The result of Alg. 3 is comparable to that of iterBM3D [16]. The peak intensity is one for all the images.

Method	Bridge	Peppers	House	USA Flag
proxBM3D (Alg. 3)	19.60	19.42	22.38	19.52
iterBM3D [16]	19.85	20.51	22.94	19.55

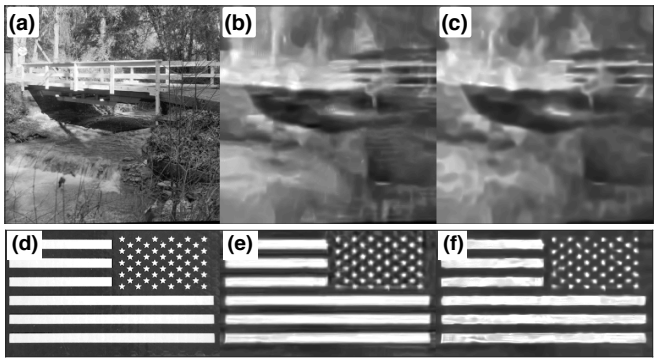

Fig. 1: Denoising of the Bridge and USA flag images at a peak intensity of one. (a) and (d) are true Bridge and USA flag images. (b) and (e) are the denoised images of iterBM3D [16], and (c) and (f) are the denoised images of proxBM3D (Alg. 3). The results of proxBM3D are comparable to that of iterBM3D.

Fig. 2 shows that the deblurring of Alg. 4 is comparable to that of the iterBM3D-Deblurring algorithm [17] when deblurring the Bridge image; the Bridge image was 1) blurred with a symmetric Gaussian function with a variance of 3 and 2) corrupted with Poisson noise. The average PSNR, over ten noise realizations, for Alg. 4 and iterBM3D-Deblurring are 22.07 dB and 22.04 dB, respectively.

Fig. 3 shows the result of denoising the Bridge image with missing pixels, where the performance of Alg. 4 is juxtaposed against the log-SPIRAL Total Variation (TV) algorithm [10]. The average PSNR, over ten noise realizations, for Alg. 4 and log-SPIRAL-TV are 20.79 dB and 19.65 dB, respectively. The vertically-striped missing pixels in Fig. 3.(b) are reminiscent of calibration problems in atmospheric lidar imaging [21], where columns of a lidar image could be missing to due to the laser frequency drift or loss of frequency stability; this anomaly was explained to us by Dr. Edwin Eloranta from the University of Wisconsin and Dr. Matthew Hayman from the National Center for Atmospheric Research.

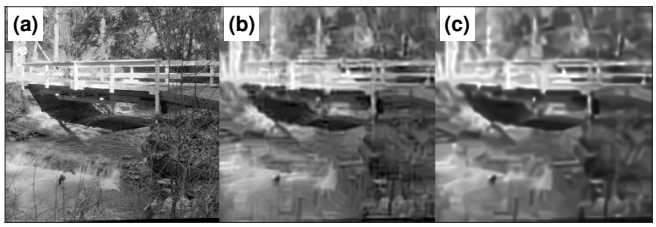

Fig. 2: Deblurring of the Bridge image at peak intensity 25.5. (a) is the true Bridge image. (b) is the deblurred image of iterBM3D-Deblurring [17] and (c) is the deblurred image of Alg. 4, where a symmetric Gaussian blur operator was used with a variance of 3. The result of proxBM3D are comparable to that of iterBM3D. The average PSNR, over ten noise realizations, for Alg. 4 and iterBM3D-Deblurring are 22.07 dB and 22.04 dB, respectively.

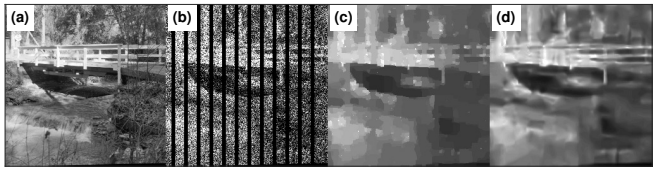

Fig. 3: Denoising of an image with missing pixels that has a maximum intensity of 5. (a) and (b) show the true image and the noisy image with missing pixels, respectively. (c) and (d) show the recovered images by the log-SPIRAL Total Variation (TV) algorithm [10] and Alg. 4, respectively. The estimate of Alg. 4 is more accurate compared to that of log-SPIRAL. Specifically, the average PSNR over ten noise realizations for Alg. 4 and log-SPIRAL are 20.79 dB and 19.65 dB, respectively. Note that there is no known generalization of iterBM3D to denoising an image with missing pixels.

6. CONCLUSION AND FUTURE WORK

We presented a regularized maximum likelihood formulation of the reconstructing a Poisson noisy image, and we demonstrated that it can be solved using a coarse-to-fine proximal gradient optimization algorithm. The algorithms we presented are highly reminiscent of the method in [16]. We provided a side-by-side comparison of our methods against the methods of [16] and [17], both from an algorithmic perspective and empirically. Experimental results show that the proposed denoising and deblurring algorithms yield comparable results when compared to that of [16] and [17]. The advantage of our proposed derivation over [17] is that it easily generalizes to inverse problem settings, which was demonstrated in the context of denoising a Poisson noisy image with missing pixels. In our future research we will study the convergence properties of the proposed algorithms, and develop a technique to select the tuning parameters of the algorithms given a noisy image.

7. REFERENCES

[1] D. L. Snyder, *Random Point Processes*, Wiley-Interscience, New York, NY, 1975.

[2] Z. T. Harmany, R. F. Marcia, and R. M. Willett, "This is SPIRAL-TAP: Sparse Poisson Intensity Reconstruction ALgorithms – Theory and Practice," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1084–1096, March 2012.

[3] X. Jiang, G. Raskutti, and R. Willett, "Minimax optimal rates for poisson inverse problems with physical constraints," To appear in *IEEE Transactions on Information Theory*, arXiv:1403:6532, 2014.

[4] J. Salmon, Z. T. Harmany, C.-A. Deledalle, and R. M. Willett, "Poisson noise reduction with non-local PCA," *Journal of Mathematical Imaging and Vision*, pp. 1–16, 2013.

[5] M. Mäkitalo and A. Foi, "Optimal inversion of the Anscombe transformation in low-count Poisson image denoising," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 99–109, 2011.

[6] M. Lebrun, M. Colom, A. Buades, and J-M. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, no. 1, pp. 475–576, 2012.

[7] C-A. Deledalle, L. Denis, and F. Tupin, "Poisson NL means: Unsupervised non local means for Poisson noise," in *ICIP*, 2010, pp. 801–804.

[8] C. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2866–2878, 2006.

[9] R. M. Willett, Z. T. Harmany, and R. F. Marcia, "Poisson image reconstruction with total variation regularization," in *Proc. IEEE International Conference on Image Processing*, 2010.

[10] A. K. Oh, Z. T. Harmany, and R. M. Willett, "Logarithmic total variation regularization for cross-validation in photon-limited imaging," in *Proc. IEEE International Conference on Image Processing*, 2013.

[11] H. Erdŏgan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Transactions on Medical Imaging*, vol. 18, no. 9, pp. 801–814, September 1999.

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

[13] R. Willett and R. D. Nowak, "Platelets: A multiscale approach for recovering edges and surfaces in photon-limited medical imaging," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 332–350, 2003.

[14] K. Dabov, A. Foi, V. Katkovnik, and K. O. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS'09)*, 2009.

[15] F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, pp. 246–254, 1948.

[16] Lucio Azzari and Alessandro Foi, "Variance stabilization for noisy+ estimate combination in iterative poisson denoising," in *IEEE Signal Processing Letters*. 2016, pp. 1086–1090, IEEE.

[17] Lucio Azzari and Alessandro Foi, "Variance stabilization in poisson image deblurring," in *Proc. 2017 IEEE Int. Sym. Biomedical Imaging (ISBI), Melbourne, Australia*, 2017.

[18] A. K. Oh, Z. T. Harmany, and R. M. Willett, "To e or not to e in Poisson image reconstruction," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, 2014, "Top 10% Paper".

[19] Amir Beck and Marc Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[20] Rui Castro, Rebecca Willett, and Robert Nowak, "Coarse-to-fine manifold learning," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on*. IEEE, 2004, vol. 3, pp. iii–992.

[21] Claus Weitkamp, *Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere*, vol. 102, Springer Science & Business, 2006.