

ANALYSIS OF p -NORM REGULARIZED SUBPROBLEM MINIMIZATION FOR SPARSE PHOTON-LIMITED IMAGE RECOVERY

Aramayis Orkusyan[†], Lasith Adhikari^{*}, Joanna Valenzuela^{*}, and Roummel F. Marcia^{*}

[†]Department of Mathematics, California State University, Fresno, Fresno, CA 93740, USA

^{*}Applied Mathematics, University of California, Merced, Merced, CA 95343, USA

ABSTRACT

Critical to accurate reconstruction of sparse signals from low-dimensional low-photon count observations is the solution of nonlinear optimization problems that promote sparse solutions. In this paper, we explore recovering high-resolution sparse signals from low-resolution measurements corrupted by Poisson noise using a gradient-based optimization approach with non-convex regularization. In particular, we analyze zero-finding methods for solving the p -norm regularized minimization subproblems arising from a sequential quadratic approach. Numerical results from fluorescence molecular tomography are presented.

Index Terms— Photon-limited imaging, non-convex optimization, sparse reconstruction, ℓ_p -norm, fluorescence molecular tomography

1. INTRODUCTION

Photon-limited data observations generally follow a Poisson distribution with a certain mean detector photon intensity, i.e.,

$$\mathbf{y} \sim \text{Poisson}(\mathbf{A}\mathbf{f}^*),$$

where $\mathbf{y} \in \mathbb{Z}_+^m$ is a vector of observed photon counts, $\mathbf{f}^* \in \mathbb{R}_+^n$ is the vector of true signal intensity, and $\mathbf{A} \in \mathbb{R}_+^{m \times n}$ is the system matrix that linearly projects the true signal to the detector photon intensity [1].

The Poisson reconstruction problem has the following constrained optimization form:

$$\begin{aligned} & \underset{\mathbf{f} \in \mathbb{R}^n}{\text{minimize}} && \Phi(\mathbf{f}) \equiv F(\mathbf{f}) + \tau \text{pen}(\mathbf{f}) \\ & \text{subject to} && \mathbf{f} \succeq 0, \end{aligned} \quad (1)$$

where $F(\mathbf{f})$ is the negative Poisson log-likelihood function $F(\mathbf{f}) = \mathbf{1}^T \mathbf{A}\mathbf{f} - \sum_{i=1}^m y_i \log(\mathbf{e}_i^T \mathbf{A}\mathbf{f} + \beta)$, where $\mathbf{1}$ is the m -vector of ones, \mathbf{e}_i is the i -th column of the $m \times m$ identity matrix, $\beta > 0$ (typically $\beta \ll 1$), $\text{pen} : \mathbb{R}^n \rightarrow \mathbb{R}$ is a sparsity-promoting penalty functional, and $\tau > 0$ is a regularization parameter.

This work was supported by NSF Grants CMMI 1333326 and DMS-1359484.

Various convex penalty techniques have previously been used as regularization terms in (1). For example, when the solution is sparse in the canonical basis, an ℓ_1 norm is simple to implement [2, 3]. A total variation seminorm with a split Bregman approach can also be used [4, 5]. Other related methods include [6, 7, 8, 9]. In this work, we consider the non-convex penalty function $\text{pen}(\mathbf{f}) = \|\mathbf{f}\|_p^p = \sum_{i=1}^n |\mathbf{f}_i|^p$ ($0 \leq p < 1$) in (1) as a bridge between the convex ℓ_1 norm and the ℓ_0 counting seminorm [10, 11, 12]. The solution to this non-convex problem can be found by minimizing a sequence of quadratic models to the function $F(\mathbf{f})$ approximated by second-order Taylor series expansion where the Hessian replaced by a scaled identity matrix $\alpha_k \mathbf{I}$ with $\alpha_k > 0$ [13, 14, 15]. Simplifying the second-order approximation yields a sequence of subproblems of the form

$$\begin{aligned} \mathbf{f}^{k+1} &= \underset{\mathbf{f} \in \mathbb{R}^n}{\arg \min} && \frac{1}{2} \|\mathbf{f} - \mathbf{s}^k\|_2^2 + \frac{\tau}{\alpha_k} \|\mathbf{f}\|_p^p \\ & \text{subject to} && \mathbf{f} \succeq 0, \end{aligned} \quad (2)$$

where $\mathbf{s}^k = \mathbf{f}^k - \frac{1}{\alpha_k} \nabla F(\mathbf{f}^k)$. Note that the subproblem (2) can be separated into scalar minimization problems of the form

$$\begin{aligned} f_s^* &= \underset{f \in \mathbb{R}}{\arg \min} && \Omega_s(f) = \frac{1}{2}(f - s)^2 + \lambda|f|^p, \\ & \text{subject to} && f \geq 0. \end{aligned} \quad (3)$$

where f and s denote elements of the vectors \mathbf{f} and \mathbf{s}^k respectively and $\lambda = \tau/\alpha_k$ [16].

Given a regularization parameter $\lambda > 0$ and p -norm for $\Omega_s(f)$ in (3), there exists a threshold value $\gamma_p(\lambda)$ (that explicitly depends on p and λ) such that if $s \leq \gamma_p(\lambda)$, the global minimum of (3) is $f_s^* = 0$; otherwise, the global minimum will be a non-zero value (see Fig. 1). When $s = \gamma_p(\lambda)$, there exists f_γ^* such that

$$\Omega_\gamma(f_\gamma^*) = \Omega_\gamma(0) \quad \text{and} \quad \Omega'_\gamma(f_\gamma^*) = 0. \quad (4)$$

By solving (4) simultaneously, we can explicitly find the threshold value $\gamma_p(\lambda)$ for given p and λ values. Specifically, $\gamma_p(\lambda)$ is given by $\gamma_p(\lambda) = (2\lambda(1-p))^{\frac{1}{2-p}} + \lambda p(2\lambda(1-$

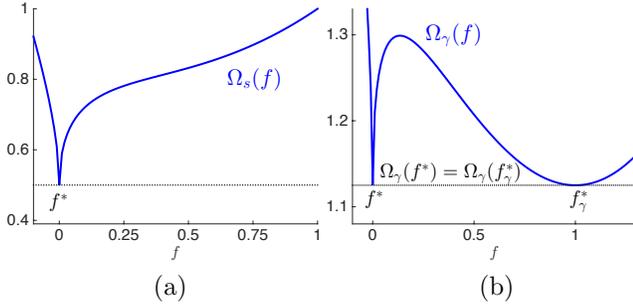


Fig. 1. The plot of the scalar quadratic function $\Omega_s(f)$, where $p = 0.5$ and $\lambda = 1.0$. (a) When s is less than the specific threshold value $\gamma_p(\lambda)$, then $f_s^* = 0$ is the unique global minimum. (b) When $s = \gamma_p(\lambda)$, there are global minima at $f^* = 0$ and f_γ^* . If $s > \gamma_p(\lambda)$, then the global minimum is uniquely at some $f_s^* > 0$.

p) ^{$\frac{p-1}{2-p}$} (see [17] for details). For any $s > \gamma_p(\lambda)$, the unique minimum f_s^* of $\Omega_s(f)$ is greater than 0 and is obtained by setting Ω'_s to 0:

$$\Omega'_s(f_s^*) = f_s^* - s + \lambda p (f_s^*)^{p-1} = 0. \quad (5)$$

We now describe zero-finding algorithms to compute the root f_s^* . To our knowledge, this work is the first careful analysis of these minimization techniques for solving the non-convex quadratic subproblem given by (3).

2. ZERO-FINDING METHODS

2.1. Fixed-Point Iteration Method

A point f^* is said to be a fixed point of a function $G(f)$ if $G(f^*) = f^*$. Setting $\Omega'_s(f)$ equal to zero, we have $s - \lambda p (f^*)^{p-1} = f^*$. The fixed-point iteration method is an iterative method for finding fixed points of a function. In particular, it defines a sequence of points $\{f_n\}$ given by $f_{n+1} = G(f_n)$. Previous methods for finding the root of $\Omega'_s(f)$ use the fixed point iteration (see, e.g., [16, 17]):

$$f_{n+1} = g(f_n) = s - \lambda p f_n^{p-1}. \quad (6)$$

2.2. Newton's Method

Note that there are various ways of defining fixed point iterations. One particular fixed-point formulation is Newton's method, which is given by the iterations

$$f_{n+1} = G(f_n) = f_n - \frac{\Omega'_s(f_n)}{\Omega''_s(f_n)}.$$

In our case, the iterations for Newton's method are given by

$$f_{n+1} = f_n - \frac{f_n - s + \lambda p f_n^{p-1}}{1 + \lambda p (p-1) f_n^{p-2}} = \frac{s + \lambda p (p-2) f_n^{p-1}}{1 + \lambda p (p-1) f_n^{p-2}}.$$

In order to simplify the computation of this iteration and avoid computing two different roots f_n^{p-1} and f_n^{p-2} , we multiply the numerator and denominator by f_n^{2-p} :

$$f_{n+1} = \frac{s f_n^{2-p} + \lambda p (p-2) f_n}{f_n^{2-p} + \lambda p (p-1)}. \quad (7)$$

The performance of fixed-point iteration and Newton's method very much depend on the choice of the initial point f_0 , which we discuss next.

2.3. Initialization

When $s = \gamma_p(\lambda)$, the solution f_γ^* such that $\Omega'(f_\gamma^*) = f_\gamma^* - \gamma_p(\lambda) + \lambda p (f_\gamma^*)^{p-1} = 0$ is given explicitly by

$$f_\gamma^* = (2\lambda(1-p))^{\frac{1}{2-p}}.$$

Then if $s = \gamma_p(\lambda) + \varepsilon$ for some $\varepsilon > 0$, we now analyze how to estimate f_s^* to initialize the zero-finding methods described previously.

First-order Taylor series approximation. To define the initial point, we can linearize $\Omega'_s(f)$ around f_γ^* and find the zero of the linearization. More specifically,

$$\begin{aligned} \Omega'_s(f_\gamma^* + \delta) &\approx \Omega'_s(f_\gamma^*) + \delta \Omega''_s(f_\gamma^*) \\ &= f_\gamma^* - (\gamma_p(\lambda) + \varepsilon) + \lambda p (f_\gamma^*)^{p-1} \\ &\quad + \delta(1 + \lambda p(p-1)(f_\gamma^*)^{p-2}) \\ &= -\varepsilon + \delta(1 + \lambda p(p-1)(f_\gamma^*)^{p-2}). \end{aligned}$$

Setting this equal to zero and solving for δ suggests the use of the initialization

$$f_s^0 = f_\gamma^* + \delta, \quad \text{where } \delta = \frac{\varepsilon}{1 + \lambda p(p-1)(f_\gamma^*)^{p-2}}.$$

Second-order Taylor series approximation. Similarly, we can use a second-order Taylor approximation to Ω'_s around f_γ^* :

$$\Omega'_s(f_\gamma^* + \delta) \approx \Omega'_s(f_\gamma^*) + \delta \Omega''_s(f_\gamma^*) + \frac{\delta^2}{2} \Omega'''_s(f_\gamma^*),$$

which yields the following approximation:

$$f_s^0 = f_\gamma^* + \delta, \quad \text{where } \delta = \frac{-b + \sqrt{b^2 - 4ac}}{2a},$$

where $a = \frac{\lambda p(p-1)(p-2)}{2} (f_\gamma^*)^{p-3}$, $b = 1 + \lambda p(p-1)(f_\gamma^*)^{p-2}$, and $c = -\varepsilon$.

The linearization and second-order Taylor approximation, however, diverge quickly from the true solution as f_s^* becomes large (see Fig. 2). We now discuss bounds on f_s^* that allow us to make more effective initial approximations to f_s^* . We first prove a lemma, which will be useful in showing bounds on f_s^* as well as other results.

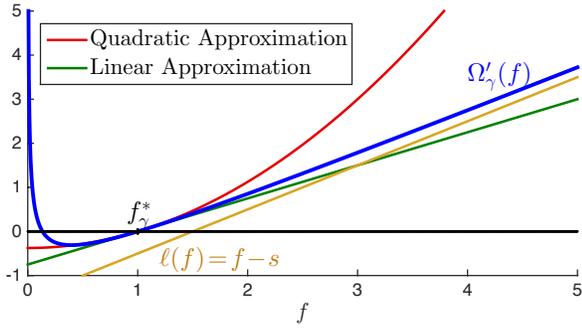


Fig. 2. Approximations to $\Omega'_\gamma(f)$ centered at f^* . As f increases, both the linear and quadratic Taylor approximation diverge from $\Omega'_\gamma(f)$. In contrast, the approximation $\ell(f) = f - s$, which are the first two terms in $\Omega'_\gamma(f)$, is more accurate for large values of f .

Lemma 1. *Let $\lambda > 0$ and $0 \leq p < 1$. Then for $s \geq \gamma_p(\lambda)$, $\lambda p(1-p)(f_s^*)^{p-2} \leq \frac{p}{2}$.*

Proof. Recall that for $s = \gamma_p(\lambda)$, there exists an $f_\gamma^* > 0$ such that (4) hold. From $\Omega_\gamma(f_\gamma^*) = \Omega_\gamma(0)$, we can obtain

$$\frac{1}{2}f_\gamma^* + \lambda(f_\gamma^*)^{p-1} = \gamma_p(\lambda), \quad (8)$$

and from $\Omega'_\gamma(f_\gamma^*) = 0$, we have

$$f_\gamma^* + \lambda p(f_\gamma^*)^{p-1} = \gamma_p(\lambda). \quad (9)$$

Setting (8) equal to (9) and with some algebraic manipulation, we have $\lambda p(1-p)(f_\gamma^*)^{p-2} = \frac{p}{2}$. For $s > \gamma_p(\lambda)$, the unique minimizer $f_s^* > f_\gamma^*$. Thus,

$$\lambda p(1-p)(f_s^*)^{p-2} < \lambda p(1-p)(f_\gamma^*)^{p-2} = \frac{p}{2},$$

which completes the proof. \square

This result allows us to prove the following theorem, which bounds the minimizer, f_s^* , of $\Omega_s(f)$:

Theorem 1. *For $\lambda > 0$ and $0 \leq p < 1$, the minimizer, f_s^* , of Ω_s is bounded by $f_s^* \leq s$. If $0 \leq p \leq \frac{1}{2}$, then the minimizer is further bounded by $\frac{2}{3}s \leq f_s^* \leq s$.*

Proof. Recall that the minimizer of Ω_s solves $\Omega'_s(f_s^*) = 0$. Solving for f_s^* , we have

$$f_s^* = \frac{s}{1 + \lambda p(f_s^*)^{p-2}}. \quad (10)$$

Rewriting the main result of Lemma 1, we obtain $\lambda p(f_s^*)^{p-2} \leq \frac{p}{2(1-p)}$. Observe that if $p \leq \frac{1}{2}$,

$$\lambda p(f_s^*)^{p-2} \leq \frac{p}{2(1-p)} \leq \frac{1}{2} \quad \text{and} \quad 1 \leq 1 + \lambda p(f_s^*)^{p-2} \leq \frac{3}{2}.$$

Using these bounds in (10) yields the desired results. \square

Note that Theorem 1 implies that as s increases, so does f_s^* . Moreover, as $s \rightarrow \infty$, $(f_s^*)^{p-2} \rightarrow 0$, and therefore, by (5), $f_s^* \rightarrow s$. Thus, a sensible initial estimate for f_s^* is s .

Fixed-point initialized Newton's Method. We can improve the initial guess from s by finding a point between f_s^* and s . The mean-value theorem guarantees the existence of $\xi \in (f_s^*, s)$ such that $\Omega'_s(\xi) = \frac{\Omega'_s(s) - \Omega'_s(f_s^*)}{s - f_s^*}$. Rearranging, we find that

$$f_s^* = s - \frac{\Omega'_s(s) - \Omega'_s(f_s^*)}{\Omega'_s(\xi)} = s - \frac{\lambda p s^{p-1}}{1 - \lambda p(1-p)\xi^{p-2}}.$$

By Lemma 1, $\frac{2-p}{2} \leq 1 - \lambda p(1-p)\xi^{p-2} \leq 1$, and thus,

$$f_s^* \approx s - \lambda p s^{p-1} \in (f_s^*, s).$$

We note that this is precisely the first fixed point iteration initialized at s .

2.4. Convergence

Guarantee of convergence. Let $e_n = f_n - f^*$ and $e_{n+1} = f_{n+1} - f^*$ represent the errors on the n -th and $n+1$ -th iterations respectively. For fixed point iteration, we have

$$\begin{aligned} e_{n+1} = f_{n+1} - f^* &= G(f_n) - f^* \\ &= G(f^* + e_n) - f^*, \\ &= G(f^*) + e_n G'(f^*) + e_n^2 G''(\xi) - f^* \\ &= f^* + e_n G'(f^*) + e_n^2 G''(\xi) - f^* \\ &= e_n G'(f^*) + e_n^2 G''(\xi). \end{aligned}$$

For small e_n , $e_{n+1} \approx e_n G'(f^*)$. In our context,

$$G(f) = s - \lambda p f^{p-1} \quad \text{and} \quad G'(f) = \lambda p(1-p)f^{p-2}.$$

By Lemma 1, $G'(f) < 1$. Therefore, the error is decreasing and the fixed point iteration method is guaranteed to converge.

To show Newton's method is guaranteed to converge, let f_c be a critical point of $\Omega'_s(f)$ i.e. $\Omega''_s(f_c) = 0$. In particular, $f_c = (\lambda p(1-p))^{\frac{1}{2-p}}$ and for any $f > f_c$, $\Omega''_s(f) = 1 + \lambda p(p-1)f^{p-2} > 0$ i.e. $\Omega'_s(f)$ is increasing in the interval (f_c, ∞) . Then, $\Omega'''_s(f) = \lambda p(p-1)(p-2)f^{p-3} > 0$ for all $f \in (0, \infty)$, which implies $\Omega'_s(f)$ is convex. Finally, we note that $f_c < (2\lambda p(p-1))^{\frac{1}{2-p}} = f_\gamma^* \leq f^*$, i.e. $\Omega'_s(f)$ has a root in (f_c, ∞) . Therefore, $\Omega'_s(f)$ is increasing, convex, and has a zero in (f_c, ∞) , and Newton's method is guaranteed to converge from any starting point in the interval (f_c, ∞) (see Theorem 2 pg. 86 in [18]).

Rate of convergence. Let ε be some set tolerance such that on the n -th iteration if $|e_n| = |f_n - f_s^*| \leq \varepsilon$ then we

will consider the algorithm to have converged to the root. For fixed point iteration, we have convergence when

$$\varepsilon \geq |e_n| = C_1 |e_{n-1}| = C_1^n |e_0| \quad (11)$$

where $C_1 = G'(f_s^*) = \lambda p(1-p)(f_s^*)^{p-2}$. Solving for n , the number of iterations required to converge, we have

$$n_{\text{Fixed Point}} \geq \frac{\ln \varepsilon - \ln |e_0|}{\ln C_1}. \quad (12)$$

For Newton's method, we have convergence when

$$\varepsilon \geq |e_n| = C_2 |e_{n-1}|^2 = C_2^{2^n - 1} |e_0|^{2^n} \quad (13)$$

where $C_2 = \frac{1}{2} \frac{\lambda p(1-p)(2-p)(f_s^*)^{p-3}}{1 - \lambda p(1-p)(f_s^*)^{p-2}}$. Solving for n in (13) yields

$$n_{\text{Newton}} \geq \frac{1}{\ln 2} \ln \left(\frac{\ln C_2 + \ln \varepsilon}{\ln C_2 + \ln e_0} \right). \quad (14)$$

Fig. 3 shows the theoretical number of iterations for fixed-point iterations and Newton's method to converge. Note that when s is near $\gamma_p(\lambda)$, fixed-point iterations take many more iterations than Newton's method. However, for large s , fixed-point iterations only require four iterations. Although this is still twice as many as the iterations for Newton's method, the number of floating point operations for fixed-point iterations is much smaller than that for Newton's method (compare (6) and (7)). Since s can take on any real value, we expect the average performance of fixed-point iteration and Newton's method will be comparable, which we see in the next section.

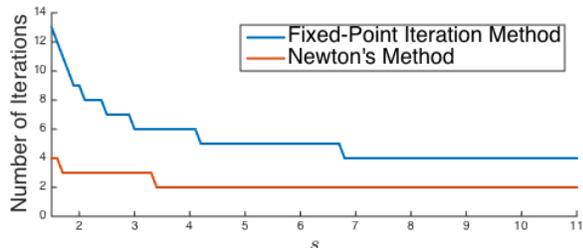


Fig. 3. Theoretical number of iterations required to converge as a function of s . Here $p = 0.5$, $\lambda = 1$, $\varepsilon = 10^{-8}$, $e_0 = s - f^*$, and $\gamma_p(\lambda) \leq s \leq 11$.

3. NUMERICAL EXPERIMENTS

We simulated a 3D cubic phantom with two embedded fluorescence capillary rod targets (see e.g., [19, 20]). For the finite element mesh, there are a total of 8,690 nodes inside the 3D cube while only 36 nodes are located inside the two rods. The fluorophore concentration of the nodes

is set to 7,000 inside the two rods and 0 outside. We chose a total of 20 excitation source positions and 1,057 detector positions on the top surface of the cube, which gives us $20 \times 1,057 = 21,140$ measurements. About one-tenth of all the measurements were used (i.e. 2,120 measurements). We assumed that the excitation wavelength is 650 nm and the emission wavelength is 720 nm in the construction of the system matrix \mathbf{A} . The tissue optical properties were $\mu_a = 0.0022 \text{ mm}^{-1}$, $\mu'_s = 1.41 \text{ mm}^{-1}$ at both 650 nm and at 720 nm. For this experiment, the simulated measurement vector \mathbf{y} is corrupted by Poisson noise with signal-to-noise ratio (SNR) of 3 dB ($\approx 57\%$ noise). In our method, we used $p = 0.74$ and $\mathbf{A}^T \mathbf{y}$ as the initial guess. Fig. 4 shows the true signal (\mathbf{f}^*) and our reconstruction.

	Time (sec)	Iterations
Fixed-point iteration	21.2829	1,281,974
Newton's method	21.0128	476,585

Table 1. Time and iteration average over 10 trials for fixed-point iteration and Newton's method to reconstruct the fluorescence molecular tomography data.

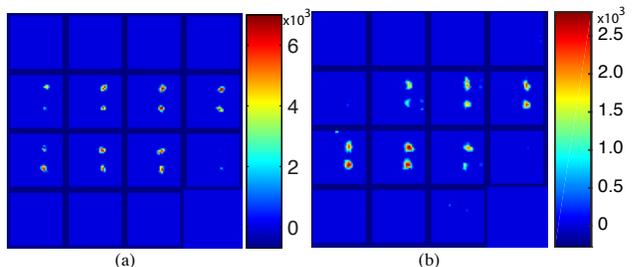


Fig. 4. (a) Horizontal slices of a simulated fluorescence capillary rod targets. (b) Reconstruction using p -norm regularized subproblem minimization.

4. CONCLUDING REMARKS

In this paper, we analyzed methods for solving the p -norm regularized subproblems arising from minimizing the Poisson-log likelihood for reconstructing sparse signals from photon-limited measurements. These non-convex subproblems do not have closed form solutions, and as such, they require numerical approaches for computing the minimizers. While Newton's method in theory should converge to the solution faster than fixed-point iterations, the number of floating-point operations needed to perform each iteration offsets the computational advantage of using derivative information.

Acknowledgments. We thank Dr. Dianwen Zhu and Prof. Changqing Li for providing the numerical codes for generating the fluorescence molecular tomography data.

5. REFERENCES

- [1] D. L. Snyder and M. I. Miller, “Random point processes in space and time,” *Springer-Verlag, New York, NY*, 1991.
- [2] D. Lingenfelter, J. Fessler, and Z. He, “Sparsity regularization for image reconstruction with Poisson data,” in *Proc. SPIE Computational Imaging VII*, 2009, vol. 7246.
- [3] Z. T. Harmany, R. F. Marcia, and R. M. Willett, “Sparse Poisson intensity reconstruction algorithms,” in *Proceedings of IEEE Statistical Signal Processing Workshop*, Cardiff, Wales, UK, September 2009.
- [4] M. Figueiredo and J. Bioucas-Dias, “Deconvolution of Poissonian images using variable splitting and augmented Lagrangian optimization,” in *Proc. IEEE Workshop on Statistical Signal Processing*, 2009.
- [5] S. Setzer, G. Steidl, and T. Teuber, “Deblurring Poissonian images by split Bregman techniques,” *J. Visual Commun. Image Represent.*, 2009, In Press.
- [6] J. A. Fessler and A. O. Hero, “Penalized maximum-likelihood image reconstruction using space-alternating generalized em algorithms,” *IEEE Trans. Image Processing*, vol. 4, no. 10, pp. 1417–1429, 1995.
- [7] S. Ahn and J. Fessler, “Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms,” *IEEE Trans. Med. Imaging*, vol. 22, no. 5, pp. 613–626, May 2003.
- [8] J. A. Fessler and H. Erdoğan, “A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction,” in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 1998.
- [9] M. Raginsky, R. M. Willett, Z. T. Harmany, and R. F. Marcia, “Compressed sensing performance bounds under Poisson noise,” *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 3990–4002, 2010.
- [10] R. Chartrand, “Exact reconstruction of sparse signals via nonconvex minimization,” *Signal Processing Letters, IEEE*, vol. 14, no. 10, pp. 707–710, Oct 2007.
- [11] R. Chartrand and V. Staneva, “Restricted isometry properties and nonconvex compressive sensing,” *Inverse Problems*, vol. 24, no. 3, pp. 035020, 2008.
- [12] R. Chartrand and W. Yin, “Iteratively reweighted algorithms for compressive sensing,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2008, pp. 3869–3872.
- [13] J. Barzilai and J. M. Borwein, “Two-point step size gradient methods,” *IMA J. Numer. Anal.*, vol. 8, no. 1, pp. 141–148, 1988.
- [14] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, “Sparse reconstruction by separable approximation,” *IEEE Trans. Signal Processing*, vol. 57, no. 7, pp. 2479–2493, 2009.
- [15] Z. T. Harmany, R. F. Marcia, and R. M. Willett, “This is SPIRAL-TAP: Sparse Poisson intensity reconstruction algorithms - theory and practice,” *IEEE Trans. Image Processing*, vol. 21, no. 3, pp. 1084–1096, 2012.
- [16] L. Adhikari and R. F. Marcia, “Nonconvex relaxation for Poisson intensity reconstruction,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Brisbane Australia, April 2015.
- [17] W. Zuo, D. Meng, L. Zhang, X. Feng, and D. Zhang, “A generalized iterated shrinkage algorithm for non-convex sparse coding,” in *2013 IEEE International Conference on Computer Vision (ICCV)*, Dec 2013, pp. 217–224.
- [18] D. R. Kincaid and E. W. Cheney, *Numerical Analysis: Mathematics of Scientific Computing*, American Mathematical Society, 3rd edition, 2002.
- [19] C. Li, G. S. Mitchell, J. Dutta, S. Ahn, R. M. Leahy, and S. R. Cherry, “A three-dimensional multispectral fluorescence optical tomography imaging system for small animals based on a conical mirror design,” *Opt. Express*, vol. 17, no. 9, pp. 7571–7585, Apr 2009.
- [20] F. Fedele, J. P. Laible, and M. J. Eppstein, “Coupled complex adjoint sensitivities for frequency-domain fluorescence tomography: theory and vectorized implementation,” *Journal of Computational Physics*, vol. 187, no. 2, pp. 597–619, 2003.