SPARSE SIGNAL RECONSTRUCTION FOR OVERDISPERSED LOW-PHOTON COUNT BIOMEDICAL IMAGING USING ℓ_P TOTAL VARIATION

Yu Lu and Roummel F. Marcia

Department of Applied Mathematics, University of California, Merced, Merced, CA, 95343 USA

ABSTRACT

The negative binomial model, which generalizes the Poisson distribution model, can be found in applications involving low-photon signal recovery, including medical imaging. Recent studies have explored several regularization terms for the negative binomial model, such as the ℓ_p quasi-norm with $0 < p < 1$, ℓ_1 norm, and the total variation (TV) quasi-seminorm for promoting sparsity in signal recovery. These penalty terms have been shown to improve image reconstruction outcomes. In this paper, we investigate the ℓ_p quasi-seminorm, both isotropic and anisotropic ℓ_p TV quasi-seminorms, within the framework of the negative binomial statistical model. This problem can be formulated as an optimization problem, which we solve using a gradient-based approach. We present comparisons between the negative binomial and Poisson statistical models using the ℓ_p TV quasi-seminorm as well as common penalty terms. Our experimental results highlight the efficacy of the proposed method.

Index Terms— Negative binomial distribution, low-count imaging, nonconvex optimization, ℓ_p total variation quasi-seminorm.

1. INTRODUCTION

In numerous photon-limited signal processing applications, such as medical imaging [1], astronomy [2], and network traffic analysis [3], the Poisson model [4] has been widely used for image reconstruction. However, the Poisson model assumes the data has the same mean and variance [5]. The negative binomial (NB) model is a more generalized statistical distribution, to which the Poisson model reduces under certain assumptions. The negative binomial model has been used for low-photon signal reconstruction [6] and matrix factorization [7]. Here, we explore the application of the negative binomial model for sparse signal recovery.

The negative binomial probability mass function (PMF) is

given by

$$P(y|r,p) = \binom{y+r-1}{y}(1-\beta)^y \beta^r,$$

where y and r represent the counts of successful and failure events, respectively. Here, β denotes the failure probability. In general, the negative binomial probability can be thought of as the probability of observing y successful events occurring before the rth failure, where each event is an independent and identically distributed Bernoulli trial with failure occurring probability β. The expectation and variance of the negative binomial PMF, denoted by μ and σ^2 respectively, are given by $\mu = r(1-\beta)/\beta$ and $\sigma^2 = r(1-\beta)/\beta^2 = \mu + \mu^2/r$ [8]. Given that the negative binomial distribution does not require identical mean and variance values. However, if we let r go to infinity, the variance will approach the expectation, which shows Poisson distribution is a limiting case of negative binomial distribution.

For optimization methods in sparse signal recovery application, sparsity-promoting penalties can improve reconstruction accuracy. Both the ℓ_p quasi-norm (for $0 < p < 1$) and the total variation (TV) quasi-seminorm have been evaluated in the contexts of both the Poisson [9, 10] and negative binomial models [11, 12]. The ℓ_p TV quasi-norm [13] has been used successfully for the Poisson model [14] for signal recovery. This paper investigates the application of the ℓ_p TV quasi-norm in the framework of the negative binomial model.

2. PROBLEM FORMULATION

Under the assumption of low-photon signal models, the observed vector of data y can be drawn from negative binomial distribution, which can be expressed as
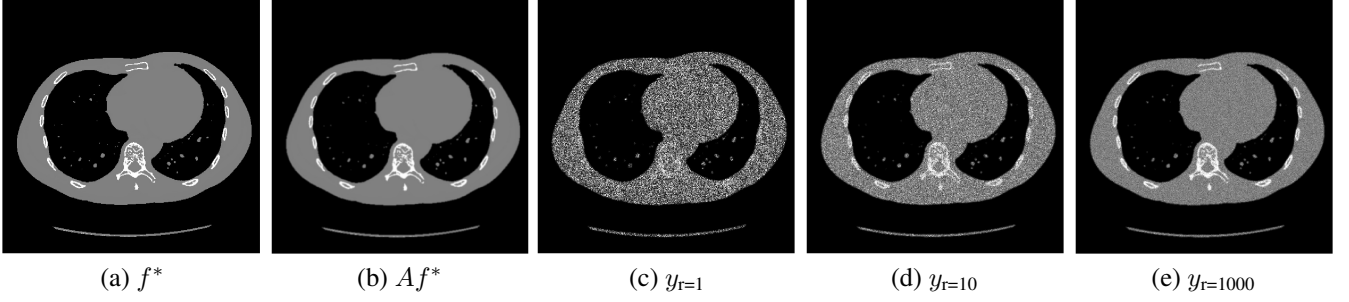
$$y \sim \text{NB}(r, \beta).$$

Replacing the probability β, the observation model becomes

$$y_i \sim \text{NB}\left(r, \frac{r}{r+(Af^*)_i}\right).$$

Here, $f^* \in \mathbb{R}_+^n$ denotes the true signal or image, while $A \in \mathbb{R}_+^{m \times n}$ represents the sensing matrix that projects the scene

---

(a) f^* (b) Af^* (c) $y_{r=1}$ (d) $y_{r=10}$ (e) $y_{r=1000}$

Fig. 1. Example of observation model. (a) The true image f^*. (b) The expected observation at the detector stage. Here, the measurement operator A is a Gaussian blur. (c) Observed measurement $y_{r=1}$ drawn from a negative binomial (NB) distribution with dispersion parameter $r = 1$. (d) Observed measurement $y_{r=10}$ drawn from a NB distribution with $r = 10$. (e) Observed measurement y_{1000} drawn from a NB distribution with $r = 1000$.

f^* onto an expected measurement set μ, given by Af^*. We denote the vector of noisy observation by $y \in \mathbb{R}_+^m$. This leads to the negative binomial statistical model

$$P(y|f) = \prod_{i=1}^{m} \binom{r + y_i - 1}{y_i} \left(\frac{r}{r + (Af)_i} \right)^r \left(\frac{(Af)_i}{r + (Af)_i} \right)^{y_i}. \quad (1)$$

See [8] for further details.

To optimize the probability of observing y in (1), we apply the maximum likelihood principle and minimize the corresponding negative binomial log-likelihood function, which is given by

$$F(f) = \sum_{i=1}^{m} (r + y_i) \log(r + (Af)_i) - y_i \log((Af)_i). \quad (2)$$

In our application of interest, the pixel intensities of the images are piece-wise constant. Consequently, the variations in these intensities are sparse. We exploit this inherent property of the signal to improve the accuracy of signal recovery. In particular, we investigate the use of the ℓ_p TV quasi-seminorm penalty with $0 < p < 1$ to promote sparsity in our solution. The completed formulation of our problem can be expressed as

$$f^* = \underset{f \in \mathbb{R}_+^n}{\arg\min} \quad \Phi(f) \equiv F(f) + \tau \|f\|_{TV}^p$$

$$\text{subject to} \quad 0 \le f, \quad (3)$$

where $\tau > 0$ is a regularization parameter. There are generally two choices for the TV quasi-seminorm: the anisotropic

$$\|f\|_{TV^{(A)}}^p = \sum_{l=1}^{m-1} \sum_{k=1}^{n} |f_{l,k} - f_{l+1,k}|^p + \sum_{l=1}^{m} \sum_{k=1}^{n-1} |f_{l,k} - f_{l,k+1}|^p$$

and the isotropic

$$\|f\|_{TV^{(I)}}^p = \sum_{l=1}^{m-1} \sum_{k=1}^{n-1} \sqrt{(f_{l,k} - f_{l+1,k})^{2p} + (f_{l,k} - f_{l,k+1})^{2p}}$$

$$+ \sum_{l=1}^{m-1} |f_{l,n} - f_{l+1,n}|^p + \sum_{k=1}^{n-1} |f_{m,k} - f_{m,k+1}|^p.$$

See [15] for further details.

3. ALGORITHM

The optimization problem defined in (2) can be solved via gradient-based methods. In particular, we use the second-order Taylor series expansion of $F(f)$ at the current iterate f^j to define a sequence of quadratic sub-problems. The first and second derivatives of $F(f)$ can be computed exactly:

$$\nabla F(f) = \sum_{i=1}^{m} \left(\frac{r + y_i}{r + (Af)_i} - \frac{y_i}{(Af)_i} \right) A^\top e_i$$

$$\nabla^2 F(f) = A^\top \left[\sum_{i=1}^{m} \left(\frac{r + y_i}{(r + (Af)_i)^2} - \frac{y_i}{(Af)_i^2} \right) e_i e_i^\top \right] A,$$

where e_i represents the ith column of the $m \times m$ identity matrix.

To reduce the computational complexity, we will apply the Barzilai-Borwein approach [16] to avoid computing the Hessian. Specifically, we approximate the second derivative of $F(f)$ with a scalar multiple of the identity matrix, i.e., $\nabla^2 F(f) \approx \alpha_j I$, where $\alpha_j > 0$ is a scalar. In the absence of the ℓ_p TV quasi-seminorm, the quadratic approximation is given by

$$F^j(f) = F(f^j) + (f - f^j)^\top \nabla F(f^j) + \frac{\alpha_j}{2} \|f - f^j\|_2^2.$$

By defining $q^j = f^j - \frac{1}{\alpha_j} \nabla F(f^j)$ and incorporating the penalty term, the corresponding quadratic subproblems and iterates can be expressed as

$$f^{j+1} = \underset{f \in \mathbb{R}^n}{\arg\min} \quad \frac{1}{2} \|f - q^j\|_2^2 + \frac{\tau}{\alpha_j} \|f\|_{TV}^p$$

$$\text{subject to} \quad 0 \le f \quad (4)$$

To deal with the ℓ_p TV quasi-seminorm, we take the reweighted approach [17, 18], which allow us to transfer an

ℓ_p TV nonconvex regularization directly to an ℓ_1 TV convex regularization:

$$\|f\|^p_{TV^{(A)}_{\gamma\omega}} = \sum_{l=1}^{m-1}\sum_{k=1}^{n}\gamma_{l,k}|f_{l,k}-f_{l+1,k}| +$$
$$\sum_{l=1}^{m}\sum_{k=1}^{n-1}\omega_{l,k}|f_{l,k}-f_{l,k+1}|,$$

and

$$\|f\|^p_{TV^{(I)}_{\gamma\omega}} =$$
$$\sum_{l=1}^{m-1}\sum_{k=1}^{n-1}\sqrt{(\gamma_{l,k}(f_{l,k}-f_{l+1,k}))^2 + (\omega_{l,k}(f_{l,k}-f_{l,k+1}))^2}$$
$$+ \sum_{l=1}^{m-1}\gamma_{l,n}|f_{l,n}-f_{l+1,n}| + \sum_{k=1}^{n-1}\omega_{m,k}|f_{m,k}-f_{m,k+1}|$$

where the weights $\gamma_{l,k}$ and $\omega_{l,k}$ are defined as

$$\gamma_{l,k} = |f^j_{l,k}-f^j_{l+1,k}|^{(p-1)},$$
$$\omega_{l,k} = |f^j_{l,k}-f^j_{l,k+1}|^{(p-1)},$$

and f^j is the output from the previous iteration j. These weights satisfy the following approximations:

$$|f^j_{l,k}-f^j_{l+1,k}|^{(p-1)}(f_{l,k}-f_{l+1,k}) \approx (f_{l,k}-f_{l+1,k})^p$$

(see [17] for details). We can now apply the fast gradient projection method [19] to solve the subproblem defined in (4).

4. EXPERIMENTS

We performed two experiments with data using various noise levels, with dispersion parameters $r = 1, 10, 25$, and 1000. All experiments involved reconstructing a blurry, noisy image (from the MATLAB Medical Imaging Toolbox) with dimensions of 512 by 512 pixels. **Each experiment was conducted 10 times, where the average of these trials are presented.** The objectives of the two experiments are:

1. To examine the effectiveness of the ℓ_p TV quasi-seminorm using different values of p.
2. To evaluate and compare the performance of the Poisson and negative binomial models when employing the ℓ_p TV quasi-seminorm.
3. To benchmark the performance of ℓ_p TV quasi-seminorm, ℓ_1 TV quasi-norm, and ℓ_p quasi-norm.

The details of each experiment are as follows:

Experiment 1: This experiment investigated the optimal solutions using negative binomial model with different p values using the ℓ_p TV quasi-seminorm. The test data for this experiment consist of 2D images that have different noise-level intensities, drawn from the negative binomial distribution.

Experiment 2: Here, we compared the results from using Poisson and negative binomial models. This experiment focuses on the relative performances of multiple types of regularization, including the ℓ_p TV quasi-seminorm, ℓ_1 TV quasi-norm, and ℓ_p quasi-norm. As with the first experiment, the test data were the same 2D images with noise drawn from a negative binomial distribution.

For the negative binomial models, several methods exist for estimating the parameter r, such as method-of-moment [20] and the maximum quasi-likelihood methods [21]. Additionally, cross-validation techniques offer a precise method for estimating the dispersion parameter [22]. However, for the sake of eliminating potential biases or inaccuracies inherent in the parameter estimation process, our experiments intentionally use the exact value of the parameter r as a prior.
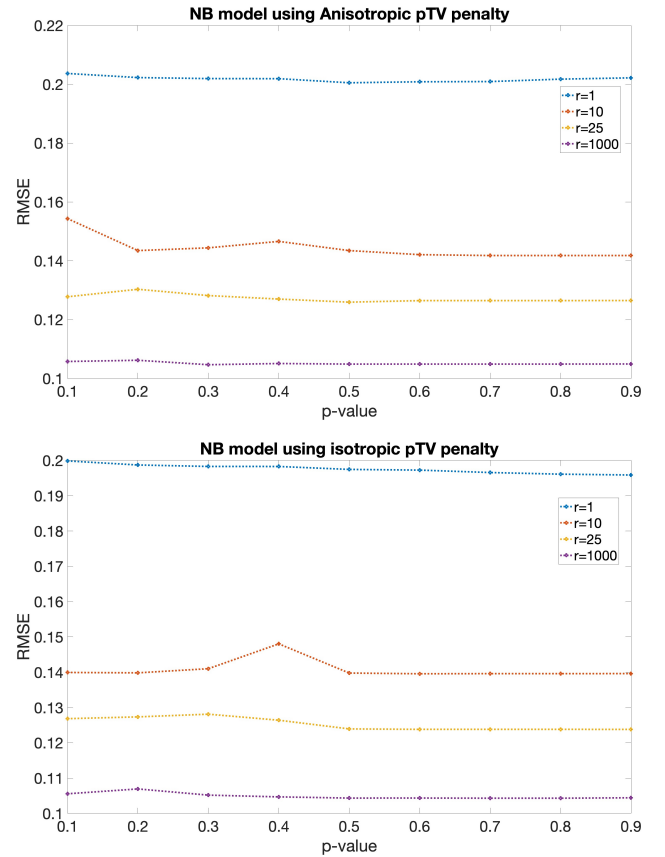
Fig. 2. **Experiment I:** RMSE analysis for 2D data reconstruction employing the negative binomial model in conjunction with the ℓ_p TV quasi-seminorm. The evaluation spans across multiple noise levels, i.e., $r = 1, 10, 25$ and $1,000$ for different p values. Observe that the RMSE values do not change significantly as the value of p changes.

4.1. Experiment I: p-value exploration

In our first experiment, we examined a set of nine distinct p values, evenly spaced between 0.1 and 0.9, using the negative binomial model with both anisotropic and isotropic ℓ_p TV quasi-seminorms. These were applied to four different noise-level images corresponding to four different dispersion parameter r values. As expected, as the parameter r increased, the corresponding RMSE decreased. When $r = 1$, the RMSE was highest. In contrast, the scenarios with the highest r value (specifically $r = 1000$) returned the lowest RMSE.

Interestingly, our findings indicate that the results of both isotropic and anisotropic ℓ_p TV quasi-seminorms are not sensitive to variations in the p value, as shown in Fig. 2. Specifically, across the range of p values, the RMSE values change by less than 1%.
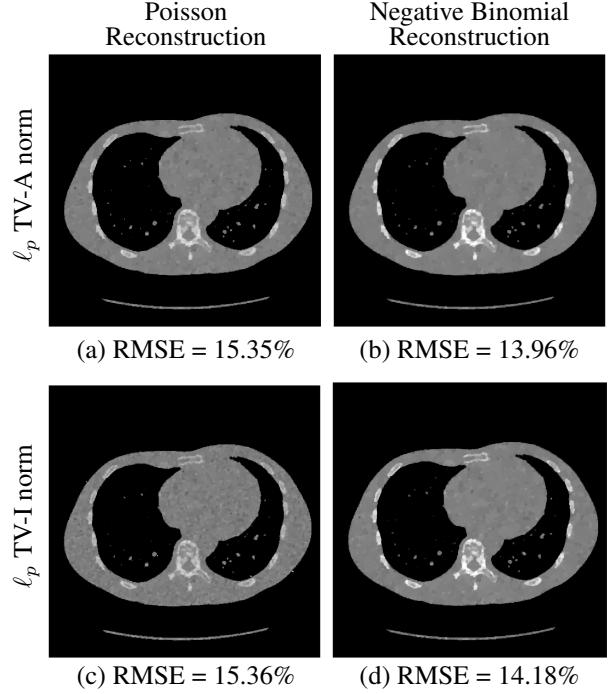
4.2. Experiment II: Comparison between Poisson and negative binomial models with mutiple regularizations

In the second experiment, we conducted a comparative analysis between the negative binomial and Poisson models using three different penalties. The results for these comparisons are presented in Table 1. Additionally, the results of $r = 10$ are visually represented in Fig. 3.

Our observations consistently show that the negative binomial model achieve lower RMSE values than the Poisson model with various r values. As r increases, the RMSE gaps between the two models diminish (see Table 1). This aligns with the fact that the Poisson distribution can be seen as a limiting case of the negative binomial distribution when the dispersion parameter r approaches infinity.

Model	Penalty	Dispersion parameter r			
		1	10	25	1,000
Negative Binomial	ℓ_p TV-A	0.2005	**0.1396**	**0.1238**	0.1047
	ℓ_p TV-I	**0.1957**	0.1418	0.1259	**0.1043**
	ℓ_1 TV-A	0.2019	0.1421	0.1267	0.1053
	ℓ_1 TV-I	0.1959	0.1399	0.1240	0.1047
	p norm	0.2104	0.1603	0.1411	0.1253
Poisson	ℓ_p TV-A	0.2364	0.1535	0.1403	0.1054
	ℓ_p TV-I	0.2231	0.1536	0.1422	0.1053
	ℓ_1 TV-A	0.2386	0.1569	0.1501	0.1062
	ℓ_1 TV-I	0.2253	0.1581	0.1484	0.1054
	p norm	0.2391	0.1739	0.1526	0.1262

Table 1. **Experiment II:** Evaluation of the Root-Mean-Square Error (RMSE) for both the negative binomial and Poisson models with four dispersion parameters ($r = 1, 10, 25,$ and $1,000$) using five regularization techniques: isotropic ℓ_p TV (ℓ_p TV-I), anisotropic ℓ_p TV (ℓ_p TV-A), isotropic ℓ_1 TV (ℓ_1 TV-I), anisotropic ℓ_1 TV (ℓ_1 TV-A), and the p quasi-norm. For each penalty, the parameter p is optimized to yield the minimum RMSE.

Poisson Reconstruction — Negative Binomial Reconstruction

ℓ_p TV-A norm

(a) RMSE = 15.35% (b) RMSE = 13.96%

ℓ_p TV-I norm

(c) RMSE = 15.36% (d) RMSE = 14.18%

Fig. 3. **Experiment II:** 2D data from a negative binomial distribution ($r = 10$). The negative binomial model yields a lower RMSE than the Poisson model with ℓ_p TV quasi-seminorm. The difference between isotropic and anisotropic versions is small.

Furthermore, our analysis reveals that the RMSE differences between anisotropic and isotropic ℓ_p TV quasi-seminorms are consistently small, regardless of whether the negative binomial or Poisson model is applied. Additionally, ℓ_p TV quasi-seminorms consistently achieve lower RMSE values than ℓ_p quasi-norms. We generally observe an improvement in RMSE values using the ℓ_p TV quasi-seminorm over the ℓ_1 TV quasi-seminorm.

5. CONCLUSION

In this study, we investigate the ℓ_p TV quasi-seminorm for signal reconstruction using 2-D data drawn from a negative binomial model with four different noise levels. Our findings demonstrate that the performance of the ℓ_p TV quasi-seminorm surpasses that of the ℓ_p quasi-seminorm in both negative binomial and Poisson model. Furthermore, our results show that the ℓ_p TV quasi-seminorm's performance remains consistent when using different p values within the range $0 < p \leq 1$. Additionally, when comparing the negative binomial model with the Poisson model, our findings verify that the negative binomial model yields lower RMSE values in signal reconstruction.

6. REFERENCES

[1] R. M. Willett and R. D. Nowak, "Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging," *IEEE Transactions on Medical Imaging*, vol. 22, no. 3, pp. 332–350, 2003.

[2] H. Lantéri and C. Theys, "Restoration of astrophysical images—the case of poisson data with additive gaussian noise," *EURASIP Journal on Advances in Signal Processing*, vol. 2005, pp. 1–14, 2005.

[3] S. B. Slimane and T. Le-Ngoc, "A doubly stochastic Poisson model for self-similar traffic," in *Proceedings IEEE International Conference on Communications ICC '95*, 1995, vol. 1, pp. 456–460 vol.1.

[4] 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, 2012.

[5] W. Gardner, E. P Mulvey, and Esther C Shaw, "Regression analyses of counts and rates: Poisson, overdispersed poisson, and negative binomial models.," *Psychological bulletin*, vol. 118, no. 3, pp. 392, 1995.

[6] Y. Lu and R. F. Marcia, "Sparse overdispersed photon-limited signal recovery with upper and lower bounds," in *Accepted to IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2023.

[7] O. Gouvert, T. Oberlin, and C. Févotte, "Negative binomial matrix factorization," *IEEE Signal Processing Letters*, vol. 27, pp. 815–819, 2020.

[8] M. H. DeGroot and M. J. Schervish, *Probability and statistics*, Pearson Education, 2012.

[9] A. Orkusyan, L. Adhikari, J. Valenzuela, and R. F. Marcia, "Analysis of p-norm regularized subproblem minimization for sparse photon-limited image recovery," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 1407–1411.

[10] A. Sawatzky, C. Brune, J. Müller, and M. Burger, "Total variation processing of images with Poisson statistics," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2009, pp. 533–540.

[11] Y. Lu and R. F Marcia, "Negative binomial optimization for low-count overdispersed sparse signal reconstruction," in *2023 31st European Signal Processing Conference (EUSIPCO)*. IEEE, 2023, pp. 1948–1952.

[12] Y. Lu and R. F. Marcia, "Overdispersed photon-limited sparse signal recovery using nonconvex regularization," in *Accepted to IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2023.

[13] X. Liu, Y. Chen, Z. Peng, J. Wu, and Z. Wang, "Infrared image super-resolution reconstruction based on quaternion fractional order total variation with lp quasinorm," *Applied Sciences*, vol. 8, no. 10, pp. 1864, 2018.

[14] L. Adhikari and R. F. Marcia, "p-th power total variation regularization in photon-limited imaging via iterative reweighting," in *2015 23rd European Signal Processing Conference (EUSIPCO)*. IEEE, 2015, pp. 1621–1625.

[15] M. Lasica, S. Moll, and P. B Mucha, "Total variation denoising in l^1 anisotropy," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1691–1723, 2017.

[16] J. Barzilai and J. M. Borwein, "Two-Point Step Size Gradient Methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 01 1988.

[17] J. Yan and W. Lu, "Image denoising by generalized total variation regularization and least squares fidelity," *Multidimensional Systems and Signal Processing*, vol. 26, no. 1, pp. 243–266, 2015.

[18] E. J Candes, M. B Wakin, and S. P Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *Journal of Fourier analysis and applications*, vol. 14, pp. 877–905, 2008.

[19] A. Beck and M. 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] S. J. Clark and J. N. Perry, "Estimation of the negative binomial parameter κ by maximum quasi-likelihood," *Biometrics*, pp. 309–316, 1989.

[21] W. W. Piegorsch, "Maximum likelihood estimation for the negative binomial dispersion parameter," *Biometrics*, pp. 863–867, 1990.

[22] C. Gu, "Cross-validating non-Gaussian data," *Journal of Computational and Graphical Statistics*, vol. 1, no. 2, pp. 169–179, 1992.