POISSON LOW-RANK MATRIX RECOVERY USING THE ANSCOMBE TRANSFORM

Pakshal Bohra[1] *and Ajit Rajwade*[2]

[1]Department of Electrical Engineering, IIT Bombay
[2]Department of Computer Science and Engineering, IIT Bombay
Email: pakshalbohra@gmail.com, ajitvr@cse.iitb.ac.in

ABSTRACT

We present an estimator, based on the Anscombe transform, for the problem of low-rank matrix recovery under Poisson noise. We derive an upper bound on the matrix reconstruction error for this estimator, considering a linear sensing operator which obeys realistic constraints like non-negativity and flux-preservation. Besides being computationally tractable (convex), our estimator also allows for principled parameter tuning. Moreover, our method is capable of handling Poisson-Gaussian noise and the case where the Poisson or Poisson-Gaussian corrupted measurements are uniformly quantized. In addition to our theoretical results, we present some numerical results for Poisson low-rank matrix recovery under varying intensity levels and number of measurements.

Index Terms— Low-rank matrix recovery, Poisson noise, Anscombe transform, reconstruction error bounds

1. INTRODUCTION

In the Low-Rank Matrix Recovery (LRMR) problem, the aim is to recover or reconstruct a (nearly) low rank matrix, from certain measurements of the matrix which may be corrupted by noise. A nearly low-rank matrix refers to one that can be approximated well by a low-rank matrix, i.e., the vector of singular values of this matrix is compressible. Typically, the number of measurements is much smaller than the number of entries in the matrix. While there exists a large body of work on the LRMR problem based on nuclear norm (i.e., sum of singular values) minimization [1, 2, 3], most of the theoretical work makes the assumption of additive signal-independent noise. However the noise in many compressive imaging systems is better modelled as non-additive and signal-dependent Poisson noise due to the photon-counting devices in the system [4, 5], while in many cases, the signal can be modelled as a nearly low-rank matrix. In Poisson LRMR, we consider the forward model $\boldsymbol{y} \sim \text{Poisson}(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))$. Here $\boldsymbol{y} \in \mathbb{Z}_+^N$ is a vector of independent measurements of the form $y_i \sim \text{Poisson}[(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))_i]$, $\boldsymbol{M} \in \mathbb{R}_+^{d_1 \times d_2}$ is a low-rank or nearly low-rank non-negative matrix and $\boldsymbol{\mathcal{A}} : \mathbb{R}^{d_1 \times d_2} \rightarrow \mathbb{R}^N (N \ll d_1 d_2)$ is the linear sensing operator.

AR acknowledges support from IITB seed grant 14IRCCSG012.

Since we are modelling a practical imaging system, the linear sensing operator $\boldsymbol{\mathcal{A}}$ must satisfy certain physical constraints. The resulting measurements from $\boldsymbol{\mathcal{A}}$ must be non-negative. Also, the total intensity of the measurements must be less than or equal to the total intensity of the original matrix (flux-preservation) [6]. Sensing operators satisfying the above two properties may not obey the Restricted Isometry Property (RIP) for matrices [2] which is often used to derive performance bounds for LRMR. To overcome this challenge, we design the sensing operator in a particular way [7] which is described in Section 2.

There exists previous work [7, 8] on Poisson LRMR where a regularized maximum likelihood estimator has been analyzed. There also exists previous work on the related problem of Poisson Matrix Completion [9, 10, 11], where the aim is to recover a matrix from a subset of its Poisson corrupted entries. However, we do not consider this problem in this paper. The work in [7], based on the penalized Poisson maximum likelihood, applies to low-rank/nearly low-rank matrices and linear operators which obey physical constraints. However, their bounds are derived for a computationally intractable estimator. Moreover, their work explores only the case of Poisson noise and not Poisson-Gaussian noise (which is a more accurate model for the noise in imaging systems [12] since the electronic fluctuations in the system give rise to a signal-independent Gaussian component of the noise as well). The regularization parameter in their implemented estimator is also difficult to choose in practice and does not have theoretical treatment.

In our work, we derive an upper bound on the reconstruction error for an estimator based on the Anscombe transform (AT) [13, 14]. The transform $f(X) = \sqrt{X + 3/8}$ converts a Poisson random variable $X \sim \text{Poisson}(\lambda)$ into one with an approximately Gaussian distribution with mean $\approx \sqrt{\lambda + 3/8}$ and variance $\approx \frac{1}{4}$. This motivates the following estimator:

$$\min_{\boldsymbol{M}} \|\boldsymbol{M}\|_* \text{ subject to } \|\sqrt{\boldsymbol{y} + c} - \sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M}) + c}\|_2 \leq \varepsilon, \quad (1)$$
$$\boldsymbol{M} \succeq \boldsymbol{0},\ \|\boldsymbol{M}\|_{1,1} = I,$$

where $\boldsymbol{y} \sim \text{Poisson}(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))$, $\boldsymbol{M}$ is a non-negative matrix which is low-rank or nearly low-rank with nuclear norm (i.e., sum of singular values) $\|\boldsymbol{M}\|_*$ and intensity $I \triangleq \|\boldsymbol{M}\|_{1,1} =$

$\sum_{i,j}|M_{ij}|$, $c \triangleq \frac{3}{8}$ and ε is a statistically motivated upper bound on $\|\sqrt{\boldsymbol{y}+c}-\sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2$. As we shall later show, our specific choice of ε does *not* rely on the exact Gaussianity of $(\sqrt{y_i+c}-\sqrt{(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))_i+c})$. The symbol $\succeq$ in $\boldsymbol{A} \succeq \boldsymbol{B}$ means that $a_{ij} \geq b_{ij}$ for all i, j in matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

This estimator (along with the performance bounds) can be easily modified for the Poisson-Gaussian case. For the Poisson-Gaussian noise model where the measurements are of the form $\boldsymbol{y} \sim \text{Poisson}(\boldsymbol{\mathcal{A}}(\boldsymbol{M})) + \boldsymbol{\eta}$ where $\eta_i \sim \mathcal{N}(0, \sigma^2)$, we can apply the Generalized AT [15] to the noisy measurements in place of the AT. This is equivalent to using $c = \frac{3}{8} + \sigma^2$ in the estimator (1), and leads to similar (though not identical) performance bounds as in Theorem 3.1.

The contribution of our work is summarized as follows:

1. We have derived performance bounds for a computationally tractable (convex) estimator, with practical linear sensing operators and for low-rank or nearly low-rank matrices.
2. We have a unified approach to handle Poisson as well as Poisson-Gaussian noise.
3. Our estimator allows for principled, statistically motivated parameter tuning since (as we show later) the term $\|\sqrt{\boldsymbol{y}+c}-\sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2$ has a bounded variance which does not depend on the original signal or the number of measurements.

2. LINEAR SENSING OPERATOR

First, we explicitly state the form of the linear operator that we have used. As shown in [2], we define the linear operator $\boldsymbol{\mathcal{A}}$ such that it can be represented by a series of matrices $\{\boldsymbol{A}_i \in \mathbb{R}^{d_1 \times d_2}\}$ such that $[\boldsymbol{\mathcal{A}}(\boldsymbol{M})]_i = \langle \boldsymbol{A}_i, \boldsymbol{M} \rangle = \sum_{k=1}^{d_1}\sum_{l=1}^{d_2}(A_i)_{kl}M_{kl}$. In other words, we can write

$$\boldsymbol{\mathcal{A}}(\boldsymbol{M}) = \begin{bmatrix} \text{vec}(\boldsymbol{A}_1)^T \\ \text{vec}(\boldsymbol{A}_2)^T \\ \vdots \\ \text{vec}(\boldsymbol{A}_m)^T \end{bmatrix} \text{vec}(\boldsymbol{M}),$$

where $\text{vec}(\boldsymbol{M})$ is a vector obtained by stacking the columns of $\boldsymbol{M}$. This kind of a linear operator actually models the measuring process in compressive imaging systems like the Rice single pixel camera [4].

The non-negativity property means that $[\boldsymbol{\mathcal{A}}(\boldsymbol{M})]_i \geq 0 \; \forall \, i$ while the flux-preservation property means that $\sum_{i=1}^{N}[\boldsymbol{\mathcal{A}}(\boldsymbol{M})]_i \leq \sum_{i=1}^{d_1}\sum_{j=1}^{d_2} M_{ij}$. We construct a non-negative and flux-preserving linear operator based on the approach in [7]. Let $\boldsymbol{Z}_i \in \mathbb{R}^{d_1 \times d_2}$ be a matrix in which each entry is either $+1$ or -1 with equal probability. We define $\boldsymbol{B}_i = \frac{\boldsymbol{Z}_i}{\sqrt{N}}$. We construct the matrix $\boldsymbol{A}_i$ as follows:

$$\boldsymbol{A}_i = \frac{\boldsymbol{B}_i}{2\sqrt{N}} + \frac{\boldsymbol{1}_{d_1 \times d_2}}{2N}. \tag{2}$$

Now each entry of $\boldsymbol{A}_i$ is either 0 or $\frac{1}{N}$. We can easily verify that the linear operator $\boldsymbol{\mathcal{A}}$ associated with the series of matrices $\{\boldsymbol{A}_i\}$ satisfies both the non-negativity and flux-preservation properties. Let $\boldsymbol{\mathcal{B}}$ be the linear operator associated with the series of matrices $\{\boldsymbol{B}_i\}$. Theorem 2.3 of [2] shows that $\boldsymbol{\mathcal{B}}$ satisfies the matrix RIP at rank r with high probability for $\mathcal{O}(dr)$ measurements where $d \triangleq \max(d_1, d_2)$.

3. THEORETICAL RESULTS

In this section, we first state a theorem from [16] which mentions some properties of $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M})) \triangleq \|\sqrt{\boldsymbol{y}+c}-\sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2$ where $c = \frac{3}{8}$. Using this result, we then formulate Theorem 3.2 below for the upper bound on the reconstruction error for Poisson LRMR.

Theorem 3.1. *Let $\boldsymbol{y} \in \mathbb{Z}_+^N$ be a vector of independent measurements such that $y_i \sim$ Poisson$[(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))_i]$ where $\boldsymbol{M} \in \mathbb{R}_+^{d_1 \times d_2}$ is a non-negative matrix and $\boldsymbol{\mathcal{A}}$ is constructed according to the discussion in Section 2. We define $\gamma_i \triangleq (\boldsymbol{\mathcal{A}}(\boldsymbol{M}))_i$. Then for $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))$, the following properties hold:*

1. $\mathbb{E}[R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))] \leq \sqrt{N/2}$
2. *Define $v \triangleq$ Var$[R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))]$. If for all $i, \gamma_i \geq 1$ and $N \geq 29$, then we have $v \leq \frac{3/4}{(2c+1)/(8(1+c)^2)} \approx 6.48$ and $\mathbb{P}\Big(R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M})) \leq \sqrt{N}(\frac{1}{\sqrt{2}}+2.55)\Big) \geq 1-1/N$.*

For completeness, we have included a proof of Theorem 3.1 in our supplemental material at [17]. This theorem does not require $(\sqrt{y_i+c}-\sqrt{(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))_i+c})$ to be a Gaussian random variable. Such an assumption would not be rigorous because as shown in [13], Gaussianity is obeyed only asymptotically when the mean of y_i tends to infinity. Theorem 3.1 provides us with a signal-independent upper bound for $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))$ and thus allows principled parameter tuning in our estimator. In our experiments, we have observed that the mean and variance of $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))$ are bounded by smaller values than those stated in the theorem above. This can be seen from Figures 1 and 2 in [16].

Theorem 3.2. *Consider corrupted measurements of the form $\boldsymbol{y} \sim$ Poisson$(\boldsymbol{\mathcal{A}}(\boldsymbol{M}))$. Let $\boldsymbol{M}^\star$ be the solution of the following optimization problem:*

$$\min_{\boldsymbol{M}} \|\boldsymbol{M}\|_* \text{ subject to } \|\sqrt{\boldsymbol{y}+c}-\sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2 \leq \varepsilon, \quad \boldsymbol{M} \succeq \boldsymbol{0}, \; \|\boldsymbol{M}\|_{1,1} = I, \tag{3}$$

where $c = \frac{3}{8}$ and $\varepsilon \triangleq \tau\sqrt{N}$ such that $\tau = (\frac{2.55}{\kappa} + \frac{1}{\sqrt{2}})$ for some $\kappa > 0$. Let $\boldsymbol{M_r}$ denote the best r rank approximation to $\boldsymbol{M}$. If $N \geq 29$, $\boldsymbol{\mathcal{B}}$ obeys the matrix RIP such that $\delta_{4r} < \sqrt{2}-1$ and the condition $\boldsymbol{\mathcal{A}}(\boldsymbol{M}) \succeq \boldsymbol{1}$ holds, then we have:

$$\mathbb{P}\left(\frac{\|\boldsymbol{M}-\boldsymbol{M}^\star\|_F}{I} \leq 8C_1\sqrt{N}\tau\sqrt{\frac{1}{I}+\frac{cN}{I^2}} + \frac{C_2 r^{-\frac{1}{2}}\|\boldsymbol{M}-\boldsymbol{M_r}\|_*}{I}\right) \geq 1-\kappa^2/N$$

where $C_1 = \frac{2\sqrt{1+\delta_{4r}}}{1-(1+\sqrt{2})\delta_{4r}}$ *and* $C_2 = \sqrt{2} + \frac{4\delta_{4r}}{1-(1+\sqrt{2})\delta_{4r}}$.

Remarks:

1. The proof for Theorem 3.2, which can be accessed in the supplemental material at [17], is inspired from [2] and [18]. These proof techniques have been modified to handle the non-negativity and flux-preserving property of the linear sensing operator, and the constraints $\|\boldsymbol{M}\|_{1,1} = I$, $\boldsymbol{M} \succeq \boldsymbol{0}$.
2. Since we are dealing with a *Poisson* inverse problem, we analyze the *relative* reconstruction error (RRE) instead of the (absolute) reconstruction error. This is because as the mean of the Poisson distribution increases, its variance also increases, causing an increase in the mean squared error but a decrease in the relative mean squared error.
3. The above bound is inversely proportional to the signal intensity I. For a fixed I, if N is increased, then the incident photon flux I is distributed across more measurements, causing a decrease in SNR per measurement and possibly degrading performance. In fact, this affects our bound.
4. As r increases, δ_{4r} of the sensing operator will increase. Hence the constants C_1 and C_2 will increase since they are monotonically increasing functions of δ_{4r}. Hence, the upper bound on the performance may actually increase with r.
5. Our experimental results in the next section show that the constraint $\|\boldsymbol{M}\|_{1,1} = I$ is not necessary, although we required it for our theoretical analysis.
6. We can extend Theorem 3.2 to handle the Poisson-Gaussian case where the measurements are $\boldsymbol{y} \sim \text{Poisson}(\boldsymbol{\mathcal{A}}(\boldsymbol{M})) + \boldsymbol{\eta}$, $\eta_i \sim \mathcal{N}(0, \sigma^2)$. For this case, we replace c in the estimator with $\tilde{c} \triangleq c + \sigma^2$. The bound for this model can be derived by following steps similar to those shown in the proofs of Theorem 3.2 and making use of Theorem 2 in [16]. In fact, this theorem can also be extended to include uniform quantization noise (with or without the Gaussian noise).

4. EXPERIMENTS

Matrix and Measurement Generation: We ran experiments for reconstruction of $Q = 50$ non-negative 50×50 low-rank matrices, from their Poisson corrupted compressive measurements. We constructed the sensing operator $\boldsymbol{\mathcal{A}}$ as per the model in Section 2. A matrix of rank $r < 50$ was generated by the multiplication of two non-negative matrices of the dimensions $50 \times r$ and $r \times 50$ respectively. The entries of these two matrices were drawn randomly from Unif$[0, 1]$. Finally, we scaled the elements of the 50×50 matrix to ensure the desired value of I (see description of experiments later in this section).

Methods Compared: We ran our simulations for the following three optimization problems:

$$\mathsf{P1}: \min_{\boldsymbol{M}} \|\boldsymbol{M}\|_* \text{ s.t. } \|\sqrt{\boldsymbol{y}+c} - \sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2 \leq \varepsilon, \boldsymbol{M} \succeq \boldsymbol{0}$$

$$\mathsf{P2}: \min_{\boldsymbol{M}} \|\sqrt{\boldsymbol{y}+c} - \sqrt{\boldsymbol{\mathcal{A}}(\boldsymbol{M})+c}\|_2 + \lambda\|\boldsymbol{M}\|_* \text{ s.t. } \boldsymbol{M} \succeq \boldsymbol{0},$$

$$\mathsf{P3}: \min_{\boldsymbol{M}} \sum_{i=1}^{N} ([\boldsymbol{\mathcal{A}}(\boldsymbol{M})]_i - y_i \log[\boldsymbol{\mathcal{A}}(\boldsymbol{M})]_i) + \lambda\|\boldsymbol{M}\|_* \text{ s.t. } \boldsymbol{M} \succeq \boldsymbol{0},$$

where $c = \frac{3}{8}$ in P1 and P2. The problem P3 is a variant of the matrix recovery optimization problem presented in [7]. Note that we have excluded the constraint $\|\boldsymbol{M}\|_{1,1} = I$ in all the three problems, as it had negligible impact on the results. All three problems are convex. We solved P1 and P2 using CVX [19] with the SDPT3 solver, and P3 using the PMLSVT algorithm [7] with parameters $K = 5000$, $t = 10^{-5}$ and $\eta = 1.1$ (see [7] for their meaning).

Study of variation of matrix/measurement parameters: We show results for two experiments described below. In both the experiments we have used the RRMSE (relative root mean-squared error) metric $\triangleq \|\boldsymbol{M} - \boldsymbol{M}^\star\|_F / \|\boldsymbol{M}\|_F$, where $\boldsymbol{M}$ and $\boldsymbol{M}^\star$ denote the true/original and reconstructed matrix respectively. In the *Experiment 1*, we studied the effect of change in intensity I on the reconstruction results. For this, we generated Poisson corrupted measurements of Q different matrices in $\mathbb{R}^{50\times 50}$, each with a fixed number of measurements $N = 1250$ and rank $r = 5$. The signal intensity was varied according to the set $\{10^4, 10^5, 2.5\times10^5, 5\times10^5, 7.5\times10^5, 10^6, 5\times10^6, 10^7, 5\times10^7, 10^8\}$. For each value of I, the median RRMSE value was recorded over the Q matrices. In the *Experiment 2*, for the Q different matrices, the rank of the matrices was fixed to $r = 5$, and their intensity was fixed to $I = 10^8$. The number of measurements was varied according to the set $\{250, 500, 750, 1000, 1250, 1500, 1750, 2000, 2250, 2500\}$. For each value of N, the median RRMSE value was recorded over the Q matrices.

In P1, the bound ε was set to $2\sqrt{N}$ based on the tail bound from Theorem 3.1 (note that $2\sqrt{N} = \sqrt{N}/\sqrt{2} + \sqrt{N}(2.55/2)$, and that this bound holds with probability $1 - (2)^2/N$, i.e. $\kappa = 2$). The same value of ε was used in both the experiments, and we refer to this variant as P1-1. Since $2\sqrt{N}$ is a conservative upper bound, we also ran both the experiments with $\varepsilon = \mu + 3\sigma$ where μ and σ are the empirical mean and standard deviation of $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))$, simulated over 100 randomly chosen matrices (termed P1-2 in Fig. 1). This choice of ε is principled since by Theorem 3.1, the statistics of $R(\boldsymbol{y}, \boldsymbol{\mathcal{A}}(\boldsymbol{M}))$ are signal-independent. For P2, in the Experiment 1, we chose the regularizer parameter λ omnisciently from set $\mathcal{S} \triangleq \{10^{-10}, 10^{-9}, ..., 10\}$, i.e. choosing the particular value of $\lambda \in \mathcal{S}$ that yielded the least squared difference between the true $\boldsymbol{M}$ (assuming it were known) and its estimate. For the Experiment 2 (where $I = 10^8$), we set

$\lambda = 10^{-3}$ as this was the value of λ which gave the smallest error corresponding to $I = 10^8$ in the Experiment 1. The problem P3 is based on the negative log-likelihood of the Poisson distribution. For P3, the regularization parameter λ was chosen omnisciently from $\mathcal{S}$ in the Experiment 1. For the Experiment 2, we chose $\lambda = 10^{-3}$ (for the same reason as in the P2 case).

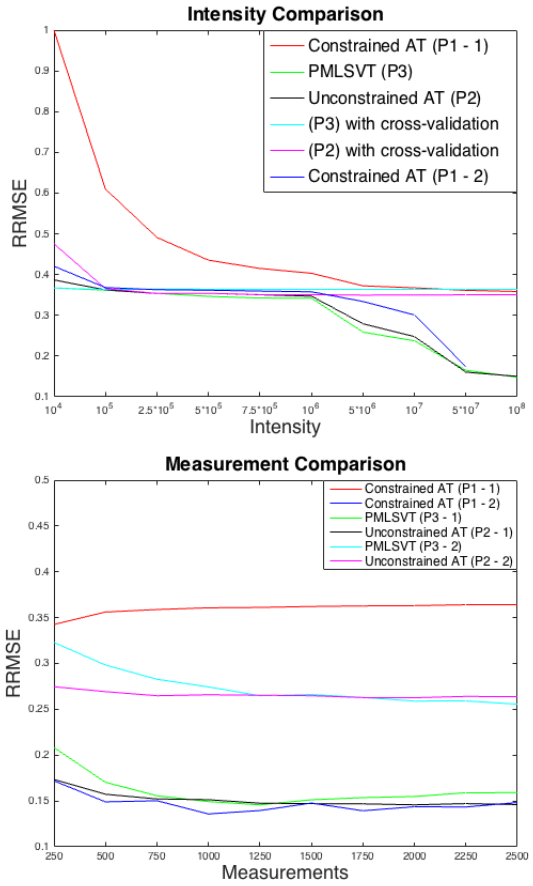

Fig. 1: Top figure: Median RRMSE comparisons (for fixed $N = 1250$ and $r = 5$, varying I) between $(P1-1)$ with $\varepsilon = 2\sqrt{N}$, $(P1-2)$ with $\varepsilon = \mu + 3\sigma$ (see text), P2 with omnisciently chosen λ, P3 with omnisciently chosen λ, P2 with cross-validation for λ, P3 with cross-validation for λ. Bottom figure: Median RRMSE comparisons (fixed $I = 10^8$ and $r = 5$, varying N) between $(P1-1)$ with $\varepsilon = 2\sqrt{N}$, $(P1-2)$ with $\varepsilon = \mu + 3\sigma$, $(P2-1)$ with $\lambda = 10^{-3}$, $(P3-1)$ with $\lambda = 10^{-3}$, $(P2-2)$ with $\lambda = 1$, $(P3-2)$ with $\lambda = 0.1$.

Comparison plots: Observing the comparison plots in Fig. 1, we see that the reconstruction results with P2 and P3 are comparable in most cases. Both showed better results than P1-1 due to the omniscient selection of λ, as against the fixed, statistically motivated, conservative $\varepsilon = 2\sqrt{N}$ in P1-1. However the results with P1-2 are comparable to and sometimes better than P2, P3. Moreover, omniscient choices are difficult and compute-intensive in practice. We saw that improper choice of λ, which is signal-dependent, led to arbitrary increase in reconstruction error in P3. To show this, we collected results on P2 and P3 via cross-validation for the Experiment 1. For this, we omnisciently chose λ which yielded the best RRMSE for $I = 10^5$ and used the same λ in P2 and P3 for all other intensity levels in the Experiment 1. For the Experiment 2, we have also shown the median RRMSE values for P2 with $\lambda = 0.1$ and P3 with $\lambda = 1$. This is just to show that the performance of P2 and P3 worsens with improper choice of λ.

Image Reconstruction: Lastly, we ran some experiments to reconstruct images from Poisson-corrupted measurements using the LRMR framework. For a $m_1 \times m_2$ image $\boldsymbol{M}$ with intensity I, we first constructed a $d_1 \times d_2$ patch matrix $\widetilde{\boldsymbol{M}}$ by taking all the 8×8 non-overlapping patches in $\boldsymbol{M}$, vectorizing them and using them as columns of $\widetilde{\boldsymbol{M}}$ (therefore, $d_1 = 64$ and $d_2 = m_1 m_2/64$). This patch matrix is a nearly low-rank matrix. We then generated measurements of the form $\boldsymbol{y} = \text{Poisson}(\boldsymbol{\mathcal{A}}(\widetilde{\boldsymbol{M}}))$ where $\boldsymbol{y} \in \mathbb{R}_+^N$ and $\boldsymbol{\mathcal{A}}$ was constructed as per the discussion in section 2. The matrix $\widetilde{\boldsymbol{M}}$ was reconstructed from $\boldsymbol{y}$ by solving P1 with $\varepsilon = 0.25\sqrt{N}$. We then reshaped this reconstructed $\widetilde{\boldsymbol{M}}$ to obtain an estimate of the original image. We ran this experiment on two images - solar flare (48×48) [7] and the flag of Texas (32×40). For each image, we ran our experiment for four cases - $(I = 10^7, N = 0.5 d_1 d_2)$, $(I = 10^7, N = 0.75 d_1 d_2)$, $(I = 10^8, N = 0.5 d_1 d_2)$ and $(I = 10^8, N = 0.75 d_1 d_2)$. The results for these experiments are shown in Fig. 2 and 3.

Fig. 2: Image reconstruction results for 'solar flare' using P1 with $\varepsilon = 0.5\sqrt{N}$. In the left to right order: original image, $(I = 10^7, N = 0.5 d_1 d_2,$ RRMSE $= 0.3899)$, $(I = 10^7, N = 0.75 d_1 d_2,$ RRMSE $= 0.3917)$, $(I = 10^8, N = 0.5 d_1 d_2,$ RRMSE $= 0.2257)$, $(I = 10^8, N = 0.75 d_1 d_2,$ RRMSE $= 0.1996)$.

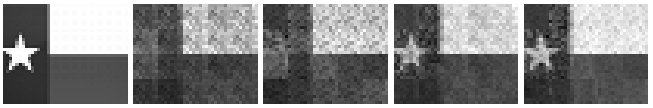

Fig. 3: Image reconstruction results for 'Texas flag' using P1 with $\varepsilon = 0.5\sqrt{N}$. In the left to right order: original image, $(I = 10^7, N = 0.5 d_1 d_2,$ RRMSE $= 0.2704)$, $(I = 10^7, N = 0.75 d_1 d_2,$ RRMSE $= 0.2461)$, $(I = 10^8, N = 0.5 d_1 d_2,$ RRMSE $= 0.1521)$, $(I = 10^8, N = 0.75 d_1 d_2,$ RRMSE $= 0.1328)$.

5. CONCLUSION

We have presented a convex estimator for Poisson LRMR with realistic sensing models and principled parameter tuning which makes it easier to use in practice. Our framework also allows for the unification of the analysis of LRMR under Poisson and Poisson-Gaussian noise. Adapting this framework for Poisson Matrix Completion and derivation of lower bounds are some directions for future work.

6. REFERENCES

[1] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM review*, vol. 52, no. 3, pp. 471–501, 2010.

[2] E. Candes and Y. Plan, "Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2342–2359, 2011.

[3] T. Cai and A. Zhang, "Sharp RIP bound for sparse signal and low-rank matrix recovery," *Applied and Computational Harmonic Analysis*, vol. 35, no. 1, pp. 74–93, 2013.

[4] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, March 2008.

[5] F. Alter, Y. Matsushita, and X. Tang, "An intensity similarity measure in low-light conditions," in *European Conference on Computer Vision*. Springer, 2006, pp. 267–280.

[6] M. Raginsky, R. Willett, Z. Harmany, and R. Marcia, "Compressed sensing performance bounds under poisson noise," *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 3990–4002, 2010.

[7] Y. Cao and Y. Xie, "Poisson matrix recovery and completion," *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1609–1620.

[8] Y. Xie, Y. Chi, and R. Calderbank, "Low-rank matrix recovery with poison noise," in *2013 IEEE Global Conference on Signal and Information Processing*, Dec 2013, pp. 622–622.

[9] S. Gunasekar, P. Ravikumar, and J. Ghosh, "Exponential family matrix completion under structural constraints," in *International Conference on Machine Learning*, 2014, pp. 1917–1925.

[10] A. Soni, S. Jain, J. Haupt, and S. Gonella, "Noisy matrix completion under sparse factor models," *IEEE Transactions on Information Theory*, vol. 62, no. 6, pp. 3636–3661, 2016.

[11] J. Lafond, "Low rank matrix completion with exponential family noise," in *Conference on Learning Theory*, 2015, pp. 1224–1243.

[12] T. Du Bosq and B. Preece, "Performance assessment of a single-pixel compressive sensing imaging system," in *Proc. SPIE*, 2016, vol. 9820.

[13] J. H. Curtiss, "On transformations used in the analysis of variance," *Ann. Math. Statist.*, vol. 14, no. 2, pp. 107–122, 06 1943.

[14] F. J. Anscombe, "The transformation of poisson, binomial and negative-binomial data," *Biometrika*, vol. 35, no. 3/4, pp. 246–254, 1948.

[15] F. Murtagh, J.-L. Starck, and A. Bijaoui, "Image restoration with noise suppression using a multiresolution support," *Astronomy and Astrophysics, Suppl. Ser*, vol. 112, pp. 179–189, 1995.

[16] D. Garg, P. Bohra, K. S. Gurumoorthy, and A. Rajwade, "Reconstruction error bounds for compressed sensing under poisson or poisson-gaussian noise using variance stabilization transforms," `https://www.cse.iitb.ac.in/~ajitvr/Poisson_PoissonGaussian_VST.pdf`, 2018.

[17] P. Bohra and A. Rajwade, "Poisson low-rank matrix recovery using the anscombe transform: Supplemental material," `https://www.cse.iitb.ac.in/~ajitvr/poisson-matrix-recovery-supplemental.pdf`.

[18] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus de Mathematiques*, vol. 346, pp. 589–592, 2008.

[19] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," `http://cvxr.com/cvx`, Mar. 2014.