BAYESIAN RESTORATION OF HIGH-DIMENSIONAL PHOTON-STARVED IMAGES

†J. Tachella, †Y. Altmann*, ‡M. Pereyra, †S. McLaughlin

†School of Engineering and Physical Sciences
‡School of Mathematical and Computer Sciences
Heriot Watt University
Edinburgh, UK

J.-Y. Tourneret

ENSEEHIT-IRIT-TeSA
University of Toulouse
Toulouse, France

ABSTRACT

This paper investigates different algorithms to perform image restoration from single-photon measurements corrupted with Poisson noise. The restoration problem is formulated in a Bayesian framework and several state-of-the-art Monte Carlo samplers are considered to estimate the unknown image and quantify its uncertainty. The different samplers are compared through a series of experiments conducted with synthetic images. The results demonstrate the scaling properties of the proposed samplers as the dimensionality of the problem increases and the number of photons decreases. Moreover, our experiments show that for a certain photon budget (i.e., acquisition time of the imaging device), downsampling the observations can yield better reconstruction results.

Index Terms— Bayesian statistics, Inverse problems, Image processing, Poisson noise, Markov chain Monte Carlo, Bouncy particle sampler

1. INTRODUCTION

Photon-limited imaging arises in many applications where the light flux changes rapidly or is extremely limited. In such scenarios, the flux is quantified at photonic levels using single-photon detectors. A wide range of applications operate in this regime, such as light detection and ranging (lidar) [1,2], medical imaging [3] and astronomy [4]. Photon count measurements generally follow Poisson statistics, which vary significantly from the ones corrupted by Gaussian noise, yielding poor results if Gaussian restoration methods are directly applied. Moreover, algorithms based on the variance-stabilizing transformation fail to provide good estimates when the number of available photons is very low. In most photon-starved imaging applications, the number of recorded photons is proportional to the acquisition time. Thus, in order to reduce the acquisition time, the aim is to provide reliable restorations using as few photons as possible.

*Part of this work was supported by the Royal Academy of Engineering under the Research Fellowship scheme.

The problem can be stated as follows: Given a set of discrete photon measurements $\boldsymbol{y} = (y_1, ..., y_K)^\intercal \in \mathbb{Z}_+^K = \{0, 1, \dots\}^K$, we are interested in recovering the underlying intensities $\boldsymbol{z} = (z_1, ..., z_N)^\intercal \in \mathbb{R}^N$ that are subjected to a linear operator $\boldsymbol{A} \in \mathbb{R}_+^{K \times N}$, such as blur, missing pixels, compressive measurements or tomographic projections, i.e.,

$$\boldsymbol{y}|\boldsymbol{z} \sim \mathcal{P}(\boldsymbol{A}\boldsymbol{z}). \tag{1}$$

This inverse problem (i.e., recovering $\boldsymbol{z}$ from $\boldsymbol{y}$) is often severely ill-posed or ill-conditioned and prior regularization is necessary to promote the solution to be in a set of feasible images $\boldsymbol{z}$. Many convex optimization algorithms based on log-concave Bayesian models have been proposed to perform maximum-a-posteriori (MAP) estimation. For example, PIDAL [5] is based on a total variation prior and solves the inverse problem using an alternating direction method of multipliers. Algorithms such as [5] are capable of computing MAP estimates relatively fast, but cannot provide uncertainty bounds for the estimates, which can be very valuable in many applications involving decision making such as self-driving cars [6]. In this paper, we formulate the inverse problem based on Eq. (1) in a Bayesian framework and investigate the efficiency of six different samplers at estimating the posterior distribution of z. Specifically, we investigate four Markov chain Monte Carlo (MCMC) methods for the Poisson restoration problem: The unadjusted Langevin algorithm (ULA) [7], the Metropolis adjusted Langevin algorithm (MALA) [7], Hamiltonian Monte Carlo (HMC) [8, Chapter 5] and the recent no U-turn Hamiltonian Monte Carlo (NUTS) scheme [9]. In addition, we evaluate the bouncy particle sampler (BPS), a non-reversible rejection-free strategy [10]. For completeness, we also include in our comparison the classical random walk Metropolis (RWM) sampler [7]. Through a series of experiments, we study the scaling properties of each sampler as the dimension of the problem increases, i.e., as the size of the image increases and we evaluate the quality of the estimates as we vary the total number of photons. The main contribution of this paper is to investigate different stochastic simulation strategies in very high dimensional and very low photon count scenarios.

ISBN 978-90-827970-1-5 © EURASIP 2018

2. BAYESIAN MODEL

Consider a vector of K noisy measurements $\boldsymbol{y} \in \mathbb{Z}_+^K$ and a vectorized intensity image $\boldsymbol{z} \in \mathbb{R}_+^N$ of n pixels. We are interested in the posterior distribution of the intensity values $\boldsymbol{z}$ given the measurements $\boldsymbol{y}$ denoted as $p(\boldsymbol{z}|\boldsymbol{y})$. Following Bayes theorem, this posterior is given by

$$p(\boldsymbol{z}|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|\boldsymbol{z})p(\boldsymbol{z})}{\int p(\boldsymbol{y}|\boldsymbol{z})p(\boldsymbol{z})d\boldsymbol{z}}. \tag{2}$$

2.1. Poisson likelihood

The photon count of measurement y_k is Poisson distributed

$$y_k|\boldsymbol{z} \sim \mathcal{P}\left(\sum_{n=1}^{N} A_{k,n} z_n\right) \quad k = 1, \ldots, K \tag{3}$$

where $\mathcal{P}(.)$ is the Poisson distribution. Assuming independent noise realizations, the likelihood of $\boldsymbol{y}$ is

$$p(\boldsymbol{y}|\boldsymbol{z}) = \prod_{k=1}^{K} p(y_k|\boldsymbol{z}) \tag{4}$$

where $p(y_i|\boldsymbol{z})$ is the probability associated with the Poisson distribution of Eq. (3). To alleviate difficulties induced by the positivity constraint on $\boldsymbol{z}$, we consider the reparametrization $z_n = e^{x_n}$ for $n = 1, \ldots, N$, where the log-intensity $x_n \in \mathbb{R}^N$ is no longer constrained to be positive, at the cost of inducing an implicit prior for $\boldsymbol{z}$. The modified log-likelihood has a bounded Lipschitz constant as $x_n \to -\infty$, yielding better stability in low-photon count scenarios. A detailed study of the trade-off involved in this change of variables can be found in [11].

2.2. Laplacian filter prior

The target log-intensity vector $\boldsymbol{x} = (x_1, ..., x_N)^\intercal$ is regularized with the prior $p(\boldsymbol{x}|\lambda) \sim \mathcal{N}(0, (\lambda \boldsymbol{P})^{-1})$, where the precision matrix is $\boldsymbol{P} = \boldsymbol{D}^\intercal \boldsymbol{D}$ and $\boldsymbol{D}$ is the block circulant matrix of the 2D Laplacian filter. This l_2 regularizer promotes smoothness in the log-image $\boldsymbol{x}$. In this work, we choose this differentiable prior to allow for a wide variety of eligible samplers. In addition, it has been shown recently in [12] that this prior can be easily modified by changing the fixed Laplacian filter by other adaptive off-the-shelf denoisers to achieve state-of-the-art restoration performance. An in-depth study of this prior in a fully Bayesian configuration is out of the scope of this paper and is subject to future work.

2.3. Posterior distribution

Following Bayes rule, the posterior distribution of $\boldsymbol{x}$ is

$$p(\boldsymbol{x}|\boldsymbol{y}, \lambda) \propto p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x}|\lambda) \tag{5}$$

which can be written as $p(\boldsymbol{x}|\boldsymbol{y}, \lambda) = \frac{1}{C}\exp(-U_{\boldsymbol{y}}(\boldsymbol{x}))$, where C is a normalizing constant and $U_{\boldsymbol{y}}(\boldsymbol{x})$ is the negative log-posterior (or energy). Precisely,

$$U_{\boldsymbol{y}}(\boldsymbol{x}) = \mathbf{1}_K^\intercal \boldsymbol{A} e^{\boldsymbol{x}} - \boldsymbol{y}^\intercal \log(\boldsymbol{A} e^{\boldsymbol{x}}) + \frac{\lambda}{2}\boldsymbol{x}^\intercal \boldsymbol{P}\boldsymbol{x} \tag{6}$$

where $[e^{\boldsymbol{x}}]_n = e^{x_n}$ and $[\log(\boldsymbol{x})]_n = \log(x_n)$ are the element-wise exponential and logarithm vectors and $\mathbf{1}_K = (1, \ldots, 1)^\intercal$ is the $K \times 1$ vector of ones. Note that $U_{\boldsymbol{y}}(\boldsymbol{x})$ is differentiable, convex and proper for any $\boldsymbol{y} \in \mathbb{Z}_+^K$. If $\boldsymbol{A}$ is injective, the energy $U_{\boldsymbol{y}}(\boldsymbol{x})$ is also strictly convex with gradient

$$\nabla U_{\boldsymbol{y}}(\boldsymbol{x}) = e^{\boldsymbol{x}} \odot \left(\mathbf{1}_N - \boldsymbol{A}^\intercal(\boldsymbol{y}^\intercal \odot \frac{1}{\boldsymbol{A}e^{\boldsymbol{x}}})\right) \odot (\boldsymbol{A}^\intercal \mathbf{1}_K) + \lambda \boldsymbol{P}\boldsymbol{x}$$

where $[\frac{1}{\boldsymbol{x}}]_n = \frac{1}{x_n}$ and $[\boldsymbol{x} \odot \boldsymbol{y}]_n = x_n y_n$ denote the element-wise vector inversion and multiplication.

3. ESTIMATORS AND INFERENCE STRATEGIES

The stochastic simulations methods presented in this work provide information of the full posterior distribution. In our experiments, we focus on two posterior statistics, the marginal posterior mean of each pixel defined as

$$E\{z_n|\boldsymbol{y}, \lambda\} = \int_0^{+\infty} z_n p(z_n|\boldsymbol{y}, \lambda) dz_n$$

which is the minimum mean squared error estimator (MMSE) of z_n [13], and the marginal posterior variance

$$\text{var}\{z_n|\boldsymbol{y}, \lambda\} = \int_0^{+\infty} (z_n - E\{z_n|\boldsymbol{y}, \lambda\})^2 p(z_n|\boldsymbol{y}, \lambda) dz_n.$$

These integrals are not analytically tractable. Thus we propose to compute numerical estimates $\hat{\mu}_n \approx E\{z_n|\boldsymbol{y}, \lambda\}$ and $\hat{\sigma}_n^2 \approx \text{var}\{z_n|\boldsymbol{y}, \lambda\}$, from samples of $p(\boldsymbol{x}|\boldsymbol{y}, \lambda)$ generated using MCMC samplers or particle trajectories (in the BPS case).

3.1. Markov chain Monte Carlo

MCMC samplers construct a Markov chain of N_{MC} samples $\{\boldsymbol{x}^{(s)}|s = 1, \ldots, N_{\text{MC}}\}$, distributed according to the posterior $p(\boldsymbol{x}|\boldsymbol{y}, \lambda)$. Discarding the first N_{bi} *burn-in* iterations, the posterior statistics of interest are then computed as

$$\hat{\mu}_n = \frac{1}{N_{\text{MC}}} \sum_{s=N_{\text{bi}}}^{N_{\text{MC}}} \exp\left[x_n^{(s)}\right] \tag{7}$$

and

$$\hat{\sigma}_n^2 = \frac{1}{N_{\text{MC}}} \sum_{s=N_{\text{bi}}}^{N_{\text{MC}}} \exp\left[2x_n^{(s)}\right] - \hat{\mu}_n^2. \tag{8}$$

We consider five different MCMC samplers: the classical RWM with a data-independent proposal and four gradient-based samplers, ULA, MALA, HMC and NUTS. In the case

ISBN 978-90-827970-1-5 © EURASIP 2018

of RWM and MALA, we adjust the parameter controlling the variance of the proposal distribution during the *burn-in* to achieve an acceptance rate close to the optimal one, as explained in [8, Chapter 4]. We used the adaptive scheme described in [14], in order to obtain an acceptance rate close to 0.23 for RWM and 0.57 for MALA. For HMC, we fix the path length and adapt the step size using the algorithm described in [9] to obtain an acceptance rate close to the optimal one of 0.65. NUTS controls both step size and number of steps adaptively to yield acceptance rates close to 0.65.

3.2. Bouncy Particle Sampler

A non-reversible rejection-free sampler was recently proposed in [10], performing better than the optimally tuned HMC for some Bayesian models. This algorithm simulates a particle that travels around the posterior with constant velocity $\boldsymbol{v} \in \mathbb{R}^N$ and performs elastic bounces with a hyperplane of normal vector $\nabla U_{\boldsymbol{y}}(\boldsymbol{x})$. If the energy is strictly convex, the sth bouncing time $\tau^{(s)}$ is computed by solving the one dimensional equation

$$U_{\boldsymbol{y}}(\boldsymbol{x}^{(s)} + \boldsymbol{v}^{(s)}\tau^{(s)}) - U_{\boldsymbol{y}}(\boldsymbol{x}^{(s)} + \boldsymbol{v}^{(s)}\tau^{*}) = -\log V \quad (9)$$

where

$$\tau^{*} = \arg\min_{t \geq 0} U_{\boldsymbol{y}}(\boldsymbol{x}^{(s)} + \boldsymbol{v}^{(s)}t) \quad (10)$$

and $V \sim \mathcal{U}(0,1)$. In our configuration, the solutions of Eq. (9) and (10) cannot be computed analytically and we use two one-dimensional Newton iterative methods to find them. The particle trajectory is ergodic, allowing estimates of μ_n and σ_n^2 to be computed as time-averages of the trajectory, i.e.,

$$\hat{\mu}_n = \sum_{s=1}^{N_{\mathrm{B}}} \frac{1}{v_n^{(s)}} \exp\left[x_n^{(s)} + v_n^{(s)}\tau^{(s)}\right] \quad (11)$$

and

$$\hat{\sigma}_n^2 = \sum_{s=1}^{N_{\mathrm{B}}} \frac{1}{2v_n^{(s)}} \exp\left[2x_n^{(s)} + 2v_n^{(s)}\tau^{(s)}\right] - \hat{\mu}_n^2 \quad (12)$$

where N_{B} is the total number of simulated bounces. To assure convergence, the particle also performs random bounces according to a user-specified refreshment rate, which was set to 1 by cross-validation (see [10] for additional details about the BPS algorithm).

3.3. Computational complexity

In a general setting, the matrix $\boldsymbol{A}$ can be dense (e.g., built from many tomographic projections or degraded by a large blurring kernel). The precision matrix $\boldsymbol{P}$ can also be dense, as the prior can account for non-local correlations between image patches [15]. Although this is not the case for the Laplacian filter, we will keep the dense assumption for the sake of generality. Under such assumptions, the conditional distribution of any pixel z_n depends on many neighbours $z_i, i \neq n$. Thus, techniques that exploit a sparse posterior factor graph (e.g., Gibbs sampling using a checkerboard scheme [16], local bouncy particle sampler [10], Hamiltonian splitting approaches [8, Chapter 5]) are not applicable. For each sample generated by an MCMC method or each bounce for NUTS, the algorithm has to compute $U(\boldsymbol{x})$ and/or $\nabla U(\boldsymbol{x})$, which are dominated by the cost of evaluating the forward operator $\boldsymbol{A}e^{\boldsymbol{x}}$ and the Laplacian prior term $\boldsymbol{P}\boldsymbol{x}$. The forward operator $\boldsymbol{A}$ requires N_A flops of order $\mathcal{O}(NK)$, which reduces to $\mathcal{O}(N \log N)$ if $\boldsymbol{A}$ is block-circulant and the Fast Fourier transform (FFT) algorithm is used. A similar cost is related to the prior term $\boldsymbol{P}\boldsymbol{x}$, as it involves one filtering operation of N_P flops with order $\mathcal{O}(N \log N)$. Note that the rest of the operations have order $\mathcal{O}(N)$. We are interested in obtaining good estimators for a small number of N_P and N_A flops. Table 3.3 summarizes the number of per-sample (or bounce) computations of each sampler.

Sampler	Complexity
RWM	$N_A + N_P$
ULA	$2N_A + N_P$
MALA	$2N_A + N_P$
HMC	$L(2N_A + N_P)$
NUTS	$L(2N_A + N_P)$
BPS	$(3k_1 + 2k_2)N_A + N_P$

Table 1. Complexity per iteration of the proposed samplers (k_1, k_2 are the numbers of Newton steps for Eq. (9) and (10)).

4. EXPERIMENTS

We investigated the performance of the samplers as functions of the total number of photon counts and the dimensionality of the problem. All the experiments were performed with the "cameraman" image. We fixed the parameter controlling the amount of regularization $\lambda = 2$ by cross-validation. Figure 1 shows the true intensity image and a noisy realization with 10^4 photons. The *burn-in* was set to 30 % of the computing time in all the simulations. First, we compared

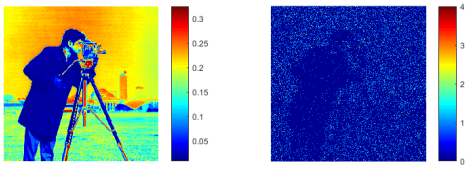

Fig. 1. Left: True intensity image. Right: Noisy observations for a 256×256 image and 10^4 photons.

ISBN 978-90-827970-1-5 © EURASIP 2018

the bias and variance of the estimates $\hat{\mu}_n$ and $\hat{\sigma}_n^2$ of a randomly chosen pixel for image sizes of 64×64, 128×128 and 256×256 pixels with a fixed photon budget of 10^4 photons and an operator $\boldsymbol{A}$ defined as a blur of 3×3 pixels. Each sampler was run 10 times using different random seeds. Figures 2 and 3 show the means and standard deviations of the estimates $\hat{\mu}_n$ and $\hat{\sigma}_n^2$ respectively. HMC achieves the best performance overall, converging without bias in less computing time. NUTS performs slightly worse, as the number of steps per iteration increases significantly faster than HMC with the image size, thus leading to less samples and worse estimates. MALA and ULA have reasonable performance, but ULA shows biased estimates as it does not converge to the exact posterior distribution [17]. The estimates of the means converge relatively fast for BPS, but the estimates of the variances show markedly slower convergence with respect to the other MCMC gradient-based samplers. RWM has the worst performance for this example.

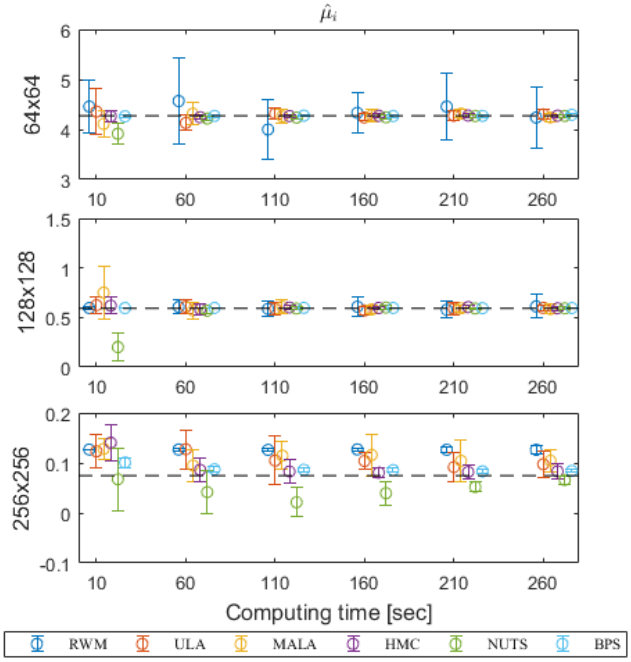

Fig. 2. Mean estimates versus the dimension of the restored image. The error bars show the variation of the estimates for 10 different runs of each algorithm. The true parameter value is denoted by the horizontal dashed line.

Second, we fixed the image size to 64×64 and compared the samplers versus the number of photon counts. Figure 4 shows the results for photon counts of 10^5 and 10^3. When the photon count is lower, the likelihood term Eq. (3) flattens, meaning longer posterior tails and more samples needed to achieve good estimation performance. Again, HMC provides the best results for a given computing time.

The sum of Poisson random variables is Poisson dis-

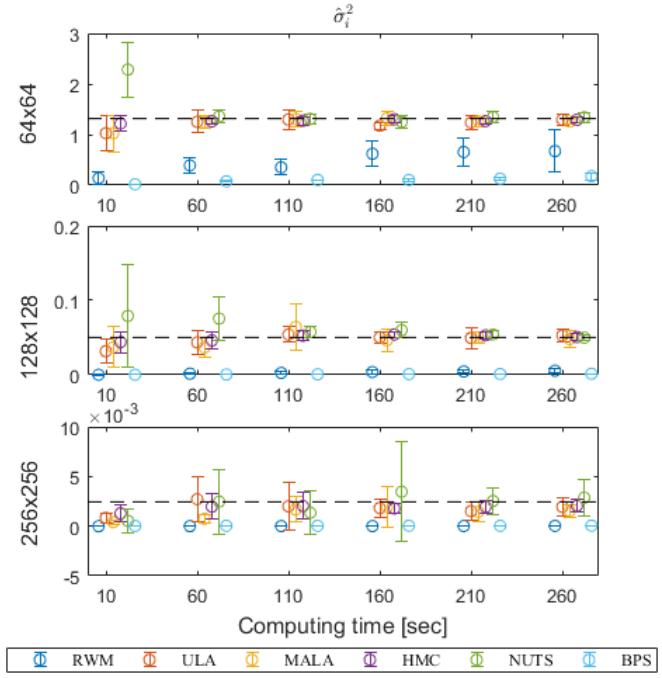

Fig. 3. Variance estimates versus the dimension of the restored image. The error bars show the variation of the estimates for 10 different runs of each algorithm. The true parameter value is denoted by the horizontal dashed line.

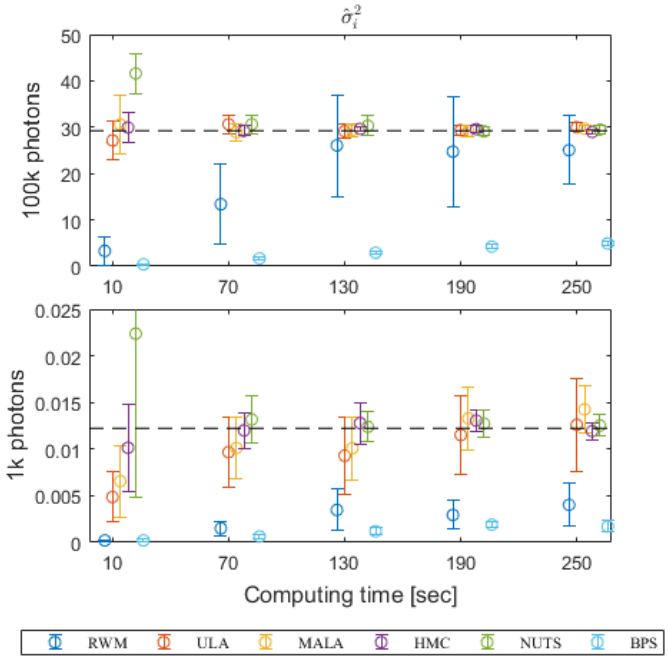

Fig. 4. Variance estimates versus the number of recorded photons. The error bars show the variation of the estimates for 10 different runs of each algorithm. The true parameter value is denoted by the horizontal dashed line.

ISBN 978-90-827970-1-5 © EURASIP 2018

tributed. Hence, binning the noisy image with a 2×2 pixels window yields a 4 times increase in signal-to-noise ratio, at the cost of lower high frequency information. Recent work in the Poisson denoising literature has shown that lower scales can add valuable information [18, 19]. We compared the normalized mean squared error

$$\text{NMSE} = \frac{\sum_{n=1}^{N}(z_n - \hat{\mu}_n)^2}{\sum_{n=1}^{N} z_n^2} \tag{13}$$

obtained by the HMC sampler for different downsampled scales. In this experiment, we removed the forward operator (i.e., the matrix $\boldsymbol{A}$ was set to the identity matrix) and fixed the computing time to 300 seconds. The coarse scale results were upsampled by simply assigning $1/4$ of the estimated intensity to each pixel in the finer scale. Table 2 shows the NMSE of each scale from 10^2 photons to $5\,10^5$ photons. In a very low-photon count scenario (less than 10^4 photons), the estimates at lower scales perform better than at higher ones, also requiring less time to converge. When the photon number increases, the results are better at finer scales, as they contain more high-frequency information.

photons/ p	16	32	64	128	256
10^2	**14.1**	16.4	29.7	67.62	200.7
10^3	10.9	**10.7**	12.08	16.48	47
10^4	7.5	6.7	**6.5**	7.3	10.7
$2\,10^4$	6.9	5.9	**5.4**	5.8	7.9
$5\,10^4$	6.5	5.2	4.5	**4.4**	6.5
$7\,10^4$	6.4	5.0	4.1	**3.9**	4.5
$5\,10^5$	6.0	4.3	2.9	2.2	**2.0**

Table 2. NMSEs at different scales (the image size is $p \times p$).

5. CONCLUSION

This paper studied a Bayesian approach for image restoration using low-photon count measurements. Six different stochastic simulation algorithms were compared to compute posterior statistics of interest, in particular the marginal pixel mean and variance, for this image restoration problem. The experiments performed with synthetic data showed that the Hamiltonian Monte Carlo sampler provides better estimates as the image size increases and the number of photon counts decreases. These results are in agreement with the theoretical scaling of the effective sample size [8] [10]. Finally, we showed experimentally the trade-off between the available photon budget and the scale associated with the best restoration. When the number of photon counts is very low compared to the image size, downsampling the image can yield better results using less computing time.

6. REFERENCES

[1] Y. Altmann, X. Ren, A. McCarthy, G. S. Buller, and S. McLaughlin, "Lidar waveform-based analysis of depth images constructed using sparse single-photon data," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 1935–1946, 2016.

[2] A. Halimi, Y. Altmann, A. McCarthy, X. Ren, R. Tobin, G. S. Buller, and S. McLaughlin, "Restoration of intensity and depth images constructed using sparse single-photon data," in *Proc. European Signal Processing Conf. (EUSIPCO), Budapest, Hungary 2016*. IEEE, 2016, pp. 86–90.

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

[4] B. Zhang, J. M. Fadili, and J.-L. Starck, "Wavelets, ridgelets, and curvelets for poisson noise removal," *IEEE Trans. Image Process.*, vol. 17, no. 7, pp. 1093–1108, 2008.

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

[6] J. Hecht, "Lidar for self-driving cars," *Optics and Photonics News*, vol. 29, no. 1, pp. 26–33, 2018.

[7] C. Robert and G. Casella, *Monte Carlo statistical methods*. Springer, 2004.

[8] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of Markov chain Monte Carlo*. CRC press, 2011.

[9] M. D. Hoffman and A. Gelman, "The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo." *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1593–1623, 2014.

[10] A. Bouchard-Côté, S. J. Vollmer, and A. Doucet, "The bouncy particle sampler: A non-reversible rejection-free Markov chain Monte Carlo method," *to appear in J. Am. Stat. Assoc.*, 2017.

[11] A. Oh and R. Willett, "Regularized non-Gaussian image denoising," *arXiv preprint arXiv:1508.02971*, 2015.

[12] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (RED)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.

[13] C. Robert, *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer, 2007.

[14] Y. F. Atchade, "An adaptive version for the Metropolis adjusted Langevin algorithm with a truncated drift," *Methodology and Computing in applied Probability*, vol. 8, no. 2, pp. 235–254, 2006.

[15] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. Conf. on Comput. Vision Pattern Recognition (CVPR)*, vol. 2, 2005, pp. 60–65.

[16] J. Gonzalez, Y. Low, A. Gretton, and C. Guestrin, "Parallel Gibbs sampling: From colored fields to thin junction trees," in *Proc. Conf. Artificial Intelligence and Statistics*, 2011, pp. 324–332.

[17] G. O. Roberts, R. L. Tweedie *et al.*, "Exponential convergence of Langevin distributions and their discrete approximations," *Bernoulli*, vol. 2, no. 4, pp. 341–363, 1996.

[18] L. Azzari and A. Foi, "Variance stabilization for noisy+ estimate combination in iterative Poisson denoising," *IEEE Signal Proc. Lett.*, vol. 23, no. 8, pp. 1086–1090, 2016.

[19] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *Journal of mathematical imaging and vision*, vol. 48, no. 2, pp. 279–294, 2014.

ISBN 978-90-827970-1-5 © EURASIP 2018