A Bayesian Approach to Denoising of Single-Photon Binary Images

Yoann Altmann, *Member, IEEE*, Reuben Aspden, Miles Padgett, and Steve McLaughlin, *Fellow, IEEE*

Abstract—This paper discusses new methods for processing images in the photon-limited regime where the number of photons per pixel is binary. We present a new Bayesian denoising method for binary, single-photon images. Each pixel measurement is assumed to follow a Bernoulli distribution whose mean is related by a nonlinear function to the underlying intensity value to be recovered. Adopting a Bayesian approach, we assign the unknown intensity field a smoothness promoting spatial and potentially temporal prior while enforcing the positivity of the intensity. A stochastic simulation method is then used to sample the resulting joint posterior distribution and estimate the unknown intensity, as well as the regularization parameters. We show that this new unsupervised denoising method can also be used to analyze images corrupted by Poisson noise. The proposed algorithm is compared to state-of-the art denoising techniques dedicated to photon-limited images using synthetic and real single-photon measurements. The results presented illustrate the potential benefits of the proposed methodology for photon-limited imaging, in particular with non photon-number resolving detectors.

Index Terms—Bayesian estimation, image denoising, Markov chain Monte Carlo methods, Photon-limited imaging, single-photon detection.

I. INTRODUCTION

SINGLE-PHOTON detectors (SPDs) are ubiquitous for applications where the light flux to be analysed is quantified at photonic levels. In particular, SPDs are particularly attractive for imaging applications where the light flux changes rapidly (of the order of picoseconds) or is extremely limited. For instance, the range resolution of SPD-based Lidar systems and their capability to resolve close objects depends on the ability of the detectors to accurately capture the time-of-arrival of photons emitted by fast laser sources [1]–[6].

Recent advances in fast SPDs and SPD arrays, coupled with efficient signal/image processing techniques have allowed the development of extreme imaging systems, including first photon [7] and single pixel [8], [9], and ghost [10]–[12] imaging systems, among others. Improving and investigating new systems however requires the development of statistical methods adapted to the discrete of sparse nature of the recorded data (photon counts or time of arrival). SPDs can be classified into two groups depending on their ability to quantify a number of detected photons within an elementary time period. Although some detectors are photon-number resolving, in this paper we consider SPDs that can generally only distinguish no detection from at least one detection, such as single-photon avalanche diodes (SPADs), photomultiplier tubes and superconducting nanowire SPDs [13]. Although potentially not too restrictive, such limitations need to be considered when developing/applying statistical methods to analyse data recorded by non photon-number resolving SPDs. Indeed, although the number of photons reaching an SPD within a time period is widely assumed to be Poisson distributed (say of mean x), the SPD saturation can have a significant influence on the distribution of the actual photon detections.

In many imaging applications involving such non photon-revolving SPDs, images are formed by summing binary detection images over several independent realizations and assuming the observed phenomenon is stationary (images identically distributed). By ensuring that the probabilities of detection per acquisition for each pixel (i.e., $1 - \exp^{-x}$ assuming the detector has unitary efficiency) are small enough (as a rule-of-thumb generally lower than 5%), the actual distribution of the total number of detected photons in each pixel using T repetitions (e.g. the binomial distribution $\mathcal{B}in\left(T, 1 - \exp^{-x}\right)$), can be approximated by a Poisson distribution with mean Tx. This approximation becomes generally less accurate as x increases (the approximation accuracy depends on x and the number of repetitions T considered). This approximation can also become less accurate when the ratio between the detection interval and the dead-time duration of the detector decreases, where the dead-time period corresponds to a period following a photon detection during which the detector cannot detect additional photons.

If this approach is well adapted to analyze fast low-intensity phenomena for which we naturally have $x << 1$, it requires 1) the intensity field to be constant across the T observations or additional assumptions about its temporal variation (e.g., intensity decay model for fluorescence microscopy [14], [15]) and 2) that $x << 1$ is valid for all the image pixels, which can be problematic when analysing scenes with a high intensity dynamic. Indeed, if the scene includes high intensity regions (i.e., where $x > 5\%$), the illumination source has to be reduced (if possible) or attenuation mechanisms (e.g. filters) used to

Y. Altmann and S. McLaughlin are with the School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K. (e-mail: Y.Altmann@hw.ac.uk; S.McLaughlin@hw.ac.uk).

R. Aspden and M. Padgett are with the School of Physics and Astronomy, University of Glasgow, Glasgow, U.K. (e-mail: Reuben.Aspden@glasgow.ac.uk; Miles.Padgett@glasgow.ac.uk).

ensure the Poisson noise approximation remains valid across all of the pixels. This automatically and artificially reduces the (already low) probability of detection in the darker regions of the image, potentially unnecessarily, to preserve a tractable observation model. This approach is counterproductive as this decrease is usually compensated for by increasing T, the number of repetitions.

In this work, and in contrast with most denoising methods developed for photon-limited data, we focus on applications for which the phenomenon can only be observed once and for which we need to infer the intensity field for each individual image. Consequently, the observation model considered assumes the observed images are binary (i.e., either no photon or at least one photon detected). We also consider the case where the detectors are photon-number resolving (i.e., data corrupted by Poisson noise). Here, we focus primarily on binary images even if the proposed methodology can be applied to Poisson data denoising, i.e., for data recorded by photon-number resolving systems. As will be seen in Section VI, we show that when using non photon-number resolving systems with relatively high detection probabilities ($x \approx 1, 1 - \exp^{-x} \approx 63\%$), it is possible to obtain similar results to photon-number resolving systems. In other words, adopting the appropriate observation model and associated estimation strategy allows for much more efficient data acquisition as it becomes possible to improve the data quality without numerous repetitions (we consider a single detection in this work). However, when saturation is significant, e.g when $x >> 1$, it becomes extremely challenging to accurately estimate the intensity field, in particular using a single frame. In this work, we limit ourselves to $\mathrm{E}\,[x] \leq 1$.

Adopting a classical Bayesian approach, we propose a flexible intensity prior model (see Section III) able to capture different sources of intensity fluctuations such as object movement and changes of the illumination conditions. Starting from the observation model of ideal detectors (Poisson likelihood), we present an alternative observation model accounting for detector limitations (Bernoulli likelihood). Both likelihoods are combined with the prior models and a stochastic simulation method (Markov chain Monte Carlo) method is investigated to exploit the resulting posteriors. An important advantage of the proposed method is that it is fully unsupervised and does not require parameter tuning, as the parameters controlling the spatial and temporal regularizations are automatically adjusted during the sampling process.

The main contributions of this paper can be summarized as follows:

1) We propose a novel, fully automated method to denoise binary, single-photon images. To the best of our knowledge, this work is the first to consider observed pixels following Bernoulli distributions, whose means are related by a nonlinear function to the underlying weak intensity values to be recovered.

2) We generalise the two-dimensional gamma-Markov random field proposed in [6] to capture the temporal correlation affecting the dynamic intensity profiles using a three-dimensional gamma-Markov random field.

3) We develop a new efficient Markov chain Monte Carlo method adapted to the Bayesian models considered, in particular to account for the binary nature of the observed images. Although the Bayesian model considered does not allow for Gibbs updates when sampling the intensity field, we propose an accept-reject procedure which allows the intensities to be updated simultaneously, with a high acceptance rate, which in turn improves the algorithm convergence and reduces its computational complexity.

4) Although dedicated to binary images, we show that the proposed method can also be applied to images corrupted by Poisson noise. In that case, the intensity updates reduce to standard Gibbs updates. Moreover, we demonstrate that for intensities close to 1, the estimation performance using binary images is similar to that obtained using images corrupted by Poisson noise.

The remained of the paper is organized as follow. Section II presents the two observation models considered. The Bayesian models are detailed in Section III and the sampling strategy proposed to exploit the resulting posterior distributions is described in Section IV. Simulation results conducted using synthetic and real single-photon data are discussed in Sections VI and VII. Conclusions and future work are finally reported in Section VIII.

II. OBSERVATION MODELS

Consider a set of T intensity images $\mathbf{X}_t$ of size $N_{\mathrm{row}} \times N_{\mathrm{row}}$ whose elements $x_{i,j,t} = [\mathbf{X}_t]_{i,j}$ are the unknown average numbers of photons reaching the detector array (composed of $N_{\mathrm{row}} \times N_{\mathrm{row}}$ detectors regularly spaced) over a given time period. The two observation models are detailed below.

A. Poisson Likelihood

In the general case where each detector can potentially detect an infinite number of photons (within a given time period Δt), it is widely acknowledged that the distribution of the photon counts $y_{i,j,t}$ in the pixel (i, j) of the tth image can be accurately modelled by a Poisson distribution, i.e.,

$$y_{i,j,t} | (\eta_{i,j}, x_{i,j,t}) \sim \mathcal{P}(\eta_{i,j} x_{i,j,t}) \tag{1}$$

where $\eta_{i,j} > 0$ is an attenuation factor that stands for the detector sensitivity/efficiency and $\mathcal{P}(\eta_{i,j} x_{i,j,t})$ denotes the Poisson distribution with mean $\eta_{i,j} x_{i,j,t}$. In this work we assume that the coefficients $\{\eta_{i,j}\}_{i,j}$ are known (they can usually be estimated during the system calibration) and do not change with time (in the case of video acquisition). Assuming independence between the detectors and the different noise realizations corrupting the images (in particular, we consider non-overlapping acquisition periods) yields the joint likelihood

$$f_0(\mathbf{Y}|\mathbf{N}, \mathbf{X}) = \prod_{i,j,t} f_0(y_{i,j,t}|\eta_{i,j} x_{i,j,t}) \tag{2}$$

where $[\mathbf{Y}]_{i,j,t} = y_{i,j,t}$, $[\mathbf{N}]_{i,j} = \eta_{i,j}$ and $f_0(\cdot|\eta_{i,j} x_{i,j,t})$ is the Poisson distribution defined in (1).

B. Bernoulli Likelihood

Although the Poisson noise assumption is relevant for many imaging applications, here we consider non photon-number resolving SPDs that can only detect at most one photon per pixel within a clock period and need to be reset to potentially detect the next photons reaching the sensor. In the remainder of this paper, we assume that Δt corresponds to this clock period (i.e., the smallest temporal sampling period), which also defines the temporal resolution of the imaging system when recording image sequences. In such cases, the detected photon counts, which become binary measurements within a period Δt, satisfy

$$y_{i,j,t} = \begin{cases} 0 & \text{if } \tilde{y}_{i,j,t} = 0 \\ 1 & \text{if } \tilde{y}_{i,j,t} \geq 1 \end{cases} \tag{3}$$

where $\tilde{y}_{i,j,t} \sim \mathcal{P}(\eta_{i,j} x_{i,j,t})$ is the photon count that would be detected by an ideal detector (able to detect an infinite number of photons). Consequently, in this case the detected photon counts are distributed according to the following Bernoulli distribution

$$y_{i,j,t} \mid (\eta_{i,j}, x_{i,j,t}) \sim \mathcal{B}er\left(1 - \exp^{-\eta_{i,j} x_{i,j,t}}\right) \tag{4}$$

whose mean is given by $1 - \exp^{-\eta_{i,j} x_{i,j,t}}$. In a similar fashion to the observation model described in Section II-A, assuming independence between the detectors and between noise realizations yields

$$f_1(\mathbf{Y}|\mathbf{N}, \mathbf{X}) = \prod_{i,j,t} f_1(y_{i,j,t}|\eta_{i,j} x_{i,j,t}), \tag{5}$$

where $f_1(\cdot|\eta_{i,j} x_{i,j,t})$ denotes the Bernoulli distribution in (4).

It is important to mention here that although Poisson and Bernoulli distributions present different shapes and supports, we have

$$\begin{aligned} f_1(y_{i,j,t} = 0|\eta_{i,j}, x_{i,j,t}) &= f_0(y_{i,j,t} = 0|\eta_{i,j}, x_{i,j,t}) \\ &= \exp^{-\eta_{i,j} x_{i,j,t}} \end{aligned} \tag{6}$$

and

$$f_1(y_{i,j,t} = 1|\eta_{i,j}, x_{i,j,t}) = \sum_{k=1}^{\infty} f_0(y_{i,j} = k|\eta_{i,j}, x_{i,j,t}), \tag{7}$$

which will be useful during the description of the proposed estimation strategy.

The next section describes the Bayesian model and associated estimation strategy proposed to solve the denoising problem considered here; that is, the estimation of the unknown and non-stationary intensity field $\mathbf{X}$ from the observed set of photon counts in $\mathbf{Y}$.

III. BAYESIAN MODEL

A. Intensity Field Modelling

As in most ill-posed inverse problems which need to be regularized, the choice of the regularization or prior model considered for image restoration is crucial both in terms of quality of image recovery and the resulting computational complexity.

In this work we investigate a Bayesian model coupled with an efficient simulation method which allows the estimation of denoised images but that can also provide information about the denoising uncertainty via measures of uncertainty from the posterior distribution of interest. Consequently, we investigate an intensity prior model which allows the use of an simple simulation strategy to exploit the posterior distribution. As will be shown in Sections VI and VII, the prior models presented in this section not only facilitate the estimation strategy but are also flexible enough to compete with standard regularizations used to denoise images corrupted by Poisson noise (e.g., total-variation [16], [17] or Gaussian MRFs [18]).

It is well known that gamma distributions are conjugate priors for the means of Poisson distributions, which makes them particularly attractive to denoise images corrupted by Poisson noise. Moreover, as will be further discussed in Section IV, gamma distributions remain conjugate priors when considering saturating sensors (i.e., assuming (4)), which is particularly convenient in simplifying the denoising problem when the data are Bernoulli distributed.

Due to the spatial organization of images, we expect the values of $x_{i,j,t}$ to vary smoothly from one pixel to another. Moreover, if an image sequence is considered, it might be relevant to capture the temporal correlation between successive images to improve the denoising performance, especially since the sampling period can be extremely short (of the order of nanoseconds or less). In order to model this behaviour, we consider an extended prior model such that the resulting prior for $\mathbf{X}$ is a hidden gamma-MRF (GMRF) [19]. In a similar fashion to [6], we introduce T auxiliary matrices $\mathbf{U}_t$ of size $(N_{\text{row}} + 1) \times (N_{\text{col}} + 1)$ with elements $u_{i,j,t} \in \mathbb{R}^+$ and $T + 1$ additional auxiliary images $\mathbf{W}_t$ of size $N_{\text{row}} \times N_{\text{row}}$. We then define a tripartite conditional independence graph between $\mathbf{X}$, $\mathbf{U} = \{\mathbf{U}_t\}$ and $\mathbf{W} = \{\mathbf{W}_t\}$ such that each $x_{i,j,t}$ of $\mathbf{X}_t$ is connected to four (spatial) neighbors of $\mathbf{U}_t$ and two temporal neighbors in $\mathbf{W}_t$ and $\mathbf{W}_{t+1}$. This 1st order neighbourhood structure is depicted in Fig. 1, where we notice that any given $x_{i,j,t}$ and $x_{i+1,j,t}$ are 2nd order neighbors via $u_{i+1,j,t}$ and $u_{i+1,j+1,t}$. Similarly, $x_{i,j,t}$ and $x_{i,j,t+1}$ are 2nd order neighbors via $w_{i,j,t+1}$. Following the general GMRF model proposed in [19] and specified here by the neighbourhood structure depicted in Fig. 1, we assign $(\mathbf{X}, \mathbf{U}, \mathbf{W})$ a (constrained) GMRF prior, and obtain the joint prior $f(\mathbf{X}, \mathbf{U}, \mathbf{W}|\alpha, \beta)$ (see [19] for the GMRF formulation adopted here). This prior model explicitly depends on the value of the hyperparameters $\alpha > 0$ and $\beta > 0$, which here act as regularization parameters that control the amount of spatial (α) and temporal (β) smoothness enforced by the GMRF. For brevity, we assume that these parameters are fixed and constant across the image sequence in the remainder of this Section. However, following an empirical Bayesian approach, in the results presented in Sections VI and VII, the value of (α, β) is adjusted automatically (either for each image or for all the images) during the early iterations of the sampler by maximum marginal likelihood estimation (the interested reader is invited to consult [6], [20] for details about the estimation of α (or (α, β))).
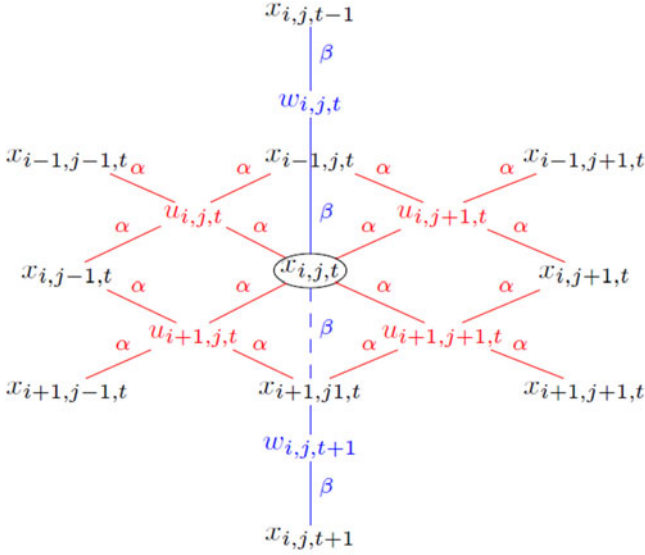
Fig. 1. Proposed 1st order GMRF neighborhood structure $\forall (i, j, t) \in \mathcal{V}_{\mathbf{X}} \times \mathcal{T}$. The red (resp. blue) sub-graph highlights the spatial (resp. temporal) neighborhood structure. For the pixels at the boundaries of the image/ image sequence, we assume the images to be cyclic spatially (e.g., $x_{0,j,t} = x_{N_{\text{row}},j,t}$) and set $x_{i,j,0} = x_{i,j,T+1} = \gamma, \forall (i, j) \in \mathcal{V}_{\mathbf{X}}$. The temporal boundary condition γ is set arbitrarily to the empirical mean of the observed images but the image sequence can also be assumed to be cyclic.

Exploiting the proposed neighbourhood structure yields

$$x_{i,j,t} | \mathbf{U}_t, \mathbf{W}, \alpha, \beta \sim \mathcal{G}_{\mathcal{X}} \left(\alpha + \beta, \tilde{x}_{i,j,t} \right) \qquad (8a)$$

$$u_{i,j,t} | \mathbf{X}_t, \alpha \sim \mathcal{IG} \left(\alpha, \alpha \tilde{u}_{i,j,t} \right) \qquad (8b)$$

$$w_{i,j,t} | \mathbf{X}, \beta \sim \mathcal{IG} \left(\beta, \beta \tilde{w}_{i,j,t} \right) \qquad (8c)$$

where

$$
\begin{aligned}
\tilde{x}_{i,j,t} &= 4/\alpha \left(u_{i,j,t}^{-1} + u_{i-1,j,t}^{-1} + u_{i,j-1,t}^{-1} + u_{i-1,j-1,t}^{-1} \right)^{-1} \\
&\quad + 2/\beta \left(w_{i,j,t}^{-1} + w_{i,j,t+1}^{-1} \right)^{-1} \\
\tilde{u}_{i,j,t} &= \left(x_{i,j,t} + x_{i+1,j,t} + x_{i,j+1,t} + x_{i+1,j+1,t} \right) / 4 \\
\tilde{w}_{i,j,t} &= \left(x_{i,j,t-1} + x_{i,j,t} \right) / 2,
\end{aligned}
$$

and $\mathcal{G}_{\mathcal{X}} (\cdot, \cdot)$ denotes a gamma distribution restricted to $\mathcal{X}$ (this distribution reduces to a standard gamma distribution with $\mathcal{X} = (0, +\infty)$) and $\mathcal{IG} (\cdot, \cdot)$ denotes an inverse-gamma distribution. If a single image or independent images are considered, the GMRF-based prior $f(\mathbf{X}, \mathbf{U}, \mathbf{W} | \alpha, \beta)$ can be simplified by removing the auxiliary variables $\mathbf{W}$ and by considering the prior model $f(\mathbf{X}, \mathbf{U} | \alpha)$, as in [6]. In that case, the neighborhood structure reduces to the red subgroup of Fig. 1 and we obtain

$$x_{i,j,t} | \mathbf{U}_t, \alpha \sim \mathcal{G}_{\mathcal{X}} \left(\alpha, \frac{\bar{x}_{i,j,t}}{\alpha} \right) \qquad (9a)$$

$$u_{i,j,t} | \mathbf{X}_t, \alpha \sim \mathcal{IG} \left(\alpha, \alpha \tilde{u}_{i,j,t} \right) \qquad (9b)$$

where

$$\bar{x}_{i,j,t} = 4 \left(u_{i,j,t}^{-1} + u_{i-1,j,t}^{-1} + u_{i,j-1,t}^{-1} + u_{i-1,j-1,t}^{-1} \right)^{-1}$$

In addition to their flexibility, one of the main motivations for considering GMRFs here is the fact that they enable a simple yet efficient sampling strategy (using the conjugacy of (8a) and (2) or (9a) and (5) (see Eq. (14))), while introducing spatial and temporal dependencies between the neighbouring intensities. As will be shown in Section IV, although we resort to accept/reject procedures when considering binary images, the resulting sampler presents high acceptance rates, which in turn induces good mixing properties and improves the convergence speed of the algorithm.

B. Joint Posterior Distributions

Now that we have defined the prior model for the unknown image or images to be recovered, we can derive the posterior distribution of $(\mathbf{X}, \mathbf{U})$ or $(\mathbf{X}, \mathbf{U}, \mathbf{W})$ (depending on whether temporal correlation is considered), given the observations $\mathbf{Y}$, and the fixed model parameters/hyperparameters $\mathbf{\Phi} = \{\mathbf{N}, \alpha, \beta\}$ and the observation model considered. Using Bayes' rule, we obtain

$$f_m (\mathbf{X}, \mathbf{U} | \mathbf{Y}, \mathbf{\Phi}) \propto f_m (\mathbf{Y} | \mathbf{N}, \mathbf{X}) f(\mathbf{X}, \mathbf{U} | \alpha) \qquad (10)$$

with $m = 0$ (Poisson noise) or $m = 1$ (Bernoulli realizations) for a single or independent images and

$$f_m (\mathbf{X}, \mathbf{U}, \mathbf{W} | \mathbf{Y}, \mathbf{\Phi}) \propto f_m (\mathbf{Y} | \mathbf{N}, \mathbf{X}) f(\mathbf{X}, \mathbf{U}, \mathbf{W} | \alpha, \beta) (11)$$

when considering temporally correlated images, where $f_m (\mathbf{Y} | \mathbf{N}, \mathbf{X})$ is given either by (2) (ideal detector) or (5) (saturated detector). The next paragraph details the Markov chain Monte Carlo (MCMC) method proposed to sample the posteriors of interest (10) and (11) and subsequently estimate the unknown intensity field $\mathbf{X}$.

IV. ESTIMATION STRATEGY

In this work we adopt a simulation based strategy to approximate, for each model, the marginal posterior mean or minimum mean square error (MMSE) estimator of $\mathbf{X}$, i.e.,

$$\widehat{\mathbf{X}} = \mathrm{E} \left[\mathbf{X} | \mathbf{Y}, \mathbf{\Phi} \right], \qquad (12)$$

where the auxiliary variables $\mathbf{U}$ (and $\mathbf{W}$ when considering correlated frames) have been marginalized. Note that by considering the marginal posterior $f_m (\mathbf{X} | \mathbf{Y}, \mathbf{\Phi})$ the corresponding measures of uncertainty automatically accounts for the uncertainty induced by the unknown auxiliary variables in $\mathbf{U}$ (and $\mathbf{W}$).

Although marginalizing analytically $\mathbf{U}$ and $\mathbf{W}$ from (10) and (11) is possible using the structure of the GMRFs $f(\mathbf{X}, \mathbf{U} | \alpha)$ or $f(\mathbf{X}, \mathbf{U}, \mathbf{W} | \alpha, \beta)$, estimating $\mathbf{X}$ directly from $f_m (\mathbf{X} | \mathbf{Y}, \mathbf{\Phi})$ is challenging due to the complexity of this non-standard and high-dimensional distribution. Fortunately, for the two observation models and the two prior models considered, (12) can be efficiently approximated with arbitrarily large accuracy by Monte Carlo integration. More precisely, it is possible to compute (12) by first using an MCMC computational method to generate samples asymptotically distributed according to (10) or (11), and subsequently using these samples to approximate the required marginal expectation.

Here we propose a Gibbs/Metropolis-within-Gibbs sampler to simulate samples from the full posterior of interest, as this type of MCMC method is particularly suitable for models involving hidden Markov random fields [21, Chap. 10]. The output of this algorithm is a Markov chain of N_{MC} samples $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(N_{MC})}$ that are asymptotically distributed according to the marginal posterior distribution $f_m(\mathbf{X}|\mathbf{Y}, \boldsymbol{\Phi})$. The first N_{bi} samples of these chains correspond to the so-called burn-in transient period and should be discarded (the length of this period can be assessed visually from the chain plots or by computing convergence tests [22]). The remaining $N_{MC} - N_{bi}$ samples are used to approximate the Bayesian estimator (12) as follows

$$\widehat{\mathbf{X}} = \frac{1}{N_{MC} - N_{bi}} \sum_{t=N_{bi}+1}^{N_{MC}} \mathbf{X}^{(t)}. \qquad (13)$$

The remainder of this section details the main steps of the proposed samplers, depending on the observation model considered. The main steps of the proposed PID-GMRF and BID-GMRF (for Poisson and Bernoulli image denoising using GMRF) are summarized in the Appendix (see Algo. VIII).

A. Sampling the Auxiliary Variables

Since the auxiliary variables $\mathbf{U}$ and $\mathbf{W}$ do not appear in the likelihoods (2) and (5), sampling from their conditional distributions reduces to sampling from (9b) (single or independent images) or (8b)-(8c) (correlated images), whatever the observation model considered (e.g., Poisson or Bernoulli model). Thanks to the structure of the GMRFs considered, the elements of $\mathbf{U}$ and $(\mathbf{U}, \mathbf{W})$ are a posteriori mutually independent (conditioned on the value of $\mathbf{X}$) and can thus be updated in a parallel manner.

B. Sampling $\mathbf{X}$

Similarly, it is easy to show that for a given realization of $\mathbf{U}$ (and $\mathbf{W}$), the elements of $\mathbf{X}$ are a posteriori independent and can be updated simultaneously. Consider a pixel (i, j, t) following the observation model (1) (Poisson noise). It is easy to obtain using the Poisson-gamma conjugacy that

$$x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \boldsymbol{\Phi} \sim \mathcal{G}_{\mathcal{X}}\left(\alpha + y_{i,j,t}, \frac{\bar{x}_{i,j,t}}{1 + \bar{x}_{i,j,t}\eta_{i,j}}\right) \quad (14)$$

for a single image or independent images and
$$x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi}$$

$$\sim \mathcal{G}_{\mathcal{X}}\left(\alpha + \beta + y_{i,j,t}, \frac{\tilde{x}_{i,j,t}}{1 + \tilde{x}_{i,j,t}\eta_{i,j}}\right) \quad (15)$$

for correlated images. These distributions, denoted as $f_0^{2D}(x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \boldsymbol{\Phi})$ and $f_0^{3D}(x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})$, respectively, can be easily sampled from via rejection sampling, that is, by sampling from non-truncated gamma distributions, in particular when the non-truncated distribution is mainly concentrated in $\mathcal{X}$.

Consider now a pixel (i, j, t), whose observation follows the observation model (4) (Bernoulli realisation). We are interested in the expression of $f_1^{2D}(x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \boldsymbol{\Phi})$ and

$f_1^{3D}(x_{i,j,t}|y_{i,j,t}, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})$, the conditional distributions of $x_{i,j,t}$ using the 2D and 3D GMRFs respectively. By recalling that (6) and (7), it can be shown that

$$f_1^{2D}(x_{i,j,t}|y_{i,j,t} = 0, \mathbf{U}_t, \boldsymbol{\Phi}) = f_0^{2D}(x_{i,j,t}|y_{i,j,t} = 0, \mathbf{U}_t, \boldsymbol{\Phi})$$

and

$$f_1^{3D}(x_{i,j,t}|y_{i,j,t} = 0, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})$$
$$= f_0^{3D}(x_{i,j,t}|y_{i,j,t} = 0, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi}) \quad (16)$$

are truncated gamma distributions. It can also be seen that $f_1^{2D}(x_{i,j,t}|y_{i,j,t} = 1, \mathbf{U}_t, \boldsymbol{\Phi})$ and $f_1^{3D}(x_{i,j,t}|y_{i,j,t} = 1, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})$ are infinite mixtures of gamma distributions which are less trivial to sample from. To tackle this problem, we introduce a Metropolis-Hastings move to update $x_{i,j,t}$ under a Bernoulli observation assumption. More precisely, for a given pixel (i, j, t) at the kth iteration of the sampler, we can use a so-called proposal distribution $q(\cdot)$ defined on $\mathcal{X}$ to generate a candidate x^*. This candidate is then accepted with probability $\mu = \min(\rho, 1)$ where

$$\rho = \frac{f_1^{2D}(x^*|y_{i,j,t}, \mathbf{U}_t, \boldsymbol{\Phi})q(x_{i,j,j}^{(k)})}{f_1^{2D}(x_{i,j,j}^{(k)}|y_{i,j,t}, \mathbf{U}_t, \boldsymbol{\Phi})q(x^*)} \quad (17)$$

and

$$\rho = \frac{f_m^{3D}(x^*|y_{i,j,t}, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})q(x_{i,j,j}^{(k)})}{f_m^{3D}(x_{i,j,j}^{(k)}|y_{i,j,t}, \mathbf{U}_t, \mathbf{W}, \boldsymbol{\Phi})q(x^*)} \quad (18)$$

using the 2D and 3D GMRFs, respectively. Otherwise, we set $x_{i,j,t}^{(k)} = x_{i,j,t}^{(k-1)}$. Here, to avoid additional algorithmic complexity (e.g., tuning the variances of Gaussian random walks), we use as proposal distributions the conditional distributions obtained under Poisson noise assumption (14) or (15), depending on the scenario considered (independent/correlated images). Using this choice of proposal, 1) when $y_{i,j,t} = 0$, we obtain $\rho = 1$ and the Metropolis-Hastings step reduces to a Gibbs step and 2) in practice we have observed that this choice leads to satisfactory acceptance rates ($\rho > 0.6$ for the pixels such such that $y_{i,j,t} = 1$) for all the results presented in Sections VI and VII.

In this Section, we considered two intensity prior models, depending on whether the $T > 1$ observed images are assumed to be temporally correlated or not. As the underlying intensity field is the same if the detectors saturate or not, a single prior model is considered when analysing Poisson or Bernoulli images. We then detailed a single sampling strategy to exploit the posterior distributions of the different scenarios. When the data are Bernoulli observations, accept/reject procedures are only required for the pixels where a detection occurs, i.e., $y_{i,j,t} = 1$, which, in the case of low illumination images ($E[y_{i,j,t}] << 1$), represent a small fraction of the pixels. In the case of Poisson data, the proposed Metropolis-within-Gibbs sampler reduces to a standard, yet highly parallelizable, Gibbs sampler. The following Sections illustrate the potential benefits of the proposed method through a series of experiments conducted using synthetic and real single-photon images.
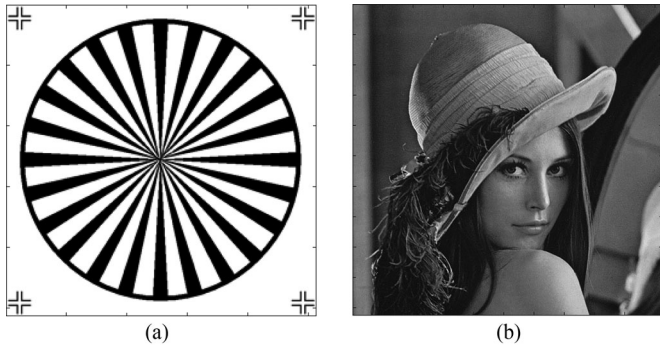
Fig. 2. (a): First image I_1 composed of a piece-wise constant intensity profile. (b) Second image I_2 considered, which presents smoother intensity variations.

TABLE I
AVERAGE NUMBER OF DETECTED PHOTONS PER PIXEL $\mathrm{E}\left[y_{i,j}\right]$ FOR THE TWO IMAGES I_1 AND I_2 CORRUPTED BY POISSON AND BERNOULLI NOISE FOR DIFFERENT VALUES OF $\mathrm{E}\left[x_{i,j}\right]$

| | | $\mathrm{E}\left[x_{i,j}\right]$ | | | | | |
		2.5%	5%	10%	50%	80%	100%
I_1	Poisson	0.025	0.05	0.10	0.50	0.80	1.00
	Bernoulli	0.025	0.05	0.09	0.36	0.48	0.54
I_2	Poisson	0.025	0.05	0.10	0.50	0.80	1.00
	Bernoulli	0.025	0.05	0.09	0.38	0.52	0.59

V. FAULTY SENSOR AND MISSING DATA

In addition to the potential of sensor saturation, faulty detectors and missing data migth occur within the array. As mentioned in the introduction, the proposed denoising framework exploits the spatial correlation between neighbouring pixels to regularize the denoising problem and the presence of spurious pixels can degrade the algorithm performance. To tackle this problem, it is possible to modify the observation models (or equivalently the likelihood functions (2) and (5) so that the spurious values are not used during the denoising process. In such cases, the associated intensities are simply drawn from their conditional prior distribution, i.e. from (9a) (single or independent images) or from (8a) (correlated images), as they no longer appear in the likelihood. We assessed the performance of the proposed methods with up to 0.1% of missing data, uniformly distributed across the field of view and we did not observe a significant degradation of the intensity estimation when compared to observing all the pixels. Although the proposed methods could be used for restoration of sparsely and/or irregularly sampled images, more informative prior models might however provide better results than the proposed GMRFs when the number of missing data becomes more significant. The very interesting and more challenging problem of sparsely sampled images constructed from extremely sparse single-photon data outwith the scope of this paper and will be the subject of future work.

VI. SIMULATIONS USING SYNTHETIC IMAGES

In this Section we investigate the performance of the proposed methods and the effect of the detector saturation on the intensity estimation performance through simulations conducted with synthetic images. First, we compare the performance of the proposed algorithms with existing methods when denoising a single image. Then we assess the benefits of the proposed 3D GMRF, when denoising a sequence of images.

A. Single Image Denoising

We evaluate the proposed methods in denoising the two test images depicted in Fig. 2. The first image of size 256×256 (circular pattern) and denoted I_1, presents a piece-wise constant intensity profile while the second image I_2, of size 512×512 presents smoother intensity variations. In all the experiments

presented in this section, for fair comparisons to methods which cannot handle missing data, we assume that all pixels are observed. We then repeated the same experiments with up to 0.1% of missing data/outliers and did not observe noticeable changes in the denoising performance of the proposed methods. Here, we used $\eta_{i,j} = 1, \forall (i,j)$. The two original images have been scaled such that the expected number of counts (averaged over the image pixels) $\mathrm{E}\left[x_{i,j}\right] \in \{2.5\%; 5\%; 10\%; 50\%; 80\%; 100\%\}$. For each value of $\mathrm{E}\left[x_{i,j}\right]$, $T = 20$ independent noisy images have been generated using the model described by (4). To compare the results with those obtained when the data are corrupted by Poisson noise, we also generate data using (1). Table I gathers details about the mean observed intensity values (averaged over all the pixels and the $T = 20$ noise realizations) for the different scenarios. In contrast to the corruption by Poisson noise, $\mathrm{E}\left[y_{i,j}\right]$ is much smaller than $\mathrm{E}\left[x_{i,j}\right]$ for large values of $\mathrm{E}\left[x_{i,j}\right]$ when considering Bernoulli noise. However, this difference (which also depends on the distribution of $x_{i,j}$ across the pixels) reduces for small values of $\mathrm{E}\left[x_{i,j}\right]$. As expected, for $\mathrm{E}\left[x_{i,j}\right] < 5\%$, the Bernoulli and Poisson distribution are very similar.

We have compared our methods with the following state-of-the art methods: First, we considered a set of methods relying on the Poisson noise assumption. Precisely, we used SPIRAL-TV [17], which solves the same optimization as PIDAL [16]; that is

$$\min_{\mathbf{X} \succeq \mathbf{0}} - \log\left(f_0(\mathbf{Y}|\mathbf{N}, \mathbf{X})\right) + \lambda_{\mathrm{TV}} \sum_{t=1}^{T} \mathrm{TV}\left(\mathbf{X}_t\right), \quad (19)$$

where $\mathrm{TV}\left(\cdot\right)$ is the total variation regularization whose influence is controlled by $\lambda_{\mathrm{TV}} \geq 0$. We also applied the other regularizations proposed in [17] but SPIRAL-TV seems to provide the best and most robust results in this very sparse photon regime, which is why we only present the results obtained with this version of SPIRAL. We also implemented an alternative algorithm which solves the following problem

$$\min_{\mathbf{X} \succeq \mathbf{0}} - \log\left(f_0(\mathbf{Y}|\mathbf{N}, \mathbf{X})\right) + \lambda_{\mathrm{Lap}} \sum_{t=1}^{T} \|\mathbf{D}\mathbf{x}_t\|_2^2, \quad (20)$$

where $\mathbf{x}_t$ is the vectorized version of $\mathbf{X}_t$ and $\mathbf{D}$ is the $N_{\mathrm{row}} \times N_{\mathrm{col}} \times N_{\mathrm{row}} N_{\mathrm{col}}$ circulant convolution matrix of the Laplacian filter [18]. In contrast to the TV regularization which promotes piece-wise constant intensity profiles, the penalization in (20) promotes smooth intensity variations. The problem (20) is solved using an ADMM scheme, similar to that used in

PIDAL, therefore, this method is referred to as "PIDAL-Lap". We also used the NL-PCA algorithm [23] which is a state-of-the-art unsupervised method for image denoising under a Poisson noise assumption (we used the parameter values recommended in [23] in all the experiments presented here).

To the best of our knowledge, there is no published algorithm to directly estimate the intensity profiles involved in (5). However, we implemented an ADMM-based algorithm based on the following convex problem

$$\min_{\mathbf{X} \succeq \mathbf{0}} - \log \left(f_1(\mathbf{Y}|\mathbf{N}, \mathbf{X}) \right) + \lambda_{\text{TV}} \sum_{t=1}^{T} \text{TV}\left(\mathbf{X}_t \right), \qquad (21)$$

This algorithm is referred to as "Ber-TV" in the remainder of this paper. Note that as described in [24], it might be possible to consider other regularizations, for both the Poisson and Bernoulli observation models. However, an extensive comparison of regularizations, potentially using changes of variables and whose regularization parameters need to be carefully adjusted, is outwith the scope of this paper and we concentrate on the TV and Laplacian-based penalizations, whose effects are easily understood and which require the adjustment of a single parameter.

We measure the performance of the different algorithms in term of normalized mean squared error (NMSE) defined by

$$\text{NMSE}_t = \frac{\sum_{i,j} (x_{i,j,t} - \hat{x}_{i,j,t})^2}{\sum_{i,j} x_{i,j,t}^2}, \qquad (22)$$

where $x_{i,j,t}$ (resp. $\hat{x}_{i,j,t}$) is the actual (resp. estimated) intensity value of the pixel (i, j, t). The lower the NMSEs, the more similar the original and reconstructed images. Note that the NMSE does not depend on the intensity dynamic of the original image. We also evaluate how the reconstruction error varies across the image pixels around the NMSE using the standard deviation of the normalized squared error

$$\sqrt{\text{Var}\left[\frac{(x_{i,j,t} - \hat{x}_{i,j,t})^2}{\sum_{i,j} x_{i,j,t}^2 / N_{\text{row}} N_{\text{col}}} \right]}. \qquad (23)$$

We have applied the proposed PID-GMRF and BID-GMRF algorithms with $N_{\text{MC}} = 2000$ (including $N_{\text{bi}} = 600$ burn-in iterations) to the data corrupted by Poisson and Bernoulli noise. We have also applied PID-GMRF to data corrupted by Bernoulli noise to simulate the performance of methods relying on a Poisson noise assumption when denoising data recorded by non photon-number resolving detectors. The PIDAL-TV, PIDAL-Lap and Ber-TV algorithms, requires tuning of a regularization parameter (λ_{TV} or λ_{Lap}). We adopted the methods proposed in [25] to automatically adjust these hyperparameters, but these methods tend to significantly overestimate the smoothness of the intensity field, due to the extreme sparsity of the observed images and thus yield poor results, visually and in terms of NMSEs. Consequently, for the results presented here, these hyperparameters have been optimized in a supervised manner in order to minimize the NMSE, which however requires knowing the actual intensity image in advance.
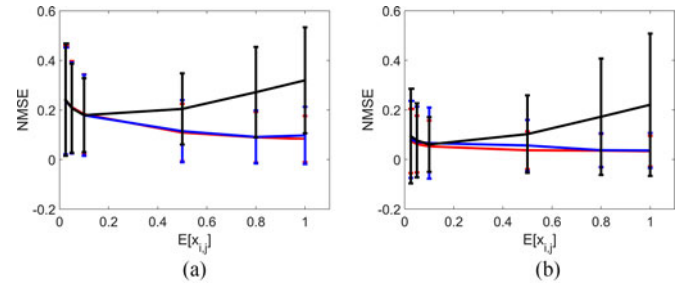

Fig. 3. Average NMSEs obtained with PID-GMRF on images corrupted by Poisson noise (red lines), with BID-GMRF on images corrupted by Bernoulli noise (blue lines) and PID-GMRF on images corrupted by Bernoulli noise (black lines). The vertical bars depict the corresponding confidence regions ($\pm$ standard deviation). The subplot (a) (resp. (b)) corresponds to I_1 (resp. I_2).

We first compare the performance of the two proposed methods when denoising images corrupted by Bernoulli and Poisson noise. Fig. 3 depicts the NMSEs and associated confidence regions ($\pm$ standard deviation), computed for each image and averaged over the $T = 20$ noise realizations. This figure shows that when $\text{E}[x_{i,j}] \to 0$ the results obtained by the two algorithms are similar for the two observation models. This can be explained by the fact that the likelihoods (1) and (4) become similar and weakly informative and that the intensity estimates are mostly driven by the intensity prior model, which is the same in all the scenarios. As expected, using a Poisson observation model when the data are Bernoulli distributed (black lines) yields less accurate intensity estimates when $\text{E}[x_{i,j}]$ increases due to the poor Poisson approximation of the Bernoulli distribution. More surprisingly, the two methods using the correct observation model (blue and red curves) present similar behaviors for all values of $\text{E}[x_{i,j}]$. Indeed, we could expect the intensity estimation to be significantly less accurate when considering Bernoulli observations since the detectors cannot detect all the photons reaching the sensors (at most one per pixel). However, these results show that although a non photon-number resolving detector is used, it is possible to obtain similar intensity estimates to those obtained by an ideal detector (provided that the appropriate observation model is used and that $\text{E}[x_{i,j}]$ is not too large).

Table II compare the NMSEs and associated standard deviations (averaged over $T = 20$ realizations) obtained by Ber-TV, BID-GMRF, NL-PCA and PID-GMRF when denoising the images I_1 and I_2 corrupted by Bernoulli noise. These results confirm that for high values of $\text{E}[x_{i,j}]$, the methods relying on Poisson noise assumption (NL-PCA and PID-GMRF) provide less accurate intensity estimates than Ber-TV and BID-GMRF. Moreover, BID-GMRF is more robust than Ber-TV for small values of $\text{E}[x_{i,j}]$ but is outperformed by Ber-TV in terms of NMSE (when appropriately tuned) when $\text{E}[x_{i,j}] \to 1$. It is important to mention that the NMSE performance has to be moderated by the relatively high standard deviations in Table II. This table also shows that even if BID-GMRF does not necessarily provide lower NMSEs, it generally provides lower standard deviations, in particular for small values of $\text{E}[x_{i,j}]$.

Figs. 4 and 5 compare examples of single image denoising using $\text{E}[x_{i,j}] = 5\%$ and $\text{E}[x_{i,j}] = 2.5\%$. These results illustrate the fact the GMRF considered is flexible enough to capture the

TABLE II
AVERAGE NORMALIZED MEAN SQUARE ERRORS (NMSEs) OBTAINED BY DIFFERENT ALGORITHMS FOR THE IMAGES I_1 AND I_2 CORRUPTED BY BERNOULLI NOISE VERSUS $\mathrm{E}\,[y_{i,j}]$

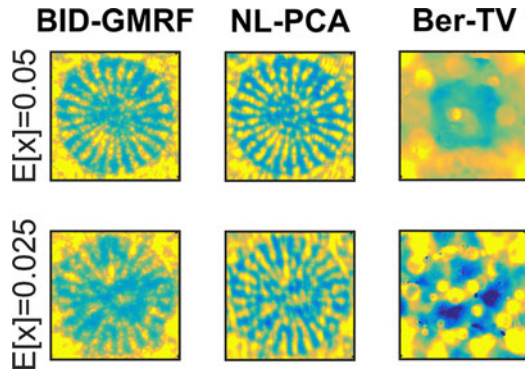| | | $\mathrm{E}\,[x_{i,j}]$ | | | | | |
		2.5%	5%	10%	50%	80%	100%
I_1	Ber-TV	0.383 (0.533)	0.309 (0.350)	0.255 (0.214)	**0.090** (0.220)	**0.068** (0.174)	**0.059** (0.157)
	BID-GMRF	**0.237** (0.215)	0.208 (0.184)	0.179 (0.164)	0.115 (0.125)	0.091 (0.107)	0.097 (0.115)
	NL-PCA	0.286 (0.888)	0.234 (0.207)	**0.159** (0.172)	0.143 (0.135)	0.206 (0.169)	0.331 (0.200)
	PID-GMRF	0.242 (0.227)	**0.207** (0.180)	0.179 (0.149)	0.204 (0.143)	0.272 (0.182)	0.320 (0.214)
I_2	Ber-TV	0.191 (0.307)	0.091 (0.157)	0.065 (0.126)	**0.042** (0.088)	**0.032** (0.077)	**0.032** (0.077)
	BID-GMRF	**0.081** (0.155)	**0.071** (0.142)	0.066 (0.143)	0.056 (0.103)	0.037 (0.068)	0.036 (0.071)
	NL-PCA	0.218 (0.757)	0.169 (0.145)	0.097 (0.092)	0.049 (0.145)	0.073 (0.228)	0.177 (0.283)
	PID-GMRF	0.095 (0.191)	0.077 (0.150)	**0.060** (0.110)	0.102 (0.156)	0.173 (0.235)	0.221 (0.287)

Fig. 4. Intensity estimates for I_1 (corrupted using (4)), using BID-GMRF, NL-PCA and Ber-TV, and for different values of $\mathrm{E}\,[x_{i,j}]$. For each row, the same scale $(0, \max(x_{i,j}))$ is used for the three methods.
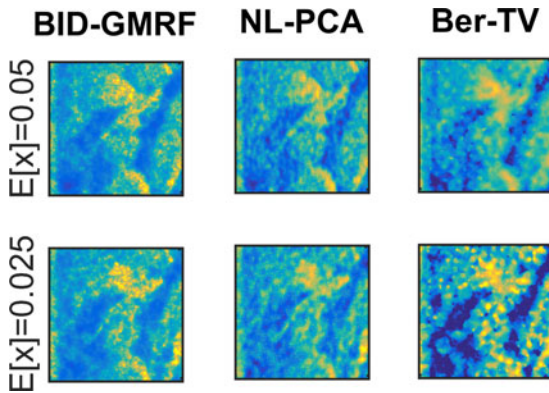

Fig. 5. Intensity estimates for I_2 (corrupted using (4)), using BID-GMRF, NL-PCA and Ber-TV, and for different values of $\mathrm{E}\,[x_{i,j}]$. For each row, the same scale $(0, \max(x_{i,j}))$ is used for the three methods.

spatial correlation of piece-wise constant (I_1) and smoother (I_2) images and is visually more robust than Ber-TV (less prominent patch-like artifacts). NL-PCA provides similar images and is able to detect spatial structure in the data but underestimate the large intensities due to the model mismatch, yielding higher NMSE when $\mathrm{E}\,[x_{i,j}] \to 1$ (see Table II).

Table III compares the NMSEs (averaged over $T = 20$ realizations) obtained by SPIRAL-TV, PIDAL-Lap, NL-PCA and PID-GMRF when denoising the images I_1 and I_2 corrupted by Poisson noise. These results shows that the proposed Bayesian approach, when assuming Poisson noise, pro-

vides more robust results than the other state-of-the-art methods when $\mathrm{E}\,[y_{i,j}] \to 0$. When $\mathrm{E}\,[y_{i,j}] \to 1$ however, the three other methods generally yield slightly better NMSEs. Although SPIRAL-TV and PIDAL-Lap need to be tuned to obtain such performance, NL-PCA does not which open routes to further improve the denoising performance of BID-GMRF, e.g., using dictionary techniques such as NL-PCA, in particular when $\mathrm{E}\,[y_{i,j}] \to 1$.

B. Denoising of Image Sequences

We now illustrate the benefits of the proposed 3D GMRF model when denoising videos constructed from single-photon data. We consider a video composed of $T = 141$ frames of size 240×310 pixels, which represents someone striking a xylophone. This video has been selected from the video library available in Matlab R2014b. In the experiments presented in this section, there is no missing data and we used $\eta_{i,j} = 1, \forall(i,j)$. The original video has been scaled such that the expected number of counts (averaged over the image pixels and frames) $\mathrm{E}\,[x_{i,j,t}] \in \{2.5\%; 5\%; 10\%; 50\%; 80\%; 100\%\}$. We have applied the proposed PID-GMRF (resp. BID-GMRF) algorithms with $N_{\mathrm{MC}} = 3000$, including $N_{\mathrm{bi}} = 1000$ burn-in iterations, to the data corrupted by Poisson (resp. Bernoulli) noise. The methods using 2D (resp. 3D) GMRFs are denoted PID-GMRF-2D and BID-GMRF-2D (resp. PID-GMRF-3D and BID-GMRF-3D). Fig. 6 compares the NMSEs, obtained by the proposed algorithms (using the correct observation model). These plots show that the NMSEs generally increase as $\mathrm{E}\,[x_{i,j,t}]$ decreases and that the NMSEs are similar across the T frames when using PID-GMRF-2D and BID-GMRF-2D. When PID-GMRF-3D and BID-GMRF-3D are used instead, the NMSEs generally decrease due to the consideration of the temporal correlation between successive frames. Moreover, the consideration of temporal correlation also reduces the standard deviations of the normalized squared errors (when compared to the 2D-GMRF). Thus the 3D-GMRF tends to prevent large deviations between the true and estimated intensities. Note that the NMSEs increase at the very beginning and the very end of the sequences due to the GMRF boundary conditions considered. This bias can however be easily reduced if we further assume that the temporal sequence is cyclic. In order not to add unnecessary assumptions in the general case, we did not present this case

TABLE III
AVERAGE NORMALIZED MEAN SQUARE ERRORS (NMSEs) OBTAINED BY DIFFERENT ALGORITHMS FOR THE IMAGES I_1 AND I_2 CORRUPTED BY POISSON NOISE VERSUS $\mathrm{E}\left[y_{i,j}\right]$

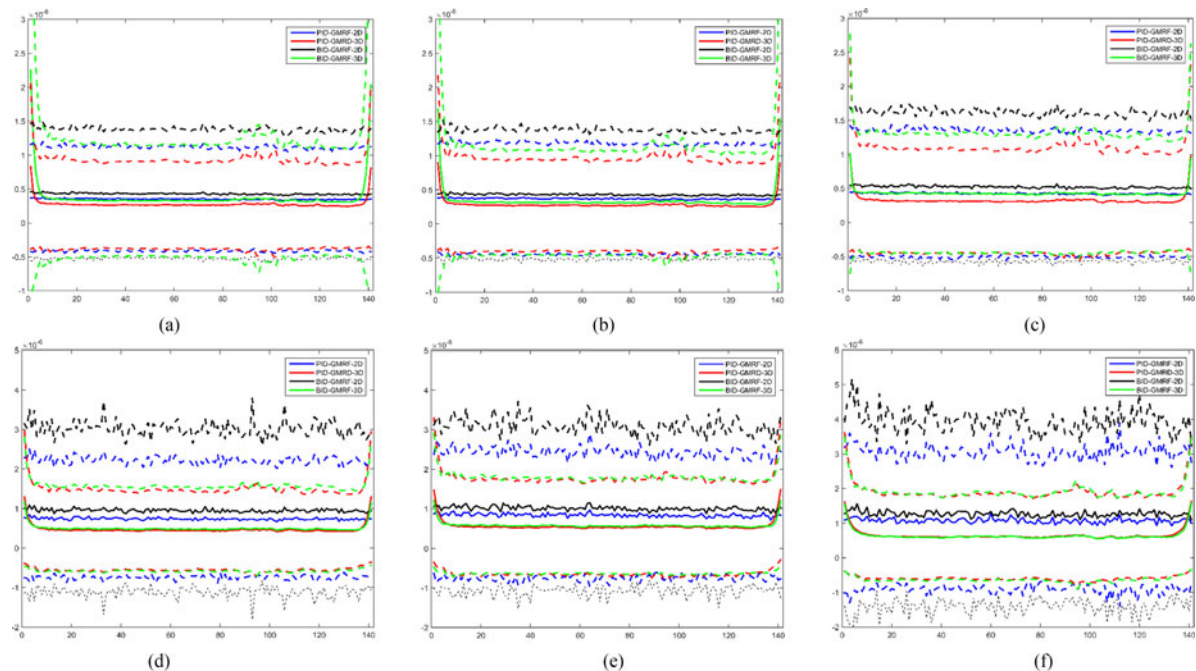| | | $\mathrm{E}\left[x_{i,j}\right]$ | | | | | |
		2.5%	5%	10%	50%	80%	100%
I_1	SPIRAL-TV	0.410 (0.564)	0.301 (0.366)	0.243 (0.360)	0.085 (0.221)	0.062 (0.176)	**0.051** (0.143)
	PIDAL-Lap	0.258 (0.289)	0.221 (0.205)	0.150 (0.209)	**0.073** (0.124)	**0.060** (0.105)	0.054 (0.098)
	NL-PCA	0.346 (1.208)	**0.210** (0.238)	**0.148** (0.174)	0.077 (0.115)	0.073 (0.110)	0.072 (0.108)
	PID-GMRF	**0.240** (0.220)	0.212 (0.186)	0.183 (0.161)	0.109 (0.116)	0.089 (0.102)	0.083 (0.093)
I_2	SPIRAL-TV	0.167 (0.258)	0.083 (0.144)	0.067 (0.128)	0.028 (0.071)	0.027 (0.070)	0.025 (0.061)
	PIDAL-Lap	**0.065** (0.124)	**0.052** (0.104)	**0.042** (0.088)	0.026 (0.064)	0.022 (0.057)	0.021 (0.052)
	NL-PCA	0.186 (0.938)	0.075 (0.174)	0.048 (0.092)	**0.022** (0.054)	**0.019** (0.050)	**0.018** (0.049)
	PID-GMRF	0.075 (0.129)	0.062 (0.114)	0.053 (0.104)	0.037 (0.076)	0.036 (0.067)	0.034 (0.062)

Fig. 6. NMSEs (solid lines) obtained with PID-GMRF and BIP-GMRF (2D and 3D versions) on the synthetic videos of the xylophone for $\mathrm{E}\left[x_{i,j}\right]=1$ (a), $\mathrm{E}\left[x_{i,j}\right]=0.8$ (b), $\mathrm{E}\left[x_{i,j}\right]=0.5$ (c), $\mathrm{E}\left[x_{i,j}\right]=0.1$ (d), $\mathrm{E}\left[x_{i,j}\right]=0.05$ (e) and $\mathrm{E}\left[x_{i,j}\right]=0.025$ (f). The dashed lines depict the corresponding confidence regions ($\pm$ standard deviation).

9which can be addressed by changing the GMRF boundary conditions (see Fig. 1)].

VII. SIMULATIONS USING REAL DATA

We illustrate the benefits of the proposed denoising framework to denoise sparse images of an object recorded by a ghost-imaging system similar to those considered in [26], [27]. The system considered here uses correlated photons at 710nm and the images were displayed on a spatial light modulator (SLM). We consider a set of 12 spatial patterns, i.e, 12 smiley faces gradually changing from a sad to happy face. The images of size 256×256 are recorded by an intensified camera with a CCD detector array (ICCD) triggered by a Perkin Elmer silicon SPAD (see [26], [27] for more details about data acquisition and setup of the ghost-imaging instrument). Each face is observed over 300 seconds with a frame rate of 1Hz, leading to 300 frames per face position. The ICCD acts here as a non photon-number resolving SPD, and thus provides binary images. The average intensity profile relates to the image of the faces formed from a polished silicon wafer onto which was patterned a microscopic gold test target. At the wavelength considered, the silicon is transparent whereas the gold layer is not. Consequently, the acquired images are darker in the region where the gold target is present. The laser source is adjusted so that the average number of detected photons per pixel and per frame is significantly lower than 5% ($\mathrm{E}\left[x_{i,j,t}\right] = 2.3 \times 10^{-4}$ and $x_{i,j,t} < 2\%, \forall (i,j,t)$). In other word, the probability of having more than one photon reaching a given pixel within a given $1s$ frame is extremely low. In this extremely sparse photon-limited imaging regime, the distributions of the photon count can be approximated by Poisson distributions. Fig. 7 depicts the accumulated photon counts obtained by summing the 300 images associated with each position.
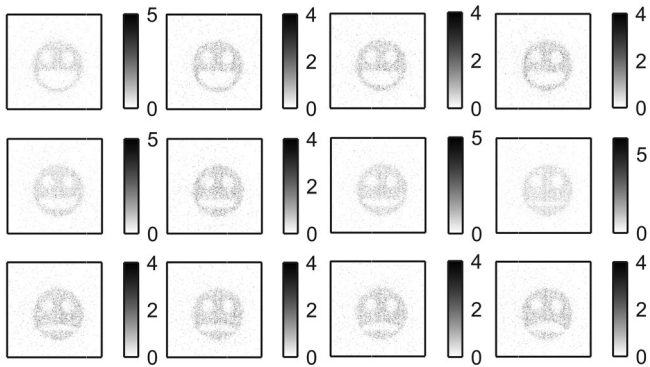
Fig. 7. Measured photon counts, obtained by integration of groups of 300 successive images.

TABLE IV
AVERAGE NORMALIZED MEAN SQUARE ERRORS (NMSEs) ($\times 10^{-2}$)
OBTAINED FOR DIFFERENT EXPOSURE TIMES/NUMBER OF FRAMES PER GROUP.

	Integration time per frame (in seconds)			
	25	50	100	300
Per frame photon counts	442.3	844.7	1689.4	5068.1
Per frame detection counts	417.3	824.0	1608.7	4379.3
Per pixel detection rate	0.64%	1.26%	2.45%	6.68%
NMSE	0.220	0.071	0.050	0.017
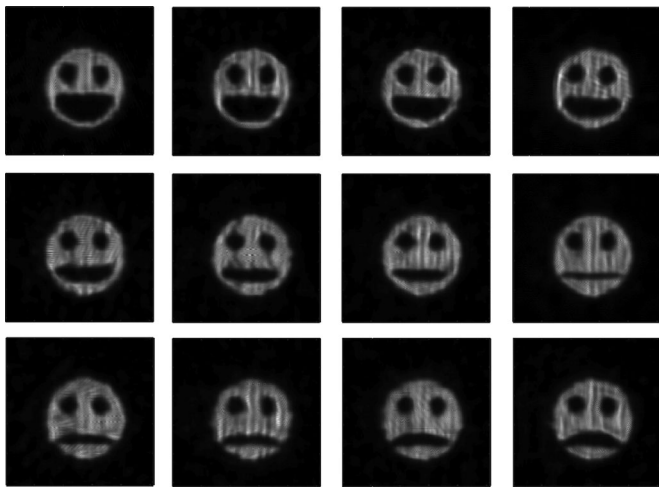
Fig. 8. Images denoised using NL-PCA and used as reference to compute NMSEs. The input images are obtained by summing groups of 300 successive original images during which the intensity field is stationary.

To illustrate the benefits of the proposed denoising method, we denoise images that would have been obtained using exposure times of 25 s, 50 s, 100 s and 300 s. Such images are obtained by integrating non-overlapping groups of 25 up to 300 images. These images (approximately corrupted by Poisson noise) are then used to produce binary images associated with the presence/absence of detected photons within successive 25s, 50s, 100s or 300 s periods. The top rows of Table IV show the average number of photons in the images to be enhanced. In particular, less than 500 photons per frame are available when considering the shortest exposure, which corresponds to a per pixel detection
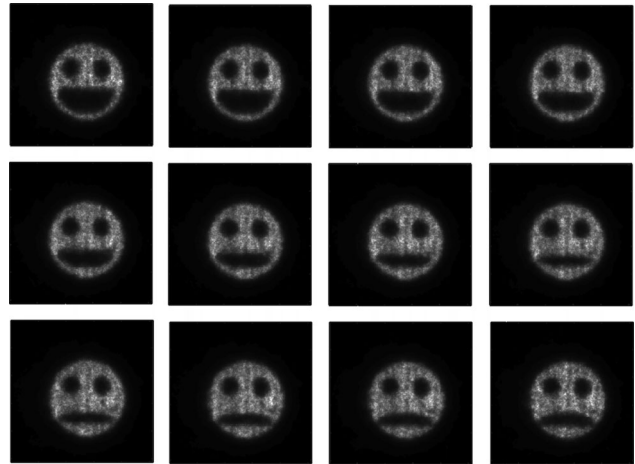

Fig. 9. Examples of images using BID-GMRF-3D. The input images are obtained by summing groups of 300 successive original images, during which the underlying intensity field is stationary. The images are then thresholded (presence/absence of detected photons) to simulate longer integration times (300 s here).
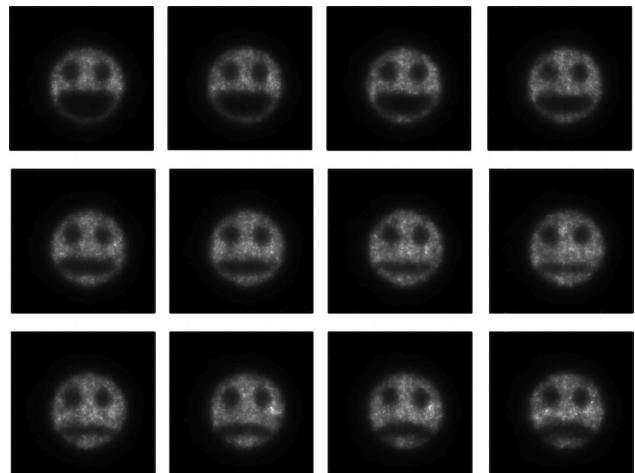

Fig. 10. Examples of images using BID-GMRF-3D. The input images are obtained by summing groups of 50 successive original images, during which the underlying intensity field is stationary. The images are then thresholded (presence/absence of detected photons) to simulate longer integration times (50s here). Each face position is thus visible in six successive images and the images presented correspond to the first image of each position.

rate less than 1%. The bottom row of Table IV compares the NMSEs obtained using BID-GMRF-2D as denoising method, where the reference intensities, depicted in Fig. 8, are those obtained with NL-PCA on the integrated groups of 300 images. Note that the images depicted in Fig. 8 present some vertical artifacts and should not be considered as absolute ground truth. However, since NL-PCA provides the visually most accurate enhanced images (over all the existing methods considered in Section VI), this algorithm has been used as reference. Note that given the reduced numbers of frames considered (only 1 frame per pattern for the longest exposure of 300 s), we did not use BID-GMRF-3D which would have exhibit significantly estimation bias at the beginning and at the end of the sequence, as in Fig. 6. Moreover, due to the low photon counts depicted in Table IV, PID-GMRF-2D provides similar results which are thus not omitted. For completeness, examples of denoised
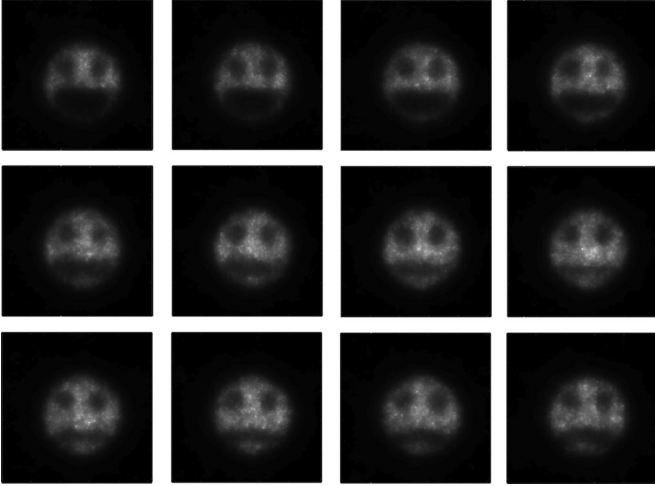
Fig. 11.   Examples of images using BID-GMRF-3D. The input images are obtained by summing groups of 25 successive original images, during which the underlying intensity field is stationary. The images are then thresholded (presence/absence of detected photons) to simulate longer integration times (25s here). Each face position is thus visible in twelve successive images and the images presented correspond to the first image of each position.

images extracted from the whole image sequences are provided in Figs. 9, 10, and 11. These results show that it is possible to use much lower frame rates (combined with lower overall exposures) and still obtain satisfactory intensity field estimates, which can be particularly useful to reduce the amount of data (divided by up to 300 here) to be stored, transmitted and/or processed. These results also show that the proposed method can be used to enhance image sequences of dynamic scenes constructed from extremely sparse single-photon data.

VIII. CONCLUSION

Here we have proposed a new Bayesian method for binary image denoising. The model considered assumed that each pixel measurement follows a Bernoulli distribution whose mean is related by a nonlinear function to the underlying intensity value to be recovered. In contrast with classical Poisson noise models, this model is particularly adapted for data recorded single-photon detectors which are not photon-number resolving, especially when the unknown mean intensity value tends to 1. A gamma Markov random field was proposed to design an intensity prior model able and capture the spatial and temporal structures of the unknown intensity field. A Markov chain Monte Carlo method was then developed to exploit the resulting posterior distribution and estimate the parameters of interest, including the regularization parameters of the Markov random field (thus avoiding parameter tuning via cross-validation). By including a minor modification of the algorithm, we have shown that the proposed method can also be applied to data corrupted by Poisson noise. A series of simulations conducted on synthetic data demonstrated the benefits (robustness) of the proposed method, especially for extremely sparse data. Moreover, we have demonstrated that the proposed version assuming Poisson noise is able to compete with state-of-the art denoising methods (based on a Poisson

noise assumption). We have shown that for average intensities close to 1, it is possible to obtain from saturating sensors, an estimation accuracy close to that obtained using non-saturating sensors. For instance, the results of simulations conducted using real sparse single-photon measurements illustrated how one can reduce the amount of data (by reducing the frame rate here but one could also adjust the laser source and reduce the overall acquisition time when possible) while being able to estimate the intensity profile without a significant performance degradation.

Here we used a hidden gamma Markov random field to build a prior model. However, we noticed that dictionary learning techniques (such as NL-PCA) can significantly improve the denoising performance in the presence of sparse single-photon images. Including such considerations in future binary image denoising methods is clearly interesting. Moreover, the generalization of the proposed methodology for images following binomial distributions (e.g., sum of binary images) is currently under investigation.

APPENDIX: PROPOSED SAMPLING STRATEGY

The main steps of the proposed PID-GMRF and BID-GMRF methods detailed in Section IV are summarized in Algo. VIII below.

Algorithm 1: Poisson/Bernoulli image denoising (PID-GMRF/BID-GMRF)

1: *Fixed input parameters:* α, β, number of burn-in iterations N_{bi}, total number of iterations N_{MC}, temporal correlation binary label $z_{3D} \in (0,1)$, observation model m (0 for Poisson and 1 for Bernoulli).

2: *Initialization* ($k = 0$)

3: Set $\mathbf{X}^{(0)}, \mathbf{U}^{(0)}$ and $\mathbf{W}^{(0)}$

4: *Iterations* ($1 \leq k \leq N_{MC}$)

5: Sample $\mathbf{U}^{(k)} \sim f(\mathbf{U}^{(k)}|\mathbf{X}^{(k-1)}, \alpha)$ in (8b)

6: **if** $z_{3D} = 1$ **then**

7:    Sample $\mathbf{W}^{(k)} \sim f(\mathbf{W}|\mathbf{X}^{(k-1)}, \beta)$ in (8c)

8: **end if**

9: **for** $(i, j, t) \in V_{\mathbf{X}} \times \mathcal{T}$ **do**

10:    **if** m=0 **then**

11:       Sample $x_{i,j,t}^k$ using ($\mathbf{U}^{(k)}, \mathbf{W}^{(k)}$) and (14) ($z_{3D} = 0$) or (15) ($z_{3D} = 1$)

12:    **else**

13:       Sample x^* using (14) ($z_{3D} = 0$) or (15) ($z_{3D} = 1$)

14:       Compute ρ and μ using (17) ($z_{3D} = 0$) or (18) ($z_{3D} = 1$)

15:       Sample $\nu \sim \mathcal{U}_{(0,1)}(\nu)$

16:       **if** $\nu < \mu$ **then**

17:          Set $x_{i,j,t}^{(k)} = x^*$

18:       **else**

19:          Set $x_{i,j,t}^{(k)} = x_{i,j,t}^{(k-1)}$

20:       **end if**

21:    **end if**

22: **end for**

23: Optional: Update α or (α, β) using [20].

24: Set $k = k + 1$.

REFERENCES

[1] A. McCarthy, R. J. Collins, N. J. Krichel, V. Fernández, A. M. Wallace, and G. S. Buller, "Long-range time-of-flight scanning sensor based on high-speed time-correlated single-photon counting," *Appl. Opt.*, vol. 48, no. 32, pp. 6241–6251, Nov. 2009.

[2] N. J. Krichel, A. McCarthy, and G. S. Buller, "Resolving range ambiguity in a photon counting depth imager operating at kilometer distances," *Opt. Express*, vol. 18, no. 9, pp. 9192–9206, Apr. 2010.

[3] A. M. Wallace, J. Ye, N. J. Krichel, A. McCarthy, R. J. Collins, and G. S. Buller, "Full waveform analysis for long-range 3d imaging laser radar," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, pp. 896–708, 2010.

[4] A. McCarthy *et al.*, "Kilometer-range depth imaging at 1550 nm wavelength using an InGaAs/InP single-photon avalanche diode detector," *Opt. Express*, vol. 21, no. 19, pp. 22 098–22 113, Sep. 2013.

[5] A. Maccarone *et al.*, "Underwater depth imaging using time-correlated single-photon counting," *Opt. Express*, vol. 23, no. 26, pp. 33 911–33 926, Dec. 2015.

[6] 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, May 2016.

[7] A. Kirmani *et al.*, "First-photon imaging," *Science*, vol. 343, no. 6166, pp. 58–61, 2014.

[8] Y. Zhang, M. P. Edgar, B. Sun, N. Radwell, G. M. Gibson, and M. J. Padgett, "3d single-pixel video," *J. Opt.*, vol. 18, no. 3, 2016, Art. no. 035203.

[9] M.-J. Sun, M. P. Edgar, D. B. Phillips, G. M. Gibson, and M. J. Padgett, "Improving the signal-to-noise ratio of single-pixel imaging using digital microscanning," *Opt. Express*, vol. 24, no. 10, pp. 10 476–10 485, May 2016.

[10] R. S. Aspden, D. S. Tasca, R. W. Boyd, and M. J. Padgett, "EPR-based ghost imaging using a single-photon-sensitive camera," *New J. Phys.*, vol. 15, no. 7, 2013, Art. no. 073032.

[11] R. S. Aspden *et al.*, "Photon-sparse microscopy: Visible light imaging using infrared illumination," *Optica*, vol. 2, no. 12, pp. 1049–1052, Dec. 2015. [Online]. Available: http://www.osapublishing.org/optica/abstract.cfm?URI=optica-2-12-1049

[12] P. A. Morris, R. S. Aspden, J. E. C. Bell, R. W. Boyd, and M. J. Padgett, "Imaging with a small number of photons," *Nature Commun.*, vol. 6, 2015, Art. no. 5913.

[13] M. D. Eisaman, J. Fan, A. Migdall, and S. V. Polyakov, "Invited review article: Single-photon sources and detectors," *Rev. Sci. Instrum.*, vol. 82, no. 7, 2011, Art. no. 071101.

[14] J. M. Zwier, G. J. V. Rooij, J. W. Hofstraat, and G. J. Brakenhoff, "Image calibration in fluorescence microscopy," *J. Microsc.*, vol. 216, no. 1, pp. 15–24, 2004.

[15] A. Jezierska, H. Talbot, C. Chaux, J.-C. Pesquet, and G. Engler, "Poisson-Gaussian noise parameter estimation in fluorescence microscopy imaging," in *Proc. IEEE Int. Symp. Biomed. Imaging*, Barcelona, Spain, May 2012, pp. 1663–1666.

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

[17] 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 Process.*, vol. 21, no. 3, pp. 1084–1096, Mar. 2012.

[18] F. Orieux, O. Feron, and J. F. Giovannelli, "Sampling high-dimensional gaussian distributions for general linear inverse problems," *IEEE Signal Process. Lett.*, vol. 19, no. 5, pp. 251–254, May 2012.

[19] O. Dikmen and A. Cemgil, "Gamma markov random fields for audio source modeling," *IEEE Trans. Audio, Speech,, Lang. Process.*, vol. 18, no. 3, pp. 589–601, Mar. 2010.

[20] M. Pereyra, N. Whiteley, C. Andrieu, and J.-Y. Tourneret, "Maximum marginal likelihood estimation of the granularity coefficient of a Potts-Markov random field within an MCMC algorithm," in *Proc. IEEE-SP Workshop Statist. Signal Process.*, Gold Coast, Australia, Jul. 2014, pp. 121–124.

[21] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York, NY, USA: Springer-Verlag, 2004.

[22] C. P. Robert and D. Cellier, "Convergence control of MCMC algorithms," in *Discretization and MCMC Convergence Assessment*, C. P. Robert, Ed. New York, NY, USA: Springer-Verlag, 1998, pp. 27–46.

[23] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson noise reduction with non-local PCA," *J. Math. Imaging Vis.*, vol. 48, no. 2, pp. 279–294, 2014.

[24] A. Oh and R. Willett, "Regularized non-Gaussian image denoising," *Math.-Opt. Control*, Aug. 2015. [Online]. Available: http://adsabs.harvard.edu/abs/2015arXiv150802971O.

[25] M. Pereyra, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Maximum-a-posteriori estimation with unknown regularisation parameters," in *Proc. Eur. Signal Process. Conf.*, Aug. 2015, pp. 230–234.

[26] D. S. Tasca, R. S. Aspden, P. A. Morris, G. Anderson, R. W. Boyd, and M. J. Padgett, "The influence of non-imaging detector design on heralded ghost-imaging and ghost-diffraction examined using a triggered ICCD camera," *Opt. Express*, vol. 21, no. 25, pp. 30 460–30 473, Dec. 2013.

[27] R. S. Aspden, P. A. Morris, R. He, Q. Chen, and M. J. Padgett, "Heralded phase-contrast imaging using an orbital angular momentum phase-filter," *J. Opt.*, vol. 18, no. 5, 2016, Art. no. 055204. [Online]. Available: http://stacks.iop.org/2040-8986/18/i=5/a=055204

Yoann Altmann (M'14) was born in Toulouse, France, in 1987. He received the Engineering degree in electrical engineering from Ecole Nationale Supérieure d'Electronique, d'Electrotechnique, d'Informatique, d'Hydraulique et des Téécommunications, Toulouse, France, and the M.Sc. degree in signal processing from the National Polytechnic Institute of Toulouse (INP Toulouse), Toulouse, France, both in 2010, and the Ph.D. degree from INP Toulouse, Toulouse, France, in 2013. From 2014 until 2016, he was with the Heriot-Watt University, Edinburgh, as a Post-Doctoral Researcher. Since 2017, he has been an Assistant Professor in the Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, U.K. His current research interests include statistical signal and image processing, with a particular interest in Bayesian inverse problems with applications to remote sensing and biomedical imaging.

Reuben Aspden received the Graduate degree and the MSci. degree in physics in 2011 from the University of Glasgow, Glasgow, Scotland, and the Ph.D. degree in 2015. He continued his work on imaging using single photons as part of the Quantic UK Quantum Technology hub following his Ph.D. Since leaving the University of Glasgow at the end of 2016, he has been involved in community development and educational outreach.

Miles Padgett is the Kelvin Chair of natural philosophy at the University of Glasgow, Glasgow, Scotland. He is fascinated by light both classical and quantum—specifically light's momentum. In 2001, he was elected to Fellowship of the Royal Society of Edinburgh and in 2014 the Royal Society, the U.K.'s National Academy. In 2009, with Les Allen, he received the IoP Young Medal, in 2014 the RSE Kelvin Medal, in 2015 the Science of Light Prize from the EPS, and in 2017 the Max Born Award of the OSA.

Steve McLaughlin (F'11) was born in Clydebank, Scotland, in 1960. He received the B.Sc. degree in electronics and electrical engineering from the University of Glasgow, Glasgow, Scotland, in 1981, and the Ph.D. degree from the University of Edinburgh, Edinburgh, Scotland, in 1990. From 1981 to 1984, he was a Development Engineer in industry involved in the design and simulation of integrated thermal imaging and fire control systems. From 1984 to 1986, he worked on the design and development of high-frequency data communication systems. In 1986, he joined the Deptartment of Electronics and Electrical Engineering, University of Edinburgh as a research fellow, where he studied the performance of linear adaptive algorithms in high noise and nonstationary environments. In 1988, he joined the academic staff at Edinburgh, and from 1991 until 2001 he held a Royal Society University Research Fellowship to study nonlinear signal processing techniques. In 2002, he was awarded a personal Chair in Electronic Communication Systems at the University of Edinburgh. In October 2011, he joined Heriot-Watt University as a Professor of Signal Processing and Head of the School of Engineering and Physical Sciences. His research interests include the fields of adaptive signal processing and nonlinear dynamical systems theory and their applications to biomedical , energy and communication systems. He is a Fellow of the Royal Academy of Engineering, Royal Society of Edinburgh, and Institute of Engineering and Technology.