

Article

Bayesian Activity Estimation and Uncertainty Quantification of Spent Nuclear Fuel Using Passive Gamma Emission Tomography

Ahmed Karam Eldaly ^{1,*}, Ming Fang ², Angela Di Fulvio ², Stephen McLaughlin ¹, Mike E. Davies ³, Yoann Altmann ^{1,*} and Yves Wiaux ¹

¹ Institute of Sensors, Signals and Systems, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK; s.mclaughlin@hw.ac.uk (S.M.); y.wiaux@hw.ac.uk (Y.W.)

² Department of Nuclear, Plasma, and Radiological Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; Mingf2@illinois.edu (M.F.); difulvio@illinois.edu (A.D.F.)

³ Institute for Digital Communications & The Joint Research Institute for Signal and Image Processing, The University of Edinburgh, Edinburgh, EH9 3JL, UK; Mike.Davies@ed.ac.uk

* Correspondence: a.karam@hw.ac.uk (A.K.E.); Y.Altmann@hw.ac.uk (Y.A.)

Abstract: In this paper, we address the problem of activity estimation in passive gamma emission tomography (PGET) of spent nuclear fuel. Two different noise models are considered and compared, namely, the isotropic Gaussian and the Poisson noise models. The problem is formulated within a Bayesian framework as a linear inverse problem and prior distributions are assigned to the unknown model parameters. In particular, a Bernoulli-truncated Gaussian prior model is considered to promote sparse pin configurations. A Markov chain Monte Carlo (MCMC) method, based on a split and augmented Gibbs sampler, is then used to sample the posterior distribution of the unknown parameters. The proposed algorithm is first validated by simulations conducted using synthetic data, generated using the nominal models. We then consider more realistic data simulated using a bespoke simulator, whose forward model is non-linear and not available analytically. In that case, the linear models used are mis-specified and we analyse their robustness for activity estimation. The results demonstrate superior performance of the proposed approach in estimating the pin activities in different assembly patterns, in addition to being able to quantify their uncertainty measures, in comparison with existing methods.



Citation: Eldaly, A.K.; Fang, M.; Di Fulvio, A.; McLaughlin, S.; Davies, M.; Altmann, Y.; Wiaux, Y. Bayesian Activity Estimation and Uncertainty Quantification of Spent Nuclear Fuel Using Passive Gamma Emission Tomography. *J. Imaging* **2021**, *7*, 212. <https://doi.org/10.3390/jimaging7100212>

Academic Editors: Fabiana Zama and Elena Loli Piccolomini

Received: 31 August 2021

Accepted: 7 October 2021

Published: 14 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: inverse problems; imaging; Bayesian inference; uncertainty quantification; tomography; Markov chain Monte Carlo

1. Introduction

In order to deter the proliferation of nuclear weapons, safeguards provide various technical measures that are used for the verification and the declarations made by the signatories to the Treaty on the Non-Proliferation of Nuclear Weapons, regarding their nuclear material and activities [1]. An important task within these safeguards is monitoring of spent fuel assemblies (SFAs) from nuclear power plants (NPPs), for detecting any eventual diversion of spent nuclear fuel for non-declared purposes. The detection of a single fuel pin missing from SFA should be reported. For any safeguards investigation of SFAs, it is important to use a minimum amount of a priori information on the SFA under study, in order to avoid biasing and potentially misleading the investigation. Eventually, IAEA approved the use of the passive gamma emission tomography (PGET) instrument [2–7] in inspections.

Passive gamma emission tomography is an imaging modality that can be used for the verification of spent nuclear fuel stored in water pools [4]. Spent nuclear fuel is highly radioactive; hence, the tomography can be conducted in passive mode, without the need of an external X-ray source, as in traditional tomography. The detection unit considered

in this work consists of two collimated CdTe detector arrays, on opposite sides of the fuel assembly, each encompassing 91 detectors. The detector arrays rotate and scan the fuel bundle in steps of one degree, which generates a sinogram by detecting gamma rays emitted by the fuel pins. A cross-sectional image of the spent fuel bundle, using image reconstruction algorithms, can then be reconstructed. Based on these images, the fuel pins can be identified and their activity estimated as the sum of pixel values inside each pin region. Image reconstruction techniques with the capability of estimating the fuel pins' activity accurately are crucial in nuclear safeguards, as they allow inspectors to monitor nuclear material and promptly identify its diversion.

Due to the nature of acquisition process (e.g., the discrete nature of the measured photon counts), the observation noise of PGET sinograms is expected to be well-approximated by Poisson noise. However, if the photon counts in a sinogram pixel is high, the Poisson distribution becomes more symmetric, and it can be well-approximated using a Gaussian distribution, whose mean and variance change across pixels. While the sinogram formation process can be approximated by a linear forward model (which simplifies the inversion), this first-order approximation does not take into account the attenuation of the measured radiation of a pin, due to the presence of other pins (or shielding materials) within the assembly. Consequently, the activity levels, estimated by inversion of a linear problem, can present biases and artefacts, depending on the noise model considered. Hence, in this work, we provide a comparison of both the Poisson and isotropic Gaussian noise models and investigate their robustness. Moreover, we also investigate different linear operators corresponding to different levels of prior knowledge about the assembly configuration.

Most of state-of-the-art algorithms considered to solve linear inverse image restoration problems (irrespective of the noise model) are either optimization or simulation-based methods. Optimization-based approaches primarily rely on log-concave Bayesian models, such as [8–15], and have been proposed to perform maximum-a-posteriori (MAP) estimation. For example, PIDAL, which stands for *Poisson image deblurring using augmented Lagrangian* [8], and SALSA, which stands for *split augmented Lagrangian shrinkage algorithm* [10], are Poisson and Gaussian image restoration algorithms based on a total-variation loss of sparsity-promoting prior, which solves the restoration problem using an alternating direction method of multipliers (ADMM). Moreover SARA, which stands for *sparsity averaging reweighted analysis* [11,12], assumes that the ill-posed problem is regularized by the assumption of average signal sparsity over representations in multiple wavelet bases. Although such algorithms are efficient in computing maximum *a posteriori* (MAP) estimates relatively quickly, they cannot provide uncertainty maps for the estimates, which can be very valuable in applications involving subsequent decision making, such as pin identification and activity estimation in the PGET context. Alternatively, many studies have considered hierarchical Bayesian models to solve the deconvolution and restoration problems akin to those in [16–18]. This class of methods assumes that the unknown model parameters are unknown stochastic quantities by assigning them suitable prior distributions, based on prior beliefs. The joint posterior distribution can then be computed using Bayes theorem and exploited using Markov chain Monte Carlo methods. These models offer a flexible and consistent methodology to deal with uncertainty in ill-posed inverse problems. Moreover, additional unknown parameters can be jointly estimated within the algorithm, such as regularization parameters. As such, they represent an attractive way to tackle ill-posed problems, such as the one considered in this work, where an automated hyperparameter setting is challenging. As an example, the authors in [17] proposed a Bayesian approach that samples from a posterior distribution built on a Poissonian likelihood and standard convex and possibly non-smooth regularizers. That study proposed an approach that relied on the split-and-augmented Gibbs sampler (SPA) [18] to tackle a Poisson/Gaussian reconstruction problem by sampling from an approximate joint posterior distribution. In particular, the log-concave property of the prior distributions allows the derivation of log-concave conditional distributions that are decoupled from the likelihood and can be sampled by proximal MCMC methods [19,20]. In this context, the SPA method

provides both image estimates, such as MAP and MMSE, and their corresponding posterior uncertainty measures.

The PGET instrument is essentially a single photon emission computed tomography (SPECT) system that allows the reconstruction of axial cross-sections of the emission map of the SFA. Industrial or medical tomographic imaging systems work in a similar manner as PGET systems, in terms of data acquisition. Furthermore, the PGET concept can be applied to perform fast neutron imaging by replacing gamma ray detectors with neutron detectors [21]. However, the PGET of spent fuel poses some unique challenges, due to the high activity of the sources (a single-pin activity is of the order of 10^{13} Bq) and the high self-attenuation of the fuel pins. Hence, robust image reconstruction algorithms are crucially required. For instance, in [7], the authors proposed a method to simultaneously reconstruct the activity and attenuation maps by formulating the reconstruction as a constrained minimization problem with a least squares data fidelity term and quadratic regularization terms with different smoothness matrices. However, the forward model proposed in [7] considered only monoenergetic gamma rays and did not account for the scattering effects of gamma rays in the PGET system that make non-negligible contributions to the total counts, which added additional uncertainty to the activity estimation results. Similarly, in [22], the authors formulated the problem of pin activity estimation as an optimization problem where a sparse regularization function is used. Although the advantages of these methods in providing fast estimates, they can provide only point estimates; hence, they can neither quantify the uncertainty of the estimates nor provide probability of presence of the pins. Moreover, associated regularisation parameters need to be manually tuned by the user, which can affect the estimated activity levels. In this work, to improve the quality of the estimated activity profiles and reduce the computational cost of the inversion procedure, we propose to only estimate the activity in a reduced number of pixels, leveraging the prior knowledge that the activity profile is expected to be null outside pin locations. For instance, for a cross-sectional image of the field of view discretized into 182×182 pixels, the pins, whose maximum number is known, form groups of 10 to 16 connected pixels, which form a small subset ($\approx 15\%$) of the 182^2 pixels. The method proposed in this paper assumes that a mask, identifying where pins are expected to be present, is available. This mask is used (i) to account for pin self-attenuation when building the linear approximate forward operator and (ii) as an upper bound for the support of the pins present in the assembly. While the geometry of the assembly is generally known, the exact presence map is often unknown (this is one feature we want to estimate), but several initial guess can be considered, as will be discussed in Section 6. The main contributions of this work can be summarized as follows.

- We formulate the pin activity estimation problem within a Bayesian framework and assign a Bernoulli truncated-Gaussian (BtG) prior model to the intensity field to be estimated. To the best of our knowledge, no work attempted to sample from such a highly-multimodal joint posterior distribution using an SPA sampler. This allows for estimating the activity of spent-fuel, including the assessment of fuel rod presence/absence.
- We compare the performance of two different noise models for pin activity estimation from PGET sinograms simulated while accounting for pin self-attenuation (i.e., not simulated using a linear model).
- In addition to estimating the activity profile, the proposed algorithms allow the automated estimation of the crucial model hyperparameters, like regularization parameters, which might affect the resulting estimated activity.

The remaining sections of the paper are organized as follows. Section 2 formulates the problem of PGET image restoration, followed by Section 3, which summarizes the likelihoods and the prior distributions assigned to the unknown parameters of the models. The resulting joint posterior distribution and the partially collapsed Gibbs sampler, used to sample that distribution, are discussed in Section 4. Simulations conducted using synthetic data (following a linear models used for inversion) are presented in Section 5 and the

analysis of more realistic data is discussed in Section 6. Conclusions and future work are, finally, reported in Section 7.

2. Problem Formulation

Figure 1 shows examples of different pin configurations (i.e., presence maps) akin to those expected to be studied using a PGET system. We can observe that each cross-sectional image consists of pins formed by groups of connected pixels (white), which will emit radiations. Black pixels do not emit measurable radiations. In this work, we assume that we know, *a priori*, a mask such as those in Figure 1 such that the activity of the black pixels is assumed to be null and thus not estimated. However, the mask can be designed conservatively and select pixels where no pins are present (e.g., the mask in Figure 1a can be selected although the actual assembly configuration is that depicted in Figure 1b). Using this user-defined mask, we defined as $\mathbf{x} = (x_1, \dots, x_N)^T \in \mathbb{R}^N$, the concatenation of the pixel values to be estimated.

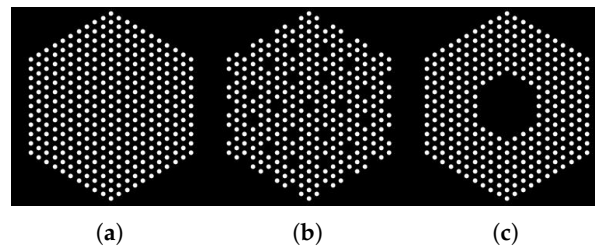


Figure 1. Examples of binary masks of the simulated fuel assemblies. (a) Case “1”, (b) Case “2”, and (c) Case “3”. In Case “1”, the mock-up fuel assembly hosted 331 ^{60}Co pins of 8.879 g/cm^3 density, which emitted 1.17 and 1.33 MeV gamma rays. In Cases “2” and “3”, 10% ^{60}Co pins were removed at the center or uniformly.

The pin activity estimation problem is formulated as follows. Given a set of measurements (sinogram) $\mathbf{y} = (y_1, \dots, y_M)^T \in \mathbb{R}^M$, we aim at recovering the underlying pin intensities $\mathbf{x} = (x_1, \dots, x_N)^T \in \mathbb{R}^N$, which are related to the observation \mathbf{y} by $\mathbf{y} = \mathcal{G}(\mathbf{x}) + \mathbf{w}$, where \mathcal{G} is a function representing the system’s response, and \mathbf{w} represents random observation noise. Due to photon scattering and attenuation effects between pins during the data formation process, the function \mathcal{G} is non-linear and difficult to model analytically. However, it can be approximated by a linear model based on the user-defined mask discussed above. Indeed, the presence map can be used to build a linear mapping from \mathbf{x} and \mathbf{y} whereby attenuation involving neighbouring pins is accounted for (the attenuation is mostly due to the presence of pins, not their activity). Hence, in this work, the function \mathcal{G} is approximated by a linear operator $\mathbf{A} \in \mathbb{R}^{M \times N}$, which is built using the simulator recently considered in [22]. Consequently, the forward model considered in this work can be expressed as:

$$\mathbf{y}|\mathbf{x} \sim \mathcal{F}(\mathbf{Ax}), \quad (1)$$

where $\mathcal{F}(\cdot)$ denotes a probability distribution and \sim reads “is distributed according to”. While the observation noise corrupting PGET sinograms is expected to be Poisson distributed, we will consider two different noise models as the linear approximation can introduce mis-modeling errors. The problem investigated in this paper is the estimation of the pin intensities \mathbf{x} from the observation vector \mathbf{y} . This inverse problem is severely ill-conditioned because of \mathbf{A} and prior regularization is necessary to promote the solution to be in a set of feasible intensities \mathbf{x} . To solve this problem, we propose a hierarchical Bayesian model and a sampling method to estimate the unknown model parameter.

3. Hierarchical Bayesian Model

This section introduces the hierarchical Bayesian model proposed to estimate the unknown parameter \mathbf{x} . This model is based on the likelihood function of the observations

and on prior distributions assigned to the unknown parameters, i.e., \mathbf{x} and potential hyperparameters.

3.1. Likelihood

Under the Poisson noise model assumption, each individual observation y_m , $m \in [1, M]$ corresponds to an independent realization of a Poisson random variable, that is, for all $m \in [1, M]$:

$$f(y_m|\mathbf{x}) \sim \mathcal{P}(\{\mathbf{Ax}\}_m), \quad m = 1, \dots, M, \quad (2)$$

where $\mathcal{P}(\cdot)$ denotes the Poisson distribution and $\{\mathbf{Ax}\}_m$ denotes the m th element of \mathbf{Ax} . The full likelihood reduces to $f(\mathbf{y}|\mathbf{x}) = \prod_{m=1}^M f(y_m|\mathbf{x})$.

The second model considered assumes independent and identically distributed (idd) additive Gaussian noise. In that case, the full likelihood reduces to:

$$f(\mathbf{y}|\mathbf{x}) = \left(\frac{1}{2\pi\sigma^2}\right)^{N/2} \exp\left(-\frac{\|\mathbf{y} - \mathbf{Ax}\|_2^2}{2\sigma^2}\right), \quad (3)$$

where σ^2 is the (unknown) noise variance.

3.2. Prior Distributions

As mentioned earlier, $\mathbf{x} \in \mathbb{R}^N$ is a concatenation of pixel intensities. If the user-defined mask is perfect, then the entries of \mathbf{x} are expected to be positive, as long as a pin present has a measurable activity level. However, if the mask is designed conservatively, a pin which is assumed to be present can have a null activity level in practice, i.e., a fraction of the elements of \mathbf{x} are equal to 0. In order to model such belief while ensuring the positivity of the activity levels, a classical choice is the following exponential prior distribution:

$$f(\mathbf{x}|\beta) = \frac{\beta}{2} \exp(-\beta\|\mathbf{x}\|_1) \times 1_{\mathbb{R}^+}(\mathbf{x}), \quad (4)$$

where β is a regularization parameter controlling the mean value of \mathbf{x} and $1_{\mathbb{R}^+}(\mathbf{x})$ is the indicator function defined on the positive orthant of \mathbb{R}^N . The resulting joint posterior distribution $f(\mathbf{x}|\mathbf{y}, \beta)$ (for a fixed β) is relatively easy to sample from using an SPA sampling strategy, as in [17,18], resulting in an algorithm that is referred to as SPA-P- ℓ_1 (Split and Augmented Gibbs sampler for Poisson noise) and SPA-G- ℓ_1 (Split and Augmented Gibbs sampler for Gaussian noise), assuming σ^2 is known. Moreover, one can maximize the logarithm of the resulting joint posterior distribution and compute the maximum *a posteriori* (MAP) estimates as in [8], resulting in the PIDAL- ℓ_1 algorithm using a Poisson image model and FISTA- ℓ_1 for Gaussian image model.

However, in this work, to model the intensity field vector sparsity, we introduce:

$$\mathbf{x}_g = z_g \mathbf{t}_g, \quad (5)$$

where $z_g \in \{0, 1\}$ is a binary label for pin g and $\mathbf{t}_g = [t_1^{(g)}, \dots, t_{K_g}^{(g)}] \in \mathbb{R}^{K_g}$ is the corresponding amplitude vector for the K_g pixels forming pin g , $\mathbf{x}_g \in \mathbb{R}^{K_g}$ and $\mathbf{x} = [\mathbf{x}_1, \dots, \mathbf{x}_G]$, where G is the maximum number of pins in the assembly, as defined by the user-defined mask. If $z_g = 0$, the activity of any pixel of the pin g is 0. If $z_g = 1$, then $\mathbf{x}_g = \mathbf{t}_g$. Note that K_g and G are known a priori, however, the number of active pins forming a specific assembly (i.e., the value of $\{z_g\}_g \in \{0, 1\}^G$) is not known and must be estimated. The decomposition described above allows one to decouple the location of the sparse components from their values. In a similar fashion to [23], to model the sparsity of \mathbf{x} , we utilize a structured spike-and-slab prior, which stems from classical spike-and-slab priors that have been used in Bayesian regression and factor models [24–27]. Following this, we assume that the

intensities in $\{t_g\}_g$ are drawn independently from the same truncated zero-mean Gaussian distribution, i.e.,

$$f([t_1, \dots, t_N] | s^2) = \prod_{n=1}^N \mathcal{N}_{\mathbb{R}^+}(t_n; 0, s^2), \quad (6)$$

where s^2 controls the prior variance of the intensity vector. The pin presence labels are assumed *a priori* mutually independent and assigned a shared Bernoulli prior distribution, leading to:

$$f(z_g | \omega) = \omega^{z_g} (1 - \omega)^{(1-z_g)}, \quad z_g \in \{0, 1\}, \quad (7)$$

and

$$f(\mathbf{z} | \omega) = \prod_{g=1}^G f(z_g | \omega). \quad (8)$$

The prior models of \mathbf{t} and \mathbf{z} depend on the hyperparameters s^2 and ω which can be difficult to fix automatically. Thus, we include them within the inference process using hyperpriors. To reflect the lack of prior knowledge about the variance s^2 in Equation (6), a weakly informative inverse-gamma prior is assigned to s^2 :

$$s^2 \sim \mathcal{IG}(s^2; \eta, \nu), \quad (9)$$

where (η, ν) are fixed to $(\eta, \nu) = (10^{-3}, 10^{-3})$. Similarly, we assign ω a conjugate beta prior distribution:

$$\omega \sim \mathcal{Be}(\omega; \alpha, \beta), \quad (10)$$

where (α, β) is fixed to $(\alpha, \beta) = (1, 1)$.

Note that we did not observe significant changes in the results when changing the hyperparameters (η, ν) to $(10^{-3}, 10^{-3})$, $(10^{-2}, 10^{-2})$ and $(1, 1)$. Similarly for (α, β) to values $(1, 1)$, $(1/9, 1)$ and $(0.1, 0.9)$, for means to the \mathcal{Be} distribution that are equal to 0.5, 0.1 and 0.9 respectively.

Note that we ran the algorithms over a range of hyperparameter values (e.g., $(\eta, \nu) = (XX, YY)$, (xx, yy) and $(xx1, xx2)$) and did not observe significant performance degradation

3.3. Joint Posterior Distribution

Assuming the parameters \mathbf{z} and \mathbf{t} are *a priori* mutually independent, the joint posterior of the parameter vector $\Delta = \{\mathbf{z}, \mathbf{t}\}$ and hyperparameters $\Phi = \{s^2, \omega\}$ can be expressed as:

$$f(\Delta, \Phi | \mathbf{y}) \propto f(\mathbf{y} | \Delta) f(\Delta | \Phi) f(\Phi), \quad (11)$$

where $f(\Delta | \Phi) = f(\mathbf{z} | \omega) f(\mathbf{t} | s^2)$ and $f(\Phi) = f(s^2) f(\omega)$. The joint posterior distribution reduces to:

$$\begin{aligned} p(\Delta, \Phi | \mathbf{y}) &\propto \mathcal{F}(\mathbf{AZt}) \times \prod_{g=1}^G \left(\left[\omega^{z_g} (1 - \omega)^{(1-z_g)} \right] \times \prod_{i=1}^{K_g} \mathcal{N}_{\mathbb{R}^+}(t_i; 0, s^2) \right) \\ &\times \mathcal{Be}(\omega; \alpha, \beta) \times \mathcal{IG}(s^2; \eta, \nu), \end{aligned} \quad (12)$$

where $\mathbf{Z} = \text{diag}([z_1 \mathbf{1}_{K_1}, z_2 \mathbf{1}_{K_2}, \dots, z_g \mathbf{1}_{K_g}])$. The next paragraph describes the sampling strategy adopted to estimate the unknown parameter vector Δ and the hyperparameters Φ .

4. Bayesian Inference

To overcome the challenging derivation of Bayesian estimators associated with $f(\Delta, \Phi | \mathbf{y})$, we use an efficient Markov chain Monte Carlo (MCMC) method to generate samples asymptotically distributed according to Equation (12). More precisely, we consider a variable-splitting inspired MCMC algorithm, namely the split-and-augmented Gibbs sampler (SPA, presented in the next paragraph) that was recently proposed in [18]. For SPA, we assume that $\mathbf{Y} = \{\mathbf{b}, \mathbf{c}\}$, where $\mathbf{b} \approx \mathbf{Ac}$ and $\mathbf{c} \approx \mathbf{Zt}$. This splitting decouples locally each vector of

variables from the rest of the variables, which will in turn make the inference easier. The joint posterior distribution in Equation (11) can be extended as:

$$\begin{aligned}
 p(\Delta, \mathbf{Y}, \Lambda, \Phi | \mathbf{y}) \propto & \mathcal{F}(\mathbf{AZt}) \times \prod_{g=1}^G f(\mathbf{z}_g, \mathbf{t}_g) \times \exp\left(-\frac{\|\mathbf{Ac} - (\mathbf{b} - \mathbf{o}_1)\|_2^2}{2\rho^2}\right) \\
 & \times \exp\left(-\frac{\|\mathbf{Zt} - (\mathbf{c} - \mathbf{o}_2)\|_2^2}{2\rho^2}\right) \times \mathcal{Be}(\omega; \alpha, \beta) \times \mathcal{IG}(s^2; \eta, \nu) \\
 & \times \exp\left(-\frac{\|\mathbf{o}_1\|_2^2}{2\alpha^2}\right) \times \exp\left(-\frac{\|\mathbf{o}_2\|_2^2}{2\alpha^2}\right),
 \end{aligned} \quad (13)$$

where $f(\mathbf{z}_g, \mathbf{t}_g) = \left[\omega^{z_g}(1-\omega)^{(1-z_g)}\right] \times \prod_{i=1}^{K_g} \mathcal{N}_{\mathbb{R}^+}(t_i; 0, s^2)$ and the new terms are divergence functions such that the joint extended posterior distribution defines a proper probability distribution. The auxiliary variables $\Lambda = \{\mathbf{o}_1, \mathbf{o}_2\}$ associated with the splitting variables $\mathbf{Y} = \{\mathbf{b}, \mathbf{c}\}$, which allow to decrease the correlation between the variables by giving an additional degree of freedom to each of the former variables (see [18] for more details). Moreover, $\rho, \alpha > 0$ are user defined parameters. In Equation (13), it is possible to use a block Gibbs sampler (with proximal MCMC step in the Poisson noise case) to sample according to the conditional distributions of each of the unknown model parameters. In practice, strong correlations appear between \mathbf{z} and \mathbf{t} . Moreover, as \mathbf{z} can be sparse, sampling $f(\mathbf{t}, s^2 | \mathbf{y}, \Delta_{\setminus \mathbf{t}}, \Phi_{\setminus (s^2)})$, where $\mathcal{H}_{\setminus \mathbf{u}}$ denotes the parameter vector \mathcal{H} whose parameter \mathbf{u} is omitted, via Gibbs sampling results in very slow convergence. Hence, we use a partially collapsed Gibbs sampler (PCGS) which provides better mixing and convergence properties [28–33]. The PCGS used here samples groups of variables (e.g., (\mathbf{z}, \mathbf{t})) from their joint posterior distribution rather than from their conditional distributions. Sampling from the joint distribution is achieved by first marginalising some variables which are then sampled from their full conditional distribution [34,35]. Precisely, we sample sequentially the elements of $\Delta, \mathbf{Y}, \Lambda$ and Φ using moves that are summarised in Algorithm 1.

Algorithm 1 Split and augmented—partially collapsed Gibbs sampling algorithm for activity estimation in PGET—version I.

- 1: **Fixed input parameters:** Number of burn-in iterations N_{bi} , total number of iterations N_{MC}
 - 2: **Initialization** ($k = 0$)
 - Set $\mathbf{z}^{(0)}, \mathbf{t}^{(0)}, \mathbf{b}^{(0)}, \mathbf{c}^{(0)}, \omega^{(0)}, s^{2(0)}, \mathbf{o}_1^{(0)}, \mathbf{o}_2^{(0)}$
 - 3: **Repeat** ($1 \leq k \leq N_{\text{MC}}$)
 1. Sample $(s^{2(k)}, \mathbf{t}_{z_g=0}^{(k)}) | (\mathbf{y}, \omega^{(k-1)}, \Delta_{\setminus \mathbf{t}_{z_g=0}}^{(k-1)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)})$
 2. Sample $(\mathbf{z}^{(k)}, \mathbf{t}^{(k)}) | (\mathbf{y}, s^{2(k)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)})$
 3. Sample $\omega^{(k)} | (\mathbf{y}, s^{2(k)}, \Delta^{(k)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)})$
 4. Sample $\mathbf{b}^{(k)} | (\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{c}^{(k-1)}, \Lambda^{(k-1)})$
 5. Sample $\mathbf{c}^{(k)} | (\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{b}^{(k)}, \Lambda^{(k-1)})$
 6. Sample $\mathbf{o}_1^{(k)} | (\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{Y}^{(k)}, \mathbf{o}_2^{(k-1)})$
 7. Sample $\mathbf{o}_2^{(k)} | (\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{Y}^{(k)}, \mathbf{o}_1^{(k)})$
 - 4: **Set** $k = k + 1$.
-

In Algorithm 1, $\mathbf{t}_{z_g=0}$ denotes the elements of \mathbf{t} whose corresponding labels in \mathbf{z} are null. Similarly, $\mathbf{t}_{z_g=1}$ denotes the elements of \mathbf{t} whose labels are equal to 1. We introduce the details of sampling each step of the algorithm above in the Appendix A.

The algorithm is stopped after N_{MC} iterations, including the N_{bi} burn-in iterations, which correspond to the transient period of the sampler (determined visually from preliminary runs). The first N_{bi} samples are discarded and the remaining samples are used to approximate the following coupled Bayesian estimators that are particularly suitable for pin activity estimation problems. For the support of the activity vector or “presence maps”, which provides the probability of presence of pins, we use the marginal maximum a posteriori (MMAP) estimator [36,37]

$$z_g^{MMAP} = \arg \max_{z_g \in \{0,1\}} f(z_g | \mathbf{y}, \Delta_{\setminus z_g}, \Phi, \mathbf{Y}, \Lambda). \quad (14)$$

As we are interested in estimating the pin intensities $\mathbf{x}_g = z_g \mathbf{t}_g = z_g [t_{1,g} \cdots t_{i,g}, \cdots, t_{K_g,g}]$, the empirical averages of the generated samples (conditional minimum mean squared error “CMMSE” estimates) of each pin pixel $x_{i,g} = z_g t_{i,g}$, conditionally on the estimated supports, is given by:

$$x_{i,g}^{CMMSE} = E[x_{i,g} | z_g^{MMAP}, \mathbf{y}, \Phi, \mathbf{Y}, \Lambda], \quad (15)$$

where if $z_g^{MMAP} = 0$, $x_{i,g}^{CMMSE} = 0$, otherwise $x_{i,g}^{CMMSE}$ is given by:

$$x_{i,g}^{CMMSE} = \frac{1}{\sum_{k>N_{bi}} z_g^{(k)}} \sum_{k>N_{bi}} (z_g t_{i,g})^{(k)}, \quad (16)$$

The variance of active pin pixels can be computed by:

$$v_{i,g}^2 = \frac{1}{\sum_{k>N_{bi}} z_g^{(k)}} \sum_{k>N_{bi}} ((z_g t_{i,g})^{(k)} - x_{i,g}^{CMMSE})^2, \quad (17)$$

Moreover, the average of pin pixel intensities can be computed as:

$$\hat{x}_g = \frac{1}{K_g} \sum_{i=1}^{K_g} x_{i,g}^{CMMSE}, \quad (18)$$

which can then assigned for the pixels forming each pin. The variance can be computed in same manner as in Equation (17).

5. Simulations Using Synthetic Datasets

In this section, we assess the performance of the proposed algorithm for cases where the data, given \mathbf{x} , are generated according to the model used to perform Bayesian inference. More precisely, we define the linear operator \mathbf{A} , compute $\mathbf{A}\mathbf{x}$, and generate sinograms corrupted by additive iid Gaussian noise and by Poisson noise. We then use the corresponding algorithm to infer \mathbf{x} . Cases where \mathbf{y} is generated using the more realistic simulator will be discussed later in Section 6.

5.1. Data Creation

To assess the performance of the proposed approaches, we created a simulated activity profile akin to those expected in actual assemblies. This image is of size (182×182) and contains 331 pins spanning a total number of 5329 pixels. The image, depicted in Figure 2c presents three different pin activity levels, and 10% of pins are missing at the center of the assembly. The response matrix of the system $\mathbf{A} \in \mathbb{R}^{65,520 \times 5329}$ was constructed by simulating the detector array response to a source of unit activity inside each pixel, using the ground truth presence maps depicted in Figure 2b. Here, the admissible support in Figure 2a contains all the possible pin locations in the assembly. This choice was made to confirm whether the algorithms are prone to underestimating or overestimating the number of pins actually present. The forward model in Equation (1), with the different noise models

in Equations (2) and (3), is then used to generate the sinograms with the system response matrix \mathbf{A} . Different activity peak values in $\{220, 160, 100, 40, 10\}$ (accounting for different signal to noise ratios), are tested for the Poisson noise case. For the Gaussian noise case, the activity peak value was set to 220, and the noise variance in the Gaussian case is assumed independent and identically distributed with different noise variances, belonging to $\{1, 2, 4, 6, 8\} \times 10^{-3}$, were considered. Note that the linear operator in the Poisson noise case is scaled to $\mathbf{A} = \mathbf{A} \times 10^3$ to keep the same activity profile as of the Gaussian noise case. Figure 2d,e show examples of the corresponding sinograms generated using the Poisson and the Gaussian noise models.

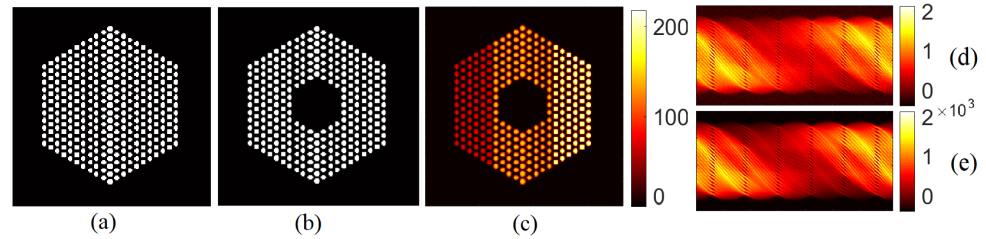


Figure 2. Simulated dataset. (a) The admissible support, (b) original binary mask, (c) ground truth activity profile when the peak value is set to 220, (d) sinogram generated with an iid Gaussian noise model with peak value = 220 and noise variance $\sigma^2 = 2 \times 10^{-3}$, and (e) sinogram generated with a Poisson noise model with peak value = 220.

5.2. Quantitative Analysis

The proposed approach is run on the simulated datasets described above. In all cases, a Markov chain of length $N_{bi} = 1.2 \times 10^4$ and a burn-in period of length $N_{MC} = 4 \times 10^3$ are used. The hyperparameters (ρ, α) of the split-and-augmented Gibbs sampler are set to $(\rho, \alpha) = (10, 1)$ and those of the P-MYULA are set to $(\lambda, \gamma) = (\rho^2, \rho^2/4)$, as in [17]. The quantitative measure used to assess the quality of estimated activity profiles is the normalized mean square error (NMSE), defined as

$$\text{NMSE} = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2}{\|\mathbf{x}\|_2^2}, \quad (19)$$

where $\hat{\mathbf{x}}$ is an estimate of \mathbf{x} . The proposed approach is compared against four of the existing approaches described earlier, namely, SPA-P- ℓ_1 and PIDAL- ℓ_1 , for the Poisson noise model, and SPA-G- ℓ_1 and FISTA- ℓ_1 for the Gaussian noise model. For both SPA-P- ℓ_1 and SPA-G- ℓ_1 , the Markov chain length N_{MC} , burn-in period N_{bi} , the hyperparameters (ρ, α) of the split-and-augmented Gibbs sampler, and those of the P-MYULA (λ, γ) are set as in the proposed approach. For the four methods, several values for the ℓ_1 regularization parameter β are tested, from which we pick the one providing the lowest NMSE. Figure 3 shows plots of NMSE versus different activity profile peak values for the Poisson noise case, and different noise variances for a fixed activity profile peak value for the Gaussian noise case. We can observe that SPA- ℓ_1 provides the highest NMSE, whereas SPA-BtG and FISTA/PIDAL- ℓ_1 provide close results. This figure shows that, when comparing MMSE estimates, the proposed Bernoulli-based prior enforcing group sparsity is more efficient than the Exponential-based prior, for both noise models. Although MAP estimation based on the Exponential-based prior is appealing, the corresponding algorithms do not allow directly uncertainty quantification.

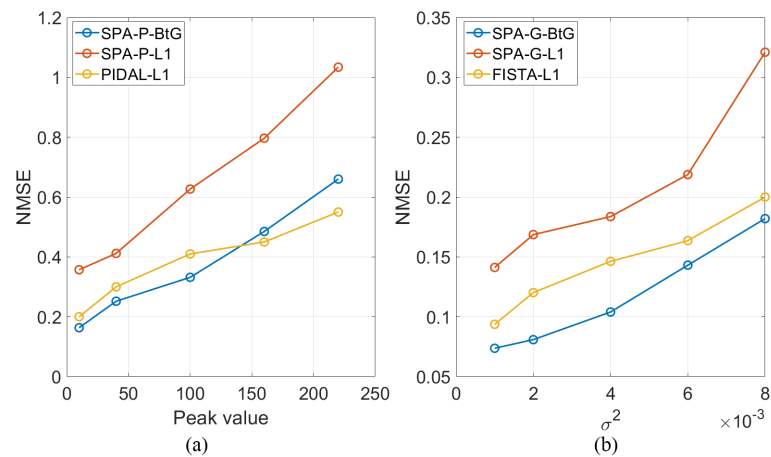


Figure 3. Normalized mean squared error (NMSE) plots for the (a) Poisson noise model with different peak values, (b) the Gaussian noise model when the maximum intensity is set to 220 and different noise variances.

5.3. Qualitative Analysis

5.3.1. The Proposed Method

In this section, we present visual results for activity estimation using the proposed approaches. For each case, we plot both the individual pin pixel activities and the average activity of pixels spanning each pin. Figures 4 and 5 show examples of results for pin activity estimation and uncertainty quantification for the Gaussian and Poisson noise models, respectively. The peak value was set to 220 in both, and the noise variance of the Gaussian noise case was set to $\sigma^2 = 2 \times 10^{-3}$. We can observe that the algorithm could successfully reconstruct the activity profile with both noise models. Moreover, it can be seen that the differences in pin activity levels are clearer when considering the average of pin pixel intensities as an activity estimate, and the uncertainty is much lower, in comparison with considering all pixel intensities. This is because averaging naturally contributes to reducing the noise in the activity estimates, equivalently enhancing the local signal to noise ratio. Consequently, the ratio of marginal standard deviation divided by the posterior mean shows clearer activity level differences. Hence, for the rest of the experiments in this paper, we will consider pin pixel averages as activity estimates.

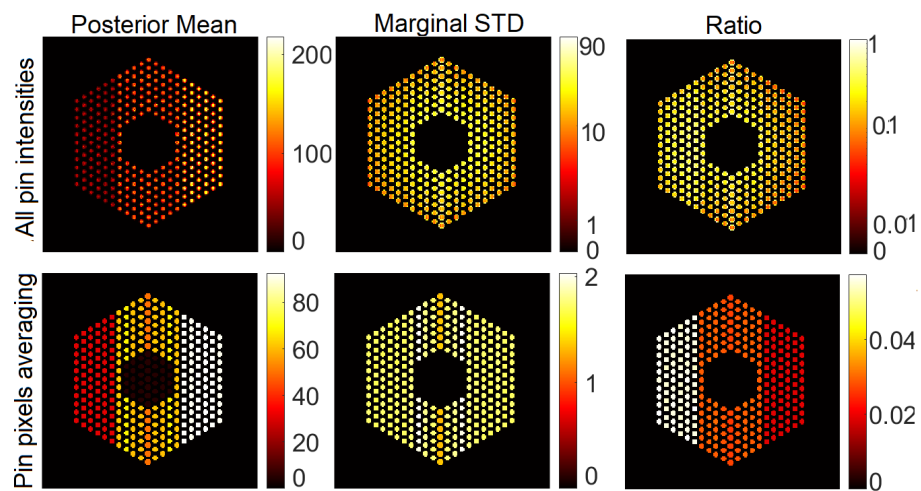


Figure 4. Results of the proposed approach using the Gaussian noise model, when the peak value is set to 220 and noise variance is set to $\sigma^2 = 2 \times 10^{-3}$. Column 1: posterior mean, Column 2: marginal standard deviation and Column 3: ratio of marginal standard deviation divided by posterior mean. Row 1: All pixels, and Row 2: pin pixels averaging.

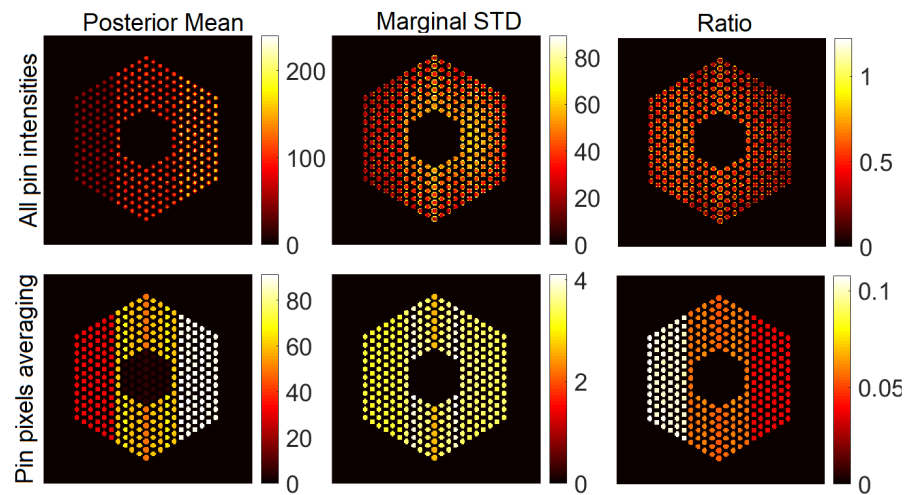


Figure 5. Results of the proposed approach using the Poisson noise model when the peak value is set to 220. Column 1: posterior mean; column 2: marginal standard deviation; and column 3: ratio of marginal standard deviation divided by posterior mean. Row 1: all pixels; row 2: pin pixels averaging.

Figure 6 depicts examples of estimated probabilities of pin presence in both the Poisson noise case, using different activity peak values, and the Gaussian noise case for two levels of signal quality. We can observe that for both the Gaussian and Poisson noise cases, when the peak value is reasonably high, the algorithm could identify with high confidence levels the 10% missing pins near the center of the assembly. However, when the peak value in the Poisson noise case decreases, and the noise variance in the Gaussian noise case increases, it becomes more difficult to identify the assembly configuration. In Figure 6a for instance, the algorithm tends to overestimate the number of pins present while pins in the outer ring of the assembly seem more difficult to identify with high confidence. While the comparison between Poisson and iid Gaussian noise is difficult as in the Poisson case the signal-to-noise ratio is pixel-dependent, we observed that for similar average SNR, the activity estimation is more challenging in the Poisson case, probably because of the non-stationary SNR.

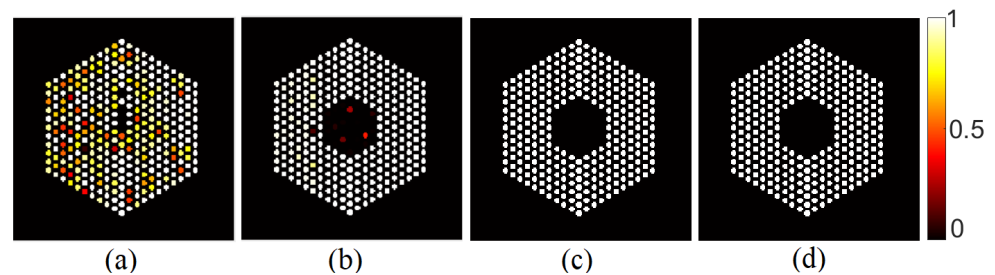


Figure 6. Probability of pin presence for (a) the Poisson noise case when peak value is set to 10, (b) the Gaussian noise case for peak value of 220 and noise variance $\sigma^2 = 1 \times 10^{-2}$, and (c) the Poisson noise case when peak value is set to 220, and (d) the Gaussian noise case when the peak value is set to 220 and the noise variance is set to $\sigma^2 = 2 \times 10^{-3}$.

5.3.2. Comparison with Existing Methods

For completeness, Figure 7 shows examples of results with the SPA-P- ℓ_1 approach for pin activity estimation using the Poisson noise model. The results for the Gaussian noise model (using SPA-G- ℓ_1) present similar trends and hence are not included here. For all pin pixels and pin pixels averaging results, we can observe that the posterior means look similar to those of the proposed approach. However, from the ratio of marginal standard deviation divided by posterior mean, the SPA-P- ℓ_1 approach is not confident about the 10% missing pins in the center. On the other hand, Figure 8 shows the results of the PIDAL- ℓ_1 approach for pin activity estimation using Poisson noise model. Those of the Gaussian

noise model using FISTA- ℓ_1 are similar and hence are not presented here. We can observe that the maximum *a posteriori* estimate is similar to the MMSE using both the proposed approach and SPA-P- ℓ_1 . However, these MAP-based methods cannot provide uncertainty measures about the estimates.

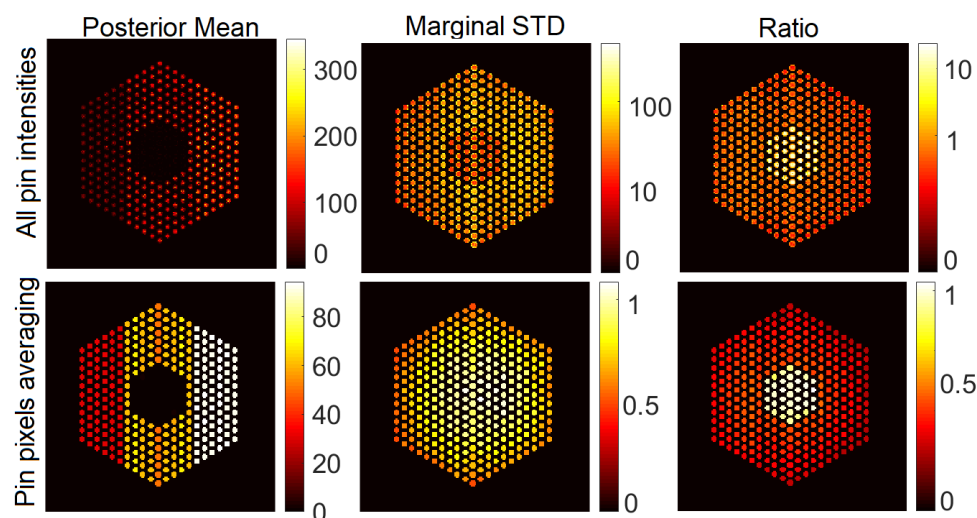


Figure 7. Pin activity estimation using SPA-P- ℓ_1 , using the Poisson noise model when peak value is set to 220. The results using the Gaussian noise model (SPA-G- ℓ_1) using peak value = 220 and noise variance $\sigma^2 = 2 \times 10^{-3}$ present similar trends and hence are not presented here.

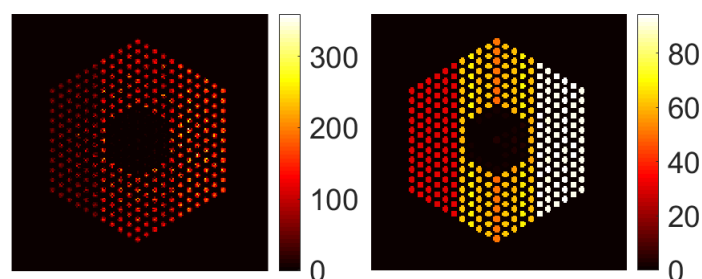


Figure 8. Pin activity estimation using PIDAL- ℓ_1 , using the Poisson noise model when the peak value is set to 220. (Left): all pin pixels; (right): average of pin pixels.

6. Simulations Using Realistic Datasets

In this section, we simulated the measurement of five fuel pin configurations in a mock-up water-water energetic reactor (WWER) fuel assembly using the software MCNP 6.2 [38], as shown in Figure 9. In this figure, we show the ground truth assembly configuration of each case. As in Case “1”, the mock-up fuel assembly hosted 331 ^{60}Co pins, which emitted 1.17 and 1.33 MeV gamma rays. We simulated ^{60}Co pins because they are used for the testing of the PGET system. In Cases “2” and “3”, we mimicked the scenario where fuel pins were missing by removing 10% fuel pins in the assembly. Each pin in the three cases is formed by a number of pixels belonging to {10–16}. In Cases “4” and “5”, 10% fuel pins were replaced by depleted uranium pins of the same size as the ^{60}Co ones. These pins are not emitting measurable gamma rays. In the five cases, the activity in each white pixel is constant. In contrast to Section 5 where we used a linear operator \mathbf{A} constructed using physical model of collimator-detector system response and simulated data using the linear model in Equation (1), here the data are generated using the MCNP 6.2 software that simulates all photon interactions in the PGET system [38]. More details on the simulation set up can be found in [22]. The simulated detection unit is a high-fidelity model of the PGET and consists of two collimated CdTe detector arrays on the top and bottom sides of the fuel assembly. The detector arrays rotate and scan the fuel bundle generating a 182×360 sinogram, as shown in Figure 10. Due to the mismatch between the

observations and \mathbf{Ax} , the robustness of the linear model is assessed using the two noise models mentioned earlier; the iid Gaussian and the Poisson noise models.

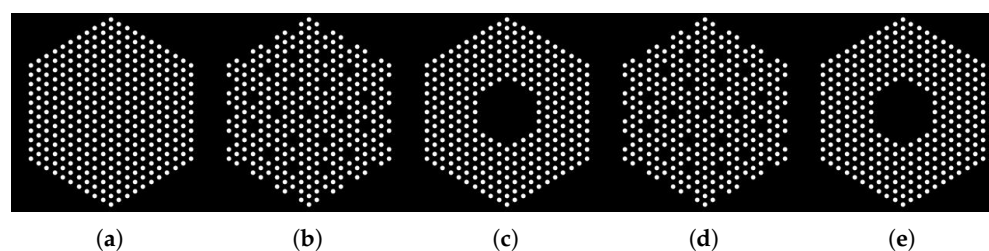


Figure 9. Actual binary masks of the simulated fuel assemblies. (a) Case “1”, (b) Case “2”, (c) Case “3”, (d) Case “4”, (e) Case “5”. In Case “1”, the mock-up fuel assembly hosted 331 ^{60}Co pins of 8.879 g/cm^3 density, which emitted 1.17 and 1.33 MeV gamma rays. In Cases “2” and “3”, 10% ^{60}Co pins were removed at the center or uniformly. In Cases “4” and “5”, 10% ^{60}Co pins were replaced by depleted uranium pins of the same size but higher density (10.4 g/cm^3) and no activity.

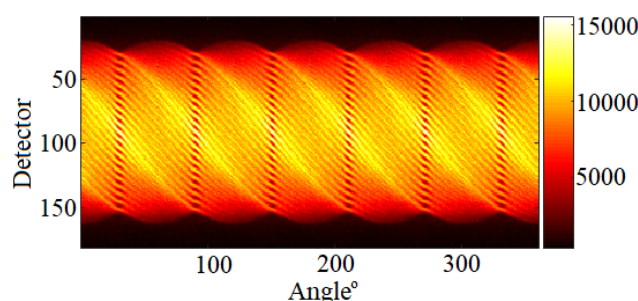


Figure 10. Simulated sinogram of Case “1”. The colour scale represents the detected photon counts.

Different response matrices, namely the IDEAL, the FULL and the EMPIRICAL, were tested to estimate the pin activity. To calculate a response matrix \mathbf{A} , pin locations are needed as input parameters to identify the pixels inside the pins and account for scattering and attenuation effects. To construct the IDEAL response matrices, the exact pin positions in each case (ground truth) are used. With such response matrices, we run the proposed algorithms to estimate the pin activity levels and to confirm whether the methods tend to underestimate the number of pins. For Case “1”, where no pins are missing, $\mathbf{A} \in \mathbb{R}^{65,520 \times 5329}$. However, for the rest of the cases (Cases “2” to “5”) where 10% of the pins are missing, the response matrix of the system is of size $\mathbf{A} \in \mathbb{R}^{65,520 \times 4732}$. In practice, these IDEAL response matrices are not available as we aim precisely at identify which pins are present. A more realistic scenario consists of assuming that all the pins are *a priori* present (although they may not). This assumption is accurate in Case 1, but it is expected to introduce errors for the other cases, in particular if a large number of pins is missing. The resulting matrix is of size $65,520 \times 5329$ and it coincides with the IDEAL matrix of Case 1. The last strategy considered here to build \mathbf{A} is referred to as EMPIRICAL and consists of identifying, as a pre-processing step, the likely pin support using a fast inversion algorithm (filtered backprojection (FBP), as in [22]). FBP provides a coarse activity map which is then thresholded to identify potential pin locations. This approach is more evolved than using the FULL \mathbf{A} and can discard correctly potential pin locations, but it can also remove legit pin locations if the threshold is set incorrectly. Using the empirical matrices \mathbf{A} , a matrix is created for each of the five cases.

6.1. The IDEAL Response Matrix

6.1.1. Results Using the Proposed Approach

In this section, we test the performance of the proposed approach using the IDEAL response matrices. In this context, the linear model in Equation (1) is expected to be a good approximation of the actual forward model. In all cases, the hyperparameters of the

model are set like those described in Section 5. Figures 11–13 show the results of activity estimation and the associated uncertainty measures, for Case 1 (full assembly), Cases 2 and 3 (missing pins) and Cases 4 and 5 (replaced pins), respectively. In these figures, we can observe that the Gaussian noise model generally provides more uniform pin activities compared to when using the Poisson noise model. Moreover, the activity profiles using the Gaussian noise model are lower, compared to the Poisson noise model, as it fitted the data better and hence improved the convergence and mixing properties of the MCMC chains. Figure 14 shows the posterior probabilities of presence of pins in the five investigated assembly patterns, using the Poisson noise model. We can observe that existing pins have a probability of presence of almost 1, thus the proposed approach is very confident in assessing the presence of pins. Similar results have been obtained using the Gaussian model and are thus not presented here. Both noise models allow correct pin identification for the five assembly patterns tested, as no additional pins are identified as missing.

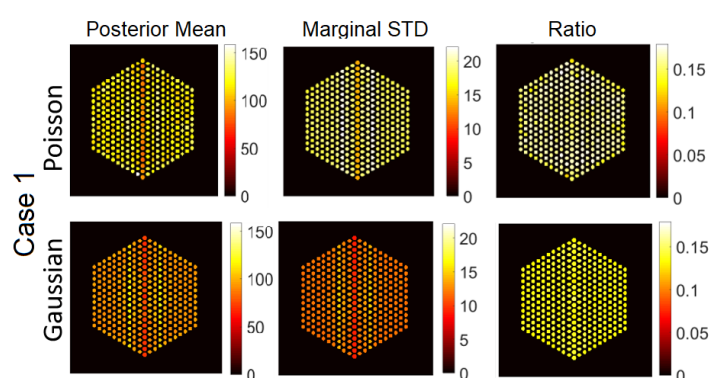


Figure 11. The IDEAL response matrix: Pin pixels averaging of the proposed approach for Case “1”. (Left-hand column): posterior mean, (middle column): marginal standard deviation, and (right-hand column): the ratio of marginal standard deviation divided by posterior mean.

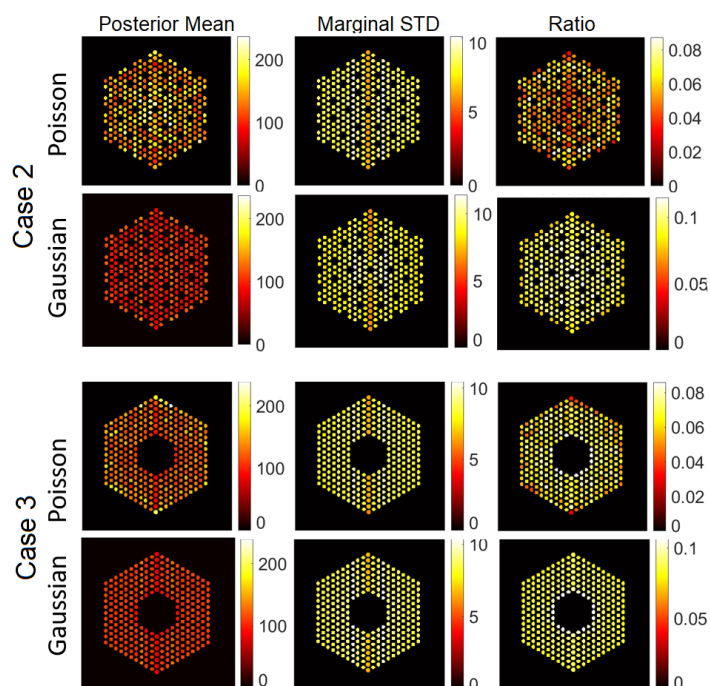


Figure 12. The IDEAL response matrix: Pin pixels averaging of the proposed approach for Case “2” and Case “3”. (Left-hand column): posterior mean, (middle column): marginal standard deviation, and (right-hand column): the ratio of marginal standard deviation divided by posterior mean.

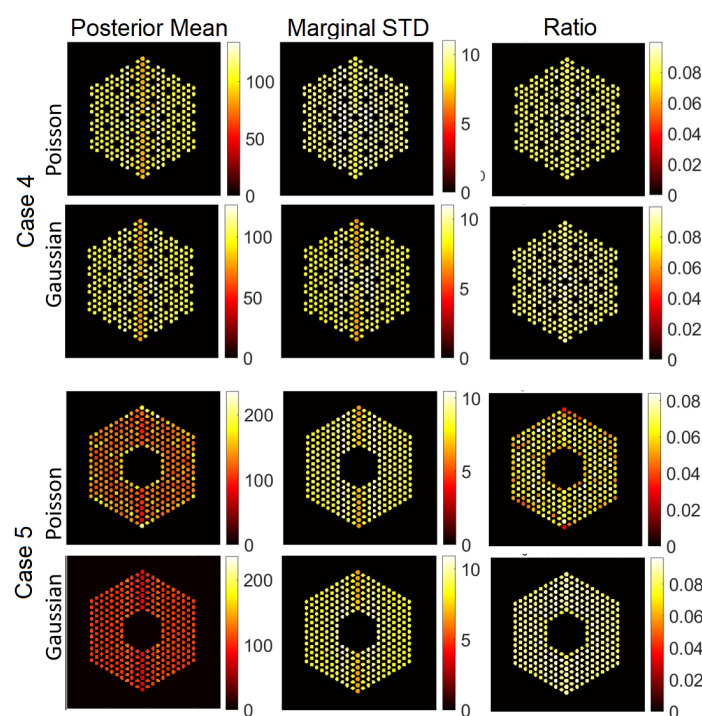


Figure 13. The IDEAL response matrix: Pin pixels averaging of the proposed approach for Case “4” and Case “5”. (Left-hand column): posterior mean, (middle column): marginal standard deviation, and (right-hand column): the ratio of marginal standard deviation divided by posterior mean.

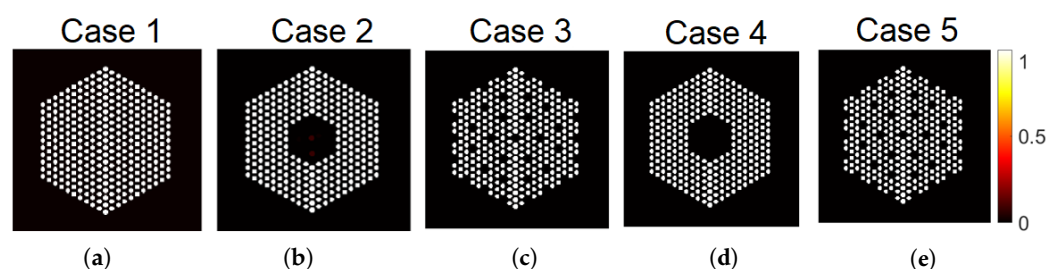


Figure 14. The IDEAL response matrix: Probability of presence of pins using the proposed approach for the Poisson noise model. (a) Case “1”; (b) Case “2”, (c) Case “3”, (d) Case “4” and (e) Case “5”. The Gaussian noise model showed similar trends and hence not presented.

6.1.2. Comparison with Existing Methods

In this part, we compare the proposed approach against the four existing methods described earlier that are SPA-G- ℓ_1 and FISTA- ℓ_1 for the Gaussian noise case, and SPA-P- ℓ_1 and PIDAL- ℓ_1 for the Poisson noise case. Figure 15 shows the activity estimation results of the five assembly patterns using these methods. We can observe that the results are similar to those obtained by the proposed approach. Moreover, all of the methods correctly identify the assembly pattern of all datasets, as in addition to the pins already identified as missing, no extra pins are. It is also worth mentioning here that the convergence of PIDAL- ℓ_1 , which considers a Poisson likelihood, was much slower than FISTA- ℓ_1 which considers a Gaussian likelihood. On the other hand, in Table 1, we show the number of particles (NPS) for the proposed approaches and the four existing methods. NPS is the mean of active pin pixels in each case. For each case, we estimated the activity of each pin by summing up pin pixel intensities and created a histogram of pin activities. The mean of this histogram is then computed, which represents the average fuel activity of each case. We can observe that, in general, the methods considering a Gaussian likelihood (FISTA- ℓ_1 , SPA-G-BtG and SPA-G- ℓ_1) provide fairly close results to the ground truth NPS.

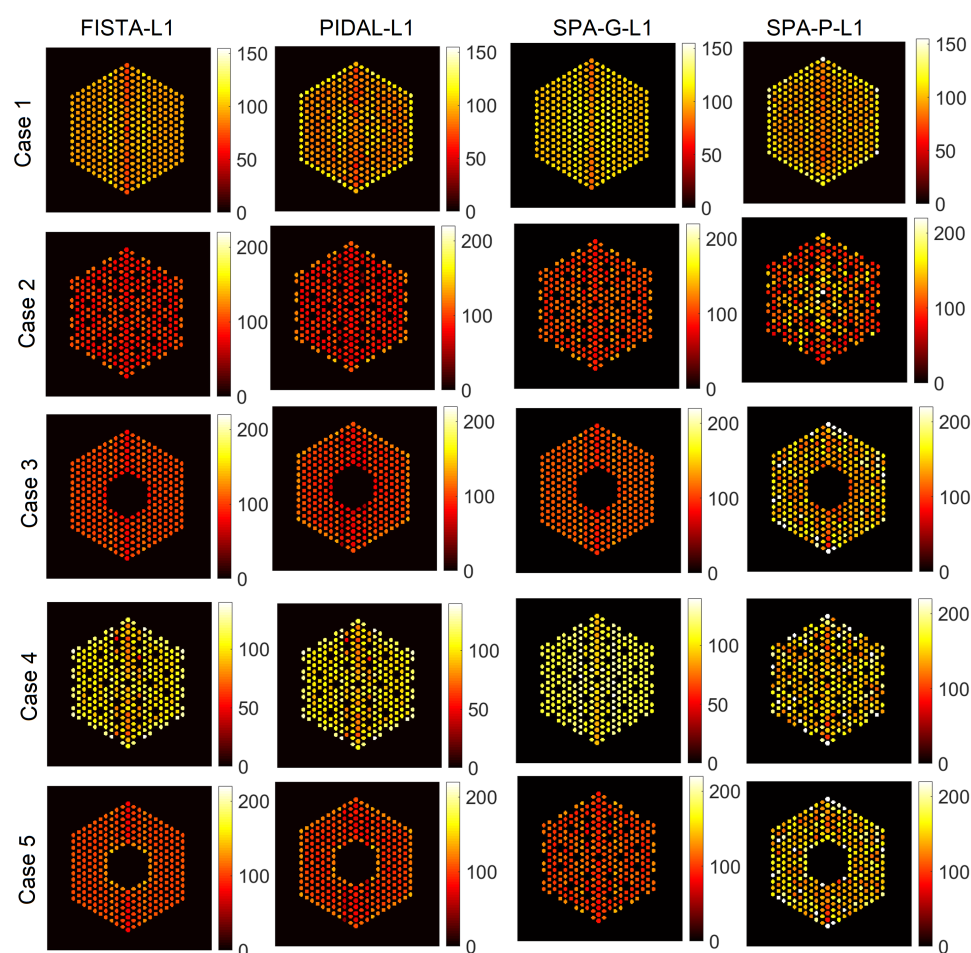


Figure 15. The IDEAL response matrix: Activity estimation using pin pixels averaging using existing methods. The regularization parameter β was set to 1×10^{-5} in the four methods.

Table 1. The IDEAL response matrix: Comparison of activity estimations ($\times 10^7$), represented by number of particle (NPS), for the proposed approaches and the four existing methods. NPS is the mean of active pin pixels in each case. Estimated activities close to the ground truth are highlighted in bold. Second close values are underlined.

	SPA-P-BtG	SPA-G-BtG	SPA-P- ℓ_1	SPA-G- ℓ_1	PIDAL- ℓ_1	FISTA- ℓ_1	Ground Truth
Case 1	2.00	1.49	2.244	1.65	1.512	<u>1.524</u>	1.51
Case 2	2.44	<u>1.54</u>	2.89	1.72	1.47	1.513	1.70
Case 3	2.10	1.58	2.37	1.77	1.49	<u>1.624</u>	1.70
Case 4	2.25	<u>1.61</u>	1.68	1.83	1.57	1.68	1.70
Case 5	2.16	<u>1.65</u>	2.47	1.80	1.69	1.69	1.70

Table 2 provides the average computation time of the proposed approach and the existing methods. The algorithms are implemented in MATLAB and the experiments are carried out on a laptop with a 2.8 GHz processor CPU, with 16 GB of RAM, under Microsoft Windows 10. It is clear that FISTA- ℓ_1 provides the lowest computation cost and PIDAL- ℓ_1 provides the highest computation complexity which is because it requires more iterations to converge. Despite the relatively high computation time of the proposed algorithms, they do not require fine tuning of the regularization parameter as in PIDAL/FISTA- ℓ_1 , SPA-P- ℓ_1 and SPA-G- ℓ_1 . Moreover, it can provide the probability of presence of pins and quantify their uncertainty measures in contrast to the other algorithms, as outlined in Table 2.

Table 2. Computation time (in minutes), using CPU for the proposed approach and state of the art algorithms.

Method	SPA-P-BtG	SPA-G-BtG	SPA-P- ℓ_1	SPA-G- ℓ_1	PIDAL- ℓ_1	FISTA- ℓ_1
Computation time (min)	50	60	45	55	75	3
Uncertainty maps	Yes	Yes	Yes	Yes	No	No

6.2. The FULL Response Matrix

6.2.1. Results Using the Proposed Approach

The IDEAL response matrices deployed in the previous subsection are not available, as we aim precisely at identify which pins are present. In this subsection, we investigate a more realistic scenario, consisting of the assumption that all the pins are *a priori* present (although, they may not be). This assumption is accurate in Case 1, but it expected to introduce errors for the other cases, as we will see below. The resulting matrix is of size $\mathbf{A} \in \mathbb{R}^{65,520 \times 5329}$. The hyperparameters of the proposed methods are set to same values as in the experiments in the Section 6.1. Figures 16–18 show the results activity estimation results obtained by our methods for the five cases. We can see that the algorithms could fairly accurately identify the assembly pattern of cases “1”, “2” and “3” but not for cases “4” and “5”. This can also be observed in the uncertainty maps of these cases. Moreover, the Gaussian noise model seems more robust in missing pins identification than using the Poisson noise model. On the other hand, Figure 19 shows the probability of presence of pins in each assembly. We can observe that the assembly patterns of cases “1”, “2”, and “3” are correctly identified, but not for cases “4” and “5”. In cases “4” and “5”, the cobalt pins are replaced by depleted UO2 and the Compton scattering cross-section of UO2 is larger than cobalt’s. UO2 pins are simulated as non-radioactive, as their radioactivity is negligible and UO2 could be used as a high-density fuel replacement material in diversion scenarios. However, UO2 pins are scattering centers of the gamma rays emitted by the surrounding fuel pins and the use of the FULL response matrix induces the algorithm to mis-classify UO2 pins as low-activity cobalt pins. One strategy to mitigate this effect could rely on narrowing the energy detection window to exclude the low-energy contribution from scattering reactions.

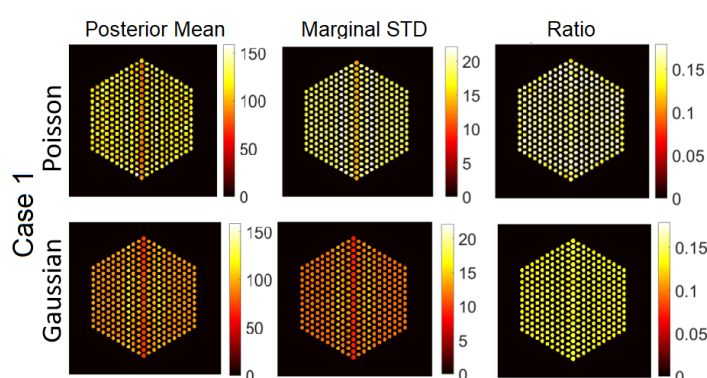


Figure 16. The FULL response matrix: pin pixels, averaging of the proposed approach for Case 1 “Full assembly”. (Left-hand column): posterior mean; (middle column): marginal standard deviation; and (right-hand column): the ratio of marginal standard deviation divided by posterior mean.

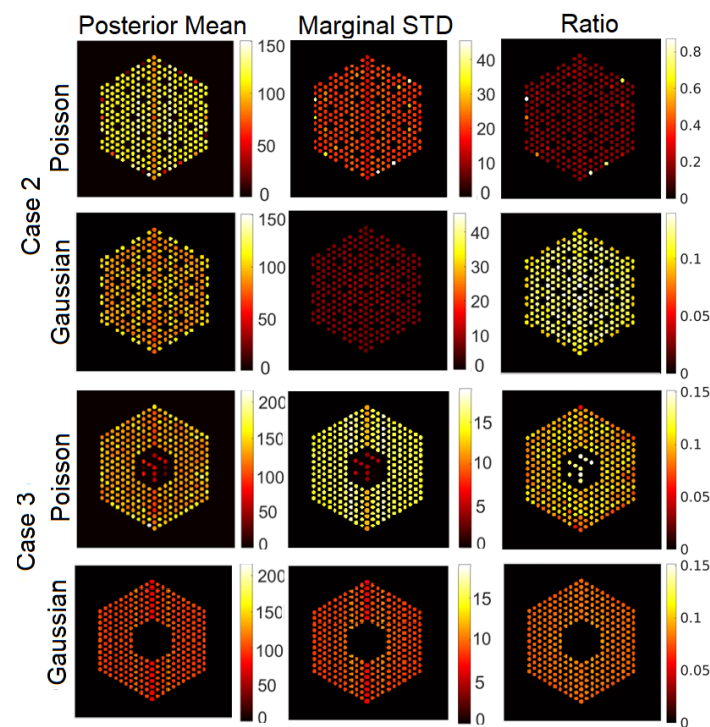


Figure 17. The FULL response matrix: pin pixels averaging of the proposed approach for Case 2 “Uniform Missing Assembly” and Case 3 “Center Missing Assembly”. (**Left-hand column**): posterior mean; (**middle column**): marginal standard deviation (in log-scale); and (**right-hand column**): the ratio of marginal standard deviation divided by posterior mean.

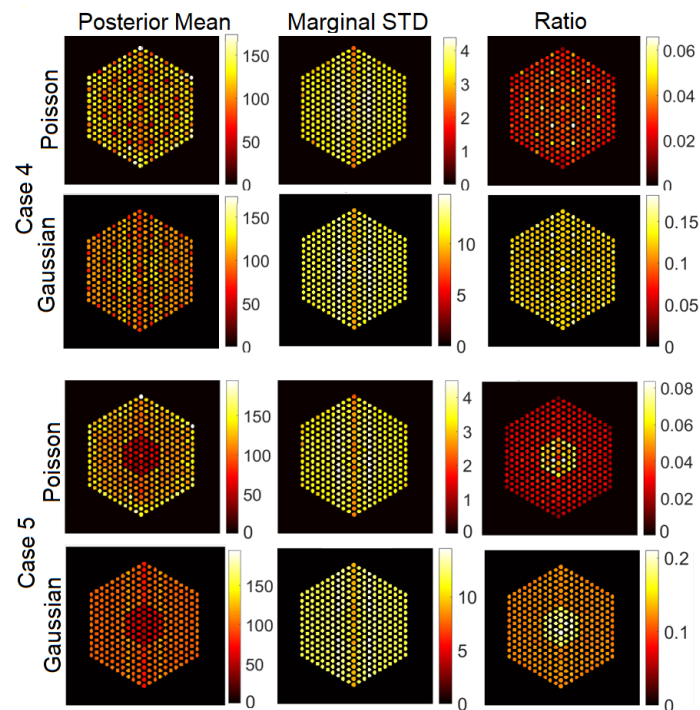


Figure 18. The FULL response matrix: pin pixels averaging of the proposed approach for Case “4” and Case “5”. (**Left-hand column**): posterior mean; (**middle column**): marginal standard deviation; and (**right-hand column**): the ratio of marginal standard deviation divided by posterior mean.

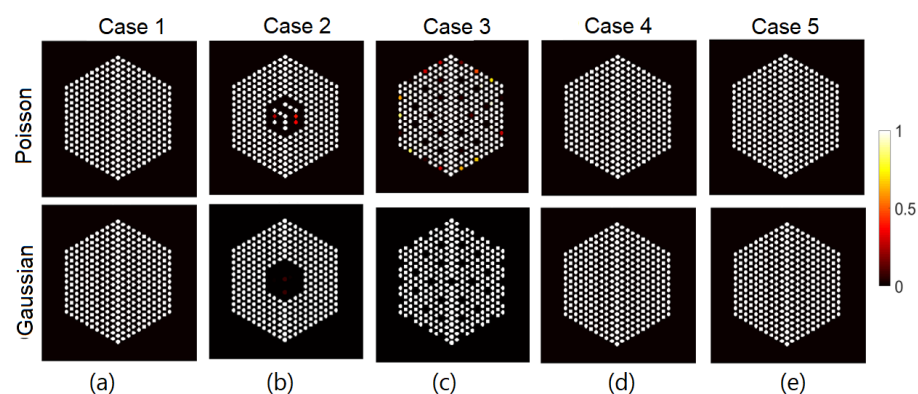


Figure 19. The FULL response matrix: probability of presence of pins using the proposed approach. (a) Case “1”; (b) Case “2”, (c) Case “3”, (d) Case “4”, and (e) Case “5”.

6.2.2. Comparison with Existing Methods

As in the previous section, the proposed approach is compared against the four approaches described earlier. Figure 20 shows the activity estimation results of the five assembly configurations. We can observe that for the missing center case, there are still some low amplitude identified pins at the center. Moreover, the proposed approach provides better results, in addition to being able to quantify the uncertainty measures of the estimates and also localize the pins.

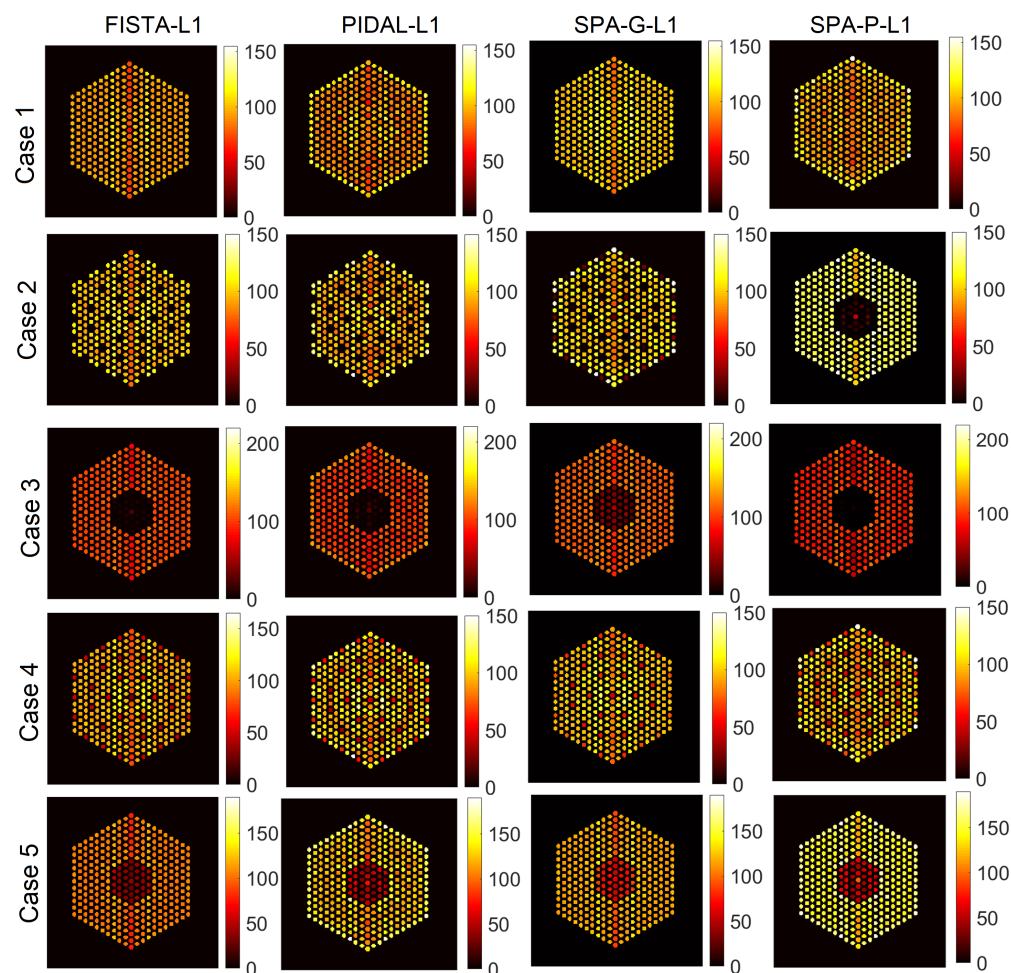


Figure 20. The FULL response matrix: activity estimation using pin pixels averaging using existing methods for the five cases. The regularization parameter β was set to 1×10^{-5} in the four methods.

6.3. The EMPIRICAL Response Matrix

As outlined in Section 6.2, the proposed approach, and existing methods, find it difficult to provide reliable estimates for cases “4” and “5”, which is probably because of the mismatch between the linear operator and the observations. Hence, the last strategy considered here is to build \mathbf{A} by identifying as a pre-processing step, the likely pin support using a fast inversion algorithm (filtered backprojection (FBP), as in [22]). FBP provides a coarse activity map which is then thresholded to identify potential pin locations. This approach is more evolved than using the FULL \mathbf{A} and can discard correctly potential pin locations. Using the empirical matrices \mathbf{A} , a matrix is created for each of the last two cases (Case 4 and 5), which is of size $\mathbf{A} \in \mathbb{R}^{65,520 \times 4899}$. Figures 21 and 22 show pin activity estimation and presence maps using both the Poisson and Gaussian noise models. We can observe that the assembly patterns of both cases are correctly identified, although in the Case “4”, there are still some missing pixels identified as existing but with low probability of presence, as in Figure 22. On the other hand, Figure 23 shows a comparison with the existing approaches, which are similar to those obtained using the proposed approach but without associated uncertainty measures and probability of presence.

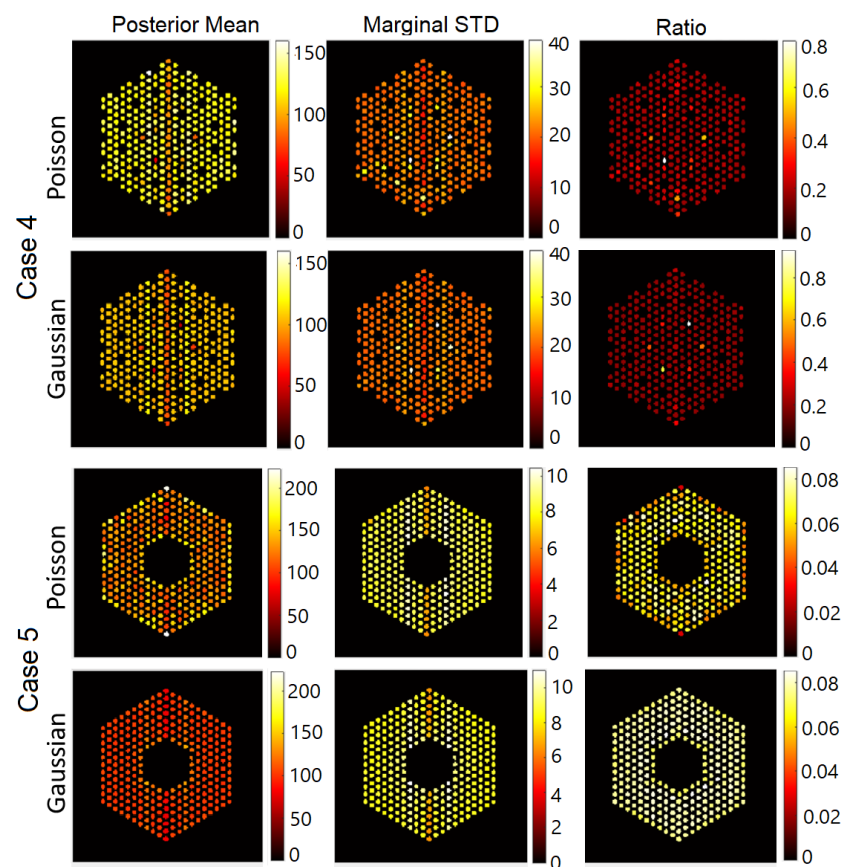


Figure 21. The EMPIRICAL response matrix: pin pixels averaging of the proposed approach for Case “4” and Case “5” using the “EMPIRICAL” response matrix. (Left-hand column): posterior mean; (middle column): marginal standard deviation; and (right-hand column): the ratio of marginal standard deviation divided by posterior mean.

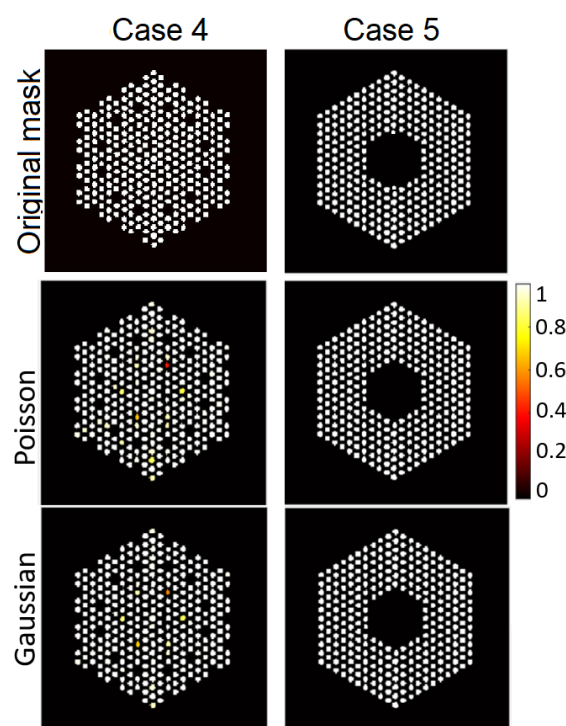


Figure 22. The EMPIRICAL response matrix: probability of presence of pins using the proposed approach. (Left): Case “4”; (right): Case “5”.

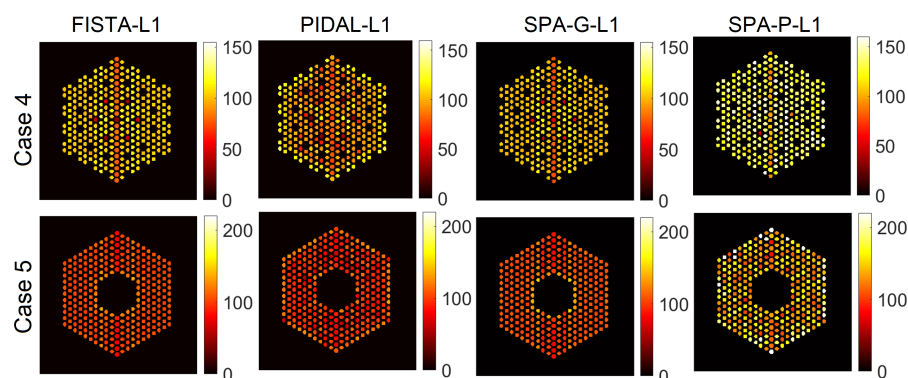


Figure 23. The EMPIRICAL response matrix: activity estimation using pin pixels averaging using existing methods, for Cases “4” and “5”. The regularization parameter β was set to 1×10^{-5} in the four methods.

7. Conclusions and Future Work

This paper introduced a hierarchical Bayesian model for pin activity estimation in PGET images, using a linear forward model. Due to the linear approximation, the classical Poisson noise assumption is likely to not hold and we assessed the robustness of the model tested using two different noise models. Prior distributions were assigned to the unknown model parameters and Bayesian inference was performed using a split and augmented—partially collapsed Gibbs sampler. The proposed approaches can provide uncertainty measures to the estimates and are fully automatic in the sense that they can estimate the model associated hyperparameters. Different response matrices in the linear approximation were also investigated. We observed that the Gaussian noise model provided better results, as it seemed more robust to model mismatch. The proposed approaches can provide uncertainty measures and posterior probabilities of presence, which cannot be obtained using classical optimisation methods. However, this comes at a higher computation cost. Future work should investigate alternative approaches to model the forward model.

Author Contributions: Conceptualization, A.K.E., Y.A. and Y.W.; methodology, A.K.E., Y.A. and Y.W.; software, A.K.E.; validation, Y.A. and Y.W.; formal analysis, A.K.E.; investigation, A.D.F., S.M., M.E.D., Y.A. and Y.W.; resources, A.D.F., S.M., M.E.D., Y.A. and Y.W.; data curation, A.K.E. and M.F.; writing—original draft preparation, A.K.E.; writing—review and editing, M.F., A.D.F., S.M., M.E.D., Y.A. and Y.W.; visualization, A.K.E., Y.A. and Y.W.; supervision, A.D.F., S.M., M.E.D., Y.A. and Y.W.; project administration, S.M., M.E.D., Y.A. and Y.W.; funding acquisition, S.M., A.D.F., M.E.D., Y.A. and Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Engineering and Physical Sciences Research Council of the UK (EPSRC) Grant number EP/S000631/1 (M.E.D.), the UK MOD University Defence Research Collaboration (UDRC) in Signal Processing (A.K.E., S.M., M.E.D., Y.A. and Y.W.), the Royal Academy of Engineering under the Research Fellowship scheme RF201617/16/31 (Y.A.), EPSRC under Grant number EP/T028270/1 (Y.W.), Nuclear Regulatory Commission (NRC) Faculty Development Grant 31310019M0011 (M.F. and A.D.F.), and Department of Energy STTR DE-SC0020733 Grant (M.F. and A.D.F.).

Institutional Review Board Statement: The authors declare research with no studies involving humans or animals.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding authors, A.K.E. and Y.A., upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

We now detail each sampling step of Algorithm 1 as follows:

Sampling $(s^2, \mathbf{t}_{z_g=0}) | (\mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda)$: We can notice that:

$$f(s^2, \mathbf{t}_{z_g=0} | \mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda) = f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda) f(\mathbf{t}_{z_g=0} | \mathbf{y}, \Phi, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda),$$

where:

$$f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda) = \int f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta, \mathbf{Y}, \Lambda) d\mathbf{t}_{z_g=0}. \quad (\text{A1})$$

Thus, sampling from $(s^2, \mathbf{t}_{z_g=0}) | (\mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda)$ can be achieved by sampling sequentially s^2 from $f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda)$, and then $\mathbf{t}_{z_g=0}$ from $f(\mathbf{t}_{z_g=0} | \mathbf{y}, \Phi, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda)$. To compute $f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta, \mathbf{Y}, \Lambda)$ in Equation (A1), we keep only the parameters that depend on s^2 in the joint posterior distribution in Equation (11), this leads to:

$$f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta, \mathbf{Y}, \Lambda) \propto \prod_{n=1}^N \mathcal{N}_{\mathbb{R}^+}(t_n; 0, s^2) \times \mathcal{IG}(s^2; \eta, \nu). \quad (\text{A2})$$

Due to the conjugacy of the distributions in Equation (A2), sampling from $f(s^2 | \mathbf{y}, \Phi_{s^2}, \Delta, \mathbf{Y}, \Lambda)$ can be achieved by sampling from the following inverse-gamma distribution [39]:

$$s^2 | (\mathbf{y}, \Phi_{s^2}, \Delta, \mathbf{Y}, \Lambda) \sim \mathcal{IG}\left(\eta + \frac{N}{2}, \nu + \frac{\sum_{n=1}^N t_n^2}{2}\right). \quad (\text{A3})$$

However, marginalizing $\mathbf{t}_{z_g=0}$ out the conditional distribution in Equation (A3), according to Equation (A1), gives:

$$s^2 | (\mathbf{y}, \Phi_{s^2}, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda) \sim \mathcal{IG}\left(\eta + \frac{N_1}{2}, \nu + \frac{\sum_{n \in I_1} t_n^2}{2}\right), \quad (\text{A4})$$

where $N_1 = \text{card}(\mathbf{t}_{z_g=1})$ and $I_1 = \{n | z_n = 1\}$.

On the other hand, sampling $f(\mathbf{t}_{z_g=0} | \mathbf{y}, \Phi, \Delta_{\mathbf{t}_{z_g=0}}, \mathbf{Y}, \Lambda)$ reduces to sampling its elements independently from their Gaussian prior distribution defined in Equation (6).

Sampling $(\mathbf{z}, \mathbf{t}) | (\mathbf{y}, \Phi, \mathbf{Y}, \Lambda)$: Updating (\mathbf{z}, \mathbf{t}) is achieved using the same marginalization trick as in Step (a) of Algorithm 1, i.e.:

$$f(\mathbf{z}, \mathbf{t} | \mathbf{y}, \Phi, \mathbf{Y}, \Lambda) = f(\mathbf{z} | \mathbf{y}, \Phi, \mathbf{Y}, \Lambda) f(\mathbf{t} | \mathbf{y}, \Delta_{\mathbf{t}}, \Phi, \mathbf{Y}, \Lambda).$$

It can be seen from Equation (11) that:

$$f(\mathbf{z} | \mathbf{y}, \Phi, \mathbf{t}, \mathbf{Y}, \Lambda) \propto \prod_{g=1}^G \left(\left[\omega^{z_g} (1 - \omega)^{(1-z_g)} \right] \times \left(2\pi\rho^2 \right)^{-\frac{K_g}{2}} \exp \left(-\frac{\|z_g \mathbf{t}_g - (\mathbf{c}_g - \mathbf{o}_{2g}\|_2^2)}{2\rho^2} \right) \right), \quad (\text{A5})$$

where $\mathbf{t}_g \in \mathbb{R}^{K_g}$ and that

$$f(z_g = e | \mathbf{y}, \Delta_{(z_g, \mathbf{t}_{g'})}, \Phi, \mathbf{Y}, \Lambda) \propto \bar{\omega}_g^{(e)} \quad \forall (g), \quad (\text{A6})$$

where $e \in \{0, 1\}$ and

$$\begin{aligned} \log(\bar{\omega}_g^{(e)}) &= -\frac{K_g}{2} \log(2\pi(\rho^2 + es^2)) \\ &- \frac{\|\mathbf{c}_g\|_2^2}{2(\rho^2 + es^2)} + \log(p(e|\omega)) + e \log \left[\prod_{i=1}^{K_g} \text{erfc} \left(-\frac{s^2 c_i^2}{\sqrt{2s^2\rho^2(\rho^2 + s^2)}} \right) \right], \end{aligned} \quad (\text{A7})$$

where $\text{erfc}(\cdot)$ is the complementary error function (e.g., $\text{erfc}(\cdot) = 1 - \text{erf}(\cdot)$). Consequently, the label z_g can be drawn from its conditional distribution (where \mathbf{t}_g has been marginalised) by drawing randomly from $\{0, 1\}$ with probabilities given by:

$$f(z_g = e | \mathbf{y}, \Delta_{(z_g, \mathbf{t}_{g'})}, \Phi, \mathbf{Y}, \Lambda) = \frac{\bar{\omega}_g^{(e)}}{\bar{\omega}_g^{(0)} + \bar{\omega}_g^{(1)}}. \quad (\text{A8})$$

Moreover, the elements of \mathbf{z} can be updated in a parallel manner using the fact that $f(\mathbf{z} | \mathbf{y}, \Phi, \Delta_{\mathbf{z}}, \mathbf{Y}, \Lambda) = \prod_g f(z_g | \mathbf{y}, \Delta_{\mathbf{t}}, \Phi, \mathbf{Y}, \Lambda)$.

Once \mathbf{z} is sampled from (A8), sampling $\mathbf{t} | (\mathbf{y}, \Delta_{\mathbf{t}}, \Phi, \mathbf{Y}, \Lambda)$ can be achieved by sampling from the following multivariate Gaussian distribution:

$$\begin{aligned} f(\mathbf{t} | \mathbf{y}, \Delta_{\mathbf{t}}, \Phi, \mathbf{Y}, \Lambda) &\propto \left(\prod_{n=1}^N \mathcal{N}_{\mathbb{R}^+}(t_n; \mathbf{0}, \mathbf{s}^2) \right) \times \exp \left(-\frac{\|\mathbf{Z}\mathbf{t} - (\mathbf{c} - \mathbf{o}_2)\|_2^2}{2\rho^2} \right) \\ \mathbf{t} | (\mathbf{y}, \Delta_{\mathbf{t}}, \Phi, \mathbf{Y}, \Lambda) &\sim \mathcal{N}_{\mathbb{R}^+}(\mathbf{t}; \boldsymbol{\mu}_{\mathbf{t}}, \boldsymbol{\Sigma}_{\mathbf{t}}), \end{aligned} \quad (\text{A9})$$

where:

$$\begin{cases} \boldsymbol{\mu}_{\mathbf{t}} = \left(\frac{\mathbf{Z}\mathbf{c}}{\rho^2} \right) \boldsymbol{\Sigma}_{\mathbf{t}}, \\ \boldsymbol{\Sigma}_{\mathbf{t}} = (\text{diag}(\mathbf{s}^{-2}) + \rho^{-2} \mathbf{Z}^T \mathbf{Z})^{-1}. \end{cases}$$

Note that $\mathbf{Z}^T \mathbf{Z} = \mathbf{Z}$ as \mathbf{Z} is a diagonal matrix. Hence, $\boldsymbol{\Sigma}_{\mathbf{t}}$ is a diagonal covariance matrix, which simplifies the sampling process.

Sampling $\omega | (\mathbf{y}, \Delta, \Phi_{\omega}, \mathbf{Y}, \Lambda)$: Due to the conjugacy of the hierarchical prior model $f(\mathbf{z}|\omega)f(\omega)$, the full conditional distribution of ω reduces to the following beta distribution:

$$\omega | (\mathbf{y}, \Delta, \Phi_{\omega}, \mathbf{Y}, \Lambda) \sim \mathcal{B}e \left(\alpha + \sum_{g=1}^G K_g z_g, \beta + N - \sum_{g=1}^G K_g z_g \right). \quad (\text{A10})$$

Sampling $\mathbf{b} | (\mathbf{y}, \Delta, \Phi, \mathbf{Y}_{\mathbf{b}}, \Lambda)$: It can be seen from Equation (13) that the full conditional distribution of \mathbf{b} reduces to:

$$f(\mathbf{b} | \mathbf{y}, \mathbf{b}, \mathbf{c}, \mathbf{o}_1) \propto \mathcal{F}(\mathbf{b}) \times \exp \left(-\frac{\|\mathbf{Ac} - (\mathbf{b} - \mathbf{o}_1)\|_2^2}{2\rho^2} \right). \quad (\text{A11})$$

When $\mathcal{F}(\mathbf{b}) = \mathcal{P}(\mathbf{b})$, due to the non-differentiability of this conditional distribution, we consider the proximal Moreau-Yoshida-unadjusted Langevin algorithm (P-MYULA) proposed in [20] to generate samples asymptotically distributed according to Equation (A11). Note also that [40] can be used which can provide faster convergence. The Markov chain of P-MYULA can be written as:

$$\mathbf{b}^{(k+1)} = (1 - \frac{\gamma}{\lambda})\mathbf{b}^{(k)} - \gamma \nabla f(\mathbf{b}^{(k)}) + \frac{\gamma}{\lambda} \text{prox}_g^\lambda(\mathbf{b}^{(k)}) + \sqrt{2\gamma} \mathcal{N}_{\mathbb{R}^+}(\mathbf{0}, \mathbf{I}), \quad (\text{A12})$$

where

$$\begin{aligned} f(\mathbf{b}) &= \frac{\|\mathbf{Ac} - (\mathbf{b} - \mathbf{o}_1)\|_2^2}{2\rho^2}, & g(\mathbf{b}) &= -\log \mathcal{P}(\mathbf{b}), \\ \nabla f(\mathbf{b}) &= \frac{\mathbf{b} - \mathbf{Ac} - \mathbf{o}_1}{\rho^2}, & \text{prox}_g^\lambda(\mathbf{b}) &= \frac{1}{2} \left(\mathbf{b} - \lambda + \sqrt{(\mathbf{b} - \lambda)^2 + 4\lambda \mathbf{y}} \right). \end{aligned} \quad (\text{A13})$$

However, when $\mathcal{F}(\mathbf{b}) = f(\mathbf{y}|\mathbf{b}) = (2\pi\sigma^2)^{-N/2} \exp \left(-\frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{b}\|_2^2 \right)$, the full conditional distribution of \mathbf{b} reduces to sampling from the following multivariate Gaussian distribution

$$p(\mathbf{b} | \Delta, \Phi, \mathbf{Y}_{\mathbf{b}}, \Lambda) \propto \mathcal{N}(\mathbf{b}; \mu_{\mathbf{b}}, \Sigma_{\mathbf{b}}), \quad (\text{A14})$$

where:

$$\mu_{\mathbf{b}} = (\sigma^{-2}\mathbf{y} + \rho^{-2}\mathbf{Ac})\Sigma_{\mathbf{b}}^{-1}, \text{ and } \Sigma_{\mathbf{b}} = \frac{\sigma^2\rho^2}{\sigma^2 + \rho^2}\mathbf{I}_N. \quad (\text{A15})$$

A simpler splitting strategy using only one split can be adopted when the noise model is Gaussian, that is $\mathbf{h} = \mathbf{Zt}$, thus eliminating one extra variable from the algorithm in Algorithm 1. We have tested the performance of the two splittings and did not find noticeable differences in the results. So, we keep the two splittings strategy described above for consistency with the Poisson noise model. Note that the noise variance σ^2 in the Gaussian noise case can be estimated using the method proposed in [41].

Sampling $\mathbf{c} | (\Delta, \Phi, \mathbf{Y}_{\mathbf{c}}, \Lambda)$: By cancelling out the terms that do not depend on \mathbf{c} in Equation (13), the full conditional distribution of \mathbf{c} can be written as:

$$p(\mathbf{c} | \Delta, \Phi, \mathbf{Y}_{\mathbf{c}}, \Lambda) \propto \exp \left(-\frac{\|\mathbf{Ac} - (\mathbf{b} - \mathbf{o}_1)\|_2^2}{2\rho^2} \right) \times \exp \left(-\frac{\|\mathbf{Zt} - (\mathbf{c} - \mathbf{o}_2)\|_2^2}{2\rho^2} \right), \quad (\text{A16})$$

which reduces to sampling from the following multivariate Gaussian distribution:

$$p(\mathbf{c} | \Delta, \Phi, \mathbf{Y}_{\mathbf{c}}, \Lambda) = \mathcal{N}(\mathbf{c}; \mu_{\mathbf{c}}, \Sigma_{\mathbf{c}}), \quad (\text{A17})$$

where:

$$\mu_{\mathbf{c}} = (\mathbf{A}^T(\mathbf{b} - \mathbf{o}_1) + \mathbf{Zt} - \mathbf{o}_2)(\mathbf{A}^T\mathbf{A} + \mathbf{I})^{-1} \text{ and } \Sigma_{\mathbf{c}} = \rho^2(\mathbf{A}^T\mathbf{A} + \mathbf{I})^{-1}. \quad (\text{A18})$$

Sampling directly from (A17) can be computationally intensive since, e.g., it requires inversion of the precision matrix, sampling $\mathbf{c} | (\Delta, \Phi, \mathbf{Y}_{\setminus \mathbf{c}}, \Lambda)$ can be efficiently achieved by the exact perturbation-optimization (E-PO) algorithm proposed in [42].

Sampling $\mathbf{o}_1 | (\Delta, \Phi, \mathbf{Y}, \Lambda_{\setminus \mathbf{o}_1})$: It can be seen that the full conditional distribution of \mathbf{o}_1 can be written as:

$$p(\mathbf{o}_1 | \Delta, \Phi, \mathbf{Y}, \Lambda_{\setminus \mathbf{o}_1}) \propto \exp\left(-\frac{\|\mathbf{A}\mathbf{c} - (\mathbf{b} - \mathbf{o}_1)\|_2^2}{2\rho^2}\right) \times \exp\left(-\frac{\|\mathbf{o}_1\|_2^2}{2\alpha^2}\right), \quad (\text{A19})$$

which reduces to sampling from the following multivariate Gaussian distribution:

$$p(\mathbf{o}_1 | \Delta, \Phi, \mathbf{Y}, \Lambda_{\setminus \mathbf{o}_1}) \propto \mathcal{N}(\mathbf{o}_1; \mu_{\mathbf{o}_1}, \Sigma_{\mathbf{o}_1}), \quad (\text{A20})$$

where:

$$\mu_{\mathbf{o}_1} = \frac{\alpha^2}{\rho^2 + \alpha^2}(\mathbf{A}\mathbf{c} - \mathbf{b}) \text{ and } \Sigma_{\mathbf{o}_1} = \frac{\rho^2\alpha^2}{\rho^2 + \alpha^2}\mathbf{I}_N. \quad (\text{A21})$$

In a similar fashion to sampling \mathbf{o}_1 , sampling \mathbf{o}_2 reduces to sampling from a multivariate Gaussian distribution with the following mean and covariance:

$$\mu_{\mathbf{o}_2} = \frac{\alpha^2}{\rho^2 + \alpha^2}(\mathbf{Z}\mathbf{t} - \mathbf{c}) \text{ and } \Sigma_{\mathbf{o}_2} = \frac{\rho^2\alpha^2}{\rho^2 + \alpha^2}\mathbf{I}_N. \quad (\text{A22})$$

Algorithm A1 is a detailed version of Algorithm 1 following the sampling steps explained above. As $\mathbf{t}_{z_g=0}$ generated in step (a) is not used during the following steps, it can thus be omitted.

Algorithm A1 Split and augmented—partially collapsed Gibbs sampling algorithm for activity estimation in PGET—version II.

- 1: **Fixed input parameters:** Number of burn-in iterations N_{bi} ; total number of iterations N_{MC}
 - 2: **Initialization** ($k = 0$)
 - Set $\mathbf{z}^{(0)}, \mathbf{t}^{(0)}, \mathbf{b}^{(0)}, \mathbf{c}^{(0)}, \omega^{(0)}, s^{(0)}, \mathbf{o}_1^{(0)}, \mathbf{o}_2^{(0)}$
 - 3: **Repeat** ($1 \leq k \leq N_{\text{MC}}$)
 - (1a) Sample $s^{(k)} | \left(\mathbf{y}, \omega^{(k-1)}, \Delta_{\setminus \mathbf{t}_{z_g=0}}^{(k-1)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)} \right)$ from Equation (A4).
 - (2a) Sample $\mathbf{z}^{(k)} | \left(\mathbf{y}, \mathbf{t}^{(k)}, s^{(k)}, \omega^{(k-1)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)} \right)$ from Equation (A8).
 - (2b) Sample $\mathbf{t}^{(k)} | \left(\mathbf{y}, \mathbf{z}^{(k)}, s^{(k)}, \omega^{(k-1)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)} \right)$ from Equation (A9).
 - (3) Sample $\omega^{(k)} | \left(\mathbf{y}, s^{(k)}, \Delta^{(k)}, \mathbf{Y}^{(k-1)}, \Lambda^{(k-1)} \right)$ from Equation (A10).
 - (4) Sample $\mathbf{b}^{(k)} | \left(\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{c}^{(k-1)}, \Lambda^{(k-1)} \right)$ from Equation (A14).
 - (5) Sample $\mathbf{c}^{(k)} | \left(\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{b}^{(k)}, \Lambda^{(k-1)} \right)$ from Equation (A9).
 - (6) Sample $\mathbf{o}_1^{(k)} | \left(\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{Y}^{(k)}, \mathbf{o}_2^{(k-1)} \right)$ from Equation (A20).
 - (7) Sample $\mathbf{o}_2^{(k)} | \left(\mathbf{y}, \Phi^{(k)}, \Delta^{(k)}, \mathbf{Y}^{(k)}, \mathbf{o}_1^{(k)} \right)$ from Equation (A22).
 - 4: **Set** $k = k + 1$.
-

References

1. Firmage, E.B. The treaty on the non-proliferation of nuclear weapons. *Am. J. Int. Law* **1969**, *63*, 711–746. [\[CrossRef\]](#)
2. White, T.; Mayorov, M.; Deshmukh, N.; Miller, E.; Smith, L.E.; Dahlberg, J.; Honkamaa, T. SPECT reconstruction and analysis for the inspection of spent nuclear fuel. In Proceedings of the 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Atlanta, GA, USA, 21–28 October 2017; pp. 1–2.
3. White, T.; Mayorov, M.; Lebrun, A.; Peura, P.; Honkamaa, T.; Dahlberg, J.; Keubler, J.; Ivanov, V.; Turunen, A. Application of passive gamma emission tomography (PGET) for the verification of spent nuclear fuel. In Proceedings of the INMM 59th Annual Meeting, Baltimore, MD, USA, 22–26 July 2018.
4. Mayorov, M.; White, T.; Lebrun, A.; Brutscher, J.; Keubler, J.; Birnbaum, A.; Ivanov, V.; Honkamaa, T.; Peura, P.; Dahlberg, J. Gamma emission tomography for the inspection of spent nuclear fuel. In Proceedings of the 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Atlanta, GA, USA, 21–28 October 2017; pp. 1–2.
5. Bélanger-Champagne, C.; Peura, P.; Eerola, P.; Honkamaa, T.; White, T.; Mayorov, M.; Dendooven, P. Effect of gamma-ray energy on image quality in passive gamma emission tomography of spent nuclear fuel. *IEEE Trans. Nucl. Sci.* **2018**, *66*, 487–496. [\[CrossRef\]](#)
6. Virta, R.; Backholm, R.; Bubba, T.A.; Helin, T.; Moring, M.; Siltanen, S.; Dendooven, P.; Honkamaa, T. Fuel rod classification from Passive Gamma Emission Tomography (PGET) of spent nuclear fuel assemblies. *arXiv* **2020**, arXiv:2009.11617.
7. Backholm, R.; Bubba, T.A.; Bélanger-Champagne, C.; Helin, T.; Dendooven, P.; Siltanen, S. Simultaneous reconstruction of emission and attenuation in passive gamma emission tomography of spent nuclear fuel. *arXiv* **2019**, arXiv:1905.03849.
8. Figueiredo, M.A.; Bioucas-Dias, J.M. Restoration of Poissonian images using alternating direction optimization. *IEEE Trans. Image Process.* **2010**, *19*, 3133–3145. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Afonso, M.V.; Bioucas-Dias, J.M.; Figueiredo, M.A. Fast image recovery using variable splitting and constrained optimization. *IEEE Trans. Image Process.* **2010**, *19*, 2345–2356. [\[CrossRef\]](#)
10. Afonso, M.; Bioucas-Dias, J.; Figueiredo, M.A. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.* **2010**, *20*, 681–695. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Carrillo, R.E.; McEwen, J.D.; Wiaux, Y. Sparsity averaging reweighted analysis (SARA): A novel algorithm for radio-interferometric imaging. *Mon. Not. R. Astron. Soc.* **2012**, *426*, 1223–1234. [\[CrossRef\]](#)
12. Carrillo, R.E.; McEwen, J.D.; Van De Ville, D.; Thiran, J.P.; Wiaux, Y. Sparsity averaging for compressive imaging. *IEEE Signal Process. Lett.* **2013**, *20*, 591–594. [\[CrossRef\]](#)
13. Eldaly, A.K.; Altmann, Y.; Akram, A.; Perperidis, A.; Dhaliwal, K.; McLaughlin, S. Patch-based sparse representation for bacterial detection. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 657–661.
14. Eldaly, A.K.; Altmann, Y.; Perperidis, A.; Krstajić, N.; Choudhary, T.R.; Dhaliwal, K.; McLaughlin, S. Deconvolution and restoration of optical endomicroscopy images. *IEEE Trans. Comput. Imaging* **2018**, *4*, 194–205. [\[CrossRef\]](#)
15. Eldaly, A.K.; Altmann, Y.; Perperidis, A.; McLaughlin, S. Deconvolution of irregularly subsampled images. In Proceedings of the 2018 IEEE Statistical Signal Processing Workshop (SSP), Freiburg im Breisgau, Germany, 10–13 June 2018; pp. 303–307.
16. Tachella, J.; Altmann, Y.; Pereyra, M.; McLaughlin, S.; Tourneret, J.Y. Bayesian restoration of high-dimensional photon-starved images. In Proceedings of the 2018 IEEE 26th European Signal Processing Conference (EUSIPCO), Rome, Italy, 3–7 September 2018; pp. 747–751.
17. Vono, M.; Dobigeon, N.; Chainais, P. Bayesian image restoration under Poisson noise and log-concave prior. In Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 1712–1716.
18. Vono, M.; Dobigeon, N.; Chainais, P. Split-and-augmented Gibbs sampler—Application to large-scale inference problems. *IEEE Trans. Signal Process.* **2019**, *67*, 1648–1661. [\[CrossRef\]](#)
19. Pereyra, M. Proximal markov chain monte carlo algorithms. *Stat. Comput.* **2016**, *26*, 745–760. [\[CrossRef\]](#)
20. Durmus, A.; Moulines, E.; Pereyra, M. Efficient bayesian computation by proximal markov chain monte carlo: when langevin meets moreau. *SIAM J. Imaging Sci.* **2018**, *11*, 473–506. [\[CrossRef\]](#)
21. Iyengar, A.S.; Hausladen, P.; Yang, J.; Fabris, L.; Hu, J.; Lacy, J.; Athanasiades, A. *Detection of Fuel Pin Diversion via Fast Neutron Emission Tomography*; Technical Report; Oak Ridge National Lab.(ORNL): Oak Ridge, TN, USA, 2017.
22. Fang, M.; Altmann, Y.; Della Latta, D.; Salvatori, M.; Di Fulvio, A. Quantitative imaging and automated fuel pin identification for passive gamma emission tomography. *Sci. Rep.* **2021**, *11*, 1–11. [\[CrossRef\]](#)
23. Newstadt, G.E.; Hero, A.O.; Simmons, J. Robust spectral unmixing for anomaly detection. In Proceedings of the 2014 IEEE Workshop on Statistical Signal Processing (SSP), Gold Coast, QLD, Australia, 29 June–2 July 2014; pp. 109–112.
24. Ishwaran, H.; Rao, J.S. Spike and slab variable selection: frequentist and Bayesian strategies. *Ann. Stat.* **2005**, *33*, 730–773. [\[CrossRef\]](#)
25. George, E.I.; McCulloch, R.E. Variable selection via Gibbs sampling. *J. Am. Stat. Assoc.* **1993**, *88*, 881–889. [\[CrossRef\]](#)
26. Chipman, H. Bayesian variable selection with related predictors. *Can. J. Stat.* **1996**, *24*, 17–36. [\[CrossRef\]](#)
27. Carvalho, C.M.; Chang, J.; Lucas, J.E.; Nevins, J.R.; Wang, Q.; West, M. High-dimensional sparse factor modeling: applications in gene expression genomics. *J. Am. Stat. Assoc.* **2008**, *103*, 1438–1456. [\[CrossRef\]](#)

28. Kail, G.; Tourneret, J.Y.; Hlawatsch, F.; Dobigeon, N. Blind deconvolution of sparse pulse sequences under a minimum distance constraint: A partially collapsed Gibbs sampler method. *IEEE Trans. Signal Process.* **2012**, *60*, 2727–2743. [\[CrossRef\]](#)
29. Lin, C.; Mailhes, C.; Tourneret, J.Y. P- and T-wave delineation in ECG signals using a Bayesian approach and a partially collapsed Gibbs sampler. *IEEE Trans. Biomed. Eng.* **2010**, *57*, 2840–2849.
30. Kail, G.; Tourneret, J.Y.; Hlawatsch, F.; Dobigeon, N. A partially collapsed Gibbs sampler for parameters with local constraints. In Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, USA, 14–19 March 2010; pp. 3886–3889.
31. Ge, D.; Idier, J.; Le Carpentier, E. A new MCMC algorithm for blind Bernoulli-Gaussian deconvolution. In Proceedings of the 2008 IEEE 16th European Signal Processing Conference, Lausanne, Switzerland, 25–29 August 2008; pp. 1–5.
32. Robert, C.; Casella, G. *Monte Carlo Statistical Methods*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2013.
33. Eldaly, A.K.; Altmann, Y.; Akram, A.; McCool, P.; Perperidis, A.; Dhaliwal, K.; McLaughlin, S. Bayesian Bacterial Detection Using Irregularly Sampled Optical Endomicroscopy Images. *Med. Image Anal.* **2019**, *57*, 18–31. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Liu, J.S. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. *J. Am. Stat. Assoc.* **1994**, *89*, 958–966. [\[CrossRef\]](#)
35. Van Dyk, D.A.; Park, T. Partially collapsed Gibbs samplers: Theory and methods. *J. Am. Stat. Assoc.* **2008**, *103*, 790–796. [\[CrossRef\]](#)
36. Doucet, A.; Godsill, S.J.; Robert, C.P. Marginal maximum a posteriori estimation using Markov chain Monte Carlo. *Stat. Comput.* **2002**, *12*, 77–84. [\[CrossRef\]](#)
37. Robert, C. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007.
38. Goorley, J.T.; James, M.R.; Booth, T.E.; Brown, F.B.; Bull, J.S.; Cox, L.J.; Durkee, J.W., Jr.; Elson, J.S.; Fensin, M.L.; Forster, R.A., III; et al. *Initial MCNP6 Release Overview-MCNP6 Version 1.0*; Technical Report; Los Alamos National Lab. (LANL): Los Alamos, NM, USA, 2013.
39. Gelman, A.; Stern, H.S.; Carlin, J.B.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2013.
40. Pereyra, M.; Miele, L.V.; Zygalakis, K.C. Accelerating proximal Markov chain Monte Carlo by using an explicit stabilized method. *SIAM J. Imaging Sci.* **2020**, *13*, 905–935. [\[CrossRef\]](#)
41. Vidal, A.F.; De Bortoli, V.; Pereyra, M.; Durmus, A. Maximum Likelihood Estimation of Regularization Parameters in High-Dimensional Inverse Problems: An Empirical Bayesian Approach Part I: Methodology and Experiments. *SIAM J. Imaging Sci.* **2020**, *13*, 1945–1989. [\[CrossRef\]](#)
42. Papandreou, G.; Yuille, A.L. Gaussian sampling by local perturbations. *Adv. Neural Inf. Process. Syst.* **2010**, *23*, 1858–1866.