2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

A Model-Based Framework for Fast Dynamic Image Sampling

G.M. Dilshan Godaliyadda*, Gregery T. Buzzard†, and Charles A. Bouman*
*School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 47907
†Department of Mathematics, Purdue University, West Lafayette, IN, 47907

***Abstract*—In many applications, it is critical to be able to sample the most informative pixels of an image first; and then once these pixels are sampled, the highest fidelity image can be reconstructed. Optimized sampling strategies generally fall into two categories: static and dynamic. In dynamic sampling, each new sample is chosen by using information obtained from previous samples. In this way, dynamic sampling offers the potential of much greater fidelity, but at the cost of greater complexity. Existing methods for dynamic non-uniform sampling of images are based on the intuition that sampling rates should be greatest in locations of greatest variation, but recent developments in the theory of optimal experimental design offer a theoretical framework for optimal sampling based on the use of a formal Bayesian prior model.**

In this paper, we introduce a fast dynamic image sampling framework based on Bayesian experimental design (BED). The method, which we call model-based dynamic sampling (MBDS) allows for the use of a general prior distribution for the image, and it incorporates a pixel-wise sampling constraint in the BED framework. The MBDS works by first generating L stochastic samples (i.e., images) from the posterior distribution given the current measurements, and then selecting the pixel with the greatest posterior variance. We also introduce a computationally efficient method for computing the stochastic samples through a local updating technique.

I. Introduction

Many applications can benefit from image sampling strategies that can select a relatively small set of measurements to accurately reconstruct the image. For example, scanning electron microscopy (SEM) and computed tomography (CT) are applications in which it is advantageous to minimize the number of measurements [1].

Optimized sampling strategies fall into two categories: static and dynamic. Static sampling methods can be used to preselect the measurements to achieve the best image fidelity. These methods include random sampling strategies such as in [2], methods based on an *a priori* knowledge of the object geometry as in [3], and methods based on optimal experimental design (OED) [4].

Alternatively, dynamic sampling methods use all previous samples to determine each new measurement. Therefore, dynamic sampling offers the potential for greater fidelity of the reconstructed image, but at the cost of greater complexity. In [5], [6] Kovačević et al. proposed methods for dynamic sampling of image pixels designed to speed acquisition for fluorescence microscopy applications. This work was designed to track features of a time-varying image with the use of a particle filter. In [7], initially different sets of pixels are measured to estimate the image, and further measurements are made where the estimated signal is non-zero. Additionally, application specific dynamic sensing methods have been proposed in [8] for selecting optimal K-space spiral and line measurements for magnetic resonance imaging (MRI), and in [9] for selecting measurement angles for binary CT. Apart from these methods, dynamic compressive sensing (DCS) methods have been proposed in [10], [11] and [12]. However, DCS is based on the assumption that the measurement is formed by the projection of the signal in an unconstrained direction. This differs fundamentally from the constrained problem of sampling a single pixel at a time. Also, even though DCS methods are based on Bayesian statistics, the existing methods are limited in the selection of the prior distribution.

In this paper, we propose a general framework for model-based dynamic image sampling (MBDS) based on Bayesian experimental design (BED). Our algorithm allows the use of a broad class of posterior distributions so that an application specific model can be selected. It also allows for the incorporation of a general class of constraints in the measurement projection, which is essential in many applications. So for example, in conventional spatial sampling, each measurement must be enforced to be the projection of a single pixel; or in tomographic projection, each view must be enforced to be the integration of the image along projection lines. In practice, this constraint changes the BED problem substantially because with each new measurement, the eigenvector structure of the posterior distribution must be re-estimated.

In order to work with a general prior and projection constraints, our MBDS method is based on direct stochastic sampling of the posterior distribution. In particular, it works by maintaining L stochastic samples, or images, generated from the posterior distribution, and then uses this set of L images to compute an empirical covariance, from which the optimal sample is determined. In [13], a similar approach is proposed to design measurements for a biochemical network with relatively low dimension. However, for a high-dimensional image, direct Monte Carlo sampling of the posterior would require too much computation for most applications. So in order to make our approach computationally practical, we introduce a technique for locally updating the stochastic sample in the neighborhood of each new measurement. This technique dramatically reduces computation as compared to brute-force posterior sampling.

G.M. Dilshan Godaliyadda and C. A. Bouman were supported by AFOSR/MURI grant #FA9550-12-1-0458. Gregery T. Buzzard was partially supported by NSF grant #DMS-0900277

II. Bayesian Experimental Design (BED) Overview

The objective of BED is to obtain a relatively small set of measurements that allow for accurate reconstruction of an unknown signal x. Let $y^{(k)}$ denote the vector composed of the first k measurements, and let x denote the unknown signal. Then on the k^{th} measurement, the entire vector of past and present measurements is given by

$$y^{(k)} = A^{(k)}x + w^{(k)}, \tag{1}$$

where $A^{(k)}$ is the projection matrix, and $w^{(k)}$ is Gaussian measurement noise that is assumed to be both independent of x and to have independent components, with variance σ^2_{noise}. Each row of $A^{(k)}$ is assumed to be a vector m of unit length so that $\|m\| = 1$. This restriction to unit length vectors is assumed so that the signal-to-noise ratio of a single measurement is fixed.

Our objective is to then select each new measurement vector, $m^{(k)}$, to be in the direction of maximum variation of the posterior distribution. More specifically, if the posterior mean and covariance is denoted by

$$\mu^{(k)}_{x|y} \triangleq \mathbb{E}\left[x|y^{(k)}\right], \tag{2}$$

$$R^{(k)}_{x|y} \triangleq \mathbb{E}\left[\left(x-\mu^{(k)}_{x|y}\right)\left(x-\mu^{(k)}_{x|y}\right)^t \middle| y^{(k)}\right], \tag{3}$$

then the measurement projection in the direction of maximum variation, $m^{(k)}$, is given by

$$m^{(k)} = \arg\max_{m\in\mathcal{D}} \left(m^t R^{(k)}_{x|y} m\right), \tag{4}$$

where $\mathcal{D} = \{m \in R^N : ||m||_2 = 1\}$ constrains each measurement vector to be of unit length. The solution to equation (4) is the normalized principal eigenvector of $R^{(k)}_{x|y}$. Once $m^{(k)}$ is found it is appended to $A^{(k)}$ to form $A^{(k+1)}$:

$$A^{(k+1)} = \begin{pmatrix} A^{(k)} \\ m^{(k)t} \end{pmatrix}. \tag{5}$$

In the next iteration x is measured using the measurement projection $m^{(k)}$ to form $y^{(k+1)}$.

We will primarily be interested in the case when $\mathcal{D}$ incorporates additional constraints. We define the set of measurements that incorporate such constraints as $\mathcal{M} \subset \mathcal{D}$.

III. Unconstrained Dynamic Sampling with a Gaussian Prior

From equation (4), it is clear that selecting a model for the posterior distribution is critical. If we assume that x is a zero mean Gaussian random vector with covariance matrix B^{-1}, then we know that its distribution must have the form

$$p_k(x) = \frac{|B|^{\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}x^t B x\right\}, \tag{6}$$

and therefore that the posterior distribution must have the form

$$p_k(x|y^{(k)}) = \frac{1}{z}\exp\left\{-\frac{1}{2}\|y^{(k)} - A^{(k)}x\|^2_{\Lambda^{(k)}} - \frac{1}{2}x^t B x\right\}, \tag{7}$$

where z is a normalizing constant, and $\Lambda^{(k)}$ is the noise covariance matrix.

Then $R^{(k)}_{x|y} = \left[(A^{(k)})^t \Lambda^{(k)} (A^{(k)}) + B\right]^{-1}$. Notice that in this case, the posterior covariance $R^{(k)}_{x|y}$ is not a function of the data $y^{(k)}$, and therefore the recursion in equations (3), (4), and (5) does not depend on the measurements. Consequently, when the prior is Gaussian, the measurement projections can be computed in advance. It should also be mentioned that in this case, each new measurement is D-optimal, and therefore results in a D-optimal sequential experimental design [4].

For the case when the measurements are unconstrained, $m^k \in \mathcal{D}$, the eigen-structure of the covariance does not change after each measurement selection. So then it can be shown that the K best measurements are the K principal eigen-vectors of the covariance matrix, $R_{x|y}$ [11].

However, we are interested in the case when the measurements are constrained, $m^{(k)} \in \mathcal{M}$, where $\mathcal{M} \subset \mathcal{D}$, and the prior is non-Gaussian. For this case, the covariance matrix must be re-estimated after each iteration and equation (4) becomes

$$m^{(k)} = \arg\max_{m\in\mathcal{M}} \left(m^t R^{(k)}_{x|y} m\right). \tag{8}$$

Furthermore, we would like a framework that can incorporate any posterior distribution, so that an application specific prior distribution can be used.

IV. Model-Based Dynamic Sampling (MBDS)

The MBDS method is designed to work with a wide range of priors and sampling constraints by directly generating stochastic samples from the posterior distribution. Figure 1 specifies the MBDS method in pseudo-code. For each new sample, L images are generated from the posterior distribution using Monte Carlo (MC) methods, and then these L images are used to compute an estimated covariance for the posterior distribution.

The estimated sample covariance is given by

$$\hat{R}^{(k)}_{x|y} = \frac{1}{L-1}\sum_{i=1}^{L}\left(x^{(k,i)} - \hat{\mu}\right)\left(x^{(k,i)} - \hat{\mu}\right)^t, \tag{9}$$

where $x^{(k,i)}$ is the i^{th} image out of L that are generated before the k^{th} sample is taken. With this covariance, the measurement vector is then selected with the constraint that $m \in \mathcal{M}$, where $\mathcal{M} \subset \mathcal{D}$. In our examples, we constrain each measurement to be of a single pixel; however, other choices are possible. Then, $\mathcal{M} = \left\{e_i \in \mathbb{R}^N : e_i(i) = 1; e_i(j) = 0\ \forall j \neq i\right\}$, and the new measurement will be the pixel location with the largest posterior variance.

Generating sample vectors from the posterior distribution $p_k(x|y^{(k)})$ can be computationally expensive, particularly when x is a high-dimensional image. To counter this problem, we introduce a strategy of localized stochastic sample updates in which we only update a block surrounding the measured pixel.

Instead of performing computationally expensive (MC) sampling for the entire image $x \in \mathbb{R}^N$, we only perform it for a window $w_s \in \mathbb{R}^b$ from x, where $b << N$.

function $\mathcal{M}_* \leftarrow$ MBDS

Generate samples from prior distribution $p(x)$ $\left\{x^{(0,1)}, x^{(0,2)}, \dots x^{(0,L)}\right\}$

%**Outputs**: $\mathcal{M}_*$ - Selected set of measurements

for $(k = 1, k \leq K, k++)$ **do**

Estimate $\hat{R}^{(k)}_{x|y}$ using equation (9)

$m^{(k)} = \arg\max_{m \in \mathcal{M}} \left(m^t R^{(k)}_{x|y} m\right)$

$\tilde{y}^{(k)} = m^{(k)} x + w^{(k)}$

$y^{(k)} = \begin{pmatrix} y^{(k-1)} \\ \tilde{y}^{(k)} \end{pmatrix}$

Generate L samples from $p_k(x|y^{(k)})$ $\left\{x^{(k,1)}, x^{(k,2)}, \dots x^{(k,L)}\right\}$

end for

end function

Fig. 1. Pseudo-code for MBDS. K is the number of total measurements to be taken; L is the number of sample vectors generated from the posterior; $\mathcal{M}$ is a constrained subset of all possible measurements; $x^{(k,j)}$ refers to the j^{th} sample vector drawn from $p_k(x|y^{(k)})$. Note that $y^{(0)}$ refers to the case when no measurements have been made.

Here, w_s includes the measured pixel location and a block surrounding it. Therefore, we maintain L stochastic samples from the posterior distribution and update them locally once a measurement is made. The block-posterior distribution is then, $p_k(w_s|y^{(k)}, w_{\sim s})$, where $w_{\sim s}$ are the pixel locations outside of the window w_s.

Consider that the samples from the previous iteration are given by $\left\{x^{(k-1,1)}, x^{(k-1,2)}, \dots x^{(k-1,L)}\right\}$. Then we stochastically sample for the block surrounding the measured pixel to generate sample vectors, $\left\{w_s^{(k,1)}, w_s^{(k,2)}, \dots w_s^{(k,L)}\right\}$, from the block-posterior. Next, we use these stochastically generated samples to replace these corresponding windows of pixels in the L images $\left\{x^{(k-1,1)}, x^{(k-1,2)}, \dots x^{(k-1,L)}\right\}$, and then this forms the new set of sample images $\left\{x^{(k,1)}, x^{(k,2)}, \dots x^{(k,L)}\right\}$. This procedure is illustrated in Figure 2.

A. Generating Samples from a Block-Posterior Distribution

Given that the block posterior distribution has the form of a Gibbs distribution, well known methods such as the Metropolis algorithm [14] or the Metropolis-Hastings (MH) algorithm [15], [16] can be used to draw samples from it. In our implementation we use the MH algorithm, where a multivariate Gaussian distribution is used as the proposal distribution.

The proposal distribution we use is a second order Taylor series approximation to $\log p_k(w_s|y^{(k)}, w_{\sim s})$. In particular, a Gaussian proposal distribution, $q_k(w_s|y^{(k)}, w_{\sim s})$, is selected so that its mean and covariance can be fit using a Taylor series expansion of the log posterior distribution.

V. Experiments Conducted

In this section, we compare results from MBDS with two sampling strategies - uniformly spaced sampling (US) and random sampling (RS). We begin by presenting details of the posterior distribution and reconstruction algorithm that we use.

A. Posterior Distribution and Image Reconstruction

We model the distribution of the unknown x using a q-GGMRF [17] since it has been used for accurate image

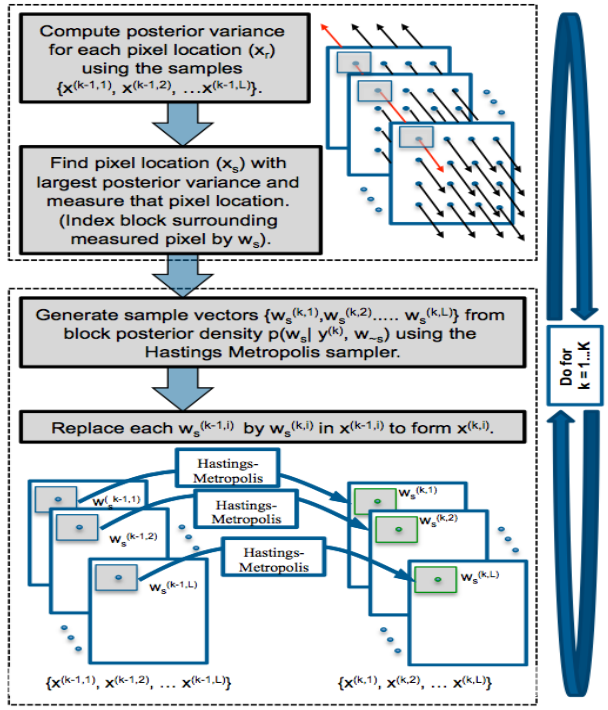


Fig. 2. The MBDS algorithm with localized posterior sample updates, to obtain K measurements. Here L sample images are kept in memory and updated locally after each new measurement is made.

reconstruction in medical [18] and materials imaging [19]. The q-GGMRF has the form

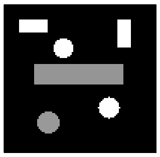
$$p_k(x) = \frac{1}{z} \exp\left\{ - \sum_{\{i,j\} \in \mathcal{P}} \frac{1}{2} \left(\frac{\left|\frac{x_i - x_j}{\sigma_x}\right|^q}{c + \left|\frac{x_i - x_j}{\sigma_x}\right|^{q-p}} \right) \right\}. \quad (10)$$

Here, p, q, c_i and σ_x are parameters of the distribution, $\mathcal{P}$ is the set of all unique pairs defined according to the neighborhood, and z is the normalizing partition function of the distribution. The resulting posterior is then

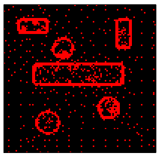
$$p_k(x|y^{(k)}) = \frac{1}{z} \exp\left\{ -\frac{1}{2} \|y^{(k)} - A^{(k)} x\|^2_{\Lambda^{(k)}} - \sum_{\{i,j\} \in \mathcal{P}} \frac{1}{2} \left(\frac{\left|\frac{x_i - x_j}{\sigma_x}\right|^q}{c + \left|\frac{x_i - x_j}{\sigma_x}\right|^{q-p}} \right) \right\}. \quad (11)$$

We define the neighborhood as the 8 pixels surrounding the pixel considered.

For image reconstruction, any method that can reconstruct the image from a sparse set of measurements can be used. For our experiments we use maximum *a posteriori* (MAP) estimation. Since we use the distribution in equation (11) as our posterior, the resulting cost function is non-quadratic, and a closed form solution for the maximum of this function cannot be analytically calculated. Therefore, we convert this problem into an iterative quadratic optimization problem by using Majorization techniques [20]–[22]. In conjunction with Majorization, we use the Iterative Coordinate Descent (ICD) optimization method [23], [24] to solve the optimization problem.

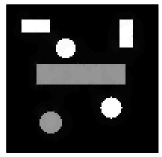


(a) Original image

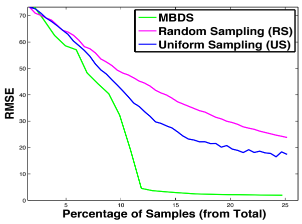


(b) MBDS - measured pixels (13%)

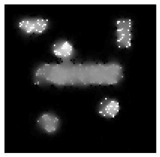
(c) MBDS - measured image (13%)



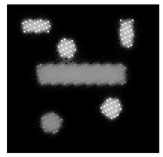
(d) MBDS - reconstruction (13%)



(e) RMSE comparison for sampling the image Geometric shapes

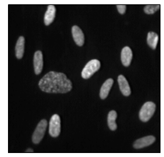


(f) RS - reconstruction (13%)

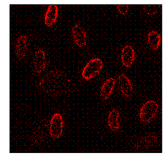


(g) US - reconstruction (13%)

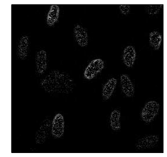
Fig. 3. Dynamic sampling simulation on Geometric Shapes (100 × 100). (b): First 13% of measured pixel locations (red) selected using MBDS. (c): Measured image (13%). (d): Image reconstructed using (c). (f) and (g): Reconstructed images when sampling locations (13%) are chosen using RS and US.



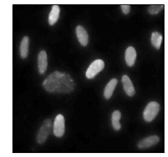
(a) Original image



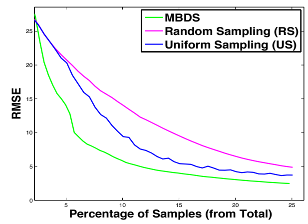
(b) MBDS - measured pixels (8%)



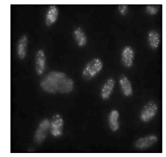
(c) MBDS - measured image (8%)



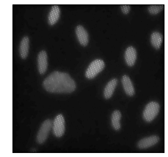
(d) MBDS - reconstruction (8%)



(e) RMSE comparison for sampling the image Geometric shapes



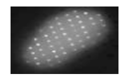
(f) RS - reconstruction (8%)



(g) US - reconstruction (8%)

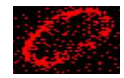
(h) Original

(i) RS - reconstruction



(j) US - reconstruction

(k) MBDS - reconstruction



(l) MBDS - samples

Fig. 4. Dynamic sampling simulation on Fluocel.pgm (256 × 256). (b): First 8% of measured pixel locations (red) selected using MBDS. (c): Measured image (8%). (d): Image reconstructed using (c). (f) and (g): Reconstructed images when sampling locations (8%) are chosen using RS and US. (h): patch extracted from (a); (i):patch extracted from (d); (j): patch extracted from (f); (k): patch extracted from (g); (l): patch extracted from (b).

Since we assume that a measurement only affects a block of pixels surrounding the measured pixel, we only perform MAP estimation for the window w_s. Therefore, after each measurement is made, we only estimate $\hat{w}_s^{(k)}$, the reconstruction for block w_s. We then insert $\hat{w}_s^{(k)}$ into $\hat{x}^{(k-1)}$, the reconstruction of the whole image before the k^{th} measurement is made, to form $\hat{x}^{(k)}$.

B. Experimental Setup and Evaluation of Results

The measurement noise for each pixel was simulated to be independent and Gaussian with a variance of $\sigma^2_{noise} = 9$, and the pixel values are between $0 - 255$. The resolutions of the two images used were 100×100 and 256×256. The block-size we used for localized stochastic sampling is 16×16. The parameters we used for the prior distribution were, $p = 1.2$, $q = 2$, $\sigma_x = 6$ and $c = 1$. For both cases we used $L = 20$ samples from the posterior distribution to estimate the sample variance. In MBDS the first 1.5% of measurement locations are uniformly spaced apart. Then, each new measurement location is selected according to the MBDS algorithm. In both these experiments, when using MBDS we select a new measurement in approximately 0.6 seconds.

In the first experiment, the image shown in Figure 3(a) was measured using RS, US and MBDS. This image, Geometric Shapes (GS), was a simulated image we created. Figure 3(b) shows the first 13% of measurement locations selected by MBDS and Figure 3(c) shows the corresponding measured image. The reconstructed image is shown in Figure 3(d). Figures 3(f) and 3(g) show the reconstructed images for random sampling (RS) and uniformly spaced sampling (US) when the same percentage (13%) of measurements are acquired. From Figure 3(d) we observe that the edges are better preserved when MBDS was used for measurement selection. Furthermore, from Figure 3(e) where the root mean squared error (RMSE) versus the percentage of measurements is plotted, we observe that MBDS outperforms US and RS quantitatively as well. From Figures 3(b) and 3(c) we observe that our algorithm concentrates measurements on the most informative pixels, the feature edges, while sparsely measuring other regions of the image.

For the second experiment we used a real image (Figure 4(a)) provided by the University of Granada (http://decsai.ugr.es/cvg/dbimagenes/). Figure 4(b) shows the first 8% of measurement locations selected by MBDS, Figure 4(c) the corresponding measured image and Figure 4(d) the reconstructed image. Figures 4(f) and 4(g) show the reconstructed images for RS and US respectively. Figures 4(i), 4(j) and 4(k) are patches extracted from the reconstructions, corresponding to the patch shown in Figure 4(h). Here we observe that by using MBDS for measurement selection, the edges of the feature as well as the details within the feature are preserved in the reconstructed patch. Figure 4(l) further illustrates this by showing the measurement locations selected by MBDS.

VI. Conclusion

In this paper, we presented a general framework for constrained dynamic sampling, which can incorporate a broad class of posterior models. The method is based on stochastic sampling of the posterior distribution using a computationally efficient algorithm; experimental results show that it can substantially improve reconstruction quality given a fixed number of measurements.

References

[1] R. Smith-Bindman, J. Lipson, and R. Marcus, "Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer," *Archives of Internal Medicine*, vol. 169, no. 22, pp. 2078–2086, 2009.

[2] H. S. Anderson, J. Ilic-Helms, B. Rohrer, J. Wheeler, and K. Larson, "Sparse imaging for fast electron microscopy," *IS&T/SPIE Electronic Imaging*, pp. 86 570C–86 570C, 2013.

[3] K. Mueller, "Selection of optimal views for computed tomography reconstruction," Patent WO 2,011,011,684, Jan. 28, 2011.

[4] A. C. Atkinson, A. N. Donev, and R. D. Tobias, *Optimum experimental designs, with SAS*. Oxford University Press, Oxford, 2007, vol. 34.

[5] T. E. Merryman and J. Kovacevic, "An adaptive multirate algorithm for acquisition of fluorescence microscopy data sets," *IEEE Trans. on Image Processing*, vol. 14, no. 9, pp. 1246–1253, 2005.

[6] C. Jackson, R. F. Murphy, and J. Kovacevic, "Intelligent acquisition and learning of fluorescence microscope data models," *IEEE Trans. on Image Processing*, vol. 18, no. 9, pp. 2071–2084, 2009.

[7] J. Haupt, R. Baraniuk, R. Castro, and R. Nowak, "Sequentially designed compressed sensing," in *Statistical Signal Processing Workshop (SSP)*. IEEE, 2012, pp. 401–404.

[8] M. Seeger, H. Nickisch, R. Pohmann, and B. Schölkopf, "Optimization of k-space trajectories for compressed sensing by bayesian experimental design," *Magnetic Resonance in Medicine*, vol. 63, no. 1.

[9] K. Joost Batenburg, W. J. Palenstijn, P. Balázs, and J. Sijbers, "Dynamic angle selection in binary tomography," *Computer Vision and Image Understanding*, 2012.

[10] M. W. Seeger and H. Nickisch, "Compressed sensing and bayesian experimental design," in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 912–919.

[11] W. R. Carson, M. Chen, M. R. Rodrigues, R. Calderbank, and L. Carin, "Communications-inspired projection design with application to compressive sensing," *SIAM Journal on Imaging Sciences*, vol. 5, no. 4, pp. 1185–1212, 2012.

[12] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, 2008.

[13] J. Vanlier, C. Tiemann, P. Hilbers, and N. van Riel, "A bayesian approach to targeted experiment design," *Bioinformatics*, vol. 28, no. 8, pp. 1136–1142, 2012.

[14] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The Journal of Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[15] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. pp. 97–109, 1970.

[16] P. H. Peskun, "Optimum monte-carlo sampling using markov chains," *Biometrika*, vol. 60, no. 3, pp. pp. 607–612, 1973.

[17] J.-B. Thibault, K. Sauer, C. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Med. Phys.*, vol. 34, pp. 4526–4544, 2007.

[18] J. Wang, K. Sauer, J.-B. Thibault, Z. Yu, and C. Bouman, "Spectrally focused markov random field image modeling in 3d ct," in *Proc. Intl Conf. on Fully 3D Reconstruction in Radiology and Nuclear Medicine*, 2011.

[19] S. Venkatakrishnan, L. Drummy, M. Jackson, M. De Graef, J. Simmons, and C. Bouman, "A model based iterative reconstruction algorithm for high angle annular dark field-scanning transmission electron microscope (haadf-stem) tomography," *Image Processing, IEEE Transactions on*, vol. 22, no. 11, pp. 4532–4544, Nov 2013.

[20] J. Ortega and W. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, ser. Classics in Applied Mathematics.

[21] A. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. on Medical Imaging*, vol. 14, no. 1, pp. 132–137, 1995.

[22] C.-S. Foo, C. B. Do, and A. Y. Ng, "A majorization-minimization algorithm for (multiple) hyperparameter learning," in *Proceedings of the 26th Annual International Conference on Machine Learning*, ser. ICML '09, 2009, pp. 321–328.

[23] K. Sauer and C. Bouman, "Bayesian Estimation of Transmission Tomograms Using Segmentation Based Optimization," *IEEE Trans. on Nuclear Science*, vol. 39, pp. 1144–1152, 1992.

[24] Z. Yu, J. Thibault, C. Bouman, K. Sauer, and J. Hsieh, "Fast model-based X-ray CT reconstruction using spatially nonhomogeneous ICD optimization," *IEEE Trans. on Image Processing*, vol. 20, no. 1, pp. 161 –175, Jan. 2011.