

Supervised Learning in image enhance from electron microscopy inversion results

Georgios Stagakis (✉ georgstag@gmail.com)

Method Article

Keywords: Bayesian Learning, Supervised Learning, Gaussian Stochastic Processes, Material Science, Discontinuous data, Markov Chain Monte Carlo methods, Metropolis-Hastings algorithm

Posted Date: January 10th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1236935/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Supervised Learning in image enhance from electron microscopy inversion results

Georgios Stagakis

Department of Mathematical Sciences

Loughborough University

Loughborough, England

g.stagakis@lboro.ac.uk

Abstract—In Nondestructive testing there is a variety of applications in Material Science, where the specimen is imaged by an Electron Microscope and then by image inversion, information is extracted for the material interior. This type of information might contain noise either by the imaging procedure or by the numerical part of the inversion. We present a method that can improve the interior density results of an inversed material from a series of Scanning Electron Microscope (SEM) images. For this method, the material density can contain some discontinuity, such as regions where it is dense and regions where there are voids. The proposed method directly stands on the Bayesian learning framework, adopting Gaussian Stochastic Processes (GSPs). Two test sample cases that contain some discontinuities in the density are tested. We also provide a comparison between two different GSP modelling approaches; one is a typical GSP and the other accounts for discontinuity, by introducing hyperparameters. The GSP method gives reconstructed data in reasonable agreement with the known original density distribution, giving confidence that the method can be applied to experimentally obtained SEM images.

Index Terms—Bayesian Learning, Supervised Learning, Gaussian Stochastic Processes, Material Science, Discontinuous data, Markov Chain Monte Carlo methods, Metropolis-Hastings algorithm.

I. INTRODUCTION

Scanning Electron Microscopy (SEM) is a standard tool to study the nanoscale surface structure of materials. The SEM image results from the interactions of an electron beam with the atoms at various depths within a study sample. Once the electron beam penetrates the sample surface, it produces signals from its interactions with the atoms beneath the incident point, a three-dimensional region, namely the interaction volume. An SEM image produced from the detected signals is thus a projection of the three-dimensional (3D) interaction volume onto a two-dimensional (2D) image. If the projection process can be inverted, taking the 2D image back into a 3D space, the inversion process is called image inversion [3], [8], [17], [19] and has the potential to reconstruct the interior material density of the sample. Image inversion techniques are of great use for nondestructive inspection, a crucial task in materials science; for example, device engineers need such nanostructure information for electronic devices [9].

Recent research [6] proposed such an image inversion technique that produces a discretised form of interior material

densities from SEM images, adopting a Bayesian approach. The technique dissects a sample material into voxels, each of which corresponds to a pixel of an SEM image, and then obtains a density value for each voxel. Although the voxel-based discretisation offers a practical approach, it faces some shortcomings, such as resolution issues. For example, if a voxel happens to contain both regions where it is dense and other regions where there are voids, the material density of the voxel as a whole would then be averaged out. In other words, the calculation of voxel-based material density largely depends on the extent to which these voxels are placed. Since material densities are, in general, discontinuous with voids that randomly appear in the interior of the material, constructing an unknown material structure in a non-discretised manner becomes a crucial task. However, up to our knowledge, such an approach has been little exploited in practice.

In this paper, we provide for the first time the methodological framework to construct non-discretised unknown material structures based on the information conveyed from the image inversion. Our approach is based on the image inversion technique [6] that calculates from SEM images the material density value for the centre of each voxel. A key idea here is to utilise Gaussian Stochastic Processes (GSPs) to predict the density value at an arbitrary location within the material sample, given those density values from the image inversion technique. In addition, we take advantage of the inversion variance in GSP modelling, in order for the GSP to modify the density in cases of inversion noise.

To test our framework, first, we model a material with a known discontinuous density function and simulate the corresponding SEM images—the projection of the density. Second, we then invert the SEM images to a 3D space—the image inversion—and reproduce the material density function from it, using GSPs. We demonstrate how our methodology works through two examples; one is a simple case with one void region therein, and the other is a highly discontinuous material which is an extreme case. We compare the results with the true situations to evaluate the efficiency of the proposed method.

The paper is laid out as follows. In the next section, we describe how we simulate two example materials and generate their SEM image data. This projection process is not necessary for real-world situations, as we usually have SEM images

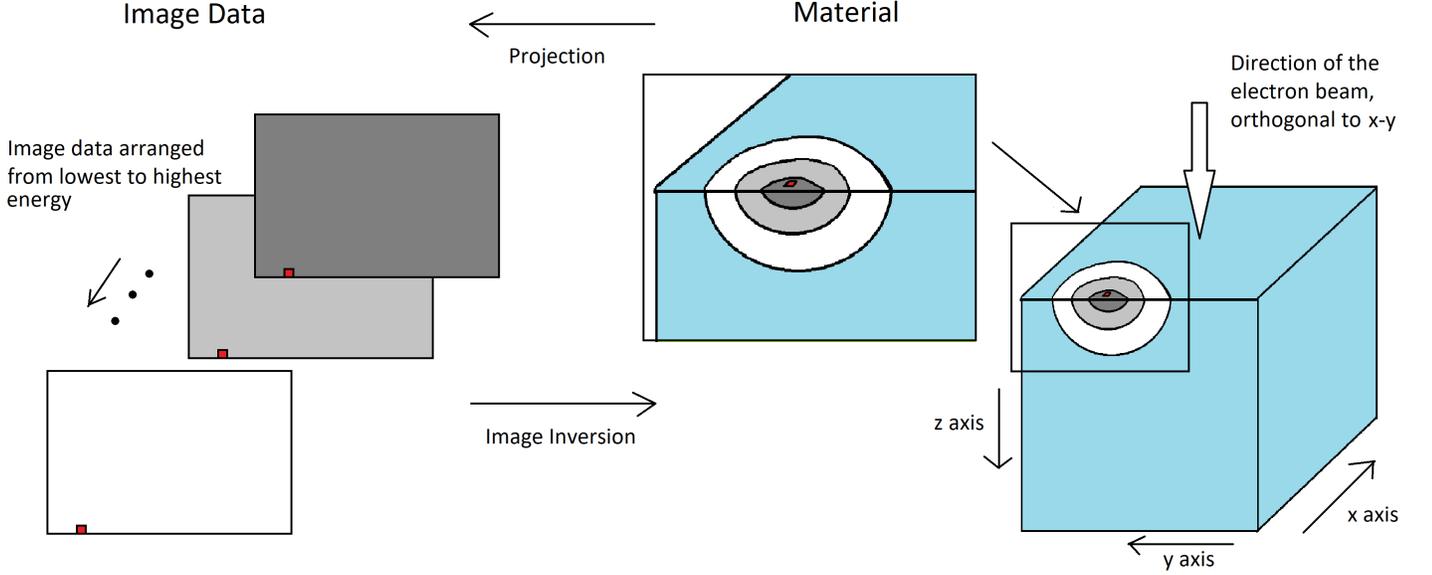


Fig. 1: Schematic figure that shows the projection and image inversion processes. The red square represents a corresponding pixel in the image data. The coordinates of the pixel can be seen with red colour in the material. The grey colour scaled areas are projected onto each corresponding SEM image.

first, but is necessary for method evaluation purposes. The third section delineates how our GSP modelling approach accommodates the image inversion method for predicting the interior material density of arbitrary locations. A particular focus is given upon two different parameterisation approaches in GSP modelling. In the fourth section, we evaluate the performance of our methodology, comparing the predicted material density value to the true one, followed by discussion and conclusion.

II. MATERIALS

As illustrative examples, we here present two simulated material samples, whose true interior density is thus known, and create their SEM images. In practice, this projection process is unnecessary as we usually have the SEM images of an actual material sample first. However, to evaluate the performance of the methodology, we numerically mimic the projection process and create SEM images, adopting a convolution technique with a microscopy correction function [10], [16].

Here, the interior of a material sample is specified in Cartesian coordinates $s = (x, y, z) \in \mathbb{R}^3$. The x - y surface is orthogonal to the incoming electron beam, and the z -axis represents the depth in the interior of the sample (Fig. 1). Note that we treat z as an absolute depth from the topmost surface throughout the paper. The material density is then specified as a function of location s as $d(s)$.

In practice, an SEM image consists of pixels, the positions of which correspond to the coordinates of the voxels dissecting the sample material. The voxel size on the x - y surface is thus the same as the pixels of the image. The discretised material

density of each voxel, $i = 1, 2, \dots, n_p$, in each depth, $z^{(m)}$, is then specified as

$$d_i^{(m)} := d(x_i, y_i, z^{(m)}). \quad (1)$$

Note that the depth is defined as the length at which the m -lowest energy beam reaches ($z^{(0)} := 0$ and $m = 1, 2, \dots, n_d$). In other words, the voxel length along the z -axis should equal to the difference in depth reached by two sequential energy electron beams. For example, the first x - y layer of voxels should expand from $z = 0$ (the topmost surface of the material sample) and reach up to the length the electron beam reached in the SEM image produced with the lowest energy; the second layer should expand from z equal to the depth the electron beam with the lowest energy reached, up to the depth the electron beam with the second-lowest energy reached.

The actual depth in z -axis can be approximated by the following formula [13],

$$z = \frac{0.0276A}{dZ^{0.89}} E^{1.67}, \quad (2)$$

where d is the material mass density, Z is the atomic number of the material, E is the energy of the electron beam, and A is the atomic weight. The electron beam with higher energy penetrates deeper and cover a wider hemispherical region inside the material (Fig. 1). In other words, the image created by a higher energy beam (kV) covers a larger region that encompasses ones created by the lower energy beams.

An SEM image can be created as the projection of the convolution of density d with kernel k [6], namely the mi-

croscopy correction function [10], [16]. A pixel $(x_i, y_i, z^{(m)})$ of an image is then given as

$$C[d * k]_i^{(m)} = \frac{1}{(z^{(m)})^2} \sum_{q=1}^m \frac{(z^{(q)})^2 - (z^{(q-1)})^2}{2} \times \sum_{t=1}^q (z^{(t)} - z^{(t-1)}) \sum_{\ell=1}^t d(x_i, y_i, z^{(\ell)}) k(z^{(m-\ell)}), \quad (3)$$

where the microscopy function is given in the form of folded normal distribution [15]:

$$k(z) = Q \left[\exp \left\{ -\frac{(z-c)^2}{2v^2} \right\} + \exp \left\{ -\frac{(z+c)^2}{2v^2} \right\} \right], \quad (4)$$

where Q, c and v are parameters that specify the microscopy function. [6] suggests three different numerical methods for the projection. The choice of these methods largely depends on the resolution of the imaging technique. In this paper, the simplest option is chosen.

An SEM image $\mathbf{I} = \{I_i^{(m)}\}$ results from the projection $C[\cdot]$, with some noise in practice [23] so that each pixel becomes,

$$I_i^{(m)} = C[d * k]_i^{(m)} + e_i^{(m)}. \quad (5)$$

In the following examples, we assume Gaussian noise whose standard deviation equals to $pI_i^{(m)}$, where p is a percentage associated with the acquisition time between images [23]. For our applications, p is 5%, which is a reasonable choice—upper bound—when there is at least 50 seconds gap between the acquisition of sequential images.

A. Sample 1: simple density function

We consider a size of material slab:

$$\Omega = \{s : [-5, 5] \times [-5, 5] \times [0, 945]\}$$

that contains a void so that its material density is

$$d(s) = \begin{cases} 0, & s \in V \subset \Omega \\ 1, & \text{otherwise,} \end{cases} \quad (6)$$

where

$$V = \left\{ s : \left(\frac{x}{2}\right)^2 + \left(\frac{y}{2}\right)^2 + \left(\frac{z-0.5}{0.2}\right)^2 \leq 1 \right\}, \quad (7)$$

is an ellipsoid within which the material density is zero (void).

We assume that the sample slab consists of $10 \times 10 \times 10$ voxels, each of which is $1 \times 1 \times (0.1 - 0.001m)$, where m is the enumerated location of each x - y layer of voxels from top to bottom, keeping the situation close to a real application. We then produce an SEM image by projection (3) with the microscopy correction function k (4) whose parameters are $Q = 0.2$, $c = 0.5$ and $v = 0.1$. The choice of k here resembles an SEM correction function often used in practice.

We will evaluate the performance of our methodology by learning the density at 6×6 different coordinates,

$$x, y \in \{-4, -2.4, -0.8, 0.8, 2.4, 4\},$$

and $z = 0.4$.

B. Sample 2: highly discontinuous density function

We define a material:

$$\Omega = \{s : [-10, 10] \times [-10, -3.34] \times [0, 0.945]\},$$

the extent of which represents a size often used in actual applications. The material density function, d , here is a positive discontinuous function studied by [6],

$$d(s) = R \frac{A}{20 + \frac{x^2}{B^2} + \frac{y^2}{B^2} + \frac{z^2}{B^2(1-D^2)}}, \quad (8)$$

where the values R, A, D and B are location dependent random components sampled from a uniform distribution, $\text{Uni}[\cdot, \cdot]$; they are specifically,

- $R(x, y, z) = \lfloor 18U \rfloor$ where $U \sim \text{Uni}[0, 1]$;
- $A(x, y) \sim \text{Uni}[0, 1]$;
- $D(x, y) \sim \text{Uni}[0, 1]$; and
- $B(x, y) \sim \text{Uni}[0, 2.66]$.

The multiplier 18, inside the floor function of R , is associated with the constant 20 in the denominator of the fraction (8). The combination of these two values assures the density value lower than 1 but not too close to 0, while U is not close to 0. Note that random variables $A(x, y)$, $B(x, y)$ and $D(x, y)$ only depend on location (x, y) , e.g. $A(x, y, z) = A(x, y, z')$ for $z \neq z'$. Regarding the range of B , $2.66 = 2 \times 1.33$ is associated with the distance over which the fixed values change.

The density function (8) is continuous over specific intervals but also makes some jumps between voxels. Such discontinuity describes voids and happens mainly due to the randomness in R , making the density of some locations very close to 0, independent of the coordinates. The formulation chosen here can generate a different discontinuous density each time, which fact allows in principle to simulate a wide range of different materials.

We assume that the material consists of $15 \times 5 \times 18$ voxels, each of which is $1.33 \times 1.33 \times (0.1 - 0.001m)$, where m is the order number of each x - y layer from the top to bottom surfaces, likewise the sampled values in (8). In total, we sampled the random values of the function (8), for $15 \times 5 \times 18$ different locations. An SEM image is then obtained by projection (3) with the microscopy correction function k (4) whose parameters are $Q = 0.2$, $c = 0.4$ and $v = 0.15$.

We will evaluate the performance of our methodology by learning the density of the x - z slice at $y = -7.2$ as an illustrative example.

III. METHODS

Once we obtain the material density estimate at the centre of each voxel from SEM images, using the image inversion technique, we can estimate the interior material density at arbitrary locations though GPS modelling.

A. Image inversion

Given SEM images $\mathbf{I} = \{I_i^{(m)}\}$ with some noises (5), we can estimate the discretised material density for the voxels, $\mathbf{d} = \{d_i^{(m)}\}$ via the Bayesian framework,

$$P(\mathbf{d}|\mathbf{I}) \propto P(\mathbf{I}|\mathbf{d})\pi(\mathbf{d}), \quad (9)$$

where, the likelihood is specified as

$$P(\mathbf{I}|\mathbf{d}) = \prod_{i,m} \frac{1}{\sqrt{2\pi}\sigma_i^{(m)}} \exp \left\{ -\frac{(I_i^{(m)} - C[d * k]_i^{(m)})^2}{2(\sigma_i^{(m)})^2} \right\}, \quad (10)$$

with known microscopy correction function k (3). The prior distribution $\pi(\mathbf{d})$ is assumed to be an independent half-normal distribution to make sure the material density to be positive. Although uniform priors could be an option here, it may cause a misspecification problem for cases in which substantial amounts of voxels (1,000 or more) are involved in the numerical computation.

The parameters for the half-normal independent priors were assigned after experimenting under uniform priors. For all the variables, the priors' parameters were the same. Since the mean was given as a highly likely value from the sample of every variable and as variance a value for which the 99% interval was higher than the maximum interval noticed under uniform priors.

Applying Metropolis-Hastings algorithm to (9), we can numerically yield the posterior distribution $P(\mathbf{d}|\mathbf{I})$, from which we can calculate the expected discretised material density for the centre of each voxel, $\tilde{d}(x_i, y_i, z^{(m)})$. By marginalising the posterior distribution, we have a posterior mean for each voxel,

$$\tilde{d}_i^{(m)} = \int d_i^{(m)} P(d_i^{(m)}|\mathbf{I}) P(\mathbf{d}_{-i}|d_i^{(m)}, \mathbf{I}) \partial \mathbf{d}, \quad (11)$$

where \mathbf{d}_{-i} is a vector omitting $d_i^{(m)}$ element. The posterior second moment, $(\sigma_i^{(m)})^2$, can also be calculated from the posterior distribution. Eq. 11 is the output obtained from the image inversion technique and will be used in the following subsection to estimate the material density for in-between locations.

B. Gaussian stochastic process (GSP) modelling

Once the expected discretised material density values, $\tilde{\mathbf{d}} = \{\tilde{d}_i^{(m)}\}$, are obtained via the image inversion, we can estimate the interior material density for arbitrary in-between locations, say $\mathbf{d}^* = \{d_i^{(m)*}\}$, using GSPs. Without loss of generality, we hereafter regard that the material density is standardised by the average and standard deviation of $\tilde{\mathbf{d}}$.

With locations $\{s_i^{(m)}\}_{i,m}$ in a 3D space and length scale parameters (lsp), $\mathbf{l} = (l_1, l_2, l_3)$, the GSP modelling framework assumes that $\tilde{\mathbf{d}}$ and \mathbf{d}^* respectively follow a multivariate normal distribution,

$$P(\tilde{\mathbf{d}}|\mathbf{l}) = N(\mathbf{0}_{\tilde{n}}, \mathbf{C}(\mathbf{l})), \quad (12)$$

$$P(\mathbf{d}^*|\mathbf{l}) = N(\mathbf{0}_{n^*}, \mathbf{C}_*(\mathbf{l})), \quad (13)$$

with the dimension $\tilde{n} = \dim(\tilde{\mathbf{d}})$ and $n^* = \dim(\mathbf{d}^*)$. Their covariance matrices $\mathbf{C}(\mathbf{l})$ and $\mathbf{C}_*(\mathbf{l})$ are specified by the locations and lsp, *i.e.*

$$\{\mathbf{C}(\mathbf{l})\}_{i,m,j,m'} = \exp \left\{ -(s_i^{(m)} - s_j^{(m')})^\top \mathbf{M} (s_i^{(m)} - s_j^{(m')}) \right\} + (\sigma_i^{(m)})^2 I_{\{i=j\}} I_{\{m=m'\}}, \quad (14)$$

where,

$$\mathbf{M} = \begin{bmatrix} \frac{1}{l_1^2} & 0 & 0 \\ 0 & \frac{1}{l_2^2} & 0 \\ 0 & 0 & \frac{1}{l_3^2} \end{bmatrix}. \quad (15)$$

Note that the specification above is exactly the same for covariance matrix $\mathbf{C}_*(\mathbf{l})$.

For making an inference, it is convenient to deal with (12) and (13) simultaneously. Consider a combined vector of material density values

$$\mathbf{d}' := (\tilde{\mathbf{d}}, \mathbf{d}^*). \quad (16)$$

It is then also follows a multivariate normal distribution whose dimension is $n = \tilde{n} + n^*$,

$$P(\mathbf{d}'|\mathbf{l}) = N(\mathbf{0}_n, \mathbf{C}'(\mathbf{l})), \quad (17)$$

with the covariance matrix in the block matrix form,

$$\mathbf{C}' = \begin{bmatrix} \mathbf{C} & \mathbf{C}_* \\ \mathbf{C}_*^\top & \mathbf{C}_{**} \end{bmatrix}.$$

Here, we now develop two different estimation approaches for material density at arbitrary locations, \mathbf{d}^* . The inference stands on the Bayesian framework,

$$P(\mathbf{d}^*, \mathbf{l}|\tilde{\mathbf{d}}) = P(\mathbf{d}^*|\tilde{\mathbf{d}}, \mathbf{l}) P(\mathbf{l}|\tilde{\mathbf{d}}) \propto P(\mathbf{d}^*|\tilde{\mathbf{d}}, \mathbf{l}) P(\tilde{\mathbf{d}}|\mathbf{l}) \pi(\mathbf{l}) = P(\mathbf{d}^*, \tilde{\mathbf{d}}|\mathbf{l}) \pi(\mathbf{l}) \quad (18)$$

Note that the common multiplier to both the sides, $P(\mathbf{d}^*|\tilde{\mathbf{d}}, \mathbf{l})$, is the posterior predictive distribution.

1) *Fixed length scale parameters*: The approach estimates \mathbf{d}^* , adopting a point estimate for lsp parameters \mathbf{l} , *i.e.* the maximum a posteriori (MAP) probability estimate that equals the mode of the posterior probability, from (18),

$$P(\mathbf{l}|\tilde{\mathbf{d}}) \propto P(\tilde{\mathbf{d}}|\mathbf{l}) \pi(\mathbf{l}).$$

Here, we assume an independent uniform distribution for the prior distribution $\pi(\mathbf{l})$. The MAP estimates are obtained by Metropolis-Hastings algorithm, the proposed distribution for which is the half-normal distribution. The mean of the proposed distributions was set equal to the last accepted value, and the variance was set after experimentation. The Higham algorithm [11] was implemented to make sure the covariance matrix, $\mathbf{C}(\mathbf{l})$, positive definite.

With the MAP estimates, say $\tilde{\mathbf{l}}$, the posterior predictive density function is given in a closed form, that is, a multivariate normal distribution [4],

$$P(\mathbf{d}^*|\tilde{\mathbf{d}}, \tilde{\mathbf{l}}) = N \left(\mathbf{C}_*^\top \mathbf{C}^{-1} \tilde{\mathbf{d}}, \mathbf{C}_{**} - \mathbf{C}_*^\top \mathbf{C}^{-1} \mathbf{C}_* \right), \quad (19)$$

suggesting that the predicted material density at an arbitrary location, $d_i^{(m)*}$, is given as an element of the mean vector,

$$\mathbf{d}^* = \mathbf{C}_*^\top \mathbf{C}^{-1} \tilde{\mathbf{d}}.$$

2) *Random length scale parameters*: This approach allows some randomness in lsp, accounting for the joint posterior probability (18),

$$P(\mathbf{d}^*, \mathbf{l} | \tilde{\mathbf{d}}) \propto P(\tilde{\mathbf{d}}, \mathbf{d}^* | \mathbf{l}) \pi(\mathbf{l}). \quad (20)$$

Prior information $\pi(\mathbf{l})$ is assumed to be uniform. The proposal distributions for the algorithm were set as normal distributions with the mean equal to the last accepted value and the variance assigned after experimentation. The proposal distributions for the lsp were truncated to be positive numbers. The posterior predictive distribution is given as

$$P(\mathbf{d}^* | \tilde{\mathbf{d}}) = \int P(\mathbf{d}^*, \mathbf{l} | \tilde{\mathbf{d}}), \partial \mathbf{l} \quad (21)$$

and the predicted material density is then given as its marginal,

$$d_i^{(m)*} = \int d_i^{(m)*} P(d_i^{(m)*} | \tilde{\mathbf{d}}) P(\mathbf{d}_{-i}^{(m)*} | d_i^{(m)*}, \tilde{\mathbf{d}}) \partial \mathbf{d}^*, \quad (22)$$

where \mathbf{d}_{-i}^* is a vector omitting $d_i^{(m)*}$ element.

The computational cost for this approach becomes expensive compared to the fixed lsp approach, as the predictive posterior probability is not given in a closed form. Due to the computational cost, the total density values of \mathbf{d}^* is learned through parallel chains, aiming to learn 30 to 40 different density values each.

C. Evaluation of learning efficiency

We evaluate the efficiency of the two methods described above through the following three aspects. First, we quantify the discrepancy between the posterior mean of material density and its true value, using the mean squared error (MSE),

$$\text{MSE} = \frac{1}{n^*} \sum_{i,m} (d_i^{(m)*} - d(s_i^{(m)}))^2. \quad (23)$$

Second, we investigate the extent to which the method allows the discontinuity in the predicted material density. Since the GSP tends to smooth out discontinuities and predicts values close each other for those locations within a short distance, we calculate the difference (Df) of density in sequential locations,

$$\text{Df} = d(s) - d(s'),$$

for the true density, $d(s)$, and predicted ones \mathbf{d}^* (posterior means of the GSP) by each method. For sample 1, the Df is calculated for the voxels aligned with the y -axis ($Df = d_i^m - d_{i+1}^m$); for sample 2 it was with the z -axis in the ascending direction ($Df = d_i^m - d_i^{m+1}$). The methods that can capture such discontinuity should demonstrate a closer Df value to that of the true density.

Third, we compare the posterior variances of the predicted density values for each method, through which we can judge the adaption of the model to the data.

IV. RESULTS

A. Sample 1

Fig. 2 presents contour plots for the x - y surface at depth $z = 0.4$. Panel (2a) is the true density, and Panel (2b) shows the output from the image inversion, $\tilde{\mathbf{d}}^{(m)}$, the posterior mean (11) of material density for each voxel. The colour gradient suggests lower material density for the region where the void is set (7). There are misestimations (deeper or lighter blue areas) that could be associated with the lack of information while we go deeper in the interior of the material or with the posterior variances which show that the true values have high posterior probability, even though there is short distance from the posterior means.

Panels (2c) and (2d) show the predicted material density, $\mathbf{d}^{(m)*}$, by fixed and random lsp approach respectively. The predicted material density is close to 1 almost everywhere in Panel (2c). The quadrilateral shapes appeared are only due to the discrete colour scheme of the plot; the actual numerical difference between these two is subtle. Note that the fixed lsp method has clearly failed to identify the void region. Due to the prior given variance, the GSP has treated the lower density within the void area as a product of noise and then has smoothed over all these “trends”, bringing them close to the most common value.

In contrast, the random lsp approach (Panel 2d) captures the void with two contour levels. The approximate material density of the centre circle is 0.2, and that of the outer circle is 0.9, close to 1. Since the density of the outer circle was calculated from a mix of voxels with some or void density, the predicted value reflects such a mixed proportion.

As to the learning efficiency between the two methods, the random lsp method shows a smaller MSE value (Table I). The posterior variances were lower for the fixed lsp (Table II). For both approaches, the difference (Df) values are closer to the truth (Table III) for the locations where there is no consecutive change in the density values (thus the true Df is zero). However, for the locations that contain a sudden switch in the density zero to one, or the other way around, the Df values tend to be far from the true ones for both the approaches.

TABLE I: Sample 1 — Mean squared errors.

<i>Method</i>	<i>Mean squared error</i>
Fixed lsp	0.114
Random lsp	0.039

TABLE II: Sample 1 — Posterior Variances of density values.

<i>Density value</i>	<i>Fixed lsp post var</i>	<i>Random lsp post var</i>
$d(-4, -4, 0.4)$	0.283	0.516
$d(-2.4, 0.8, 0.4)$	0.259	0.433
$d(-0.8, 0.8, 0.4)$	0.277	0.369

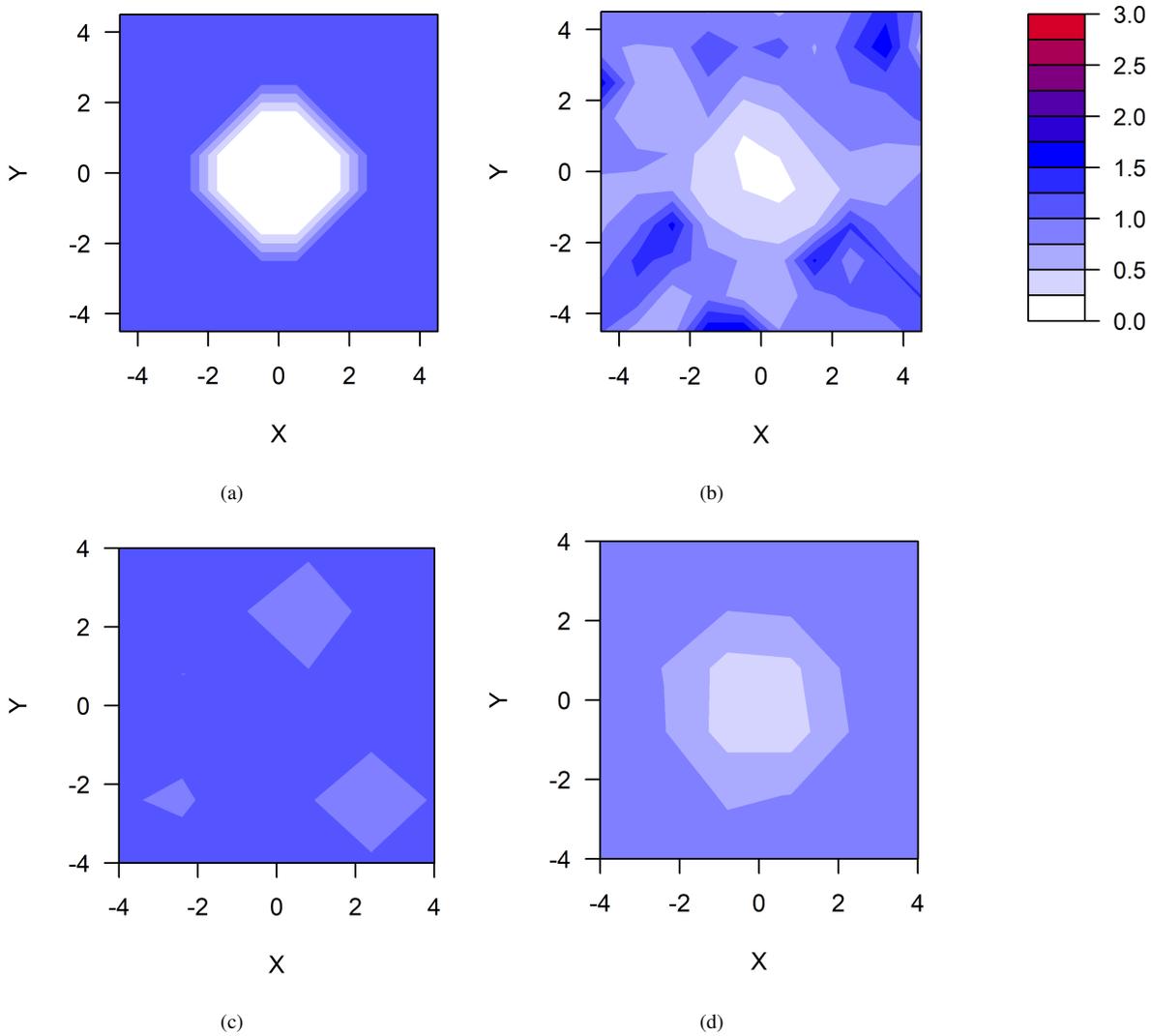


Fig. 2: Results for Sample 1. The true material density is everywhere equal to 1, except for the interior of the void density (6). (a): the true form of the material density function $d(x, y, 0.4)$; (b): the discretised form of the material density function $\tilde{d}(x, y, 0.4)$ via the image inversion; (c): the estimated material density $d^*(x, y, 0.4)$ using the fixed lsp approach; and (d) the estimated material density $d^*(x, y, 0.4)$ using the random lsp approach.

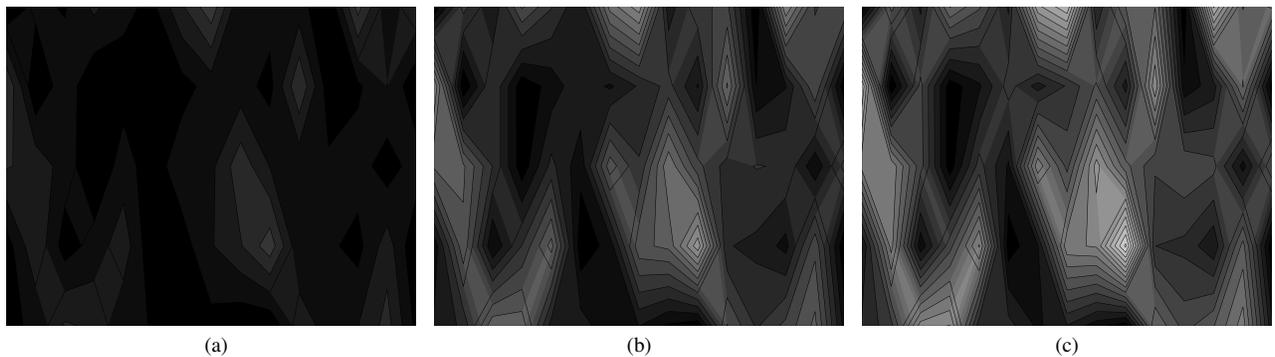


Fig. 3: Simulated SEM images of Sample 2. Panels show three of the 18 simulated images, each of which is projected by different power level of the beam: the 7th lowest (a), the 12th lowest (b) and highest (c).

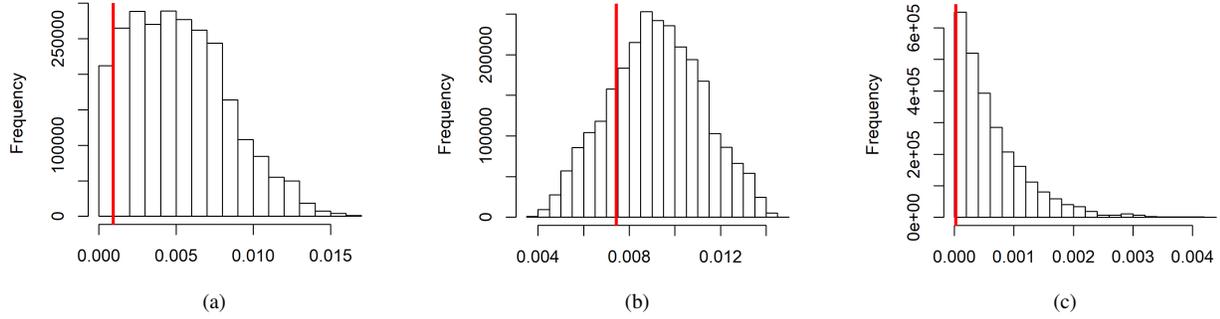


Fig. 4: Histograms of the inverted density of Sample 2. Each panel shows $\tilde{d}(x, y, z)$ for different locations. The red line indicates the true density value.

TABLE III: Sample 1 — Difference (Df) of arbitrary sequential locations.

True Df	Fixed lsp means Df	Random lsp means Df
0	0.021	0.053
1	0.026	0.331
1	0.108	0.385
0	0.009	0.032

B. Sample 2

Fig. 3 a–c show three examples from the 18 simulated images of Sample 2. The panel images (3a)–(3c) are respectively obtained from the 7th, 12th lowest and the highest energy beam.

Fig. 4 d–f show histograms of $\tilde{d}(x, y, z)$ for three different locations, the approximated posterior distribution via the image inversion. The red line represents the true value at the corresponding location. In general, the posterior means provided by the image inversion are close to the true values.

Fig. 5 presents contour plots of the material density in the x – z slice at $y = -7.2$. Panel (5a) shows the true density. Panel (5b) is the output of the image inversion, \tilde{d} . There are some misestimations compared to the true values (Panel 5a), for example, missing a void from where it should be $(0, -7.2, 0.5)$ and $(5, -7.2, 0.4)$, or a dense area $(5, -7.2, 0.2)$. However, the inversion manage captures relatively well the key features of the original image.

The learning results of each GSP, \mathbf{d}^* , are presented in Panel (5c) for the fixed lsp approach and Panel (5d) for the random lsp approach. The fixed lsp approach (5c) has resembled well the inversion image, \tilde{d} , (5b). The most noticeable difference is the shrunken void regions in (5c) with some separations from one another, for example, location $(-5, -7.2, 0.7)$. This result implies that, for a highly discontinuous case like this, the learning improvement is minuscule when the lsp is fixed. The results of the random lsp approach (Panel 5d) exhibit more divergent from the inversion image (5b). The learning has successfully toned down the red areas at the centres of the darker blue quadrilaterals. However, it has significantly underestimated the void regions compared to the true image

(5a), which is due to the information from the neighbouring x – z slice, in which the corresponding coordinates possess non-void density.

TABLE IV: Sample 2 — Mean squared errors.

Method	Mean squared error
Fixed lsp	1.94×10^{-5}
Random lsp	2.32×10^{-5}

TABLE V: Sample 2 — Posterior Variances of density values.

Density value	Fixed lsp post var	Random lsp post var
$d(-10, -7.34, 0)$	7×10^{-3}	4×10^{-5}
$d(-8.67, -7.34, 0.185)$	7×10^{-3}	3.6×10^{-5}
$d(-4.68, -7.34, 0.350)$	7×10^{-3}	4.3×10^{-5}
$d(3.3, -7.34, 0.675)$	7×10^{-3}	3.6×10^{-5}
$d(7.34, -7.34, 0)$	7×10^{-3}	4.8×10^{-5}

TABLE VI: Sample 2 — Difference (Df) for arbitrary sequential locations.

True Df	Fixed lsp means Df	Random lsp means Df
-0.006	-0.017	-0.009
0.003	-6×10^{-5}	-0.004
-0.002	-0.014	-0.008
-0.008	0.016	0.013
-2×10^{-6}	-0.019	-0.011

For the learning efficiency between two approaches, the fixed lsp shows a smaller MSE (Table IV). The difference in MSE between these two approaches is due to the misspecification of the void region. Table V shows the posterior variances for the same arbitrary location. Under the fixed lsp approach, the variance is the same for all the coordinates and at least 100 times greater than for the random lsp. Since the density has remained close to the inversion image by the fix lsp approach, the higher variance works, covering a wider range for possible values. The random lsp approach exhibits such lower posterior variance, showing a better adaption to the function. In terms of Dfs between the two approaches, Table VI indicates a better

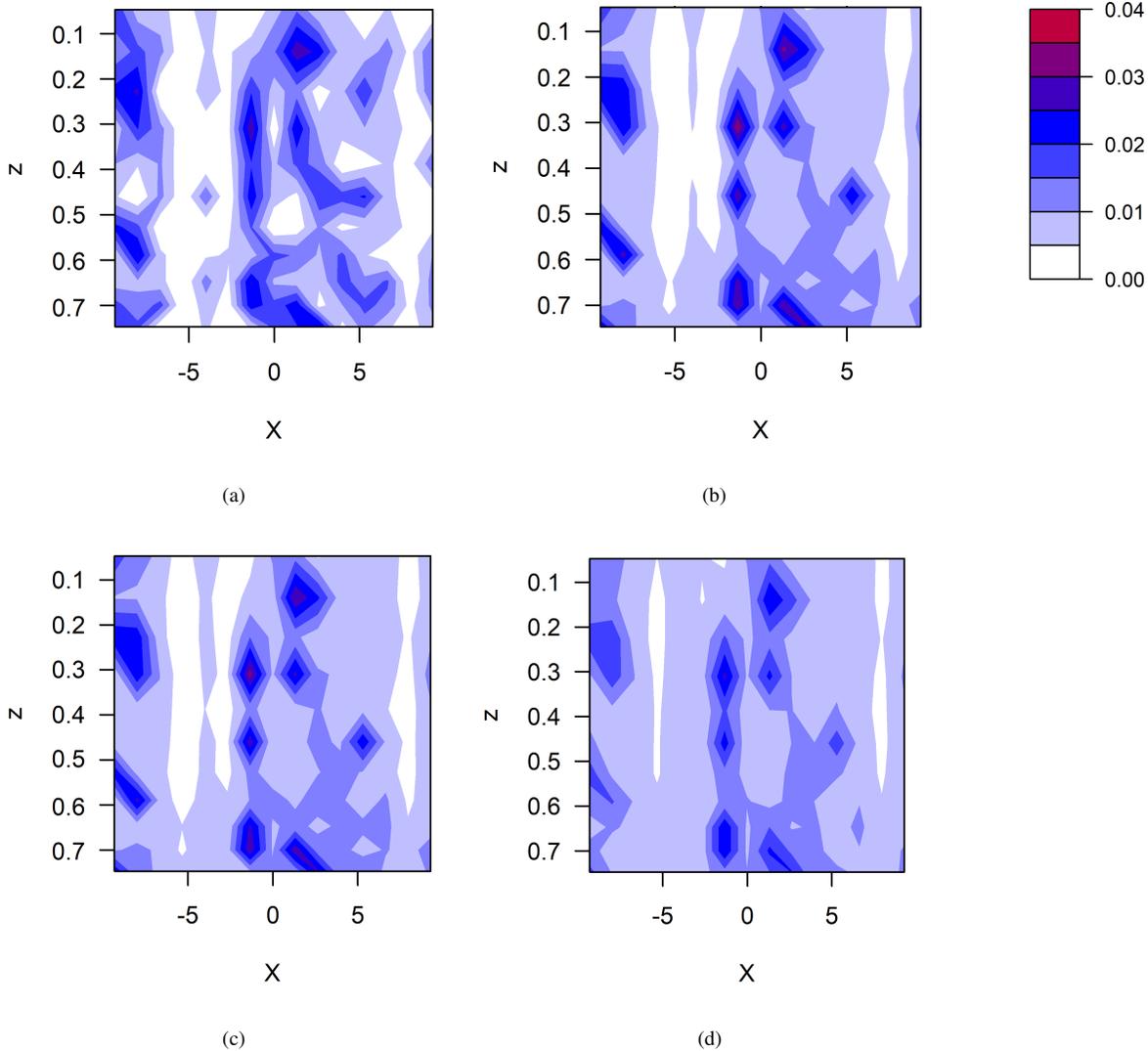


Fig. 5: Results for Sample 2. (a): the true form of the material density function $d(x, -7.2, z)$; (b): the discretised form of the material density function $\tilde{d}(x, -7.2, z)$ via the image inversion; (c): the estimated material density $d^*(x, -7.2, z)$ using the fixed lsp approach; and (d) the estimated material density $d^*(x, -7.2, z)$ using the random lsp approach.

performance of the random lsp approach that adapts well to the random variation of the material density. The cases that the fixed lsp prevails in Dfs, they are significantly less and probably due to randomness in the definition of the true density function.

V. DISCUSSION

The interior material density often contains discontinuities due to some random voids within the material. We have, for the first time, proposed here a methodological framework and demonstrated how it could estimate the non-discretised material density structure from SEM images, using the two simulation examples. Although GSPs perform better for smooth functions [4], our comparison between the two different parametrisation approaches has highlighted that the

random lsp approach, an extension of the non-nested model [14], can cope better with some discontinuities in the material structure.

As we have shown, the level of discontinuity within the material plays a pivotal role in the success of GSP modelling. Sample 1 has illustrated better results, particularly for the random lsp approach, as the interior material density is continuous everywhere except for the boundary of the dense and the spherical void regions. Contrary, Sample 2 is discontinuous almost everywhere among all these voxels, which is rather an extreme scenario. In normal circumstances, materials exhibit more locally homogeneous structures in the voxel scale, meaning that neighbouring voxels share the same or similar density. Despite such an unusual scenario of Sample 2, our modelling approaches have partially managed to capture

the key feature of the material structure. The posterior mean values have stayed more less the same as the inversion with high variance, allowing the values to vary independently from the coordinates. These results are not surprising since the use of GSP with the squared-exponential covariance kernel is usually suggested in Machine Learning for learning smooth continuous functions. Importantly, allowing the randomness to the l_{sp} , the GSP has managed to delineate the interior material density efficiently.

A key aspect of GSP is covariance kernel modelling as it specifies the process. For example, [4] suggests the radial bias function (RBF) kernel, which simply exerts a multiplicative coefficient, namely the amplitude parameter, to our kernel (14) like,

$$A \exp \left\{ - (s_i^{(m)} - s_j^{(m')})^\top \mathbf{M} (s_i^{(m)} - s_j^{(m')}) \right\} + (\sigma_i^{(m)})^2 I_{\{i=j\}} I_{\{i \leq \tilde{n}\}}, \quad (24)$$

as a typical choice and also recommends that the prior variance should be used only for $\tilde{d}_i^{(m)}$. Here, the amplitude parameter A will regulate the variations. However, we have found that, in our case, the amplitude parameter induces possible model misspecification, due to the discontinuity in the samples, and our modelling approach (14) works better.

Further, the posterior probability distribution (20) can be multimodal, suggesting the need for different l_{sp} values for different coordinates rather than globally fixed values. Our results underline the fact that the non-nested model [14] is sufficient in order to capture the key features in the interior material density. However, the random l_{sp} approach asks a computational cost, for which some forms of parallelisation is required. For example, we recommend learning the density values in a group of 30 variables parallel at the most.

REFERENCES

- [1] Andrew G. Wilson, Zhiting Hu, Russ R. Salakhutdinov, Eric P. Xing, "Stochastic Variational Deep Kernel Learning", *Advances in Neural Information Processing Systems* 29 (NIPS 2016).
- [2] Andrew Gelman, Donald Rubin, John B. Carlin, David Dunson, and Hal S. Stern, "Bayesian Data Analysis 3rd edition", CRC Press, Boca Raton, 1995.
- [3] Bertero Mario, and Boccacci Patrizia, "Introduction to inverse problems in imaging", Taylor and Francis, 1998.
- [4] Carl Edward Rasmussen, and Christopher K. I. Williams, "Gaussian Processes for Machine Learning", The MIT Press, Cambridge, Massachusetts, London England, 2006.
- [5] Christian Robert, and George Casella, "Monte Carlo Statistical Methods", Springer, New York, 2004.
- [6] Dalia Chakrabarty, Fabio Rigat, Nare Gabrielyan, Richard Beanland, and Shashi Paul, "Bayesian Density Estimation via Multiple Sequential Inversions of 2-D Images with Application in Electron Microscopy", *Technometrics*, vol. 57 (2), 2015, pp. 217-233.
- [7] Dalia Chakrabarty, Munmun Biswas, and Sourabh Bhattacharya, "Bayesian nonparametric estimation of Milky Way parameters using matrix-variate data, in a new Gaussian Process based method", *Electronic Journal of Statistics*, vol. 9 (1), 2015, pp. 1378-1403.
- [8] Davis G. R., Dowker S. E. P., Elliott J. C., Anderson P., Wassif H. S., Boyde A., Goodship A. E., Stock S. R., and Ignatiev, K., "Non-destructive 3D structural studies by X-ray microtomography", *Advances in X-ray Analysis*, vol. 45, 2002, pp. 485.

- [9] Gabrielyan Nare, Konstantina Saranti, Krishna Nama Manjunatha, and Shashi Paul, "Growth of low temperature silicon nano-structures for electronic and electrical energy generation applications", *Nanoscale Research Letters*, vol 8 (1), 2013, pp. 83.
- [10] Goldstein J., Newbury D. E., Joy D. C., Lyman C. E., Echlin P., Lifshin E., Sawyer L., and Michael J., "Scanning Electron Microscopy and X-ray Microanalysis", Springer, New York, 2003.
- [11] Higham Nick, "Computing the nearest correlation matrix - a problem from finance", *IMA Journal of Numerical Analysis*, vol. 22, 2002, pp. 329-343.
- [12] James Hensman, Nicolo Fusi, Neil D. Lawrence, "Gaussian Processes for Big Data James", *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence (UAI2013)*.
- [13] Kanaya K., and Okamaya S., "Penetration and energy-loss theory of electrons in solid targets", *Journal of Physics D: Applied Physics*, vol. 5 (1), 1972, pp. 43-58.
- [14] Kangrui Wang, and Dalia Chakrabarty, "Deep Bayesian Supervised Learning given Hypercuboidally-shaped, Discontinuous Data, using Compound Tensor-Variate & Scalar-Variate Gaussian Processes", in press.
- [15] Leone F. C., Nottingham R. B., and Nelson L. S., "The Folded Normal Distribution", *Technometrics*, vol. 3 (4), 1961, pp. 543.
- [16] Merlet Claude, "An Accurate Computer Correction Program for Quantitative Electron Probe Microanalysis", *Mikrochim. Acta*, vol. 114/115, 1994, pp. 363-376.
- [17] Natterer Frank, and Wübbeling Frank, "Mathematical methods in image reconstruction", Philadelphia: SIAM., 2001.
- [18] Paciorek Chris, Schervish Mark, "Spatial modelling using a new class of nonstationary covariance functions", *Environmetrics*, 17, 2006, pp. 483-506.
- [19] Panaretos M. Victor, "On Random Tomography with Unobservable Projection Angles", *Annals of Statistics*, vol. 37 (6), 2009, pp. 3272.
- [20] Stuart Andrew, "Inverse problems: a Bayesian perspective", in *Acta Numerica*, Vol. 19 Cambridge University Press, 2010, pp. 451.
- [21] Sawyer L., Grubb D. T., Meyers Gregory, "Polymer Microscopy", New York: Springer Science, 2008.
- [22] Tarantola Albert, *Inverse Problem Theory and Methods for Model Parameter Estimation*, Philadelphia: SIAM, 2004.
- [23] Vanderlinde William, Caron James, "Blind Deconvolution of SEM Images", *ISTFA 2007, Conference Proceedings of the 33rd International Symposium for Testing and Failure Analysis*, ASM International, 2007, pp. 97-102.
- [24] Williams Chris, "Prediction with Gaussian processes: From linear regression to linear prediction and beyond", In: Jordan, M.I. (ed.) *Learning in Graphical Models*, pp. 599-621. Kluwer Academic, Dordrecht (1998).
- [25] Williams Chris, Barber David, "Bayesian classification with Gaussian processes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, vol 20 (12), pp. 1342-1351.