UNCERTAINTY QUANTIFICATION OF INCLUSION BOUNDARIES IN THE CONTEXT OF X-RAY TOMOGRAPHY[*]

BABAK MABOUDI AFKHAM[†], YIQIU DONG[†], AND PER CHRISTIAN HANSEN[†]

Abstract. In this work, we describe a Bayesian framework for reconstructing the boundaries of piecewise smooth regions in the X-ray computed tomography (CT) problem in an infinite-dimensional setting. In addition to the reconstruction, we are also able to quantify the uncertainty of the predicted boundaries. Our approach is goal oriented, meaning that we directly detect the discontinuities from the data, instead of reconstructing the entire image. This drastically reduces the dimension of the problem, which makes the application of Markov Chain Monte Carlo (MCMC) methods feasible. We show that our method provides an excellent platform for challenging X-ray CT scenarios (e.g., in case of noisy data, limited angle, or sparse angle imaging). We investigate the performance and accuracy of our method on synthetic data as well as on real-world data. The numerical results indicate that our method provides an accurate method in detecting boundaries of piecewise smooth regions and quantifies the uncertainty in the prediction.

Key words. X-ray CT, Bayesian framework, inverse problems, Whittle-Matérn field, goal-oriented UQ

AMS subject classifications. 60G10, 60G15, 60G50, 60G60, 60J20, 28-08, 46B99, 65C05, 65C20, 65C40

1. Introduction. Computed tomography (CT) imaging is the task of reconstructing a positive attenuation field (in the form of an image) from a finite number of projections (e.g., sinograms). CT reconstruction is often formulated as an inverse problem [30]. *Filtered back-projection* [37] is a classic reconstruction approach to solve this inverse problem for some CT settings. However, the quality of the reconstructed image is compromised by imaging challenges such as noise, sparse angle imaging (to reduce the amount of harmful radiation), or limited angle imaging (to avoid obstacles in the imaging site or due to application set-up, e.g., in mammography). Therefore, finding alternative approaches for CT reconstruction has attracted attention in the past few decades [8, 12, 41, 48, 49, 55].

In many CT reconstruction methods, such as algebraic iterative methods and regularization methods [26], the goal is to identify objects in the image with approximately homogeneous attenuation coefficient distinct from the background attenuation [15, 19, 24, 59]. Therefore, CT reconstruction is often followed by an image segmentation step to partition the image into piecewise smooth/constant regions. The boundaries between such regions often carry valuable information [56]. Error propagation from reconstruction to segmentation, due to concatenation of such methods, can introduce artifacts in the segmentation. Such effects are also amplified in case of noisy or incomplete data.

To avoid such artifacts many methods attempt to combine the reconstruction and the segmentation (and subsequently boundary extraction) steps. One common approach is to describe the images as a level set of a smooth functions, see, e.g., [3, 8, 33, 41, 55, 60, 61] for selected examples. Lambda Tomography [20, 21, 52, 58] is another approach in identifying boundaries. These methods are based on filtered back-projection method where the filters are chosen to emphasize boundaries. An-

---

   [†]Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, Denmark (bmaaf@dtu.dk, yido@dtu.dk, pcha@dtu.dk).

other common approach is to construct a deforming/parametric curve which evolves to fit the boundaries between partitions, see, e.g., [12, 48, 49, 50] as a non-exhaustive selection. Large parameter spaces, dependency on discretization, noisy data, and limited angle imaging can be challenging for some of the mentioned methods. Moreover, all the methods mentioned above lack the quantification of the uncertainties in the reconstructed/segmented images with respect to noisy/perturbed data. Uncertainty quantification could be particularly important in applications where images are used to determine the location and the size of objects. For example, in medical imaging, CT reconstruction is used to track the evolution of tumor boundaries over time and make treatment decisions.

A popular approach to characterize uncertainties for inverse problems is within the Bayesian framework. In this approach, all quantities of a model are represented as random variables. The solution to the inverse problem is then the probability distribution, *the posterior distribution*, of the quantity of interest after all given information (e.g., the *prior* belief in the model) is incorporated into the model. Qualities of the posterior distribution can then be interpreted as the degree of confidence in predicting the quantity of interest. In general, Markov Chain Monte Carlo (MCMC) methods [38] are applied to explore the posterior distribution. In the context of X-ray CT, the Bayesian framework has been successfully applied, especially for reconstructing the attenuation field, see, e.g., [9, 10, 51, 55].

In the past decade, a formulation of a Bayesian inversion theory in an infinite dimensional setting has attracted attention, see [14] and the references therein. In this setting, model parameters are modeled as random functions rather than real-valued random variables. This provides a discretization-independent numerical platform for exploring the posterior distribution. Although posterior distributions for such problems are in infinite dimensions, the MCMC methods for exploring such distributions are similar to the more traditional (finite-dimensional) ones.

Infinite-dimensional Bayesian methods for tomography problems are not new. For example, we mention recent works in the context of electrical impedance tomography (EIT), inverse scattering problem and other partial differential equation based inverse problems, see [7, 17, 16, 35, 27, 18, 44] and the references therein. The underlying partial differential equation provides a natural platform (a Hilbert space) for the infinite-dimensional random variable to be well defined. Detailed analysis of these inverse problems in a Bayesian setting [14] shows that the solution (the posterior distribution) is well-defined and bounded under perturbations of the data. Therefore, the Bayesian platform provides an alternative method to investigate the well-posedness of an inverse problem.

However, to the best knowledge of the authors, we still lack a well-established infinite-dimensional framework for the CT problem. In this paper, we show that the well-posedness results, mentioned above, can be extended to the CT problem. This provides an excellent tool to evaluate and quantify the uncertainties in the reconstruction and segmentation of images.

Here, we focus on looking for objects/inclusions in an image with a homogeneous attenuation coefficient distinct from the background. Especially, we are interested in reconstructing the boundaries of the objects and quantifying the uncertainties in this reconstruction due to noisy or incomplete data. One approach in evaluating such uncertainties is to consider a Markov random field (MRF) prior. We can estimate uncertainties using, e.g., MCMC methods. The edges are then detected using thresholding-based methods [40]. We call this approach the sampling-then-threshholding (STT) method. The advantage of STT methods is that advanced methods for reconstruction

and edge detection/segmentation have been proposed. But some of the challenges of using STT methods are: 1) error propagation from reconstruction to segmentation; 2) resolution dependency, i.e., the quality of reconstructed boundaries in STT depends on the resolution of discretized attenuation field; 3) large computational cost due to the large number of samples of a 2D reconstruction problem. These challenges make the use of STT methods unattractive. In Figure 1, we show a numerical comparison between our method, which will be introduced later, and an STT method that combines the two methods in [40, 51]. The STT method obtains the boundaries of the objects in each sample after applying the edge detection algorithm. The mean of the boundaries can give us some information on the uncertainties in the location of the edges. But note that in this way we cannot really quantify the uncertainties of the boundary curve. In addition, in Figure 1(f) we can clearly see the resolution dependency.
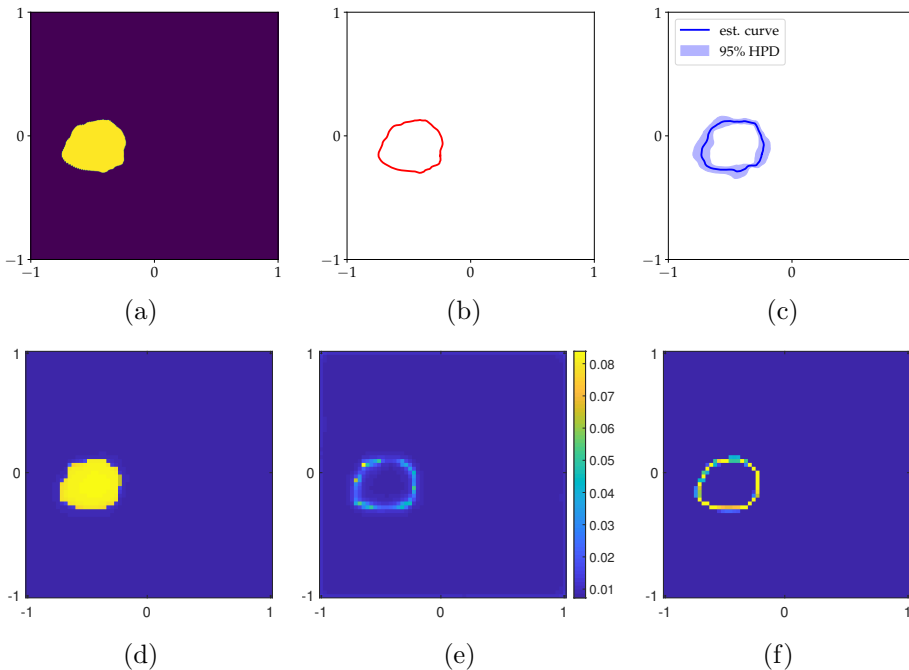

FIG. 1. *Comparing the performance of our method and the sampling-then-threshholding (STT) method which combines the methods in [40, 51]. (a) True attenuation field, (b) true boundary, (c) the estimated boundary by using our method together with the 95% highest posterior density credibility interval, (d) the mean of the attenuation reconstructions by STT, (e) the standard deviation of the reconstructions by STT, (f) the mean of the boundaries by STT.*

In this paper, we provide a goal-oriented, infinite-dimensional Bayesian framework for the CT problem. In our method, we reconstruct the boundary directly, and without the need for reconstructing the attenuation field. This avoids the error propagation and reduces the dimensionality of the problem from finding a 2D image to a 1D boundary of a region. In addition, we model the boundaries as random functions in order to avoid resolution dependency. Our method consists of two stages: in the stage 1, we approximate locations of all objects in order to maintain the well-posedness of the problem that is only valid for single inclusion; in stage 2, we identify the boundaries of each object, which can be parallelized with respect to the number

of inclusions to further improve the efficiency. Our method provides an uncertainty band around the boundary to quantify the reliability in the prediction of the boundaries. We apply the preconditioned Crank-Nicolson (pCN) method [11] as an MCMC method for exploring the posterior. The pCN ensures that the reconstruction of the boundaries is independent of the discretization of the random variables. Figure 1(c) shows an example of our method predicting boundaries of the inclusion together with the uncertainty in this prediction. It is clear that our result is resolution independent.

This paper is organized as follows: In Section 2 we formulate the CT problem and introduce the Radon transform. The infinite dimensional Bayesian framework for inverse problems is presented in Section 3 and Section 4. In Section 3, we introduce the Matérn-Whittle random field and construct the log-Gaussian, level set and the star-shaped priors. In Section 4, we construct a likelihood function for the CT problem, introduce the posterior, and discuss the the existence and the well-posedness of the posterior distribution. In Section 5 we introduce our two-stage method to identify inclusion boundaries. We evaluate the performance of the method for simulated images in Section 6 for single and multiple inclusions, as well as in sparse and limited angle imaging settings. In addition, we show the performance of our method on a real CT-scan of a tomographic X-ray data of a lotus root slice filled with different chemical elements [5]. We present conclusive remarks in Section 7.

2. Radon Transform and Forward Model. Tomography, or slice/volume imaging, comprises methods in reconstructing the internal structure of an object from external measurements. In X-ray computed tomography, the X-rays interact with the matter in the object, and we measure the attenuation of the rays on detectors placed on the other side of the object [47].

A common approach in modeling X-ray interaction with an object is described by line integrals over a density field. Let $D \subset \mathbb{R}^2$ be a bounded region with Lipschitz boundary. Let $0 < \alpha \in L^2(D)$ denote a density field defined on D. This field represents the attenuation of the X-rays [47]. A measurement can then be described by the line integral

$$(2.1) \qquad R_{\theta,s}[\alpha] := \int_{L_{\theta,s}} \alpha(x) \; dl.$$

Here, $s \in [-1, 1]$ parameterizes points on the line $L_\theta^\perp$ passing through the origin and making the angle $\theta \in [0, \theta_{\max})$, $\theta_{\max} \le \pi$, with the x-axis. Furthermore, $L_{\theta,s}$ is the perpendicular line to $L_\theta^\perp$ crossing it at point s, and dl is an infinitesimally small length on $L_{\theta,s}$. We assume that α has a compact support such that the integration in (2.1) is carried over a finite section of $L_{\theta,s}$.

Note that $R_{.,.}$ can be viewed as a linear functional from lines in $\mathbb{R}^2$ to $\mathbb{R}$ in which case it is referred to as the *Radon transform* [26, 37, 46]. The simplest set of measurement lines are a finite set of equidistant parallel lines perpendicular to $L_\theta^\perp$, for a finite set of angles θ. This setup is referred to as the *parallel-beam geometry*. The collected measurements of this type is commonly referred to as a *sinogram*.

We discretize θ into N_θ angles and s into N_s detector pixels to construct $N := N_\theta \times N_s$ measurements. We define an observation vector $y(\alpha) \in \mathbb{R}^N$ to contain elements of noisy measurement of type

$$(2.2) \qquad y_{i,j}(\alpha) = R_{\theta_i,s_j}[\alpha] + \varepsilon_{i,j}, \qquad i = 1, \ldots, N_\theta, \; j = 1, \ldots, N_s,$$

where $\varepsilon_{i,j}$ are independent and identically distributed (i.i.d.) following $\mathcal{N}(0, \sigma_n^2)$, for some $\sigma_{\text{noise}} > 0$. Note that, in general, noise in CT data is non-Gaussian. The choice

of observation noise in (2.2) is an approximation. Representing (2.2) in vector form we obtain

$$\boldsymbol{y}(\alpha) = G\alpha + \boldsymbol{\varepsilon}. \tag{2.3}$$

Here, $G : L^2(D) \to \mathbb{R}^N$ contains line-integrals R_{θ_j, s_j} referred to as the *forward map* and $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$, with $\Sigma = \mathrm{diag}(\sigma_{\mathrm{noise}}^2, \ldots, \sigma_{\mathrm{noise}}^2)$. Note that since R_{θ_j, s_j} is linear then G is locally Lipschitz. The task of finding α from $\boldsymbol{y}(\alpha)$ forms an inverse problem. This inverse problem is severely under-determined since α is infinite dimensional and $\boldsymbol{y}$ is finite dimensional.

In many tomography applications we are interested in finding the discontinuities of α and not reconstructing the entire α. In this paper, we take a *goal-oriented* approach in finding and quantifying the uncertainty of discontinuities in α. Such discontinuities can be parameterized with a significantly lower dimensionality than α.

In this paper, we take a Bayesian approach in representing the tomography inverse problem. Severe under-determination of the inverse problem is regularized by a prior knowledge/assumption on α in the form of a probability distribution, which is introduced in Section 3. In Section 4, we apply the Bayes' theorem to combine the prior distribution with the forward map G, then derive a novel posterior distribution which represents the solution to the inverse problem.

3. Prior Models. In this section we introduce various prior modelling techniques to construct a density field α, which is inspired by the works in [14, 17, 18, 1]. We first review how to construct random functions on a Hilbert space by Gaussian random variables, then present two prior models used in our method.

3.1. Review of Gaussian Random Variables. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a complete probability space with Ω a measurable space, $\mathcal{A}$ a σ-algebra defined on Ω and $\mathbb{P}$ a probability measure. Furthermore let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space. See [28] for further information on a probability space. A random element $\xi \in H$ is called an H-valued Gaussian random variable if $\langle \xi, \zeta \rangle$, for all $\zeta \in H$, is a real-valued Gaussian random variable, i.e., measurable function on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

THEOREM 3.1. *[28] Let ξ be an H-valued Gaussian random variable. Then we can find $m \in H$, the* mean function, *and a symmetric, non-negative and trace-class linear operator $Q : H \to H$, the* covariance operator, *such that*

$$\langle m, \zeta \rangle = \mathbb{E}\langle \xi, \zeta \rangle, \qquad\qquad \forall \zeta \in H, \tag{3.1a}$$
$$\langle Q\zeta, \eta \rangle = \mathbb{E}\langle \xi - m, \zeta \rangle\langle \xi - m, \eta \rangle, \qquad\qquad \forall \zeta, \eta \in H, \tag{3.1b}$$

where $\mathbb{E}$ represents expectation. We write $\xi \sim \mathcal{N}(m, Q)$ and say ξ is Gaussian with measure $\mathbb{P} \circ \xi^{-1}$.

THEOREM 3.2. *[25] Let $m \in H$ and $Q : H \to H$ be a symmetric, non-negative and trace-class linear covariance operator. Furthermore let $\{e_i\}_{i \in \mathbb{N}}$ be eigenfunctions of Q and $\{\lambda_i\}_{i \in \mathbb{N}}$ be eigenvalues of Q sorted in decreasing order. An H-valued random variable ξ is a Gaussian random variable with the distribution $\mathcal{N}(m, Q)$ if and only if*

$$\xi = m + \sum_{i \in \mathbb{N}} \sqrt{\lambda_i} \beta_i e_i, \tag{3.2}$$

where $\{\beta_i\}_{i \in \mathbb{N}}$ is a sequence of i.i.d. real-valued random variables with distribution $\mathcal{N}(0, 1)$. The expansion in (3.2) is referred to as the Karhunen-Loéve (KL) expansion of ξ.

Remark 3.3. In case all random variables are Gaussian random variables, we can construct a probability measure on H using push-forward notation $\mu_0 = \xi_\star(\mathbb{P})$. Therefore, $(H, \mathcal{B}(H), \mu_0))$, with $\mathcal{B}(H)$ the Borel σ-algebra, forms a probability space. We refer to μ_0 as the *prior measure*.

Theorem 3.2 indicates how we can construct random functions on a Hilbert space H using a covariance operator Q. We can approximate the sum in (3.2) by truncating it after N_{KL} terms. We refer the reader to [28, 25] for convergence properties of the KL expansion.

In the following subsection we introduce a family of covariance operators that we can use to construct Gaussian random functions according to (3.2).

3.1.1. Matérn-Whittle Covariance. A widely used family of covariance operators are the Matérn-Whittle covariance operators [28, 42, 44, 62] which allows control over regularity, amplitude and correlation length of samples. Here we briefly introduce this covariance operator but we refer the reader to Appendix A for a deeper discussion.

A simplified Matérn-Whittle covariance operator [44] is given by

$$(3.3) \qquad Q_{\gamma, \tau} = (\tau^2 I - \Delta)^{-\gamma}.$$

Here, Δ and I are the Laplacian and the identity operators in 1 or 2 dimensions, respectively. Furthermore, $\tau = 1/\ell > 0$ controls the correlation length and $\gamma = \nu + 1$ is the smoothness parameter (see [34] for more detail). For the covariance operator (3.3) to be well defined we need to impose proper boundary conditions. See [44] for more detail on types of boundary conditions.

We note that in one dimension, i.e., when $D \subset \mathbb{R}$, and for $\tau = 1$ the zero-mean Gaussian random variable ξ distributed according to the covariance $Q_{\gamma, 1}$ takes the form

$$(3.4) \qquad \xi(x) = c \sum_{i \in \mathbb{N}} \left(\frac{1}{k} \right)^\gamma \left(\xi_i^1 \sin(kx) + \xi_i^2 \cos(kx) \right),$$

for some constant $c > 0$. Here, $\xi_i^1, \xi_i^2 \sim \mathcal{N}(0, 1)$ are real Gaussian random variables. This type of Gaussian random variables will be used in later sections to model the boundaries of inclusions.

3.2. Prior as a Push-Forward Measure. In this section we apply a nonlinear transformation on Gaussian random variables to represent the image α. Assume that α is piecewise constant, and we use the level set [17] and the star-shaped [18] fields to construct it. Later, in Section 4, we discuss how to use these fields to construct a likelihood function.

3.2.1. Level Set Parameterization. Let ξ be an H-valued Gaussian random variable and take $c \in \mathbb{R}$. Define $D_i^{\mathrm{ls}} \subset D$, for $i = 1, 2$ as

$$(3.5) \qquad D_1^{\mathrm{ls}} := \{ x \in D \mid \xi(x) < 0 \}, \qquad D_2^{\mathrm{ls}} := D \backslash D_1^{\mathrm{ls}}.$$

We define the level set mapping $F_{\mathrm{ls}} : H \to L^p(D)$, with $2 \leq p < \infty$, as

$$(3.6) \qquad F_{\mathrm{ls}}[\xi](x) = a^- \mathbb{1}_{D_1^{\mathrm{ls}}}(x) + a^+ \mathbb{1}_{D_2^{\mathrm{ls}}}(x),$$

where $\mathbb{1}_{D_i^{\mathrm{ls}}}(x)$, for $i = 1, 2$, is the indicator function and $0 < a^- < a^+$. We define $D^0 = \overline{D}_1^{\mathrm{ls}} \cap \overline{D}_2^{\mathrm{ls}}$ which contains the points of discontinuity. Throughout this paper we assume

that $m(D^0) = 0$ where m is the Lebesgue measure defined on $\overline{D}$. This assumption is to ensure that the boundary of the inclusions has indeed a lower dimensionality than the image α. We refer the reader to [29] for more detail.

Remark 3.4. It is shown in [29] that the assumption $m(D^0) = 0$ is sufficient for F_{ls} to be continuous. This means if $\{\xi^\epsilon\}_{\epsilon>0}$ is a sequence of functions such that for any $x \in D$

$$\xi(x) - \epsilon \le \xi^\epsilon \le \xi(x) + \epsilon, \tag{3.7}$$

then $\|F_{\text{ls}}[\xi^\epsilon] - F_{\text{ls}}[\xi]\|_{L^p(D)} \to 0$, μ_0-almost surely.

3.2.2. Star-Shaped Parameterization. Let $\mathbf{T} = [0, 2\pi)$ and $H \subset L^2(\mathbf{T})$ contain period functions on $\mathbf{T}$. Furthermore, define $\vartheta : D \to [0, 2\pi)$ be the continuous map from Cartesian coordinates to the angular component of polar coordinates. Define *star-shaped inclusions* $D_i \subset D$, for $i = 1, \ldots, N_{\text{inc}}$, as

$$D_i(\xi_i, c_i) := \{x \in D \mid \|x - c_i\|_2 < F_{\text{lg}}(\xi_i(\vartheta(x - c_i)))\}, \tag{3.8}$$

where

$$F_{\text{lg}}[\xi] = \exp(\xi) \tag{3.9}$$

defines the *log-Gaussian* field $F_{\text{lg}} : H \to L^2(D)$ in order to construct an image with positive attenuation. Here ξ is an H-valued Gaussian random variable. Since D is bounded, then F_{lg} is almost surely a continuous map. In addition, $\|\cdot\|_2$ is the Euclidean norm and $c_i \in D$ for $i = 1, \ldots, N_{\text{inc}}$ are independent stochastic centers of inclusions. Note that c_i and ξ_i are not necessarily i.i.d. Let $D_0 := (D_1 \cup \cdots \cup D_{N_{\text{inc}}})^C$, where the superscript C represents the complement of a set. We define the star-shaped mapping $F_{\text{star}} : H^{N_{\text{inc}}} \times \mathbb{R}^{2N_{\text{inc}}} \to L^p(D)$ with $2 \le p < \infty$ as

$$F_{\text{star}}[(\boldsymbol{\xi}, \boldsymbol{c})](x) = a^- \mathbb{1}_{D_0}(x) + \sum_{i=1}^{N_{\text{inc}}} a^+ \mathbb{1}_{D_i}(x), \tag{3.10}$$

where $\boldsymbol{\xi} = [\xi_1, \cdots, \xi_{N_{\text{inc}}}]^T$, $\boldsymbol{c} = [c_1, \cdots, c_{N_{\text{inc}}}]^T$, $a^- < a^+ \in \mathbb{R}^+$ and we refer to a^+ as the inclusion intensity. The following assumptions are considered when drawing samples of ξ_i and c_i:

(I) D_i, for $i = 1, \ldots, N_{\text{inc}}$ are disjoint.
(II) $\|x_1 - x_2\|_2 > d_{\min}^{\partial D}$, for all $x_1 \in D_1 \cup \cdots \cup D_{N_{\text{inc}}}$ and $x_2 \in \partial D$ the boundary of D. This insures all inclusions are away from the boundary of D.
(III) $\|x_1 - x_2\|_2 > d_{\min}^D$, for all $x_1 \in D_i$ and $x_2 \in D_j$ for $i \ne j$ and $i, j \ne 0$. This insures that the inclusions are well separated.

Remark 3.5. It is shown in [18] that when $N_{\text{inc}} = 1$ and ξ is Lipschitz continuous (e.g., $\gamma > 1$ in (3.3)) then F_{star} is a continuous mapping. This means if $\{\xi^\epsilon\}_{\epsilon>0}$ and $\{c^\epsilon\}_{\epsilon>0}$ are sequences of H-valued random variables and sequence of points in D, respectively, such that $\|\xi - \xi^\epsilon\|_\infty \to 0$ and $\|c^\epsilon - c\|_2 \to 0$, then $F_{\text{star}}[(\xi^\epsilon, c^\epsilon)] \to F_{\text{star}}[(\xi, c)]$ in measure. When $N_{\text{inc}} > 1$, assumptions (I)-(III) ensures that we can divide D into subregions with only single inclusions.

Remark 3.6. The background attenuation $a^- \mathbb{1}_{D_0}$ in (3.10) is constant. However, in many applications a^- varies smoothly in the domain. To account this variation, we modify (3.10) to

$$F_{\text{star}}^{\text{noisy}}[\xi_0, (\boldsymbol{\xi}, \boldsymbol{c})](x) = F_{\text{lg}}[\xi_0](x) \mathbb{1}_{D_1}(x) + \sum_{i=1}^{N_{\text{inc}}} a^+ \mathbb{1}_{D_i}(x). \tag{3.11}$$

We distinguish between the noise term ε in (2.3) and the variation in the background $F_{\lg}[\xi_0](x)$ in our model. The latter can accounts for experiment's systematic error while the former noise models the measurement error.

We need to specify a prior measure $\boldsymbol{\mu}_0$ for $(\boldsymbol{\xi}, \boldsymbol{c})$. We define $\boldsymbol{\mu}_0 = \boldsymbol{\mu}_0^1 \otimes \boldsymbol{\mu}_0^2$, where $\boldsymbol{\mu}_0^1$ is a Gaussian measure on $H^{N_{\mathrm{inc}}}$ and $\boldsymbol{\mu}_0^2$ is a measure on $R^{2N_{\mathrm{inc}}}$ (e.g., Lebesgue measure). We assume that ξ_i and c_i for $i = 1, \ldots, N_{\mathrm{inc}}$ are independent, then $\boldsymbol{\mu}_0^1$ and $\boldsymbol{\mu}_0^2$ can be factorized further into simpler measures as $\boldsymbol{\mu}_0^1 = \mu_0^{1,1} \otimes \cdots \otimes \mu_0^{1,N_{\mathrm{inc}}}$ and $\boldsymbol{\mu}_0^2 = \mu_0^{2,1} \otimes \cdots \otimes \mu_0^{2,N_{\mathrm{inc}}}$. For each $i = 1, \cdots, N_{inc}$ the random variable (ξ_i, c_i) is in the separable Hilbert space $X^i := H^i \times \mathbb{R}^2$ equipped with the norm $\|(\xi_i, c_i)\|_{X^i} := \|\xi_i\|_{H^i} + \|c_i\|_2$, and accordingly we define $(\boldsymbol{\xi}, \boldsymbol{c}) \in X$. Note that the superscript i in H^i indicates the index of the space corresponding to the random variable ξ_i. This should not be confused with the differentiability order of the Hilbert space.

4. The posterior distribution for the CT problem. In this section we derive a novel posterior distribution for the CT problem. Before doing so, we first construct the likelihood function based on the prior models introduced in Section 3. Then, we show that this likelihood function fits in the Lipschitz-Hellinger well-posedness framework in [14, 1]. Therefore, the posterior measure exists and is unique.

4.1. The Likelihood Function. Define the probability space $(X, \mathcal{B}(X), \boldsymbol{\mu}_0)$ and the image $\alpha(\xi) = F_{\mathrm{ls}}[\xi]$ or $\alpha(\boldsymbol{\xi}, \boldsymbol{c}) := F_{\mathrm{star}}[(\boldsymbol{\xi}, \boldsymbol{c})]$. In this section we only construct the likelihood and the posterior based on the star-shaped field. The formulas regarding the level set field follows immediately by setting c as a constant in the single inclusion case. We refer the reader to [18] for a detailed discussion.

Recall that the noise $\boldsymbol{\varepsilon}$ is distributed according to $\boldsymbol{\varepsilon} \sim \mathcal{N}(\boldsymbol{0}, \Sigma)$. We assume that $\boldsymbol{\varepsilon}$ is independent of $(\boldsymbol{\xi}, \boldsymbol{c})$. With the star-shape field the *negative log-likelihood* function $\Phi : X \to \mathbb{R}$ is formulated as the least squares distance

$$(4.1) \qquad \Phi\left((\boldsymbol{\xi}, \boldsymbol{c}); \boldsymbol{y}\right) = \frac{1}{2}\|\boldsymbol{y} - \mathcal{G}(\boldsymbol{\xi}, \boldsymbol{c})\|_{\Sigma}^2,$$

where $\mathcal{G} = G \circ F_{\mathrm{star}}$, and $\| \cdot \|_{\Sigma} = \|\Sigma^{-1/2} \cdot \|_2$. For the CT problem with a single inclusion, it satisfies the following conditions. We refer the reader to Appendix B for the proof.

PROPOSITION 4.1. *Let $(X, \mathcal{B}(X), \mu_0)$ be the probability space defined above. The negative log-likelihood Φ defined in (4.1) with a single inclusion, i.e. $N_{inc} = 1$, satisfies the following conditions:*
  (i) *There is a continuous function $K : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ such that for every $\rho > 0$, $(\xi_1, c_1) \in X$, and bounded observation vector $\boldsymbol{y}$ with $\|\boldsymbol{y}\|_{\Sigma} \le \rho$,*

$$(4.2) \qquad 0 \le \Phi((\xi_1, c_1), \boldsymbol{y}) \le K(\rho, \|(\xi_1, c_1)\|_X).$$

  (ii) *For a fixed observation vector $\boldsymbol{y} \in \mathbb{R}^N$, $\Phi(\cdot; \boldsymbol{y}) : X \to \mathbb{R}$ is continuous μ_0-a.s. on the probability space $(X, \mathcal{B}(X), \mu_0)$.*
  (iii) *There exists a continuous map $M : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ such that for every pair of observation vectors $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^N$ with $\|\boldsymbol{y}_1\|_{\Sigma}, \|\boldsymbol{y}_2\|_{\Sigma} \le \rho$, and every $(\xi_1, c_1) \in X$,*

$$(4.3) \qquad |\Phi((\xi_1, c_1), \boldsymbol{y}_1) - \Phi((\xi_1, c_1), \boldsymbol{y}_2)| \le M(\|(\xi_1, c_1)\|_X, \rho)\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_{\Sigma}.$$

4.2. Posterior Distribution. In this section we present Bayes' theorem to connect the prior measure with the likelihood function and construct the posterior measure $\boldsymbol{\mu^y}$.

THEOREM 4.2. *[14] Let Φ be the negative log-likelihood defined in (4.1) satisfying Proposition 4.1 and $\boldsymbol{\mu}_0$ be the prior measure defined on $(X^{N_{inc}}, \mathcal{B}(X^{N_{inc}}))$. Then there is a posterior measure $\boldsymbol{\mu}^{\boldsymbol{y}}$ absolutely continuous with respect to $\boldsymbol{\mu}_0$, i.e., $\boldsymbol{\mu}^{\boldsymbol{y}} \ll \boldsymbol{\mu}_0$, and defined through the Radon-Nikodym derivative [28]*

$$(4.4) \qquad \frac{d\boldsymbol{\mu}^{\boldsymbol{y}}}{d\boldsymbol{\mu}^0} = \frac{1}{Z} \exp\left(-\Phi((\boldsymbol{\xi}, \boldsymbol{c}); \boldsymbol{y})\right),$$

where Z is the normalization constant and for $\boldsymbol{y}$-almost surely

$$(4.5) \qquad Z := \int_X \exp\left(-\Phi((\boldsymbol{\xi}, \boldsymbol{c}); \boldsymbol{y})\right) \, \boldsymbol{\mu}_0(d(\boldsymbol{\xi}, \boldsymbol{c})) > 0.$$

We show in Appendix B that the posterior measure $\boldsymbol{\mu}^{\boldsymbol{y}}$ constructed with the star-shaped prior in a single inclusion case is well-posed. Existence and well-posedness of the posterior measure constructed with level-set prior together with a linear forward map (e.g., the Radon transform G) is discussed in [17, 1]. In the next section we utlize both priors to develop a two-stage method for identifying the location and boundaries of inclusions in an image α.

5. Two-Stage Method for Detecting Inclusion Boundaries. In Section 3 we presented two prior models that are used in constructing a posterior distribution. Subsection 4.2 shows the Bayesian formula for modelling the CT problem. In this section we provide a two-stage method in detecting the location of inclusions in an image and also estimate the boundary of the inclusions. In the first stage we use the level set prior (3.6) to construct a posterior and use the mean to obtain the approximated centers of inclusions. In the second stage we use the star-shape prior to estimate the boundaries of the inclusions.

We remind the reader that the star-shaped prior for multiple inclusions with unknown centers can be used to construct a posterior. However, constructing a sampling method for such a posterior is challenging (e.g., the Metropolis-within-Gibbs-type method can result in highly correlated samples and increased computational cost, see [45]). The first stage of the method is to ensure that we can decompose the images into regions containing a single inclusion. Remark 3.5 then guarantees that we can converge to the right solution for each individual inclusion. This decomposition also allows parallel computation with respect to the number of inclusions for exploring the posterior.

5.1. Stage 1: Estimating Centers of Inclusions. In this section we construct the posterior measure $\mu^{\boldsymbol{y}}$ in (4.4) using the Matérn covariance (3.3) and the level set map F_{ls} introduced in (3.6).

We assume that the correlation-length parameter τ and the regularity parameter γ are known. Once the covariance $Q_{\gamma,\tau}$ is constructed we discretize (3.3) using a finite differences (FD) discretization scheme with a pixel size of $2/N_s$, where N_s is the number of detector pixels. We truncate the KL expansion after N_{KL} terms. We propose two approaches to draw samples from a truncated KL expansion.

In the first approach we consider periodic boundary conditions on the box centered at the origin with the size 2-by-2. Note that $Q_{\gamma,\tau}$ is defined via its Fourier transform. Therefore, this choice of boundary conditions allows efficient sampling from $Q_{\gamma,\tau}$ using the fast Fourier transform. For a sample $\xi \sim Q_{\gamma,\tau}$ we first compute the Fourier transform $\widehat{\xi}$ of ξ, see Appendix A. We then use the inverse fast Fourier transform

(IFFT) to obtain ξ_{ex}. To Ensure that ξ is only valid in D we define $\xi := \xi_{\mathrm{ex}}|_D$. The image α is then constructed using the map $F_{\mathrm{ls}}[\xi]$ for fixed $a^+, a^- \geq 0$. This method allows efficient generation of Gaussian samples.

An alternative sampling approach is to discretize $Q_{\gamma,\tau}$, e.g. using FD methods, to obtain a covariance matrix. We can construct efficient sampling methods by computing the Cholesky factor or the principle square root [4] of this covariance matrix. Note that this computation is carried only once for a set of parameters τ and γ. The image α is then construct by directly computing the KL expansion in (3.2) and using the map $F_{\mathrm{ls}}[\xi]$ for fixed $a^+, a^- \geq 0$. In our numerical experiments we only used the first approach.

We construct the posterior measure $\mu^{\boldsymbol{y}}$ and explore it by using the preconditioned Crank-Nicolson (pCN) method [11], which is introduced in Appendix C in detail. The algorithm of Stage 1 is given in Algorithm 5.1.

Algorithm 5.1 Detecting inclusion centers

1: Construct the posterior measure $\mu^{\boldsymbol{y}}$ using the level set prior for an observation vector $\boldsymbol{y}$.
2: Draw samples $\{\xi^{(j)}\}_{j=1}^{N_{\mathrm{sample}}}$ by using pCN with Algorithm C.1.
3: Compute $\overline{\alpha} = F_{\mathrm{ls}}[\mathbb{E}\xi]$.
4: Estimate N_{inc}, $\boldsymbol{c}$ and bounding boxes $\square^i$ with $i = 1, \cdots, N_{\mathrm{inc}}$ from $\overline{\alpha}$.

To show the performance of our method in Stage 1, we consider two types of mean for the density field suggested in [18]. The first mean is computed on the space H and then push-forward to the density space using the map F_{ls}, i.e. $\overline{\alpha} = F_{\mathrm{ls}}[\mathbb{E}\xi]$. This mean conserves the piecewise constant nature of the density field. The other type is the sample mean in the density field, i.e., $\widehat{\alpha} = \mathbb{E}F_{\mathrm{ls}}[\xi]$. This mean does not construct a piece-wise constant nature of the density field but provides an uncertainty estimate on the boundaries of the inclusions.

We use $\overline{\alpha} = F_{\mathrm{ls}}[\mathbb{E}\xi]$ to estimate the approximate location of the inclusions. We can use standard matrix/image segmentation methods (e.g., [57]) to identify individual inclusions. This also gives us an estimate of the number of existing inclusions. The center of mass for each inclusion is an approximation of the center for the star-shape inclusion.

Remark 5.1. The center defined in the star-shaped inclusions is, in general, not the center of mass. However, the center is just a modelling tool to describe inclusions and it is not explicitly needed in most applications.

In Figure 2 we show an example of Stage 1. The true densities are generated randomly using the noisy star-shaped prior (3.11). We consider a parallel-beam geometry described in Section 2 with 100 equidistantly spaced angles in $[0°, 180°)$, i.e. $\theta_{\max} = 180°$ and $N_\theta = 100$. The detector length is set to 2 and centered at the origin, thus, the size of a detector pixel is $2/N_s$ with $N_s = 100$. The number of inclusions is unknown.

Table 1 lists the estimated centers of mass in Stage 1 compared with the star-shaped centers used to generate Figures 2a and 2d. Note that the estimated centers of mass are far from the centers assigned to the star-shaped inclusions, but they are accurate estimates of the true centers of mass. We report that the method consistently had similar results for other examples tested by the authors.

We notice that if assumption (III) in Subsection 3.2.2 is satisfied the inclusions are well approximated. In the case of Figure 2d we see that $F_{\mathrm{ls}}[\mathbb{E}\xi]$ falsely identifies

(a) True density $\alpha(\boldsymbol{\xi})$ (b) $F_{\mathrm{ls}}(\mathbb{E}\boldsymbol{\xi})$ (c) $\mathbb{E}F_{\mathrm{ls}}(\boldsymbol{\xi})$

(d) True density $\alpha(\boldsymbol{\xi})$ (e) $F_{\mathrm{ls}}(\mathbb{E}\boldsymbol{\xi})$ (f) $\mathbb{E}F_{\mathrm{ls}}(\boldsymbol{\xi})$
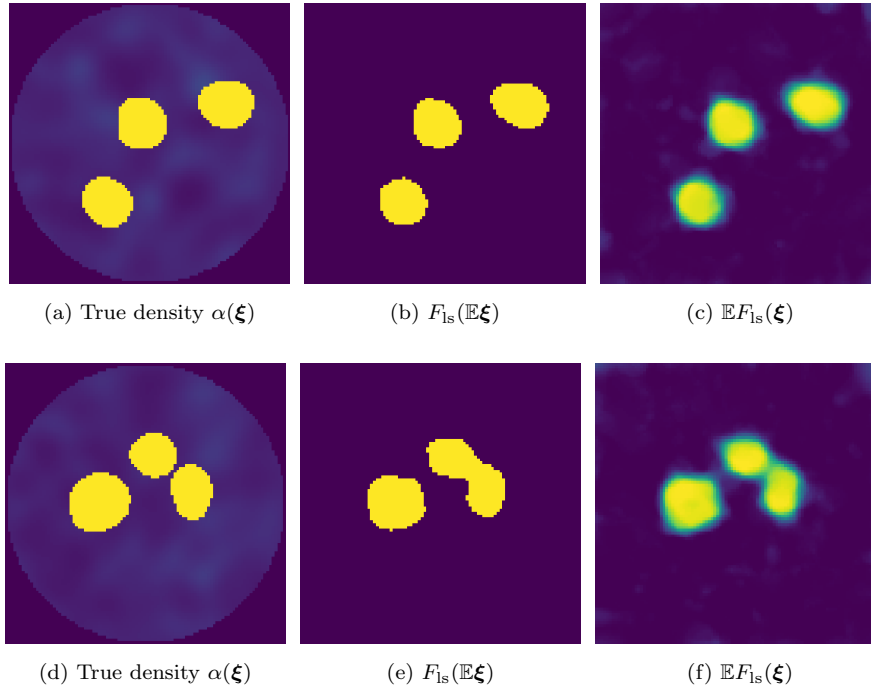
FIG. 2. *Example of Stage 1 in identifying the location of the inclusions. Each row corresponds to the same experiment.*

TABLE 1
Prediction of centers

	inclusion 1	inclusion 2	inclusion 3
star-shaped centers	$(-0.298, 0.452)$	$(-0.157, -0.143)$	$(0.382, -0.252)$
exact centers of mass	$(-0.279, 0.548)$	$(-0.139, -0.058)$	$(0.424, -0.305)$
estimated centers of mass	$(-0.272, 0.540)$	$(-0.137, -0.059)$	$(0.424, -0.300)$

the two close inclusions as one. However, Figure 2f suggests that there is uncertainty in detecting the inclusions and this can be interpreted as the violation of assumption (III).

We assign a bounding box $\square^i$ with $i = 1, \dots, N_{\mathrm{inc}}$ for each inclusion. We first find the leftmost pixel of the ith inclusion in $\overline{\alpha}$, then set the leftmost position of $\square^i$ by subtracting $d_{\min}^D/2$ pixels in order to ensure assumption (III) in Subsection 3.2.2 is satisfied. We define the right, top and bottom bounds of $\square^i$ in a similar manner. These bounding boxes are essential for Stage 2, since they decompose the domain into regions with a single inclusion each. Remark 3.5 ensures that there is a well-defined posterior distribution on such regions for the star-shaped prior.

Remark 5.2. The main purpose of Stage 1 is to estimate the bounding boxes and decompose the image into regions with a single inclusions. This can be achieved with other methods, e.g., Bayesian methods with MRF-type priors [36] or Besov-type priors [13, 32]. Although the sampling method used in Stage 1 is robust to discretization

refinement, the forward mapping G is still dimension-dependent. We can speed up Stage 1 by truncating the KL expansion with only $10 \sim 20$ modes to evaluate the bounding boxes. In our experiments the computational costs of Stage 1 is negligible compared to Stage 2.

Remark 5.3. We emphasize that the performance of Stage 2 depends on the correct detection of number of inclusions in Stage 1. If Stage 1 of the approach fails to detect the correct number of bounding boxes, the true inclusions will not be supported by the posterior in Stage 2.

5.2. Stage 2: Estimation and Uncertainty Quantification of the Boundaries of Inclusions. In this stage we construct the posterior measure μ^y using the star-shaped prior. The number of inclusions N_{inc} is estimated in Stage 1 together with the bounding boxes that decompose the domain into regions with a single inclusion each. Note that in Stage 2 we only reconstruct the boundaries of the inlcusions provided by Stage 1. We let ξ_i, for $i = 1, \ldots, N_{\mathrm{inc}}$ be the H-valued random variables with $H = \mathbf{T}([0, 2\pi))$.

We assume that the correlation parameter $\tau = 1$ and γ is known in (3.3). This choice of τ indicates that points on the boundary are highly correlated. Since ξ_i models the boundary of an inclusion, we may use the periodic boundary conditions to construct $Q_{\gamma,1}$. Furthermore, since ξ_i is 1D, we can assemble them directly from (3.4). Furthermore, we assume m in (3.2) is a known constant for all ξ_i. The centers of mass estimated in Stage 1 are used as the centers of the star-shaped inclusions.

The posterior measure μ^y is constructed following Subsection 4.2. We draw samples from the posterior distribution using the Gibbs sampling method, which is introduced in Appendix C, to estimate the two types of mean for the posterior, i.e. $\overline{\alpha}_{\mathbb{E}c} = F_{\mathrm{star}}[(\mathbb{E}\xi, \mathbb{E}c)]$ and $\widehat{\alpha}_{\mathbb{E}c} = \mathbb{E}F_{\mathrm{star}}[(\xi, c)]$, introduced in Subsection 5.1. Here, the subscript $\mathbb{E}c$ means that the expected image is drawn around the expected centers.

We consider the highest posterior density (HPD) credibility interval [22] to estimate the uncertainties in detecting the boundaries of the inclusions. A 95% HPD interval for an 1D real-valued random variable with probability density $\pi(x)$ is defined by
$$I_{\mathrm{HPD}} = \{x \in \mathbb{R} | \pi(x) > c_{\mathrm{HPD}}\}$$
where $c_{\mathrm{HPD}} > 0$ is the largest number such that

(5.1)
$$\int_{\{x \in \mathbb{R} | \pi(x) > c_{\mathrm{HPD}}\}} \pi(x) \ dx = 0.95,$$

i.e., the smallest interval with the highest credibility. Note that there are other credibility intervals that can be used, e.g., the commonly used equal-tailed credibility interval [22]. We choose the HPD credibility interval since it reveals the asymmetry and multi-modal properties of the posterior. Furthermore, it contains the maximum a posterior (MAP) estimate [22].

In Algorithm 5.2, we summarize how to use the HPD to quantify the uncertainties in the estimated boundaries. Note that we compute the HPD intervals only with respect to c, since we find that in the context of the experiments in this paper the center of the star-shaped inclusion reveals the full modality of the posterior. Furthermore, the knowledge of the exact location of the center results in a posterior with a single mode. The results for the latter is omitted for brevity.

In the last step of Algorithm 5.2 we quantify the local uncertainties for each angle ϑ, then we construct a radial credibility band by interpolating the radial HPD intervals

---

Algorithm 5.2 Visualizing the uncertainty in the estimated boundaries

---

1: Input N_{inc}, $\boldsymbol{c}$ and $\square^i$ for $i = 1, \cdots, N_{\text{inc}}$ obtained in Stage 1 by using Algorithm 5.1.

2: **for** $i = 1, \cdots, N_{\text{inc}}$ **do**

3:     Construct the posterior measure $\mu^{\boldsymbol{y}}$ using the star-shaped prior with a single inclusion.

4:     Draw samples $\{(\xi_i^{(j)}, c_i^{(j)})\}_{j=1}^{N_{\text{sample}}}$ by using the Gibbs sampling method with Algorithm C.2.

5:     Using the samples of $\{c_i^{(j)}\}_{j=1}^{N_{\text{sample}}}$, compute the HPD intervals I_{HPD}^i, which includes N_{modes}^i disjoint subintervals, $\{I_{\text{HPD}}^{i,k}\}_{k=1}^{N_{\text{modes}}^i}$, according to N_{modes}^i modes in the posterior.

6:     Compute the mean $\{(\bar{\xi}_i^k, \bar{c}_i^k)\}$ of the samples $\{(\xi_i^{(j)}, c_i^{(j)})\}$ with $j \in I_{\text{HPD}}^{i,k}$ for $k = 1, \cdots, N_{\text{modes}}^i$.

7:     Collect the HPD intervals based on the samples of $\{\xi_i^{(j)}(\vartheta)\}_{j \in I_{\text{HPD}}^{i,k}}$ for each angle ϑ to represent the uncertainty in $\bar{\xi}_i^k(\vartheta)$.

8: **end for**

---

for each inclusion. This estimate provides a local uncertainty quantification of the shape of the inclusion. In addition, we also provide the global variance $\mathbb{E}\|\xi_i - \mathbb{E}\xi_i\|_{H^i}^2$ as a global uncertainty estimator for each inclusion.

Remark 5.4. The samples from the radial random variables ξ_i are generally coupled with the samples of the centers $c_i^{(j)}$, due to the structure of the the posterior. However, by ignoring the center component of the samples we can approximately achieve independent samples of ξ_i. We plot the credibility band with respect to $\bar{c}_i^k$, for $i = 1, \ldots, N_{\text{inc}}$. This is only for illustration purposes and should not be confused with samples of the random variable $\xi|(c_i = \bar{c}_i^k)$.

6. Numerical Results. In this section we test the presented method on synthetic images with inclusions. The test images contain single and multiple inclusions with boundaries with various regularities. We also test the method on a tomographic X-ray data of a lotus root filled with attenuating objects.

6.1. Generated Inclusions. In this section, we consider synthetic images with inclusions. We construct an image (attenuation field α) using the star-shaped prior. We then apply our two-stage method to estimate the center of inclusions and quantify the uncertainty in estimating the boundaries.

We consider the domain D to be the unit disk and construct the random attenuation fields using the noisy star-shaped prior (3.11)

$$(6.1) \qquad \alpha^{\text{noisy}}(x) = F_{\text{star}}^{\text{noisy}}[\xi_0, (\boldsymbol{\xi}, \boldsymbol{c})](x),$$

where foreground and background attenuation is set to $a^+ = 1$ and $a^- = 0.1$, respectively. Recall that ξ_0 is the random variable that controls the fluctuations in the background, while $\boldsymbol{\xi}$ determines the boundaries of the inclusions. We draw samples from the noisy star-shaped density by first randomly choosing a center in the unit disk and then taking $\xi_0 \sim \mathcal{N}(0, Q_{\gamma^0, \tau^0})$ with $\gamma^0 = 2.5$, $\tau^0 = 50$ as the parameters for the background. We use a sampling-and-elimination method to ensure that the inclusion appears inside the unit disk. We truncate the KL expansion after 100 terms.

In our numerical tests, we use the scikit-image package [53] (version 0.19.0.dev0), a Python image processing toolkit, to carry out the forward calculations, i.e., 2D Radon transforms of images. The package uses the parallel beam geometry described in Section 2. We set $\theta_{\max} = 180°$ and $N_\theta = 100$ for a full set of images. The detector length is set to 2 and centered at the origin. This means that $L_\theta^\perp$ is centered at the origin and $s \in [-1, 1]$.

In Stage 2 (Algorithm C.2) of our method, we first run a warm-up phase of 20 Gibbs iterations with $N_{\mathrm{pCN}} = N_{\mathrm{MH}} = 500$. We tune the parameters in the proposal such that the loops within the Gibbs sampler provide an acceptance rate of 15% to 25%. Then, we fix the parameters, set $N_{\mathrm{pCN}} = N_{\mathrm{MH}} = 20$, and calculate 2×10^4 samples from $\boldsymbol{\mu^y}$. We note that designing efficient MCMC methods for such posteriors is an area of active research, see, e.g., [17, 18, 45].

6.1.1. Single Inclusion. In this section we consider attenuation fields α with a single inclusion. We investigate the effect of the observation noise on the estimations in our method. The results in this section are illustrated on randomly generated test examples. The authors found similar results for all other test examples with same setting.

To create ground truth images we use the noisy start-shaped prior (3.11), and set the inclusion $\xi_1 \sim \mathcal{N}(0, Q_{\gamma^1, \tau^1})$ with $\gamma^1 \in \{2, 3\}$ and $\tau^1 = 1$. We define the *signal to noise ratio (SNR)* in an observation to be

$$(6.2) \qquad \mathrm{SNR} = \frac{\|\boldsymbol{y}\|_2}{\|\boldsymbol{\varepsilon}\|_2},$$

and define the *noise level* percentage in an observation to be $1/\mathrm{SNR} \times 100$.

In Figure 3 we show the test problem with $\gamma^1 = 3$ and noise level 1%. Here, the true star-shaped center of the inclusion is $c = (0.154, -0.215)$. We provide some samples from the prior distribution and the posterior distribution according to Stage 2 of our method in Figures 4a and 4b. In addition, in Figures 4c to 4e we illustrate the performance of our method. By using HPD we are able to quantify the uncertainties of the boundaries according to each mode, which are given in Figures 4c and 4d. For each mode although the estimated center does not match the true star-shaped center, the boundary of the inclusion is still well reconstructed. The reason is that the star-shaped center is not uniquely defined and is not consistent with the center of the mass that is used in our method. Moreover, we would like to remind that the main goal of our method is to reconstruct the boundary of the inclusion. Figure 4e shows the posterior mean with the 99% HPD interval.

The posterior mean is not a good representation of the boundary due to the multi-modal nature of the posterior. This is clearly seen in Figure 4e. Therefore, in the rest of this section we let $(\bar{\xi}_i^1, \bar{c}_i^1)$, computed from the first HPD interval $I_{\mathrm{HPD}}^{i,1}$, represent the boundary. The choice of the first HPD is arbitrary. The authors confirm that in the experiments in this work the quality of the reconstructed boundaries $(\bar{\xi}_i^k, \bar{c}_i^k)$ is comparable for all HPD intervals $I_{\mathrm{HPD}}^{i,k}$.

In Figure 5 we present the diagnostics for the Gibbs sampler in Algorithm C.2 for 4 MCMC chains. The first chain is constructed by using the mass center obtained in Stage 1 as the initial center for star-shaped inclusion. The other chains use a uniformly distributed random initial guesses from the bounding box. In all our numerical results, we only focus on samples in I_{HPD}^1. We notice that the estimated boundaries detected for other HPD intervals are comparable with the ones in I_{HPD}^1 in all our experiments. Figure 5a shows the auto-correlation function (ACF) of these samples for each chain
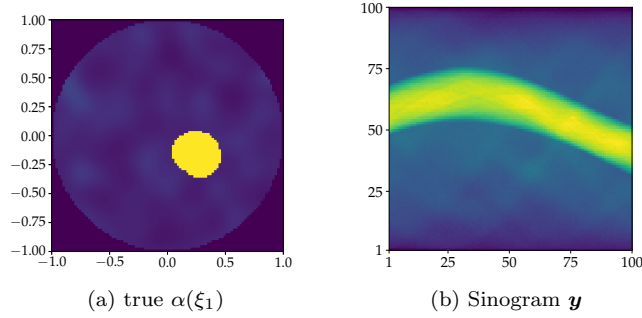
(a) true $\alpha(\xi_1)$

(b) Sinogram $\boldsymbol{y}$

FIG. 3. *Test problem for the single-inclusion case with regularity parameter $\gamma^1 = 3$ and noise level 1%.*

(a) prior samples

(b) posterior samples

(c) 1st posterior mode

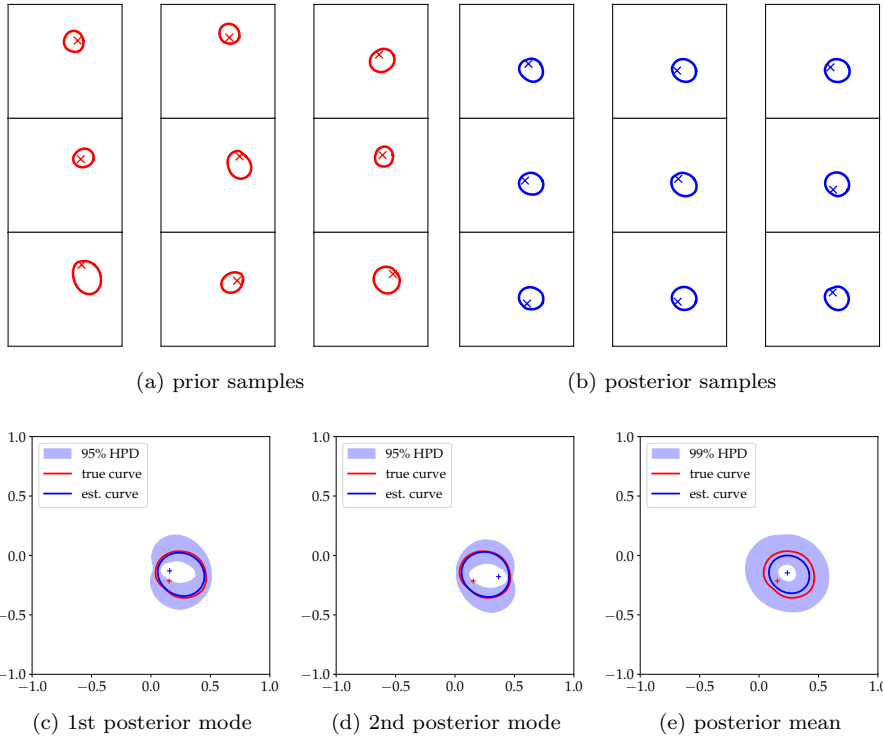(d) 2nd posterior mode

(e) posterior mean

FIG. 4. *Results of our method to the single-inclusion test problem shown in Figure 3. In Figures 4c to 4e, the red curve and the red cross represent the true boundary and the true star-shaped center. The blue cross and the blue curve represent $\mathbb{E}c_1$ and $F_{star}[(\mathbb{E}\xi_1, \mathbb{E}c_1)]$, respectively.*

as well as the mean of all chains. We can see that ACF drops to 0 after a reasonable number of samples.

We compute the effective sample size (ESS) of the MCMC samples following [54]. The ESS and the auto-correlation function is computed using the ArviZ [31] package in Python. We refer the reader to [38, 54] for more detail on ESS.

The ESS for all chains are between 200 and 1000 samples. In Figure 5b we can see that the mean curves of samples in I^1_{HPD} for all chains match. This is a necessary condition for the MCMC method to be converging [23].
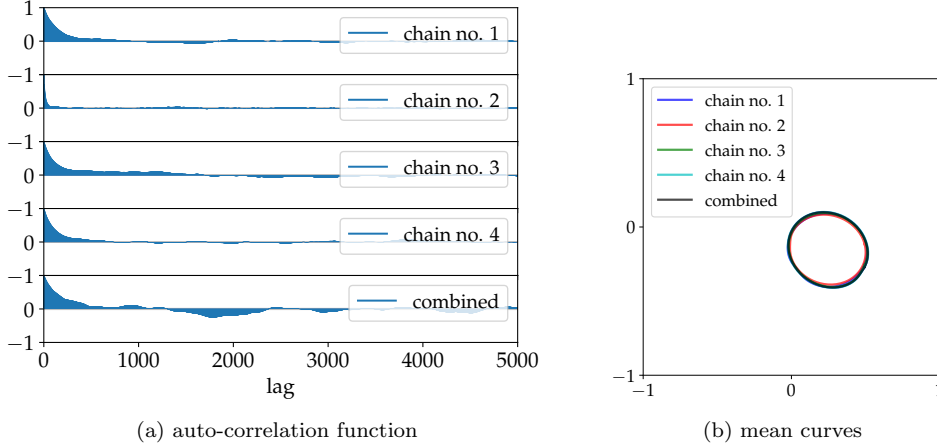

(a) auto-correlation function   (b) mean curves

FIG. 5. *Performance of the Gibbs sampler in Algorithm C.2. Here we only consider the samples in I^1_{HPD}.*

To illustrate the effect of the noise level, we apply our method to reconstruct the boundary from sinograms including 2.5% and 5% noise, respectively. The true α is the same as shown in Figure 3a, and the results are show in Figure 6. We can see that our method can provide a good estimated boundary of the inclusion, although the estimated star-shaped center increases the uncertainty in estimating the boundary. It is clear that the accuracy of estimation increases as the noise level is reduced. In Figure 6 the 99% HPD band represents the uncertainty in estimating the boundary of the inclusion. We notice that the HPD band in all test cases completely covers the true outline of the inclusions. Furthermore, it is significantly smaller for the test case with smaller noise level. The HPD band does not provide a uniform uncertainty around the boundary, as it is suggested in Figure 6. The variation in the uncertainty around the boundary is a result of the estimated center and the choice of the Whittle-Matérn covariance, i.e., the prior parameters.

To study the global uncertainty estimation, in Table 2 we list the global variances $\mathbb{E}\|\xi_1 - \mathbb{E}\xi_1\|^2_H$ for different noise levels. We notice that the relative difference in variance among the test cases confirms the uncertainty that is visualized in Figure 6. This confirms our intuition that larger observation noise results in larger uncertainty in our estimation.

TABLE 2
Estimation of the global variances.

	Figure 4c	Figure 6a	Figure 6b
$\mathbb{E}\|\xi_1 - \mathbb{E}\xi_1\|^2_H$	0.048	0.323	0.334

In the previous tests, we assume that we know the correct value of γ^1. Now we will discuss the importance of the setting of γ^1. We generate the ground truth image using $\gamma^1 = 2$, and compare the results by applying our method with $\gamma^1 = 2$ and 3, which are shown in Figure 7.

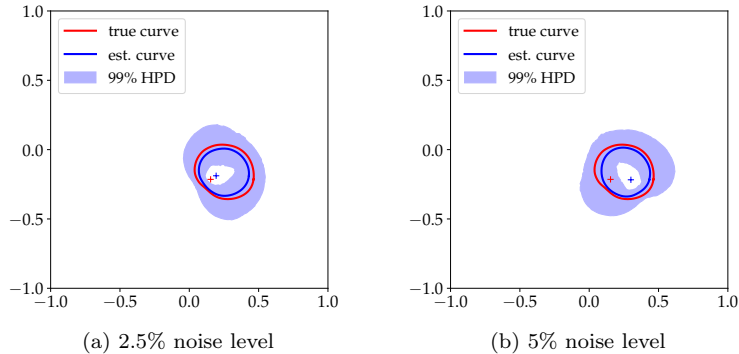(a) 2.5% noise level　　　　　　　(b) 5% noise level

FIG. 6. *Illustration of the performance of our method with respect to different noise levels. The red cross represents the true star-shaped center, and the blue cross represents $\mathbb{E}c_1$ from Stage 1.*

With the correct γ^1, we can see that the estimated curve in Figure 7c accurately represents the true boundary. The location of some of the larger indentations in boundary is also represented in the estimated boundary. For this example, we provide some of the prior and posterior samples in Figures 7a and 7b, respectively.

Figure 7d shows the results by using $\gamma^1 = 3$ in our method. The reconstructed curve (the blue curve) is smoother than the true curve due to larger regularity parameter γ^1. However, it fairly estimates the overall shape and orientation of the inclusion. We notice that the HPD band around the boundary is wider than the one in Figure 7c. A wrong choice of the prior parameters may be the reason for larger uncertainty in the boundary.

In Figure 7 we also give the global variance in both cases. The relative differences between the variances verifies the uncertainty bands indicated in the figures. We note that the global variance in Figure 7d is significantly larger than the one in Figure 7c. This can potentially be used to identify the correct regularity in case γ^1 is unknown. One approach can be to minimize the global variance over γ^1. This is left as future work.

6.1.2. Multiple Inclusions. In this section we consider attenuation fields α that contain multiple inclusions. We construct the ground truth density fields using the noisy star-shaped prior given in (3.11) with $N_{\text{inc}} = 3$. We let $\xi_0 \sim \mathcal{N}(0, Q_{\gamma^0, \tau^0})$ with $\gamma^0 = 2.5$ and $\tau^0 = 50$, and set $\xi_i \sim \mathcal{N}(0, Q_{\gamma^i, \tau^i})$ with $\gamma^i = 3$ and $\tau^i = 1$ for $i = 1, \ldots, N_{\text{inc}}$. We expect the inclusions to contain smooth boundaries. Drawing samples from $\alpha^{\text{noisy}}(x)$ requires sampling-and-elimination step to ensure that inclusions are inside the domain and they do not collide. We truncate the KL expansion for ξ_0 after 200 terms and for ξ_i, $i = 1, \ldots, N_{\text{inc}}$, after 100 terms. In this test, we use the noise level 1%. The CT scan geometry is the same as in the test problem in Subsection 6.1. In Figure 8, we show a noisy sinogram as well as the true image. In our method, Stage 1 serves for finding the bound boxe and the center to each inclusion. Then, in Stage 2 the Gibbs samplers according to Algorithm 5.2 are run parallel for each inclusion. We collect 10^4 samples from the posterior distribution with an addition of 10^3 samples in the burn-in stage.

The estimated boundaries are presented in Figure 8c. We notice that the centers of all inclusions are estimated with high accuracy. Furthermore, the posterior mean for the boundary of the inclusion provides a precise estimate for the true boundaries.
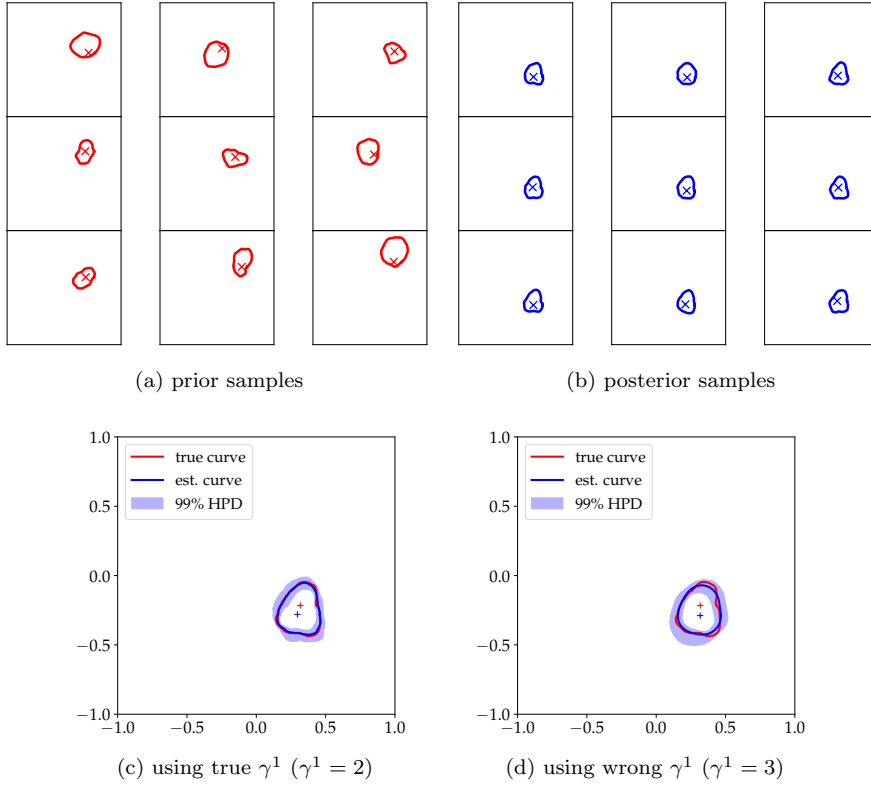
(a) prior samples   (b) posterior samples

(c) using true γ^1 ($\gamma^1 = 2$)   (d) using wrong γ^1 ($\gamma^1 = 3$)

FIG. 7. *Illustration of the performance of our method for for different values of the regularity parameter γ^1. The global variances, $\mathbb{E}\|\xi_1 - \mathbb{E}\xi_1\|_H^2$, in the estimations for Figures 7c and 7d are 0.0561 and 0.198, respectively.*
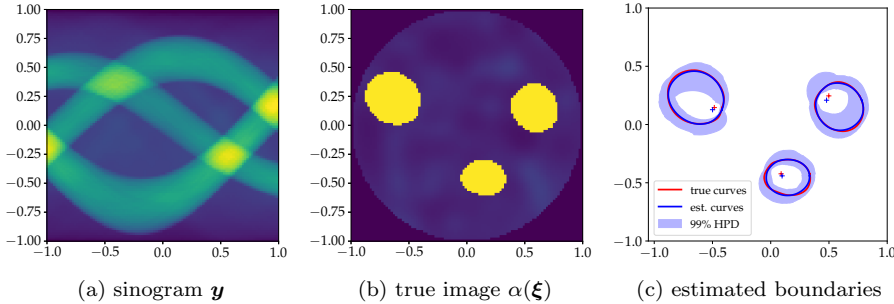

(a) sinogram $\boldsymbol{y}$ (b) true image $\alpha(\boldsymbol{\xi})$ (c) estimated boundaries

FIG. 8. *Illustration of the uncertainty in detecting the boundary for an image with multiple inclusions.*

Note that the true inclusions have different size, shape and orientation (e.g., one of the inclusions is more elongated). Although the prior distribution is set identical for all inclusions, the method is able to find such subtle differences .

We list the global variance of ξ_i, $i = 1, 2, 3$, in Table 3. The HPD band illustrated in Figure 8c is compatible with the estimated global variances. We report that we find similar results for images with at most 5 inclusions. Given a good initial guess

provided from Stage 1, we expect the method to perform as well for any number of inclusions.

	Figure 8c		
Inclusion number	$i = 1$	$i = 2$	$i = 3$
True centers	$(-0.484, 0.146)$	$(0.499, 0.246)$	$(0.090, -0.423)$
Estimated centers	$(-0.500, 0.128)$	$(0.480, 0.209)$	$(0.101, -0.439)$
Variance in ξ_i	0.083	0.050	0.065

6.1.3. Sparse and limited angle imaging. In this section we estimate the boundary of a single inclusion in a sparse and limited angle imaging configuration. We compare measurement geometries where the number of angles is $N_\theta = 10$ (compared to $N_\theta = 100$ in Subsections 6.1.1 and 6.1.2) and we set the angular range to be $[0, \theta_{\max})$ with $\theta_{\max} \in \{180°, 90°, 45°\}$. The interval $[0, \theta_{\max})$ is uniformly discretized into imaging angles θ_j, $j = 1, \ldots, N_\theta$. The regularity of the inclusion is chosen to be $\gamma^1 = 2.5$. The rest of the parameters for the forward problem and the sampling methods are identical to those in Subsection 6.1.1.
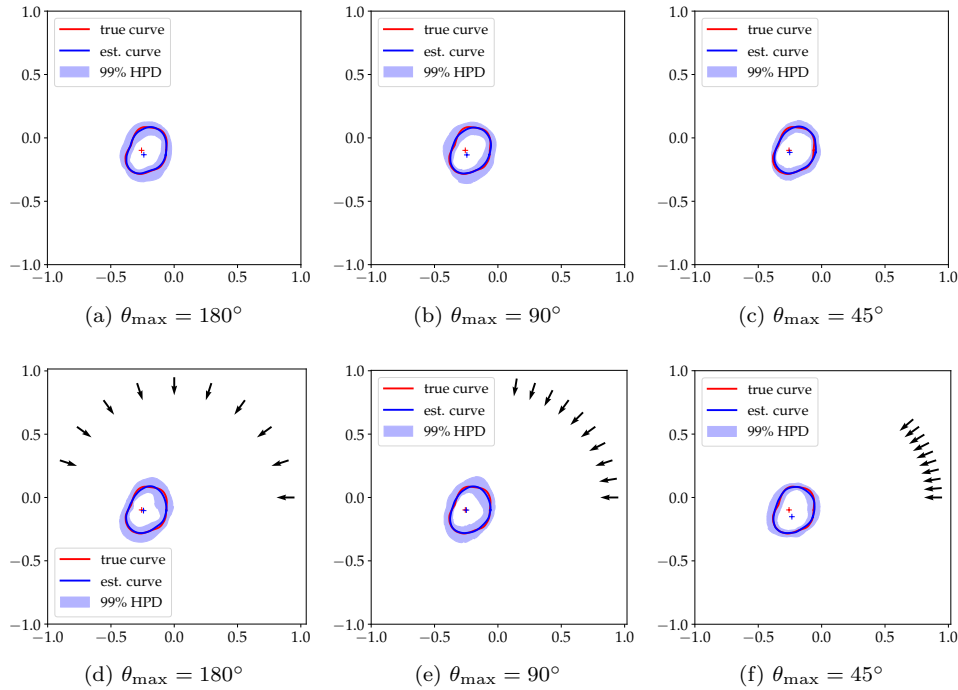

FIG. 9. *Example of estimating the boundary of a single inclusion with sparse and limited angle imaging. First row: a full set of angles with $N_\theta = 100$; second row: sparse angles with $N_\theta = 10$ as illustrated by the arrows.*

To understand the effect of the sparsity and imaging angles on estimations, we compare the results by using our method to the case where a full set of angles ($N_\theta =$

100) is available with the ones from sparse and limited angle cases, see Figure 9. We notice that the uncertainty in the estimation of the boundary of the inclusion is significantly larger than the previous test cases. As we move from left to right, we cannot notice a qualitative difference in the amount of uncertainty in the estimated boundary. We notice the increase in the width of the uncertainty band as we move from the top figures to the bottom figures. Therefore, the method is significantly more sensitive to the number of observations than the number of imaging angle.

We remark that in all test cases in this section the estimated boundary $(\bar{\xi}_i^1, \bar{c}_i^1)$ provides excellent approximations to the true boundary and the true center of the star-shaped inclusion. Furthermore, we report that the overall behavior of the global variances were comparable with the uncertainty presented in Figure 9 as seen in Table 4.

TABLE 4
Estimation of the global variance.

	Figure 9a	Figure 9b	Figure 9c	Figure 9d	Figure 9e	Figure 9f
$\mathbb{E}\|\xi_1 - \mathbb{E}\xi_1\|_H^2$	0.0538	0.0485	0.0312	0.0571	0.0493	0.0320

6.2. Lotus Root. In this section we apply our method to tomographic X-ray data of a lotus root filled with attenuating objects from the open data sets in [5, 6]. The data is taken over 360 angles around the object with the fan-beam geometry. We apply filtered back-projection (FBP) to reconstruct the image α, which is shown in Figure 10a and considered as the ground truth. The target of the test is to reconstruct the boundary of a piece of circular chalk (made of calcium) placed inside the lotus root.

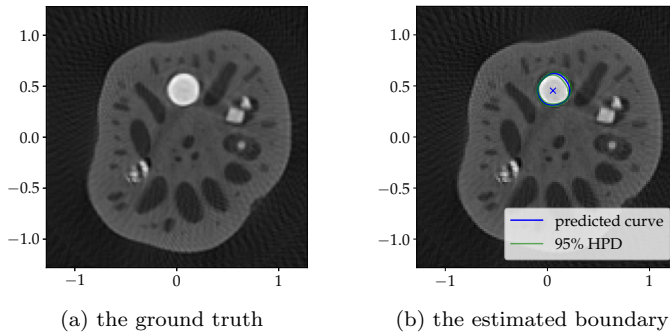(a) the ground truth   (b) the estimated boundary

FIG. 10. *Prediction of the boundary of a circular piece of chalk in a 2D slice of a lotus root.*

To test the performance of our method for real data, we down-sample the data with 100 equidistantly spaced angles in $[0, 180°)$ as $\boldsymbol{y}$. We assume that the noise level is 1%, and set the foreground attenuation to be $a^+ = 0.025$ and the background attenuation to be $a^- = 0.001$. In Stage 1 of our method we set the correlation length for the Matérn-Whittle field to be $\ell = 0.02$ and the regularity parameter $\gamma = 3$. The rest of the modelling parameters are chosen identical to the previous tests.

The estimated curve for the circular chalk is presented in Figure 10b. Note that our method does not provide a reconstructed image for the whole domain. Here, we keep the ground truth behind the curve estimation as a visualization aid. It is

clear that the curve precisely follows the outline of the circular chalk. Furthermore, the estimated center appears approximately at the center of the inclusion, which is consistent with the fact that the center of a star-shaped inclusion with a perfect circular boundary is the same as the center of the mass. The HPD band around the boundary of the inclusion suggests that the reconstruction is very accurate.

7. Conclusions. This work presents an infinite dimensional Bayesian framework for the X-ray CT problem for goal-oriented estimation and uncertainty quantification of inclusion boundaries. The proposed method reconstructs the boundaries of inclusions with constant attenuation that can be represented as a star-shaped inclusion on a smoothly varying background. Furthermore, we provide a HPD band around the boundary to quantify the uncertainty of the reconstruction. The method is carried out in two stages. In Stage 1 we identify approximate locations of the inclusions by sampling the posterior constructed with the level set prior. This stage decomposes the image into regions with a single inclusion. Stage 2 comprises sampling from the posterior distribution constructed by using the star-shaped prior. The decomposition of the image in Stage 1 guarantees the well-posedness of the reconstruction problem.

The numerical results show that our method reconstructs the boundaries of the inclusions accurately and provides a reliable tool to quantify the uncertainty in the prediction. Furthermore, the method consistently performs well in detecting inclusions in challenging X-ray CT scenarios (e.g., for sparse and limited angle imaging). Our results from applying the method to a real data, in the form of X-ray measurements of a lotus root filled with a circular pieces of chalk, suggest that this method can be extended to real world applications.

Appendix A. More on The Matérn-Whittle Covariance. Recall $D \subset \mathbb{R}^2$ is a bounded region with Lipschitz boundary. We take $H = L^2(D)$. The covariance function of two points $x, y \in D$ for the Matérn-Whittle distribution [28, 42] is given by

$$(A.1) \qquad q_{\sigma,\nu,\ell}(x_1, x_2) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|x_1 - x_2\|_2}{\ell} \right)^\nu K_\nu \left(\frac{\|x_1 - x_2\|_2}{\ell} \right).$$

Here, $\| \cdot \|_2$ is the Euclidean norm, $\ell > 0$ is the spatial correlation length, $\nu > 0$ is the smoothness parameter, $\sigma^2 > 0$ is the variance of the value of the field (amplitude scale). Furthermore, K_ν is the modified Bessel function of the second kind of order ν [34, 42].

For $\nu = 1/2$ (A.1) reduces to the exponential covariance $q_{\sigma,1/2,\ell} = \sigma^2 \exp(-d/\ell)$, with $d = |x_1 - x_2|$. For larger ν the smoothness of $q_{\sigma,\nu,\ell}$ increases. One way to see this is that for $\nu = 1/2 + p$ with $p \in \mathbb{N}^+$ (A.1) can be written as a product of an exponential with a polynomial of order p [2, 42]. Therefore, larger p contributes to higher regularity. As $p \to \infty$ the polynomial tends to an exponential function and $q_{\sigma,\nu,\ell}$ converges to the squared exponential (Gaussian) covariance function $q_{\sigma,\infty,\ell} = \sigma^2 \exp(-d^2/(2\ell^2))$.

We can construct a discrete density function by discretizing D and computing (A.1) for each pair of points. However, this approach gives in a full covariance matrix for large correlation length. Inversion of such covariance matrices is challenging for inverse problem applications [34, 44].

An alternative approach is to formulate the covariance as a differential operator of a stochastic partial differential equation (SPDE). A detailed discussion on how to construct this SPDE is beyond the scope of this paper. Below we present a brief sketch of this construction and refer the reader to [44, 62] for a detailed discussion.

The main idea is to construct an H-valued Gaussian random variable ξ from its Fourier expansion. Let ψ be white noise on D, i.e. ψ has zero mean with the covariance operator being the Dirac's delta function $q(x_1, x_2) = \delta(\|x_1 - x_2\|_2)$. Subsequently, all Fourier modes in $\widehat{\psi}$ will be present in the Fourier transform of ψ. Furthermore, let $S(w)$ be the Fourier transform of $q_{\sigma,\nu,\ell}/\sigma^2$, known as the *power spectrum*. Now we can define ξ through its Fourier transformation by rescaling $\widehat{\psi}$ with $S(w)$ as

$$(A.2) \qquad\qquad \widehat{\xi} := \sigma\sqrt{S(w)}\widehat{\psi}.$$

By definition, ξ is a Gaussian random variable distributed according to the covariance function $q_{\sigma,\nu,\ell}$. We recover ξ by applying the inverse Fourier transform to (A.2) and obtain

$$(A.3) \qquad \frac{1}{\sqrt{b(\nu)\ell^2}}(I - \ell^2\Delta)^{(\nu+1)/2}\xi = \psi, \qquad b(\nu) = \sigma^2\frac{4\pi\Gamma(\nu+1)}{\Gamma(\nu)}.$$

Here Δ is the Laplace operator and the covariance operator of ξ corresponding to the covariance function $q_{\sigma,\nu,\ell}$ is given by

$$(A.4) \qquad\qquad Q_{\sigma,\nu,\ell} = b\ell^2(I - \ell^2\Delta)^{-\nu-1}.$$

A simplification of (A.4) is presented in [17, 62] which takes the form

$$(A.5) \qquad\qquad Q_{\gamma,\tau} = (\tau^2 I - \Delta)^{-\gamma}.$$

Here, $\tau = 1/\ell > 0$ controls the correlation length and $\gamma = \nu + 1$ is the smoothness parameter (see [34] for more detail). For the covariance operator (3.3) to be well defined we need to impose proper boundary conditions. See [44] for more details on types of boundary conditions.

Appendix B. Existence and Well-posedness of the posterior measure.
In this section we show that the CT problem introduced in Sections 4 and 5 is well-posed. We assume that the attenuation field α is bounded and strictly positive, i.e., there exist $\alpha^+, \alpha^- \in \mathbb{R}^+$ such that $\alpha^- < \alpha < \alpha^+$, and thus $\alpha \in L^\infty(D)$. We use $\mathcal{S}(D)$ to denote the space of all such attenuation fields α. Recall that for an attenuation field with a single inclusion, F_{star} maps functions from X to $\mathcal{S}(D)$. Here, $X = H \times \mathbb{R}^2$ forms a separable Hilbert space. The following proposition from [18] shows that F_{star} is a continuous map. We refer the reader to [18] for the full proof.

PROPOSITION B.1. *[18] Let $F_{star} : X \to \mathcal{S}(D)$ be the star-shaped map and let $\{\xi_1^\epsilon\}_{\epsilon>0}$ and $\{c_1^\epsilon\}_{\epsilon>0}$ be a sequence of functions in H and a sequence of points in D, respectively, such that $\|\xi_1 - \xi_1^\epsilon\|_\infty \to 0$ and $\|c_1 - c_1^\epsilon\|_2 \to 0$. Then, we have $F_{star}[(\xi_1^\epsilon, c_1^\epsilon)] \to F_{star}[(\xi_1, c_1)]$ in measure.*

Let $\mathcal{A}$ be the Borel σ-algebra constructed on X with respect to the norm

$$(B.1) \qquad\qquad \|(\xi_1, c_1)\|_X := \|\xi_1\|_H + \|c_1\|_2.$$

To define a probability measure on X we assume that ξ_1 and c_1 are independent random variables such that their joint probability measure takes the form $\mu_0 = \mu_0^1 \otimes \mu_0^2$,

with $\mu_0^1 = \mathcal{N}(0, \mathcal{C})$ a Gaussian measure on H and μ_0^2 the Lebesgue measure on D. Now the triplet $(X, \mathcal{A}, \mu_0)$ forms a probability space. Furthermore, We define the probability measure on $\mathcal{S}(D)$ to be the push-forward measure $\mu \circ F_{\mathrm{star}}^{-1}$.

To show Lipschitz-Hellinger well-posedness of the CT inverse problem with the star-shaped prior we follow the framework in [1]. The structure of the proofs follows [18] closely where we modify the proofs for a forward model constructed with the Radon transform. We first need to prove that the likelihood function constructed with the Radon transform is bounded, see Proposition 4.1. We provide the proof here.

Proof of Proposition 4.1.

(i) The Radon transform G is a bounded linear operator, see Chapters 6.2 and 6.6 in [26]. Therefore, it is continuous and we can find a constant $C > 0$ such that for $\alpha \in \mathcal{S}(D)$ we have

(B.2)
$$\|G\alpha\|_2 \leq C\|\alpha\|_{L^\infty(D)} = C\alpha^+.$$

Representing α by the star-shaped mapping defined in Equation (3.10), we have $\|F_{\mathrm{star}}((\xi_1, c_1))\|_{L^\infty(D)} \leq \alpha^+$. It follows

(B.3)
$$\begin{aligned}
\Phi((\xi_1, c_1), \boldsymbol{y}) &\leq \frac{1}{2}\left(\|\boldsymbol{y}\|_{\Sigma_n}^2 + \|G \circ F_{\mathrm{star}}(\xi_1, c_1)\|_{\Sigma_n}^2\right) \\
&\leq \frac{1}{2}\left(\|\boldsymbol{y}\|_{\Sigma_n}^2 + C\|F_{\mathrm{star}}(\xi_1, c_1)\|_{L^\infty(D)}\right) \\
&\leq \frac{1}{2}\left(\rho^2 + C\alpha^+\right),
\end{aligned}$$

where $\|\boldsymbol{y}\|_{\Sigma_n} \leq \rho$. We define $K(\rho) := \frac{1}{2}\left(\rho^2 + C\alpha^+\right)$.

(ii) In Proposition B.1 we show that F_{star} is a continuous map. Furthermore, we discussed in part (i) that G is also a continuous map. Therefore, for a fixed $\boldsymbol{y} \in \mathbb{R}^N$, $\Phi(\cdot; \boldsymbol{y})$ is a composition of continuous maps. This concludes the proof.

(iii) Let $(\xi_1, c_1) \in X$ and $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^N$ such that $\|\boldsymbol{y}_1\|_{\Sigma_n}, \|\boldsymbol{y}_2\|_{\Sigma_n} \leq \rho$. We have

(B.4)
$$\begin{aligned}
|\Phi((\xi_1, c_1), \boldsymbol{y}_1) &- \Phi((\xi_1, c_1), \boldsymbol{y}_2)| \\
&= \frac{1}{2}|\langle \boldsymbol{y}_1 - \boldsymbol{y}_2, \boldsymbol{y}_1 + \boldsymbol{y}_2 - 2\mathcal{G}(\xi_1, c_1)\rangle_{\Sigma_n}| \\
&\leq \frac{1}{2}\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_{\Sigma_n}\|\boldsymbol{y}_1 + \boldsymbol{y}_2 - 2\mathcal{G}(\xi_1, c_1)\| \\
&\leq \frac{1}{2}\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_{\Sigma_n}\left(\|\boldsymbol{y}_1\|_{\Sigma_n} + \|\boldsymbol{y}_2\|_{\Sigma_n} + 2\|\mathcal{G}(\xi_1, c_1)\|_{\Sigma_n}\right) \\
&\leq (\rho + C\alpha^+)\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|_{\Sigma_n}.
\end{aligned}$$

We define $M(\rho) := (\rho + C\alpha^+)$. $\qquad\qquad\square$

Now we show that the posterior measure $\mu^{\boldsymbol{y}}$ is a well-defined measure on X.

THEOREM B.2. *The posterior measure $\mu^{\boldsymbol{y}}$ Equation (4.4) is well-defined.*

Proof. We showed in Proposition 4.1 that $\Phi(\cdot; \boldsymbol{y})$ is μ_0-a.s. continuous and $\Phi((\xi_1, c_1), \cdot)$ is locally Lipschitz. This is sufficient condition for Φ to be jointly continuous with respect to $\mu_0 = \mu_0^1 \otimes \mu_0^2$, see the proof of theorem 3.8 in [18]. Therefore, it is left to show that the normalization constant Z in Equation (4.4) is bounded away

from zero. Define $B := B_H \times D \subset X$, where B_H is an open ball in H. For $\boldsymbol{y} \in \mathbb{R}^N$, with $\|\boldsymbol{y}\|_{\Sigma_n} \leq \rho$ we have

$$
\begin{aligned}
\int_X \exp(-\Phi((\xi_1, c_1); \boldsymbol{y})) \ \mu_0(d\xi) &\geq \int_B \exp(-\Phi((\xi_1, c_1); \boldsymbol{y})) \ \mu_0(d\xi) \\
&\geq \int_B \exp(-K(\rho)) \ \mu_0(d\xi) \\
&= \exp(-K(\rho))\mu_0(B) > 0.
\end{aligned}
$$
(B.5)

Here we used positivity of the exponential function in the first inequality, and condition (i) in Proposition 4.1. Note that since μ_0^1 is Gaussian and B_H is an open set then the $\mu_0^1(B_H) > 0$. Therefore, $\mu_0(B) = \mu_0^1(B_H)\mu_0^2(D) > 0$. $\qquad\square$

Let μ_1 and μ_2 be two posterior probability measures defined on $(H, \mathcal{B}(H))$ such that they are both absolutely continuous with respect to the prior measure μ_0. The Hellinger distance [39] between μ_1 and μ_2 is given by

$$
(B.6) \qquad d_{\text{Hell}}(\mu_1, \mu_2) = \sqrt{\frac{1}{2} \int_H \sqrt{\frac{d\mu_1}{d\mu_0}(\xi)} - \sqrt{\frac{d\mu_2}{d\mu_0}(\xi)} \ \mu_0(d\xi)}.
$$

THEOREM B.3. *[14] Let Φ be the negative log-likelihood defined* (4.1) *satisfying assumptions (i)-(iii) in Proposition* 4.1 *and μ_0 be the prior measure defined on $(X, \mathcal{A})$. Furthermore, let $\boldsymbol{y}, \boldsymbol{y}' \in \mathbb{R}^N$ be two observation vector such that $\|\boldsymbol{y} - \boldsymbol{y}'\|_{\Sigma_n} \leq \rho$ for a fixed $0 \leq \rho < \infty$. Then we can find $C = C(\rho) > 0$ such that*

$$
(B.7) \qquad d_{Hell}(\mu^{\boldsymbol{y}}, \mu^{\boldsymbol{y}'}) \leq C\|\boldsymbol{y} - \boldsymbol{y}'\|_{\Sigma_n}.
$$

Bayesian well-posedness expressed in Theorem B.3 means that the posterior distribution remains bounded when the observation vector $\boldsymbol{y}$ is perturbed.

Appendix C. Markov Chain Monte Carlo (MCMC) Methods. In this section we briefly introduce the random walk Metropolis-Hastings (RWM) method [43], the preconditioned Crank-Nicolson (pCN) method [11], and the Gibbs sampling method [43]. These methods allow us to sample from the posterior measure $\boldsymbol{\mu^y}$. We choose pCN to sample infinite dimensional Gaussian random variables as it is suitable for random variables on function spaces. Furthermore, the convergence rate of this method is independent of the cut off value N_{KL} in (3.2).

Let $f : H \to \mathbb{R}^k$, be a μ_0-measurable function. We approximate $\mathbb{E}f$ as the ergodic average

$$
(C.1) \qquad \mathbb{E}f := \int f \ \mu_0(d\xi) = \lim_{N_{\text{sample}} \to \infty} \frac{1}{N_{\text{sample}}} \sum_{i=1}^{N_{\text{sample}}} f(\xi^{(i)}).
$$

Here, $\{\xi^{(i)}\}_{i=1}^{N_{\text{sample}}}$ is a Markov chain with a transition kernel $\mathbb{K}_y$. We refer the reader to [30] for properties of a transition kernel suitable for (C.1).

The aim of an MCMC method is to construct a transition function (a formula to go from $\xi^{(i)}$ to $\xi^{(i+1)}$ in (C.1)) which result in an appropriate transition kernel $\mathbb{K}_y$. As an MCMC method, pCN is suitable for drawing samples from an H-valued Gaussian random variables, and is robust under the discretization of the random variables. We first introduce pCN to sample from a single Gaussian random variable ξ, then

discuss how to generalize this to sample from multiple independent Gaussian random variables.

Let $\mu^{\boldsymbol{y}}$ be the posterior measure on a Hilbert space $(H, \langle \cdot, \cdot \rangle)$ and $\mu_0 \sim \mathcal{N}(0, Q)$ be a prior measure. Let $\mathbb{K}(\xi, \cdot)$ be the transition kernel on H and η denote a measure on $H \times H$ such that if $\xi \sim \mu^{\boldsymbol{y}}$ then $\zeta | \xi \sim \mathbb{K}_y(\xi, \cdot)$. We denote by $\eta^\perp$ the measure where the roles of ξ and ζ are reversed. If $\eta^\perp$ is equivalent to η, in the sense of measure, then the Radon-Nikodym derivative $d\eta^\perp / d\eta$ is well defined and we can define the acceptance probability

$$(\text{C.2}) \qquad a(\xi, \zeta) = \min \left\{ 1, \frac{d\eta^\perp}{d\eta}(\xi, \zeta) \right\}.$$

This means $\xi^{(i+1)} := \zeta$ with probability $a(\xi^{(i+1)}, \zeta)$ and $\xi^{(i+1)} := \xi$ otherwise.

The standard random walk proposal function results in $\eta^\perp$ that is singular with respect to η [11] when ξ is an H-valued function. This results in rejecting all proposed moves with probability 1. The pCN proposal function [11] is given by

$$(\text{C.3}) \qquad \zeta = (1 - b^2)^{1/2} \xi + b\varrho,$$

where $\varrho \sim \mathcal{N}(0, Q)$, and $b \in [0, 1]$. This choice of the proposal results in a well defined $d\eta^\perp / d\eta$ given by [11]

$$(\text{C.4}) \qquad \frac{d\eta^\perp}{d\eta}(\xi, \zeta) = \exp(\eta(\xi, \zeta) - \eta(\zeta, \xi)) = \exp(\Phi(\xi; \boldsymbol{y}) - \Phi(\zeta; \boldsymbol{y})).$$

We summarize the pCN sampling method in Algorithm C.1.

Algorithm C.1 pCN for collecting N_{sample} samples

1: Set $j = 0$ and take the initial sample $\xi^{(0)}$.
2: **for** $j \le N_{\text{sample}}$ **do**
3:     Propose $\zeta = (1 - b^2)^{1/2} \xi^{(k)} + b\varrho$, $\varrho \sim \mathcal{N}(0, Q)$.
4:     Set $\xi^{(j+1)} = \zeta$ with probability $a(\xi^{(j)}, \zeta)$ defined in (C.2) together with (C.4), otherwise $\xi^{(j+1)} = \xi^{(j)}$.
5:     $j \leftarrow j + 1$.
6: **end for**

Recall that star-shaped prior for each inclusion is defined as the joint random variable (ξ_i, c_i). To construct an MCMC method to sample from $\mu^{\boldsymbol{y}}$ we use a *Gibbs*-type [43] sampling method. Such methods alternatively sample from the random variables $\xi_i | c_i, \boldsymbol{y}$ and $c_i | \xi_i, \boldsymbol{y}$ and constructed Markov chain $[\xi_i^{(j)}, c_i^{(j)}]_{j=1}^{N_{\text{sample}}}$ contains ergodic properties as in Equation (C.1) [43]. In this paper we construct a Gibbs sampler following the structure:

$$(\text{C.5}) \qquad \xi_i^{(j+1)} \sim \mathbb{K}_{\boldsymbol{y}}^{c_i^{(j)}}(\xi_i^{(j)}, \cdot), \qquad c_i^{(j+1)} \sim \mathbb{L}_{\boldsymbol{y}}^{\xi_i^{(j+1)}}(c_i^{(j)}, \cdot).$$

Here, $\mathbb{K}_{\boldsymbol{y}}^{c_i}$ and $\mathbb{L}_{\boldsymbol{y}}^{\xi_i}$ are Metropolis-Hastings Markov kernel reversible with respect to $\xi_i | c_i, \boldsymbol{y}$ and $c_i | \xi_i, \boldsymbol{y}$, respectively. For the random variable $\xi_i | c_i, \boldsymbol{y}$ we use a pCN proposal function with the acceptance probability of

$$(\text{C.6}) \qquad a(\xi_i, \zeta) = \min\{1, \exp(\Phi((\xi_i, c_i); \boldsymbol{y}) - \Phi(\zeta_0, (\zeta, c_i); \boldsymbol{y}))\}.$$

For the random variable $c_i | \xi_i, \boldsymbol{y}$ we use the standard random walk MH proposal function with the acceptance probability of

$$(C.7) \qquad r(c_i, o) = \min\{1, \exp(\Phi((\xi_i, c_i); \boldsymbol{y}) - \Phi((\xi_i, o); \boldsymbol{y}))\}.$$

Note that we dropped the uniform prior in (C.7) as both c_i and o_i are restricted to the same interval. We summarize the Gibbs sampler for $\mu^{\boldsymbol{y}}$ constructed with the star-shaped prior in Algorithm C.2. Here, N_{pCN} and N_{MH} represent the number of within-Gibbs samples for each component.

Algorithm C.2 The Gibbs sampling method for collecting N_{sample} samples

---

1: Set $j = 0$ and take the initial sample $(\xi_i^{(0)}, c_i^{(0)})$.
2: **for** $j \leq N_{\mathrm{sample}}$ **do**
3:     Set $\xi_i^{(j,0)} = \xi_i^{(j)}$.
4:     **for** $k \leq N_{\mathrm{pCN}}$ **do**
5:         Propose $\zeta = (1 - b_1^2)^{1/2} \xi^{(j,k)} + b_1 \varrho, \ \varrho \sim \mathcal{N}(0, Q)$.
6:         Set $\xi_i^{(j,k+1)} = \zeta$ with probability $a(\xi_i^{(j,k)}, \zeta)$ defined in (C.6), otherwise $\xi_i^{(j,k+1)} = \xi_i^{(j,k)}$.
7:         $k \leftarrow k + 1$.
8:     **end for**
9:     Set $\xi_i^{(j+1)} = \xi_i^{(j,k+1)}$.
10:    Set $c_i^{(j,0)} = c_i^{(j)}$.
11:    **for** $k \leq N_{\mathrm{MH}}$ **do**
12:        Propose $o = c_i^{(j,k)} + b_2 \rho, \ \rho \sim \mathcal{N}(0, I)$.
13:        Set $c_i^{(j,k+1)} = o$ with probability $r(c_i^{(j,k)}, o)$ defined in (C.7), otherwise $c_i^{(j,k+1)} = c_i^{(j,k)}$.
14:        $k \leftarrow k + 1$.
15:    **end for**
16:    Set $c_i^{(j+1)} = c_i^{(j,k+1)}$.
17:    $j \leftarrow j + 1$.
18: **end for**

REFERENCES

[1] *Inverse problems: a Bayesian perspective*, Acta Numerica, 19, https://doi.org/10.1017/S0962492910000061.

[2] M. ABRAMOWITZ, I. A. STEGUN, AND R. H. ROMER, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, American Association of Physics Teachers, 1988.

[3] C. V. ALVINO AND A. J. YEZZI, *Tomographic reconstruction of piecewise smooth images*, in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., vol. 1, IEEE, 2004, pp. I–576–I–581.

[4] S. BANERJEE AND A. ROY, *Linear algebra and matrix analysis for statistics*, vol. 181, Crc Press Boca Raton, 2014.

[5] T. A. BUBBA, A. HAUPTMANN, S. HUOTARI, J. RIMPELÄINEN, AND S. SILTANEN, *Tomographic X-ray data of a lotus root filled with attenuating objects*, Sept. 2016, https://doi.org/10.5281/zenodo.1254204.

[6] T. A. BUBBA, A. HAUPTMANN, S. HUOTARI, J. RIMPELÄINEN, AND S. SILTANEN, *Tomographic x-ray data of a lotus root filled with attenuating objects*, 2016, https://arxiv.org/abs/arXiv:1609.07299.

[7] T. Bui-Thanh and O. Ghattas, *An analysis of infinite dimensional bayesian inverse shape acoustic scattering and its numerical approximation*, SIAM/ASA J. Uncertainty Quantification, 2 (2014), pp. 203–222, https://doi.org/10.1137/120894877.

[8] T. F. Chan and L. A. Vese, *Active contours without edges*, IEEE Trans. Image Processing, 10 (2001), pp. 266–277.

[9] C. Chapdelaine, A. Mohammad-Djafari, N. Gac, and E. Parra, *Variational Bayesian approach in model-based iterative reconstruction for 3D X-ray computed tomography with Gauss-Markov-Potts prior*, in Multidisciplinary Digital Publishing Institute Proceedings, vol. 33, 2019, p. 4.

[10] Y. Chen, D. Gao, C. Nie, L. Luo, W. Chen, X. Yin, and Y. Lin, *Bayesian statistical reconstruction for low-dose X-ray computed tomography using an adaptive-weighting nonlocal prior*, Computerized Medical Imaging and Graphics, 33 (2009), pp. 495–500.

[11] S. L. Cotter, G. O. Roberts, A. M. Stuart, and D. White, *MCMC methods for functions: Modifying old algorithms to make them faster*, Statistical Science, 28 (2013), https://doi.org/10.1214/13-sts421.

[12] V. A. Dahl, A. B. Dahl, and P. C. Hansen, *Computing segmentations directly from X-ray projection data via parametric deformable curves*, Measurement Science and Technology, 29 (2017), p. 014003.

[13] M. Dashti, S. Harris, and A. Stuart, *Besov priors for Bayesian inverse problems*, Inverse Problems and Imaging, 6 (2012), pp. 183–200.

[14] M. Dashti and A. M. Stuart, *The bayesian approach to inverse problems*, in Handbook of Uncertainty Quantification, R. Ghanem, D. Higdon, and H. Owhadi, eds., Springer, 2017, pp. 311–428, https://doi.org/10.1007/978-3-319-12385-1_7.

[15] E. Debreuve, M. Barlaud, G. Aubert, and J. Darcourt, *Attenuation map segmentation without reconstruction using a level set method in nuclear medicine imaging*, in Proc. 1998 International Conference on Image Processing, ICIP98, vol. 1, IEEE, 1998, pp. 34–38.

[16] M. M. Dunlop, T. Helin, and A. M. Stuart, *Hyperparameter estimation in Bayesian MAP estimation: Parameterizations and consistency*, The SMAI J Computational Mathematics, 6 (2020), pp. 69–100.

[17] M. M. Dunlop, M. A. Iglesias, and A. M. Stuart, *Hierarchical Bayesian level set inversion*, Statistics and Computing, 27 (2016), pp. 1555–1584, https://doi.org/10.1007/s11222-016-9704-8.

[18] M. M. Dunlop and A. M. Stuart, *The Bayesian formulation of EIT: Analysis and algorithms*, 2016, https://doi.org/10.3934/ipi.2016030.

[19] H. Elhalawani, A. S. Mohamed, A. L. White, J. Zafereo, A. J. Wong, J. E. Berends, S. AboHashem, B. Williams, J. M. Aymard, A. Kanwar, et al., *Matched computed tomography segmentation and demographic data for oropharyngeal cancer radiomics challenges*, Scientific Data, 4 (2017), p. 170077.

[20] A. Faridani, D. V. Finch, E. L. Ritman, and K. T. Smith, *Local tomography II*, SIAM Journal on Applied Mathematics, 57 (1997), pp. 1095–1127.

[21] A. Faridani, E. L. Ritman, and K. T. Smith, *Local tomography*, SIAM Journal on Applied Mathematics, 52 (1992), pp. 459–484.

[22] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, CRC Press, 2013.

[23] A. Gelman and D. B. Rubin, *Inference from iterative simulation using multiple sequences*, Statistical Science, 7 (1992), pp. 457–472.

[24] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. P. Pereira, M. J. Clarkson, and D. C. Barratt, *Automatic multi-organ segmentation on abdominal ct with dense v-networks*, IEEE Trans. Medical Imaging, 37 (2018), pp. 1822–1834.

[25] G. Grimmett and D. Stirzaker, *Probability and Random Processes*, no. 391, Oxford University Press, 2001.

[26] P. C. Hansen, J. S. Jørgensen, and W. R. B. Lionheart, eds., *Computed Tomography: Algorithms, Insight, and Just Enough Theory*, SIAM, PA, 2021.

[27] J. Huang, Z. Deng, and L. Xu, *Bayesian approach for inverse interior scattering problems with limited aperture*, Applicable Analysis, (2020), pp. 1–14.

[28] I. A. Ibragimov and Y. A. Rozanov, *Gaussian Random Processes*, Springer-Verlag, NY, 1978.

[29] M. A. Iglesias, Y. Lu, and A. M. Stuart, *A Bayesian level set method for geometric inverse problems*, 2015, https://doi.org/arXivpreprintarXiv:1504.00313.

[30] J. Kaipio and E. Somersalo, *Statistical and Computational Inverse Problems*, Springer, 2006.

[31] R. Kumar, C. Carroll, A. Hartikainen, and O. Martin, *Arviz a unified library for exploratory analysis of Bayesian models in Python*, Journal of Open Source Software, 4

(2019), p. 1143, https://doi.org/10.21105/joss.01143, https://doi.org/10.21105/joss.01143.

[32] M. LASSAS, E. SAKSMAN, AND S. SILTANEN, *Discretization-invariant Bayesian inversion and Besov space priors*, Inverse Problems & Imaging, 3 (2009), pp. 87–122, https://doi.org/10.3934/ipi.2009.3.87.

[33] O. LEY ET AL., *Lower-bound gradient estimates for first-order hamilton-jacobi equations and applications to the regularity of propagating fronts*, Advances in Differential Equations, 6 (2001), pp. 547–576.

[34] F. LINDGREN, H. RUE, AND J. LINDSTRÖM, *An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73 (2011), pp. 423–498, https://doi.org/10.1111/j.1467-9868.2011.00777.x.

[35] S. LIU, J. JIA, Y. D. ZHANG, AND Y. YANG, *Image reconstruction in electrical impedance tomography based on structure-aware sparse Bayesian learning*, IEEE Trans. Medical Imaging, 37 (2018), pp. 2090–2102.

[36] M. MARKKANEN, L. ROININEN, J. M. HUTTUNEN, AND S. LASANEN, *Cauchy difference priors for edge-preserving Bayesian inversion*, Journal of Inverse and Ill-posed Problems, 27 (2019), pp. 225–240, https://doi.org/doi:10.1515/jiip-2017-0048.

[37] F. NATTERER, *The Mathematics of Computerized Comography*, SIAM, 2001.

[38] A. B. OWEN, *Monte Carlo Theory, Methods and Examples*, 2013.

[39] L. PARDO, *Statistical Inference Based on Divergence Measures*, CRC press, 2018.

[40] J. R. PARKER, *Algorithms for Image Processing and Computer Vision*, Wiley Publishing, 2nd ed., 2010.

[41] R. RAMLAU AND W. RING, *A Mumford–Shah level-set approach for the inversion and segmentation of X-ray tomography data*, Journal of Computational Physics, 221 (2007), pp. 539–557.

[42] C. E. RASMUSSEN, *Gaussian processes in machine learning*, in Summer school on machine learning, Springer, 2003, pp. 63–71.

[43] C. ROBERT AND G. CASELLA, *Monte Carlo Statistical Methods*, Springer, 2013.

[44] L. ROININEN, , J. M. J. HUTTUNEN, AND S. LASANEN, *Whittle-Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography*, Inverse Problems & Imaging, 8 (2014), pp. 561–586, https://doi.org/10.3934/ipi.2014.8.561.

[45] A. K. SAIBABA, P. PRASAD, E. DE STURLER, E. MILLER, AND M. E. KILMER, *Randomized approaches to accelerate MCMC algorithms for Bayesian inverse problems*, Journal of Computational Physics, 440 (2021), p. 110391, https://doi.org/10.1016/j.jcp.2021.110391.

[46] K. T. SMITH AND F. KEINERT, *Mathematical foundations of computed tomography*, Appl. Opt., 24 (1985), pp. 3950–3957, https://doi.org/10.1364/AO.24.003950, http://opg.optica.org/ao/abstract.cfm?URI=ao-24-23-3950.

[47] N. B. SMITH AND A. WEBB, *Introduction to Medical Imaging*, Cambridge University Press, 2009, https://doi.org/10.1017/cbo9780511760976.

[48] C. SOUSSEN AND A. MOHAMMAD-DJAFARI, *Polygonal and polyhedral contour reconstruction in computed tomography*, IEEE Trans. Image Processing, 13 (2004), pp. 1507–1523.

[49] S. SULLIVAN, A. NOBLE, AND J. PONCE, *On reconstructing curved object boundaries from sparse sets of X-ray images*, in International Conference on Computer Vision, Virtual Reality, and Robotics in Medicine, Springer, 1995, pp. 385–391.

[50] J.-P. THIRION, *Segmentation of tomographic data without image reconstruction*, IEEE Trans. Medical Imaging, 11 (1992), pp. 102–110.

[51] F. URIBE, J. M. BARDSLEY, Y. DONG, P. C. HANSEN, AND N. A. B. RIIS, *A hybrid Gibbs sampler for edge-preserving tomographic reconstruction with uncertain view angles*, 2021, https://arxiv.org/abs/arXiv:2104.06919.

[52] E. VAINBERG, I. KAZAK, AND V. KUROZAEV, *Reconstruction of the internal three-dimensional structure of objects based on real-time integral projections*, Soviet Journal of Nondestructive Testing-USSR, 17 (1981), pp. 415–423.

[53] S. VAN DER WALT, J. L. SCHÖNBERGER, J. NUNEZ-IGLESIAS, F. BOULOGNE, J. D. WARNER, N. YAGER, E. GOUILLART, AND T. YU, *scikit-image: image processing in python*, PeerJ, 2 (2014), p. e453, https://doi.org/10.7717/peerj.453.

[54] A. VEHTARI, A. GELMAN, D. SIMPSON, B. CARPENTER, AND P.-C. BURKNER, *Rank-Normalization, Folding, and Localization: An Improved $\widehat{R}$ for Assessing Convergence of MCMC (with Discussion)*, Bayesian Analysis, 16 (2021), pp. 667 – 718, https://doi.org/10.1214/20-BA1221, https://doi.org/10.1214/20-BA1221.

[55] L. WANG, A. MOHAMMAD-DJAFARI, AND N. GAC, *X-ray computed tomography using a sparsity enforcing prior model based on haar transformation in a bayesian framework*, Fundamenta Informaticae, 155 (2017), pp. 449–480.

[56] S. WANG, A. CHEN, L. YANG, L. CAI, Y. XIE, J. FUJIMOTO, A. GAZDAR, AND G. XIAO, *Com-

prehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome, Scientific Reports, 8 (2018), pp. 1–9.

[57] J. R. WEAVER, *Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors*, The American Mathematical Monthly, 92 (1985), pp. 711–717, https://doi.org/10.1080/00029890.1985.11971719.

[58] J. W. WEBBER, E. T. QUINTO, AND E. L. MILLER, *A joint reconstruction and lambda tomography regularization technique for energy-resolved X-ray imaging*, Inverse Problems, 36 (2020), p. 074002.

[59] A. D. WESTON, P. KORFIATIS, T. L. KLINE, K. A. PHILBRICK, P. KOSTANDY, T. SAKINIS, M. SUGIMOTO, N. TAKAHASHI, AND B. J. ERICKSON, *Automated abdominal segmentation of ct scans for body composition analysis using deep learning*, Radiology, 290 (2019), pp. 669–679.

[60] S. YOON, A. R. PINEDA, AND R. FAHRIG, *Level set reconstruction for sparse angularly sampled data*, in 2006 IEEE Nuclear Science Symposium Conference Record, vol. 6, IEEE, 2006, pp. 3420–3423.

[61] S. YOON, A. R. PINEDA, AND R. FAHRIG, *Simultaneous segmentation and reconstruction: A level set method approach for limited view computed tomography*, Medical Physics, 37 (2010), pp. 2329–2340.

[62] H. ZHANG, *Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics*, Journal of the American Statistical Association, 99 (2004), pp. 250–261.