

# Structural Gaussian priors for Bayesian CT reconstruction of subsea pipes

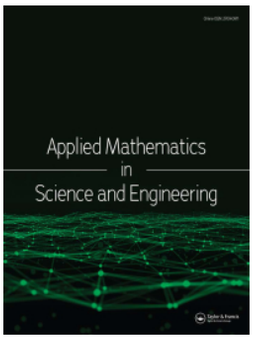
**Christensen, Silja L.; Riis, Nicolai A.B.; Uribe, Felipe; Jørgensen, Jakob S.**

*Published in:*
Applied Mathematics in Science and Engineering

*Link to article, DOI:*
10.1080/27690911.2023.2224918

*Publication date:*
2023

*Document Version*
Publisher's PDF, also known as Version of record



Applied Mathematics in Science and Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gipe21

# Structural Gaussian priors for Bayesian CT reconstruction of subsea pipes

Silja L. Christensen, Nicolai A. B. Riis, Felipe Uribe & Jakob S. Jørgensen

**To cite this article:** Silja L. Christensen, Nicolai A. B. Riis, Felipe Uribe & Jakob S. Jørgensen (2023) Structural Gaussian priors for Bayesian CT reconstruction of subsea pipes, Applied Mathematics in Science and Engineering, 31:1, 2224918, DOI: 10.1080/27690911.2023.2224918

**To link to this article:** https://doi.org/10.1080/27690911.2023.2224918



Published online: 26 Jun 2023.

APPLIED MATHEMATICS IN SCIENCE AND ENGINEERING
2023, VOL. 31, NO. 1, 2224918
https://doi.org/10.1080/27690911.2023.2224918

Taylor & Francis
Taylor & Francis Group



# Structural Gaussian priors for Bayesian CT reconstruction of subsea pipes

Silja L. Christensen [a], Nicolai A. B. Riis [a,b], Felipe Uribe [c,*] and Jakob S. Jørgensen [a]

[a]Department of Applied Mathematics and Computer Science, Technical University of Denmark. Richard Petersens Plads, Kongens Lyngby, Denmark; [b]Copenhagen Imaging ApS, Herlev, Denmark; [c]School of Engineering Science, Lappeenranta-Lahti University of Technology (LUT), Lappeenranta, Finland

**ABSTRACT**

A non-destructive testing (NDT) application of X-ray computed tomography (CT) is inspection of subsea pipes in operation via 2D cross-sectional scans. Data acquisition is time-consuming and costly due to the challenging subsea environment. While reducing the number of projections in a CT scan yields time and cost savings, this compromises the reconstruction quality when conventional reconstruction methods are used. To address this issue we take a Bayesian approach to CT reconstruction and focus on designing an effective prior that introduces additional information to the problem. We propose a new class of *structural Gaussian priors* to enforce expected material properties in different regions of the reconstructed image based on available a-priori information about the pipe's layered structure. The prior is composed from Gaussian distributions, and we utilize this to propose an efficient computational strategy to sample the resulting posterior distribution. This is essential in practical NDT applications for fast processing of the large-scale data we see in CT. Numerical experiments with synthetic and real data show that the proposed structural Gaussian prior reduces artifacts and enhances contrast in the reconstruction, compared to using only a conventional Gaussian Markov random field prior or no prior information at all. Furthermore, we obtain uncertainty estimates that indicate the certainty of the reconstructions increase when more prior information is added.

**ARTICLE HISTORY**
Received 13 July 2022
Accepted 6 June 2023

**KEYWORDS**
X-ray computed tomography; uncertainty quantification; Bayesian inversion; non-destructive testing; pipeline inspection

**MATHS**
65C20; 62F15; 65R32; 94A08

## 1. Introduction

Subsea pipelines are used to transport oil and gas around the world. It is crucial that the subsea pipes are in good condition to avoid environmentally damaging leaks. Therefore, efficient methods for non-destructive inspection are in demand. One suitable method is X-ray computed tomography (CT) that is used to obtain 2D cross-sectional images of the pipes using data from a finite number of projections [1–4]. To produce high-quality images with limited noise and artifacts, standard CT reconstruction methods require

**CONTACT** Jakob S. Jørgensen jakj@dtu.dk Department of Applied Mathematics and Computer Science, Technical University of Denmark, Richard Petersens Plads, Building 324, Kongens Lyngby 2800, Denmark
*Part of the work by F.U. was done while employed at the Technical University of Denmark.

lengthy measurements at many projection (view) angles. To reduce costs, it is desirable to limit measurement time, and this can be achieved by reducing the number of view angles. However, if standard iterative reconstruction methods are used, this compromises the reconstruction quality significantly [5]. Furthermore, due to the scarcity and noisiness of the data, the reconstructions are subject to uncertainty. Therefore, we follow a Bayesian approach, where we combine information about the CT model and projection data with a prior probability distribution of the image. This results in a posterior distribution that describes the reconstruction and its associated uncertainty. By leveraging prior knowledge of the internal structure of the pipes, we enhance the quality of image reconstruction, even in cases where measurements are taken at only a few view angles.

### 1.1. Literature review

Previous work has taken a deterministic approach to CT scanning of subsea pipes [4]. In that work, the authors use microlocal analysis to design a favourable offset scan geometry and propose a reconstruction method based on *shearlets* with a weighted sparsity penalty. However, this method does consider or quantify the uncertainty associated with the reconstructions. We expand on this work by adopting a Bayesian approach to CT reconstruction of subsea pipes, enabling probabilistic modelling of noise and prior information. Subsequently, through sampling of the resulting posterior distribution, we can obtain reconstructions along with their respective uncertainty estimates.

In Bayesian CT reconstruction [6–8], it is important to choose an appropriate prior. Typical choices in this setting are Markov random field (MRF) priors based on multivariate Gaussian and Laplace distributions. These represent a statistical parallel to the well-known Tikhonov and total variation (TV) regularization methods, which are commonly used in classical inverse problems and in particular CT [9,10]. An overview of several MRF type priors applied to 2D inverse problems (including CT) is presented in [11], while more detailed descriptions can be found in earlier publications [12,13]. A recent approach is the Cauchy difference prior (also a MRF-type prior), which is shown to be edge-preserving and discretization invariant [14,15]. All MRF priors have regularizing effects, and depending on the scanned object, a smoothness-promoting or edge-preserving prior can be chosen. However, a common characteristic for these distributions is that they do not model specific features, such as the layer structure in subsea pipes, when used as priors in Bayesian reconstruction. Furthermore, the application of non-Gaussian MRF priors requires the use of more advanced posterior sampling techniques, which can be computationally demanding and, hence, impractical for large-scale problems.

If there is prior knowledge about the specific structure of the target image, a more informative prior can be formulated. Several studies use this idea by forming a texture-preserving prior for low-dose CT reconstruction using a previous normal-dose scan [16–19]. Another study exploits the fact that the human body is composed of a finite number of tissue types, each with well-known attenuation coefficients, and designs the prior accordingly, including an inherent classification problem [20]. In cases where there is prior knowledge about the location of edges within the internal structure of the object, the 'structural priors' proposed in [21] may be applicable. They formulate a Gaussian prior with anisotropic correlation, in which the image intensity varies smoothly within predefined domains while large jumps are allowed across domain boundaries. If the location of

edges within the image are specified incorrectly, the structural prior leads to bias in the posterior, and the reconstruction will include traces of the incorrectly specified edge information as seen in [21]. The structural priors, and variants thereof, have been applied in a number of imaging problems such as electric impedance tomography [21,22], electrical resistance tomography [23], optical tomography [24], magnetic resonance imaging [25,26], and inpainting [27].

### 1.2. Contributions

The driving idea in this work is to utilize as much prior information as possible. We have knowledge of the general layout of the pipes since they are manufactured according to specific design standards. As a result, we propose a *structural Gaussian prior* that introduces information regarding the layered structure and constitutive materials of the pipes. Numerical experiments show that this prior improves the reconstruction quality compared to conventional priors, particularly when only a few view angles are available. Similarly to the structural priors in [21], our proposed structural Gaussian prior introduces information about the expected internal structure in the object of interest. However, our prior is characterized by two fundamental assumptions that differ from existing ones. First, [21] assumes the position of domain boundaries in the image are known exactly; for instance, from other imaging modalities or previous high-quality reference scans, while our approach only requires approximate positions. Second, within the predefined domains of the image the structural prior from [21] only assumes smoothness, while our approach assumes strong prior expectations (mean values) of image intensities. We remark that [28] includes a structural prior variant that assumes the presence of known values within a certain image domain, which aligns with our second assumption. However, this approach still requires exact location of the boundaries.

Our structural Gaussian prior is designed with the challenges of sampling high-dimensional distributions in mind. By choosing a Gaussian prior and likelihood, the resulting Gaussian posterior can be sampled efficiently based on principles of the *randomize-then-optimize* (RTO) [29] approach for linear inverse problems. In previous studies, we have observed the use of linear RTO sampling in cases where the prior has zero mean (e.g. [8,11]). However, to the best of our knowledge, the linear RTO method has not been applied when the prior has a non-zero mean, which is required for our structural Gaussian prior. Therefore, we generalize the linear RTO sampler, and show that posterior samples obtained using this method conform to the expected Gaussian posterior distribution.

Finally, we do not know of any previous studies that apply Bayesian methods for CT imaging of subsea pipes. By taking this approach and incorporating the informative structural Gaussian prior, we improve reconstruction quality and obtain related uncertainty estimates. In practice, this may aid in risk assessment and decision making regarding the condition of the pipes.

### 1.3. Outline

Our paper is organized as follows. In Section 2 we describe CT imaging including the assumed X-ray attenuation model, the formulation of the inverse problem, and a simple deterministic reconstruction method. In Section 3 we introduce the Bayesian linear inverse

problem, and we briefly describe the principles of posterior sampling and the advantages of using Gaussian distributions. Section 4 describes how we apply Bayesian CT reconstruction to imaging of subsea pipes; this section also includes the proposed structural Gaussian prior and the generalization of the linear RTO method for efficient posterior sampling. In Section 5 we perform numerical experiments with synthetic and real data. In Section 6 we discuss and conclude our findings.

## 2. CT imaging

X-ray CT is an imaging modality that creates contrast through the distinct linear attenuation coefficients of different materials. It is a non-destructive imaging method that is widely used in a number of applications; for instance inspection of subsea pipes in this work. By illuminating an object of interest with X-rays and recording projections from multiple angles, it is possible to reconstruct an image of the objects interior. The inverse problem in X-ray CT consists of inferring the unknown pixelized cross-section image from external measurements of X-ray attenuation; in this context the forward model describes the mapping from a known pixelized image to measurements of X-ray attenuation. Below we describe relevant physics of X-ray attenuation and use this as a forward model to formulate the CT inverse problem. We also describe the conjugate gradient least-squares algorithm, which is a commonly used iterative method. Later, we will incorporate this algorithm in our proposed methodology, and we will compare our numerical results to reconstructions obtained using this algorithm in the conventional way. Finally, we describe the physical phenomenon *buildup*, how it can lead to model error in the CT inverse problem, and how we deal with that.

### 2.1. CT inverse problem

In the CT inverse problem, we seek to reconstruct an image of linear attenuation coefficients given an X-ray attenuation model and observed X-ray absorption from many directions. In general, the attenuation model of an X-ray travelling through a non-homogeneous material obey the Lambert–Beer's law which can be described via the line integral:

$$I = I_0 \exp\left(-\int_{\mathcal{L}} \alpha(\ell)\, \mathrm{d}\ell\right), \tag{1}$$

where $I_0$ is the X-ray intensity entering the material, $I$ is the observed X-ray intensity leaving the material, $\mathcal{L}$ denotes the line path the X-ray follows through the material, and $\alpha(\ell)$ is a continuous function that represents the linear attenuation in the material along the line $\mathcal{L}$ parametrized by the parameter $\ell$. This model assumes a monochromatic pencil beam, such that scattering effects from other X-rays can be ignored.

We discretize a reconstruction domain into an image with $n = N \times N$ pixels, where $N$ is the domain discretization size, and express the discrete attenuation model for one X-ray with index $i$:

$$I_i = I_0 \exp\left(-\sum_{j=1}^{n} a_{ij} x_j\right). \tag{2}$$

Now, $x$ is the discretized version of $\alpha$, $x_j$ denotes the attenuation coefficient of the $j$'th pixel in the reconstruction domain, and $a_{ij}$ describes the path length of ray $i$ through pixel $j$, where $a_{ij} = 0$ when ray $i$ does not intersect pixel $j$.

Given $m$ observations of the attenuated X-ray intensity $I_i$, $i = 1, \ldots, m$ from different directions and using the discrete X-ray attenuation model, we can express the full CT inverse problem as a system of linear equations,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{e}, \tag{3}$$

where $\mathbf{y} \in \mathbb{R}^m$ is the observed X-ray absorption (known as a sinogram) defined for each vector element $[\mathbf{y}]_i = -\log(I_i/I_0)$, $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the forward operator matrix and contains entries $[\mathbf{A}]_{ij} = a_{ij}$ defined by the discrete attenuation model (2), $\mathbf{e} \in \mathbb{R}^m$ is additive noise, and $\mathbf{x} \in \mathbb{R}^n$ is the unknown vector of attenuation coefficients with entries $[\mathbf{x}]_j = x_j$. We obtain a reconstruction of the unknown object by solving the inverse problem (3) for $\mathbf{x}$. More details on the X-ray attenuation model and resulting CT inverse problems can be consulted in [9, Chap. 4.2].

### 2.2. CGLS reconstruction

There are numerous methods for solving (3). Here we focus on the *conjugate gradient least-squares* (CGLS) algorithm [30, Chap. 7], which obtains the least-squares solution by solving

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2. \tag{4}$$

This method is commonly used in CT reconstruction libraries such as ASTRA [31,32] and CIL [33,34], and we will refer to the least-squares solution computed by this method as the CGLS reconstruction.

CGLS is an iterative algorithm and in each iteration one application of $\mathbf{A}$ (forward projection) and its adjoint $\mathbf{A}^T$ (backprojection) are performed. For large-scale problems, CGLS is attractive as it allows for matrix-free application of the forward model and its adjoint, resulting in a computationally efficient algorithm. To obtain the best solution, the algorithm should be stopped at semi-convergence to balance reconstruction sharpness with the increasing amounts of noise introduced with more iterations [9]. Another advantage of the CGLS algorithm is that it eliminates the need for assembling and solving the set of normal equations. This process would otherwise square the conditioning number compared to (3), and make the problem unnecessarily ill-conditioned.

### 2.3. Buildup

When dealing with thick, high-density materials, scattering from other X-rays cannot be ignored, and it becomes necessary to account for a physical phenomenon referred to as *buildup* in the material. Let us consider the X-ray intensity at a specific location within a material. Some of the intensity will be due to photons that have reached the point undisturbed (primary), and some of the intensity will be due to scattered photons (secondary). The buildup is defined as the ratio of the total (primary plus secondary) intensity to the primary intensity. When the buildup effect is present, both primary and secondary intensity reach the detector, resulting in a higher intensity reading compared to the case where only primary intensity reaches the detector.

We can model the buildup action using a so-called *buildup factor*. Assuming a slab of homogeneous material with thickness $T$, the X-ray attenuation through the slab including buildup is described as [35]:

$$I = I_0 \, B \exp\left(-T\alpha_{\mathrm{A}}\right). \tag{5}$$

Here $B$ is the buildup factor, which depends on the X-ray energy, slab thickness, and $\alpha_{\mathrm{A}}$, and its value can be found in [36] for different materials. $\alpha_{\mathrm{A}}$ denotes the material's *linear attenuation coefficient* and it can be computed from the material density $\rho$ and the mass attenuation coefficient $\kappa$, which depends the X-ray energy [37]:

$$\alpha_{\mathrm{A}} = \kappa\rho, \tag{6}$$

and these material constants can be found, e.g. in the NIST database [38].

If buildup is present but ignored, model error is introduced. To quantify this model error, we introduce the *total attenuation coefficient* $\alpha_{\mathrm{tot}}$ to account for both the linear attenuation and the buildup, and we derive an expression that relates $\alpha_{\mathrm{tot}}$ and $\alpha_{\mathrm{A}}$. Consider again the slab of homogeneous material with thickness $T$. When ignoring buildup, the attenuation through the slab described by (1) simplifies to

$$I = I_0 \exp\left(-T\alpha_{\mathrm{tot}}\right). \tag{7}$$

Note, that here it is only the model that neglects buildup, in reality the effect may still be present. Therefore, we use the total attenuation $\alpha_{\mathrm{tot}}$ in (7). Since the intensities $I$ and $I_0$ measured in the slab experiment do not depend on our choice of model, only on the reality of whether or not buildup is present, we can equate the right-hand sides of (5) and (7), and solve for $\alpha_{\mathrm{tot}}$ to obtain:

$$\alpha_{\mathrm{tot}} = \alpha_{\mathrm{A}} - \frac{\log\left(B\right)}{T}. \tag{8}$$

Consequently, if buildup is present but not modelled when solving the CT inverse problem, we expect to reconstruct $\alpha_{\mathrm{tot}}$ as instead of $\alpha_{\mathrm{A}}$. Note, when buildup is negligible, the value of the buildup factor is $B = 1$, and (8) simplifies to $\alpha_{\mathrm{tot}} = \alpha_{\mathrm{A}}$ indicating, that the total attenuation is solely due to linear attenuation.

## 3. Bayesian linear inverse problems

In the Bayesian framework, the linear inverse problem (3) is approached by treating $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{e}$ as random variables. The objective is to find the probability density function (PDF) called the *posterior*, which is the Bayesian solution to the inverse problem [6]. Specifically, we seek a distribution that assigns probabilities to all possible $\mathbf{x}$ given the forward model $\mathbf{A}$ and an observation of the data random variable $\mathbf{y} = \mathbf{y}_{\mathrm{obs}}$. The posterior $p(\mathbf{x} \mid \mathbf{y})$ can be obtained from Bayes' theorem for continuous probability densities:

$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{x}, \mathbf{y})}{p(\mathbf{y})} = \frac{p(\mathbf{y} \mid \mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}. \tag{9}$$

With fixed observed data $\mathbf{y}_{\mathrm{obs}}$, the density $p(\mathbf{y} \mid \mathbf{x})$ is viewed as a function of $\mathbf{x}$ and is referred to as the *likelihood function*, denoted by $L(\mathbf{x} \mid \mathbf{y} = \mathbf{y}_{\mathrm{obs}})$. The PDF $p(\mathbf{x})$ is known as the *prior*

and $p(\mathbf{y})$ is a normalization constant. It is common practice to express the posterior as being proportional to the product of the likelihood function and the prior:

$$p(\mathbf{x} \mid \mathbf{y} = \mathbf{y}_{\text{obs}}) \propto p(\mathbf{x}) L(\mathbf{x} \mid \mathbf{y} = \mathbf{y}_{\text{obs}}). \tag{10}$$

The prior encapsulates our knowledge about the probabilistic properties of the unknown $\mathbf{x}$. The likelihood function quantifies how well $\mathbf{x}$ explains the observed data [39], and it is derived from the forward model and noise. The posterior represents the combined knowledge of the unknown $\mathbf{x}$ after incorporating the observed data and our prior beliefs, and it provides a comprehensive solution to the inverse problem. Importantly, the posterior is a distribution of possible solutions to the inverse problem, enabling the quantification of uncertainty associated with the solution.

### 3.1. Posterior sampling

We can in principle select any distribution for the prior and likelihood. However, for many cases, it is either infeasible or impractical to derive a closed-form expression for the resulting posterior distribution. Even if a closed-form solution is available, it might be difficult to evaluate in large-scale problems. Therefore, in practice, the posterior distribution is often estimated using Markov-chain Monte Carlo (MCMC) sampling.

A posterior sample is a realization from the posterior distribution obtained using a sampling algorithm. To obtain accurate statistical estimators, the algorithm should produce independent samples. Integrated autocorrelation time (IACT) is a metric used to analyse the correlation between the samples. Ideally the IACT has a value of 1, which indicates that the samples are effectively independent [11, p. 87].

Assuming that the computed MCMC samples are representative of the posterior distribution, statistical quantities that characterize the posterior can be computed. Common estimators include: the *posterior mean*, which is the optimal mean-squared error solution, and the *maximum a-posteriori* (MAP), which is the mode of the distribution [40]. It is also possible to estimate the 95% interquantile range of the posterior, defined as the difference between the 2.5% and 97.5% quantiles. This metric provides insight into the uncertainty of the solution by indicating that 95% of possible solutions to the inverse problem are within this range.

### 3.2. Gaussian distributions

A multivariate Gaussian distribution is given by the PDF

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \left( \left| \mathbf{R}^T \mathbf{R} \right| \right)^{1/2} \exp\left( -\frac{1}{2} \|\mathbf{R}(\mathbf{x} - \boldsymbol{\mu})\|_2^2 \right), \tag{11}$$

where $|\cdot|$ denotes the matrix determinant, $\boldsymbol{\mu} \in \mathbb{R}^n$ is the mean, and the product $\mathbf{R}^T\mathbf{R} \in \mathbb{R}^{n \times n}$ represents a full-rank *precision matrix* (inverse of the covariance matrix). We refer to $\mathbf{R}$ as the *square-root precision matrix*. For rank-deficient precision matrices of rank $r$, with $n \geq r > 0$, the so-called *intrinsic* Gaussian distribution is defined with

the PDF [41, Chap. 3]:

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{k/2}} \left( \left| \mathbf{R}^T \mathbf{R} \right|^* \right)^{1/2} \exp\left( -\frac{1}{2} \|\mathbf{R}(\mathbf{x} - \boldsymbol{\mu})\|_2^2 \right), \tag{12}$$

where $|\cdot|^*$ is the generalized determinant given by the product of the $k$ non-zero eigenvalues of the precision matrix.

Gaussian distributions are convenient to work with in the context of Bayesian inverse problems. Most importantly, the product of Gaussian distributions is also Gaussian. This means that choosing a Gaussian prior and likelihood will lead to a Gaussian posterior, that is fully characterized by closed-form expressions for its mean and precision/covariance [6]. We also remark that for Gaussian distributions, the mean and the MAP coincide.

For large-scale Bayesian inverse problems, it is numerically impractical to evaluate the Gaussian posterior precision matrix, while the posterior mean is easier to obtain as we will see in Section 4.4. Despite the mean being a useful estimator, it does not fully capture the potential of the Bayesian approach as it does not convey any information about the associated uncertainty. Indeed, one of the strong points of Bayesian inversion is the ability to give uncertainty estimates. Therefore, MCMC sampling is necessary to estimate posterior uncertainty. However, the closed-form expression for the posterior can be leveraged to develop an efficient sampler based on principles of *randomize-then-optimize* (RTO) sampling [29]; we describe the specifics of our approach in Section 4.4.

## 4. Structural Gaussian priors in Bayesian approach to CT

CT reconstruction is a linear inverse problem and, thus, it can be solved in the Bayesian framework described in Section 3. To perform Bayesian CT reconstruction we must specify a likelihood and prior. Given the high dimensionality of the CT problem, we have limited ourselves to using Gaussian likelihood and priors, which offer the advantage of a closed form expression for the posterior. Although the high dimensionality prohibits direct evaluation of the posterior, the closed form expression can be utilized to formulate an efficient sampler.

In this work, we are motivated by a case where CT imaging is used to inspect subsea pipes. Given prior information about the pipes' internal materials and layered structure, we use the structural Gaussian prior for Bayesian CT reconstruction of a subsea pipe. We remark, however, that the structural Gaussian prior framework is applicable in any imaging problem where the general image structure is approximately known.

This section describes the likelihood and prior that are used to define the Bayesian inverse problem, and explains how the resulting posterior distribution is derived and subsequently sampled.

### 4.1. Likelihood

Assuming the use of high-intensity X-rays in the data collection process, the data noise is approximated as additive Gaussian. Specifically, $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \lambda^{-1}\mathbf{I}_n)$, where $\lambda$ is a precision parameter and $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ is an identity matrix. Based on the CT problem defined in (3)

and the additive Gaussian noise assumption, we derive the likelihood function,

$$L(\mathbf{x} \mid \mathbf{y} = \mathbf{y}_{\text{obs}}) = \left(\frac{\lambda}{2\pi}\right)^{m/2} \exp\left(-\frac{\lambda}{2} \left\| \mathbf{A}\mathbf{x} - \mathbf{y}_{\text{obs}} \right\|_2^2\right). \tag{13}$$

### 4.2. Structural Gaussian priors

We must specify a prior in order to perform Bayesian CT reconstruction. In this work, the fundamental idea is to use prior knowledge of the pipe's layered structure and the materials comprising it, and incorporate this information into what we call *structural Gaussian priors* (SGPs). The SGP combines (via a stacking operation) priors for each of the materials in the pipe, thereby enforcing local priors in different regions. In addition to the pipe material and geometry information, the SGP also contains a global prior on the pixel differences. The purpose of this is to add a global smoothness assumption on the image.

To create the SGP, we require two components: (i) masks for each pipe layer and the image background, and (ii) prior estimates of the linear attenuation coefficients of the layer materials. In the following, we discuss the available prior information, how to define it in terms of distributions, and how to assemble the SGP.

#### 4.2.1. Layered structure of pipes

Assume the general structure of the image to be reconstructed is known and defined by approximate boundaries between different materials. Note, that no specific geometry for the approximate material boundaries are required by the SGP. They can for instance be specified by circles, ellipses, rectangles, functions, or even manually drawn. The only requirements are that regions between the boundaries are not overlapping, and that we can define image masks as described below based on the boundaries. Specific to our problem, we know subsea pipes are structured such that they consist of layers of different materials, and we assume that we can define (not necessarily concentric) circular, approximate boundaries between the materials in the pipe as illustrated in the middle image in Figure 1.

The general structure of the image to be reconstructed is included in the prior using binary image array masks indicating non-overlapping regions of different materials. The masks are defined based on the approximate boundaries, but made slightly smaller, such that there is no specific local prior near the approximate boundaries. This is done to minimize the risk of misclassifying pixels near the layer boundaries. The masks used in our application are illustrated to the right in Figure 1. The coloured regions represent a mask each, where specific local priors are used, while the white regions represent boundary regions with no local priors. Note, that steel rods holding the pipe in place during the laboratory CT scan are included in the boundary regions, because these would not be present in a subsea scan of a pipe in operation. Alternatively, assuming that we know their approximate locations, we could also have imposed local priors here.

The masks are given by binary vectors $\mathbf{m}_k \in \mathbb{R}^n$, where $k = 1, \ldots, p$ denote the region number. The value of an element is

$$[\mathbf{m}_k]_j = \begin{cases} 1 & \text{if pixel } j \text{ is in region } k \\ 0 & \text{otherwise} \end{cases}, \quad \text{for } j = 1, \ldots, n, \tag{14}$$

where $[\mathbf{m}_k]_j$ denotes the $j$'th element of the mask vector $\mathbf{m}_k$.

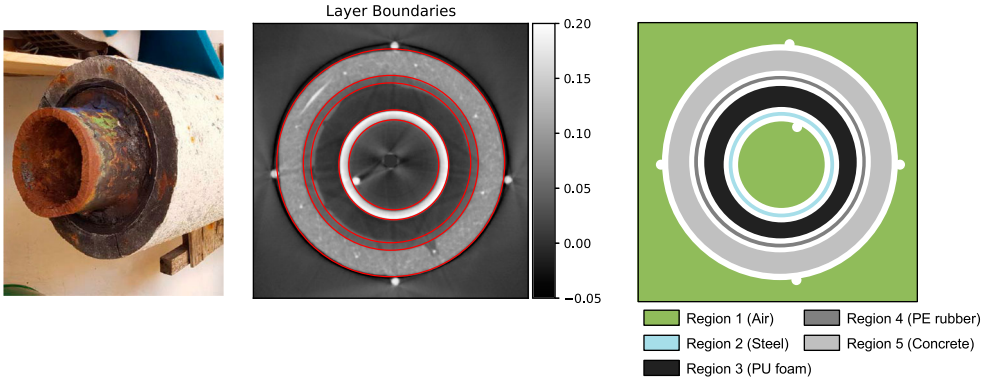



**Figure 1.** *Left:* Real pipe with four layers of different materials and a background of air. Photograph first published in [4]. © IOP Publishing. Reproduced with permission. All rights reserved. *Middle:* Reconstruction computed using CGLS with approximate layer boundaries marked by red circles. *Right:* SGP masks for $k = 1, \ldots, 5$ used to form local priors that each represent a material. The white regions do not have any local prior.

For inspection of pipes in operation, it is reasonable to assume that general information about the internal pipe structure is available, for instance from CAD drawings from the manufacturer or previous CT scans of the pipe with high-quality data. In cases where no information about the position of the pipe layer boundaries is readily available, it is possible to estimate such locations even from low-quality data. We can for instance fit circles (or ellipses) in a pre-processing step using the approach in [42]. In this work, we defined the SGP masks using circular, approximate layer boundaries. The centre coordinates and radii defining these boundaries where chosen based on a CGLS reconstruction of high-quality CT data of the real pipe and (Figure 1 left). Finally, it should be noted that the better we know the internal structure, the more informative a prior we can construct. However, if the general structure is only partially known, for instance if only the external dimensions of the pipe is known, we can still formulate a prior in the SGP framework, it will just be less informative.

#### *4.2.2. Linear attenuation coefficients of pipe materials*

Given a known homogeneous material and known X-ray energy, we can estimate its attenuation coefficient based on physics using material constants as described in Section 2.3. Note, that this is a fundamentally different from solving the CT inverse problem. Therefore, we can utilize this to compute a prior mean for the attenuation coefficients of the pipe materials in the Bayesian inverse problem. That is, for each pipe material we can compute expected attenuation coefficients $\alpha_k$, $k = 1, \ldots, p$, and use them as means in the SGP.

Recall, that our inverse problem assumed a X-ray attenuation model without buildup. For most materials this is a reasonable assumption and we can determine the prior mean for a given material as $\alpha_k = \alpha_{\mathrm{A}}$, where $\alpha_{\mathrm{A}}$ is given as in (6). For thick, high-density materials, however, buildup is not negligible and model error occurs in the inverse problem. This model error leads to the reconstructed attenuation coefficients $\mathbf{x}$ of such materials being underestimated compared to their true linear attenuation coefficients $\alpha_{\mathrm{A}}$, because secondary intensity is measured but not included in the attenuation model. Correcting

the error in the forward model would complicate the inverse problem. Instead, we accept the model error, and expect that the reconstructed attenuation coefficients $\mathbf{x}$ in thick, high-density materials correspond to total attenuation coefficients $\alpha_{\text{tot}}$ that represent both linear attenuation and buildup. Consequently, for such materials, we must compute a value for $\alpha_{\text{tot}}$ to obtain a prior mean. In Section 2.3 the model error was quantified for a slab material, and an expression (8) was found for $\alpha_{\text{tot}}$. Making the slightly crude assumption that the expression in (8) transfers to other material geometries, we estimate the prior mean for a thick, high-density material as $\alpha_k = \alpha_{\text{tot}}$.

#### 4.2.3. Assembling the SGP

We assemble the SGP from a set of Gaussian distributions with PDF's denoted $p_k(\mathbf{x})$ and given by their means $\boldsymbol{\mu}_k$ and square-root precision matrices $\mathbf{R}_k$ for $k = 0, \ldots, p$. This includes a number of local priors with information about specific attenuation coefficients in specific regions ($k = 1, \ldots, p$), and a global smoothness prior ($k = 0$). These priors are specified below:

Local priors: The independent and identically distributed (IID) Gaussian prior pushes posterior pixel values towards the prior mean. We use this property to express prior information about expected linear attenuation coefficients $\alpha_k$ in the $p$ different materials. The prior is defined as an intrinsic Gaussian (12) parameterized by mean and square-root precision:

$$\boldsymbol{\mu}_k = \alpha_k \mathbf{1} \quad \text{and} \quad \mathbf{R}_k = \sqrt{\delta_k}\mathbf{M}_k, \quad \text{for } k = 1, \ldots, p, \tag{15}$$

where $\mathbf{M}_k \in \mathbb{R}^{n \times n}$ is a diagonal matrix with $\mathbf{m}_k$, defined in (14), on the diagonal, $\mathbf{1} \in \mathbb{R}^n$ is a vector of ones, and $\delta_k$ is a prior precision parameter that controls the strength of the prior. In summary, the prior $p_k(\mathbf{x})$ expresses that we expect to find a material with linear attenuation coefficient $\alpha_k$ in the region defined by mask $\mathbf{m}_k$.

Global prior: The Gaussian Markov random field (GMRF) prior is a smoothing prior that promotes neighbouring pixels to take the same value. The pipe materials can be considered roughly homogeneous and therefore the GMRF is used as a global prior on the entire image domain. We index this prior with $k = 0$, and in this case the prior distribution is multivariate Gaussian (11) with mean and square-root precision:

$$\boldsymbol{\mu}_0 = \mathbf{0}, \quad \text{and} \quad \mathbf{R}_0 = \sqrt{\delta_0}\begin{bmatrix}\mathbf{D}_1 \\ \mathbf{D}_2\end{bmatrix}, \tag{16}$$

where $\mathbf{0} \in \mathbb{R}^n$ is a vector of zeros, $\mathbf{D}_1 = \mathbf{I}_N \otimes \mathbf{D}$, $\mathbf{D}_2 = \mathbf{D} \otimes \mathbf{I}_N$, $\mathbf{D} \in \mathbb{R}^{(N+1)\times N}$ is a finite difference matrix with backward differences and zero Dirichlet boundary conditions, and $\otimes$ denotes the Kronecker product. Note that $\mathbf{R}_0$ has full column rank, so the precision matrix $[\mathbf{R}_0^T\mathbf{R}_0]$ also has full rank, which confirms that the PDF $p_0(\mathbf{x})$ is multivariate Gaussian.

The SGP is constructed under the assumption that the priors $p_k(\mathbf{x})$ are independent. Then, the SGP is defined as a product of the individual prior PDF's:

$$p(\mathbf{x}) = \prod_{k=0}^{p} p_k(\mathbf{x}). \tag{17}$$

Since each $p_k$ is by definition Gaussian, the full SGP is also Gaussian with mean and square-root precision given by

$$\mathbf{R}_{\text{SGP}} = \begin{bmatrix} \mathbf{R}_0 \\ \vdots \\ \mathbf{R}_p \end{bmatrix} \quad \text{and} \quad \boldsymbol{\mu}_{\text{SGP}} = \left[\mathbf{R}_{\text{SGP}}^T \mathbf{R}_{\text{SGP}}\right]^{-1} \sum_{k=0}^{p} \mathbf{R}_k^T \mathbf{R}_k \boldsymbol{\mu}_k. \tag{18}$$

We remark that the full column rank of $\mathbf{R}_0$ leads to $\mathbf{R}_{\text{SGP}}$ also having full column rank. Therefore, the prior precision matrix $[\mathbf{R}_{\text{SGP}}^T \mathbf{R}_{\text{SGP}}]$ has full rank and is invertible.

#### 4.2.4. SGP configurations

The SGP should be configured according to the information available. In this work, we apply the SGP in Bayesian CT reconstruction of subsea pipes. Given the pipe illustrated in Figure 1 we can potentially have information about up to $p = 5$ image regions (four pipe layers and the background), where we can use the local priors. We define SGP configurations for three possible scenarios, with increasing levels of information:

- GMRF: Assuming we have no structural information, we can only apply the global smoothness prior: $p(\mathbf{x}) = p_0(\mathbf{x})$.
- SGP-BG: Assuming we have information about the pipe's external dimensions but not internal structure, we use the local prior on the air background ('BG') and global smoothness prior such that $p(\mathbf{x}) = \prod_{k=0}^{1} p_k(\mathbf{x})$.
- SGP-F: Assuming we approximate information about the full ('F') internal pipe structure, we apply local priors in all $p = 5$ pipe layers and the global prior for smoothness such that $p(\mathbf{x}) = \prod_{k=0}^{5} p_k(\mathbf{x})$.

### 4.3. Posterior PDF

Since the prior and likelihood are Gaussian, we can obtain a closed-form expression for the posterior, which is also Gaussian [6]. Using the likelihood in (13) and the prior in (17)–(18), the posterior PDF is given by Bayes' Theorem (10) as:

$$\begin{aligned} p(\mathbf{x} \,|\, \mathbf{y} = \mathbf{y}_{\text{obs}}) &\propto \left(\tfrac{\lambda}{2\pi}\right)^{\frac{m}{2}} \exp\left[-\tfrac{\lambda}{2} \left\|\mathbf{A}\mathbf{x} - \mathbf{y}_{\text{obs}}\right\|_2^2\right] \\ &\quad \times \prod_{k=0}^{p} \left(\left|\tfrac{\mathbf{R}_k^T \mathbf{R}_k}{2\pi}\right|^*\right)^{\frac{1}{2}} \exp\left[-\tfrac{1}{2}\|\mathbf{R}_k(\mathbf{x} - \boldsymbol{\mu}_k)\|_2^2\right] \\ &\propto \exp\left[-\frac{\lambda}{2} \left\|\mathbf{A}\mathbf{x} - \mathbf{y}_{\text{obs}}\right\|_2^2 - \frac{1}{2} \sum_{i=0}^{p} \|\mathbf{R}_k(\mathbf{x} - \boldsymbol{\mu}_k)\|_2^2\right]. \end{aligned} \tag{19}$$

The above expression defines the Gaussian random vector

$$\mathbf{x} \,|\, \mathbf{y} = \mathbf{y}_{\text{obs}} \sim \mathcal{N}(\boldsymbol{\mu}_{\text{post}}, (\mathbf{R}_{\text{post}}^T \mathbf{R}_{\text{post}})^{-1}), \tag{20}$$

where the posterior square-root precision is given by

$$\mathbf{R}_{\text{post}} = \begin{bmatrix} \sqrt{\lambda}\mathbf{A} \\ \mathbf{R}_0 \\ \vdots \\ \mathbf{R}_p \end{bmatrix} \in \mathbb{R}^{q \times n} \quad \text{with } q = \left(m + 2N\,(N+1) + pn\right), \tag{21}$$

and the posterior mean is given implicitly by

$$\mathbf{R}_{\text{post}}^T \mathbf{R}_{\text{post}} \boldsymbol{\mu}_{\text{post}} = \lambda \mathbf{A}^T \mathbf{y}_{\text{obs}} + \sum_{k=0}^{p} \mathbf{R}_k^T \mathbf{R}_k \boldsymbol{\mu}_k. \tag{22}$$

Since $\mathbf{R}_0$ and thus $\mathbf{R}_{\text{post}}$ have full column rank, the posterior precision is invertible, and the posterior distribution is Gaussian with well-defined covariance. That is, the closed-form expression for the posterior distribution is (20), which can be seen by factoring out (19) and reordering the equation into the form (11).

### 4.4. Posterior sampling

The large-scale nature of CT makes it impractical to evaluate the closed form expression of the posterior distribution, because the matrix $\mathbf{A}$ is impractical to form explicitly. We can, however, use by matrix-free implementations of the application of $\mathbf{A}$ and its adjoint. This implies that the posterior mean can be obtained by solving (22), for instance using the CGLS algorithm. However, to compute the posterior covariance $[\mathbf{R}_{\text{post}}^T \mathbf{R}_{\text{post}}]^{-1}$, which contains information about the uncertainty in the posterior, we must form $\mathbf{A}^T\mathbf{A}$, which is numerically impractical. To obtain an uncertainty estimate related to the posterior, we instead sample the posterior as described below, using an RTO-based method that only relies on applications of $\mathbf{A}$ and its adjoint, without ever forming them explicitly. The samples are then used to compute the 95% interquantile range that represents a metric for uncertainty in the posterior.

We can efficiently sample the posterior (20) without explicitly forming the matrix $\mathbf{A}$ or its adjoint using an RTO-based method. We generate one realization $\mathbf{x}^*$ of the random variable $\mathbf{x} \,|\, \mathbf{y} = \mathbf{y}_{\text{obs}}$ by solving the following linear system of equations, and we let $\mathbf{x}^*$ represent a posterior sample.

$$\mathbf{R}_{\text{post}} \mathbf{x}^* = \begin{bmatrix} \sqrt{\lambda}\mathbf{y}_{\text{obs}} \\ \mathbf{R}_0 \boldsymbol{\mu}_0 \\ \vdots \\ \mathbf{R}_p \boldsymbol{\mu}_p \end{bmatrix} + \boldsymbol{\xi}^*, \tag{23}$$

where $\boldsymbol{\xi}^*$ denotes a realization of the random variable $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_q)$. Note, that a new realization of $\boldsymbol{\xi}^*$ is required for each posterior sample. We use CGLS to obtain the least-squares solution, in a numerically efficient way. Algorithm 1 summarizes the sampling procedure.

**Algorithm 1:** Sampling of the posterior

Initial states: $\mathbf{x}^{(0)}$;
**for** $t = 1, \ldots, n_s$ **do**
  /* Step 1: */
  Obtain a sample from $\mathbf{x} \mid \mathbf{y} = \mathbf{y}_{\text{obs}}$ by solving (23) for $\mathbf{x}^*$ using CGLS with $n_{\text{CGLS}}$ iterations and warm-start, i.e. initial point is $\mathbf{x}^{(t-1)}$;
  $\mathbf{x}^{(t)} \leftarrow \mathbf{x}^*$;

If the CGLS algorithm is run until convergence when solving (23), we obtain independent samples of the posterior (20). However, in order to reduce computation time, we apply warm-starting using the previous posterior sample as a starting guess for the next one. We also fix the number of CGLS iterations to $n_{\text{CGLS}}$. This approach makes a burn-in phase necessary, i.e. some initial samples are discarded. Furthermore, the independence of the samples might be compromised, since we do not check for convergence. Instead, we will monitor the IACT of the chains to check for potential correlations.

Our sampling method is a linear RTO method, because we randomize the system of linear equations by adding $\boldsymbol{\xi}^*$ and optimize for $\mathbf{x}^*$ by solving the system in a least-squares sense. We remark that, while we have seen linear RTO sampling applied in cases where the prior has zero mean (e.g. [8,11]), to the best of our knowledge, the linear RTO method has not been applied in cases with a non-zero prior mean before. Therefore, we will show that samples $\mathbf{x}^*$ obtained by the least-squares solution to (23), are in fact distributed according to the posterior.

To show that $\mathbf{x}^*$ is distributed according to the posterior, we first form the normal equations, i.e.

$$\left(\lambda \mathbf{A}^T\mathbf{A} + \sum_{k=0}^{p} \mathbf{R}_k^T\mathbf{R}_k\right)\mathbf{x}^* = \lambda \mathbf{A}^T\mathbf{y}_{\text{obs}} + \sum_{k=0}^{p} \mathbf{R}_k^T\mathbf{R}_k\boldsymbol{\mu}_k + \mathbf{R}_{\text{post}}^T\boldsymbol{\xi}. \tag{24}$$

Inserting (21) and solving for $\mathbf{x}^*$ we find

$$\begin{aligned}\mathbf{x}^* &= \left(\mathbf{R}_{\text{post}}^T\mathbf{R}_{\text{post}}\right)^{-1}\left(\lambda \mathbf{A}^T\mathbf{y}_{\text{obs}} + \sum_{k=0}^{p} \mathbf{R}_k^T\mathbf{R}_k\boldsymbol{\mu}_k\right) + \left(\mathbf{R}_{\text{post}}^T\mathbf{R}_{\text{post}}\right)^{-1}\mathbf{R}_{\text{post}}^T\boldsymbol{\xi} \\ &= \boldsymbol{\mu}_{\text{post}} + \mathbf{R}_{\text{post}}^{\dagger}\boldsymbol{\xi},\end{aligned} \tag{25}$$

where $\mathbf{R}_{\text{post}}^{\dagger} = (\mathbf{R}_{\text{post}}^T\mathbf{R}_{\text{post}})^{-1}\mathbf{R}_{\text{post}}^T$ is the Moore–Penrose pseudoinverse. As $\mathbf{R}_0$ has full column rank, $\mathbf{R}_{\text{post}}^{\dagger}$ also has full row rank. Note that $\mathbf{x}^*$ is a linear transformation of the Gaussian random variable $\boldsymbol{\xi}$. It follows that $\mathbf{x}^*$ is Gaussian distributed with mean and covariance [43]:

$$\mathrm{E}(\mathbf{x}^*) = \boldsymbol{\mu}_{\text{post}} + \mathbf{R}_{\text{post}}^{\dagger}\mathbf{0} = \boldsymbol{\mu}_{\text{post}} \tag{26a}$$

$$\mathrm{Cov}(\mathbf{x}^*) = \mathbf{R}_{\text{post}}^{\dagger}\mathbf{I}(\mathbf{R}_{\text{post}}^{\dagger})^T = \left(\mathbf{R}_{\text{post}}^T\mathbf{R}_{\text{post}}\right)^{-1}. \tag{26b}$$

This confirms that $\mathbf{x}^*$ is distributed according to the target posterior distribution. That is, by solving (23) we obtain samples of CT reconstruction images distributed according to (20).

## 5. Numerical experiments

FORCE Technology (https://forcetechnology.com/) has provided CT data from a deep-sea oil pipe acquired in a laboratory experiment. We conduct synthetic- and real-data CT experiments to assess the performance of the proposed SGPs for obtaining useful CT reconstructions and relevant uncertainty estimates. We interpret the posterior mean as the CT reconstruction and the posterior 95% interquantile ranges for each pixel as associated uncertainty estimates. The results are compared to a CGLS reconstruction, which is a conventional deterministic reconstruction method, where no prior information is used.

### 5.1. Data description

Deep sea oil pipes are large and consist of dense materials including steel. In order to penetrate the pipe with X-rays, it is necessary to use high power and a narrow beam source. The narrow beam cannot illuminate the full pipe, but an exterior tomography setup can be used. In this case, the X-ray source is offset from the pipe centre, which is also the rotation axis. It can be shown that this setup is ideal, if we are interested in recovering the pipe layer structure and defects therein rather than the centre of the pipe [4]. We use real data acquired from a prototype set-up of an oil pipe CT scanner. A linear detector panel with 510 sensors of sizes 0.8 mm each was used, and 360 equi-angular projections were recorded in a full 360 degree rotation. The pipe, its specifications, and the offset fan-beam acquisition geometry are shown in Figure 2.

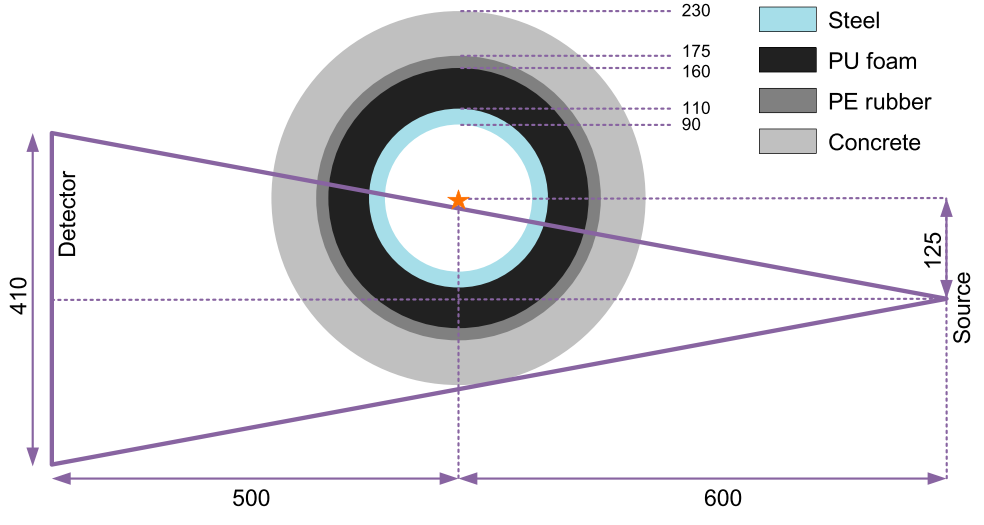



**Figure 2.** Pipe and offset scan geometry specifications; dimensions in mm. The CT rotation axis is marked with a star.

### 5.1.1. Phantom

Based on the real pipe structure (Figure 1) and material constants (Table 1), we also construct a synthetic phantom and corresponding SGP mask to perform simulation experiments (see Figure 3). To avoid inverse crime, the phantom is defined on a $1024 \times 1024$ pixel image and the reconstruction is performed on a $500 \times 500$ pixel image that represents a $55 \times 55$ cm domain leading to pixels of size $1.1 \times 1.1$ mm.

We add steel inclusions in the phantom's concrete layer to simulate steel reinforcement bars in the real pipe. The inclusions are in the radial or tangential directions and their widths are varied from 2–7 mm. These inclusions will not be accounted for in the SGP prior; this in order to test the prior's ability to reconstruct unexpected features. Furthermore, with the offset-scan geometry, no X-rays will be parallel to near-radial edges, and therefore we expect that the radial inclusions are more difficult to reconstruct than the tangential (see, e.g. [4] for more details).

### 5.1.2. Sinograms

Using the pipe phantom, offset-scan geometry, and 2% noise level (i.e. $\|\mathbf{e}\|_2/\|\mathbf{Ax}\|_2 \sim 2\%$), we simulate a sinogram with 360 view angles distributed equally over a full rotation. Assuming the real sinogram has an equivalent noise distribution, we compare the synthetic and real data (Figure 4). Five bands in the sinograms mark the centre region, steel, PE, PU, and concrete layers in that order from the bottom of the sinograms. The bands in the real-data sinogram follow sinusoidal curves because the pipe centre did not coincide with the rotation centre, and the phases of the sinusoidal bands vary because the pipe layers are not completely concentric. The overall similarity between the real- and synthetic-data

**Table 1.** Best estimates for the material constants at beam intensity 2 MeV.

| $k$ | Material | $\kappa$[cm$^2$/g] | $\rho$[g/cm$^3$] | $B$ | $T$[cm] | $\alpha_k$[cm$^{-1}$] |
|---|---|---|---|---|---|---|
| 1 | Air | 0.044 | 0.0012 | – | – | $\sim 0$ |
| 2 | Steel[a] | 0.042 | 7.9 | 2.013 | 4 | 0.16 |
| 3 | PU foam | 0.051[b] | 0.15 | – | – | 0.0077 |
| 4 | PE rubber | 0.051 | 0.94 | – | – | 0.048 |
| 5 | Concrete | 0.046 | 2.3 | – | – | 0.11 |

[a]Subject to build-up.[b]Estimated by PE rubber's $\kappa$ due to missing value in database.

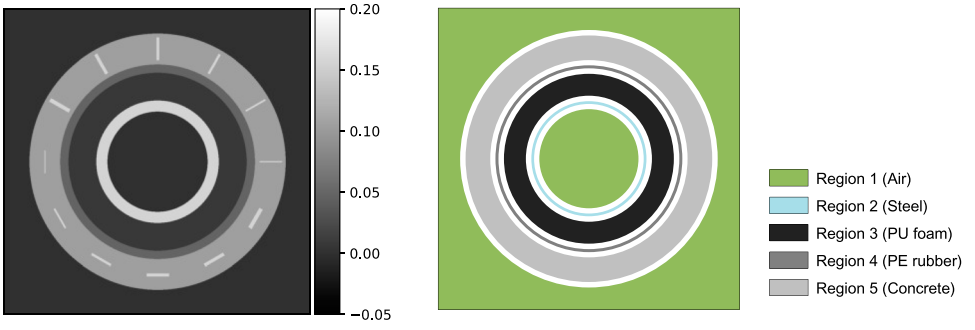



**Figure 3.** *Left:* Pipe phantom modelling the real pipe. *Right:* SGP masks for $k = 1, \ldots, 5$ used to form local priors that each represent a material with attenuation coefficient $\alpha_k$ given in Table 1. The white regions do not have any local prior.

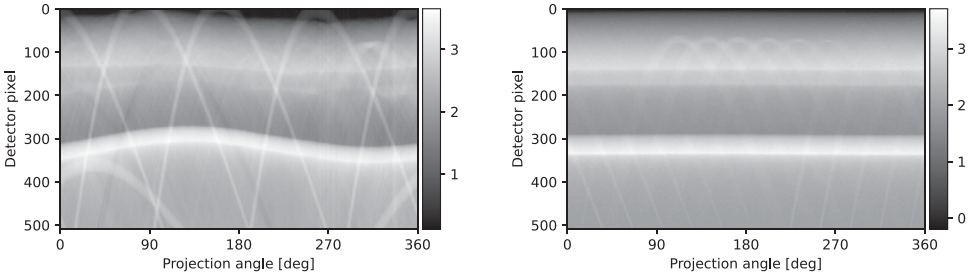



**Figure 4.** *Left:* Real subsea pipe sinogram. *Right:* Synthetic sinogram with 2% added noise.

sinograms indicates that our phantom is realistic and the scan geometry close to the real setup. For sparse view angle experiments, we simulate data with only 50%, 20% or 10% of the total view angles in the sinograms. Note that, since we reconstruct on a $500 \times 500$ pixel image, the CT problem is under-determined in all our experiments, i.e. there are fewer rows than columns in the forward operator matrix $\mathbf{A}$.

## 5.2. Implementation details and parameter choices

Forming and sampling the posterior in (19) requires choosing several parameters. In the sections below we specify our parameter choices, including SGP means and precision parameters, likelihood precision parameter, and parameters related to the sampling.

### 5.2.1. Specification of SGP local means

To specify the local means in the SGP, we use the fact that the pipe consists of four known materials and that we know the internal structure approximately from a CGLS reconstruction of high-quality data as described in Section 4.2.1. Specifically, for real-data experiments, we use the masked regions illustrated in Figure 1, where there is steel in region 2, polyurethane (PU) foam in region 3, polyethylene (PE) rubber in region 4, and concrete in region 5. The pipe is scanned with air as the background and this is denoted region 1. Similarly, we use the masked regions shown in Figure 3 for the synthetic-data experiments.

For each material, we compute the SGP mean value $\alpha_k$ as described in Section 4.2.2, using material constants from the NIST database [38] and buildup factor from [36], and assuming a monochromatic beam of energy 2 MeV. The beam energy is in fact a spectrum with 2 MeV mean energy, but around these energies the mass attenuation coefficients do not vary significantly, and therefore we expect that the monochromatic beam assumption is reasonable. The specific values used in the numerical experiments are given in Table 1. For air, concrete, polyurethane, and polyethylene, it is assumed buildup is negligible, so their respective $\alpha_k = \alpha_{\mathrm{A}}$ and they are computed according to (6). For the much denser material steel, however, buildup can not be neglected, so its $\alpha_k = \alpha_{\mathrm{tot}}$ and it is computed using (8) with the buildup factor given in Table 1.

### 5.2.2. Specification of likelihood and SGP precision parameters

We assume data noise of 2% corresponding approximately to a likelihood precision parameter $\lambda = 400$.

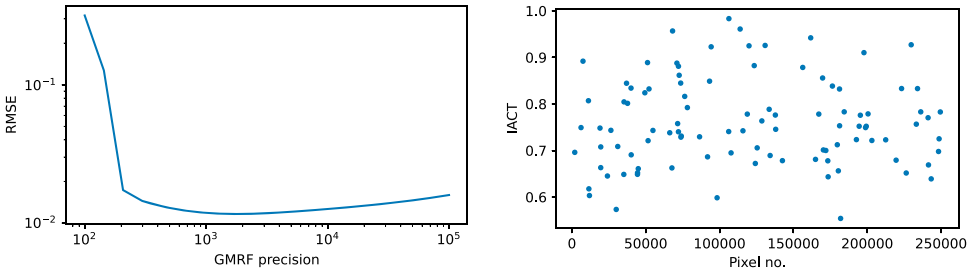



**Figure 5.** *Left:* Example of GMRF precision parameter sweep from the synthetic experiment. From this, we choose $\delta_0 = 1000$ near the minimum. *Right:* Examples of integrated autocorrelation time (IACT) for 100 randomly selected pixel chains.

For the SGP, the precision parameters for the local priors are chosen heuristically such that they cover reasonable intervals around the estimated $\alpha$, representing some uncertainty in the estimated means. We expect steel reinforcement bars in the concrete layer, which are not modelled in the prior, so the precision parameter is reduced here to allow a larger range of possible values. Hence, the chosen precision parameters are: $\delta_1 = \delta_2 = \delta_3 = \delta_4 = 1000$ and $\delta_5 = 500$. The final parameter choice is the global GMRF precision $\delta_0$. If we fix all other precision parameters, we can perform a parameter sweep for $\delta_0$ values and choose the value that minimizes the root-mean-square-error (RMSE) with respect to the synthetic ground truth. An example of $\delta_0$ tuning can be seen in Figure 5 (left). We transfer the GMRF precision choices from the synthetic-data to the real-data case, and check qualitatively that the corresponding reconstructions are acceptable.

#### 5.2.3. Sampling

We sample from the posterior using Algorithm 1. In all numerical experiments below, we fix $n_{\text{CGLS}} = 10$ and run Algorithm 1 for $n_s = 3000$ iterations and discard the first 1000 as burn-in. This leaves us with a sample size of 2000.

### 5.3. Results of synthetic-data experiments

#### 5.3.1. Example of Bayesian inversion with SGP prior

We consider a sparse-angle case with only 72 equi-angular projections out of the original 360 (i.e. 20%). We employ the SGP-F prior in the Bayesian inverse problem. First, we tune the global GMRF precision parameter as described above in Section 5.2.2. Figure 5 (left) illustrates the parameter sweep from which we choose $\delta_0 = 1000$ since the RMSE is near a minimum. Then we obtain 2000 samples from the posterior as described in Section 5.2.3. Since the reconstructed parameter $\mathbf{x}$ is high-dimensional, we compute the IACT at 100 randomly selected pixel chains and consider this as representative for the full parameter. Figure 5 (right) shows that IACT is close to or below one in all cases, indicating that the samples are quasi-independent.

Figure 6 shows the posterior mean, the 95% interquantile range, and 1D slices through the mean and 95% interquantile range. We observe that the pipe layers are well reconstructed. However, the contrast of the steel inclusions in the concrete layer are dampened due to the choice of prior. The prior promotes the expected linear attenuation coefficient

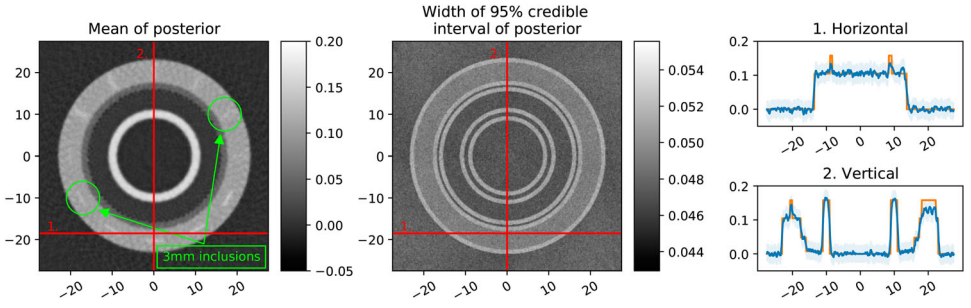



**Figure 6.** Synthetic test case with 20% view angles and SGP-F prior. *Left, center:* 2D images of the posterior mean and the 95% interquantile range. *Right:* 1D slices of posterior mean (blue, noisy line) and 95% interquantile range (blue shade), compared to ground truth (orange, piecewise constant line).

of the concrete, which is lower than that of steel. We also note higher noise in the concrete layer than in the remaining parts of the reconstruction due to $\delta_5 < \delta_{1,\ldots,4}$. The 95% interquantile range is a direct reflection of the SGP prior action. We see that imposing local priors in the pipe layers lowers the uncertainty estimate of the posterior compared to boundary areas without local priors on them. Furthermore, priors with high precision values lead to lower uncertainty in the posterior. As expected, we also see that the radial steel inclusions in the concrete layer are more difficult to reconstruct than the tangential inclusions. For instance, the 3 mm radial inclusion is barely distinguishable, while the 3 mm tangential inclusion is fairly clear.

### *5.3.2. Comparative study*

We investigate the effect of the SGP when reducing the number of view angles. Figure 7 shows posterior means from 16 experiments. In the figure, the number of view angles decreases by rows, and the prior information increases by columns. We observe that, as we add more structural information, the reconstructions improve in terms of (1) reduced streak artifacts, (2) increased layer contrast, (3) reduced RMSE, and (4) detection of the steel inclusions that are hidden in artifacts. The action of the prior is especially clear for 10% and 20% view angles, indicating that the structural priors become more effective in sparse view angle cases.

The deterministic CGLS reconstruction with 360 view angles produces an image where all the main features of the pipe are visible. However, we also see artifacts: (1) along the radial steel inclusions, because these edges have no tangential rays, (2) in the centre of the image, because there is no data, and (3) a dark ring just outside of the pipe, which is a typical artifact that appears at the fan-beam edge. As we decrease the number of view angles, we get more dominating artifacts from the fan-beam edges. When we include the GMRF prior, the artifacts are smoothed out, but they are still very dominating in the image. Thus, this standard prior is not sufficient to produce good reconstructions, especially for sparse view angle settings. Alternatively, when using the SGP-BG prior, which is more informative, we manage to reduce the artifacts in the background substantially. Furthermore, we see increased contrast between pipe materials for the 10% and 20% view angle cases. This is due to the reduced precision parameter of the smoothing GMRF prior, made possible by the added information from the SGP-BG prior. When we use the most informative SGP-F prior, we

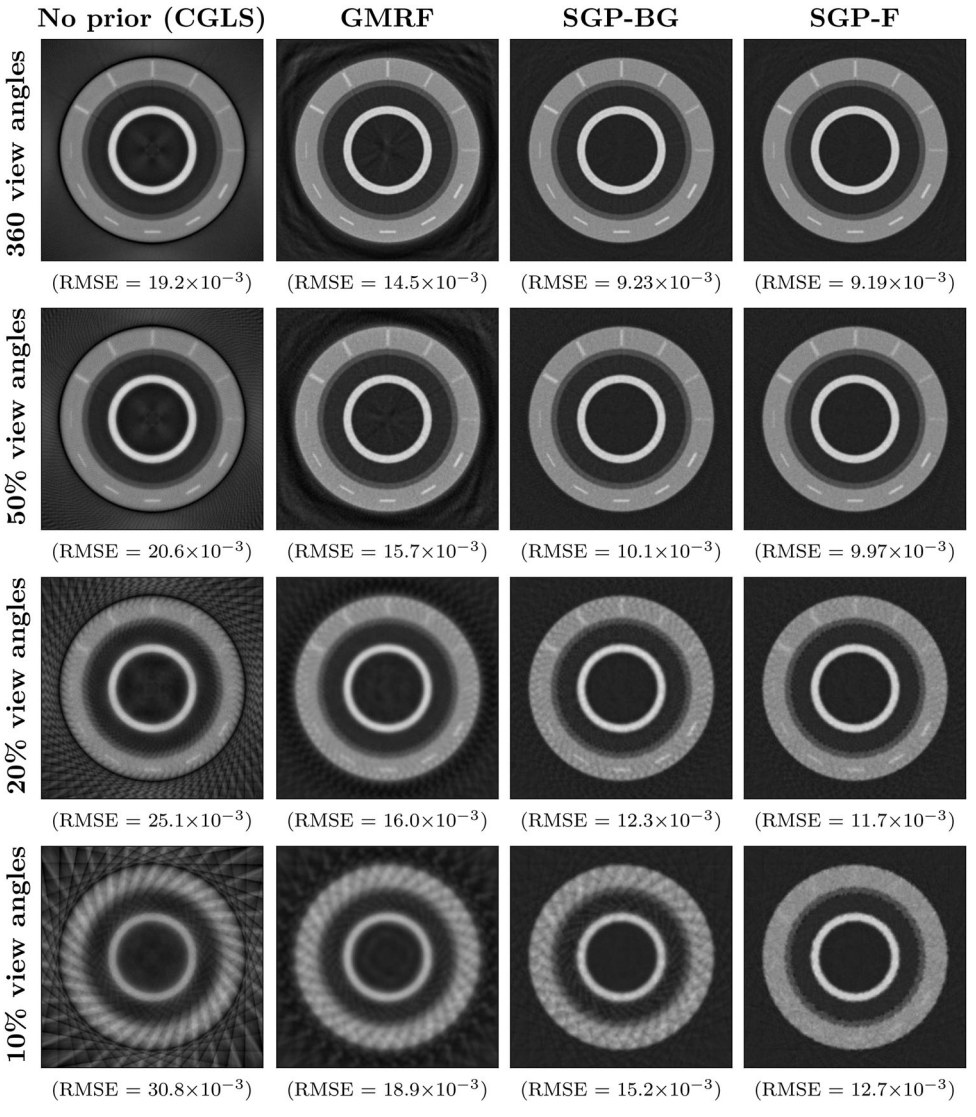



**Figure 7.** Posterior means and RMSE using synthetic data. The columns from left to right use increasingly informative priors: no prior (obtained as a CGLS reconstruction), Gaussian Markov Random Field (GMRF) prior, structural Gaussian prior on the background (SPG-BG), and structural Gaussian prior on the full image domain (SGP-F). The rows from top to bottom use decreasing numbers of view angles: 360 (or 100%), 50%, 20%, and 10% view angles. The colour range is [−0.05, 0.20] for all images.

can suppress artifacts in the background and pipe layers. In the concrete layer, however, we note that unexpected inclusions (here steel reinforcements) that are not modelled in the prior, are also removed. Recall, that we did not impose any local prior in boundary areas, and therefore the artifacts are still present there. This we see as a 'ragged' boundary, especially for sparse view angle cases.

We see that, with the SGP-F prior, we can resolve tangential steel inclusions down to widths of 2 mm using only 20% of the view angles. The radial inclusions are well

reconstructed down to 4 mm. The smaller 2-3 mm radial inclusions are still visible but faint and blurred. Compared to the deterministic reconstruction with no prior, or to the reconstruction with a GMRF prior alone, the results still show an improvement. Note that, in those cases, the narrow radial steel inclusions disappear among the artifacts. Moreover,

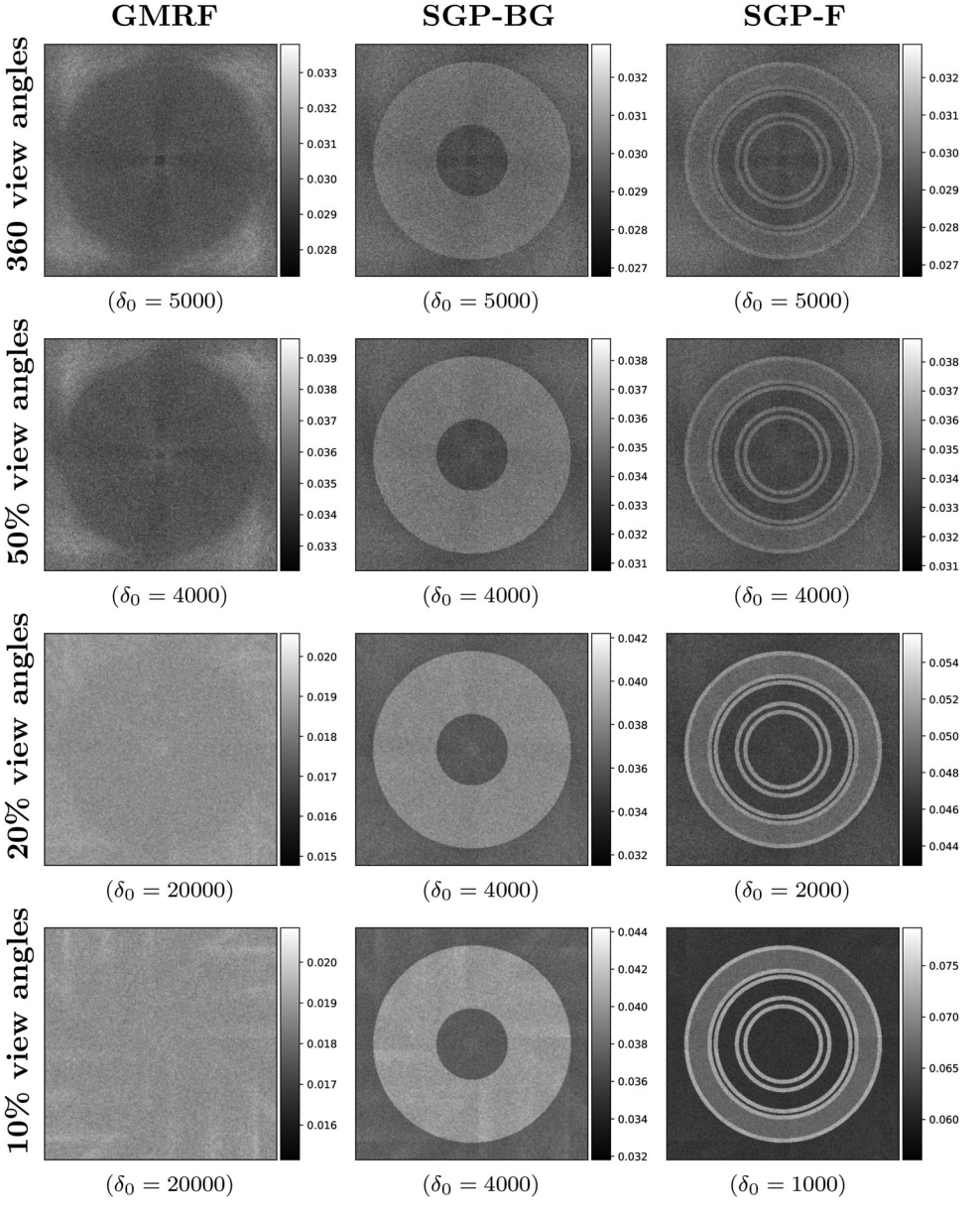



**Figure 8.** Posterior 95% interquantile ranges using synthetic data. The columns from left to right use increasingly informative priors: Gaussian Markov Random Field (GMRF) prior, structural Gaussian prior on the background (SPG-BG), and structural Gaussian prior on the full image domain (SGP-F). The rows from top to bottom use decreasing numbers of view angles: 360 (or 100%), 50%, 20%, and 10% view angles. The colour range is $[-0.05, 0.20]$ for all images.

for the 10% view angles case, it becomes difficult to reconstruct any of the steel inclusions using any of the priors. Only the widest tangential inclusions appear faintly in the reconstructions.

In Figure 8, we show the 95% interquantile ranges of the posteriors which reflect the uncertainty in the reconstructions. Note that the colourbar for each image is different. The GMRF prior yields a posterior uncertainty pattern that reflects the scan-geometry and that comes from the uncertainty of the likelihood. For the SGP priors, the posterior uncertainty pattern reflects the structure of the prior. For SGP-F, we observe lowest uncertainty in the background and in the steel, PU, and PE layers (regions 1-4), where we find the highest prior precision parameters. The concrete layer (region 5) has a lower prior precision,

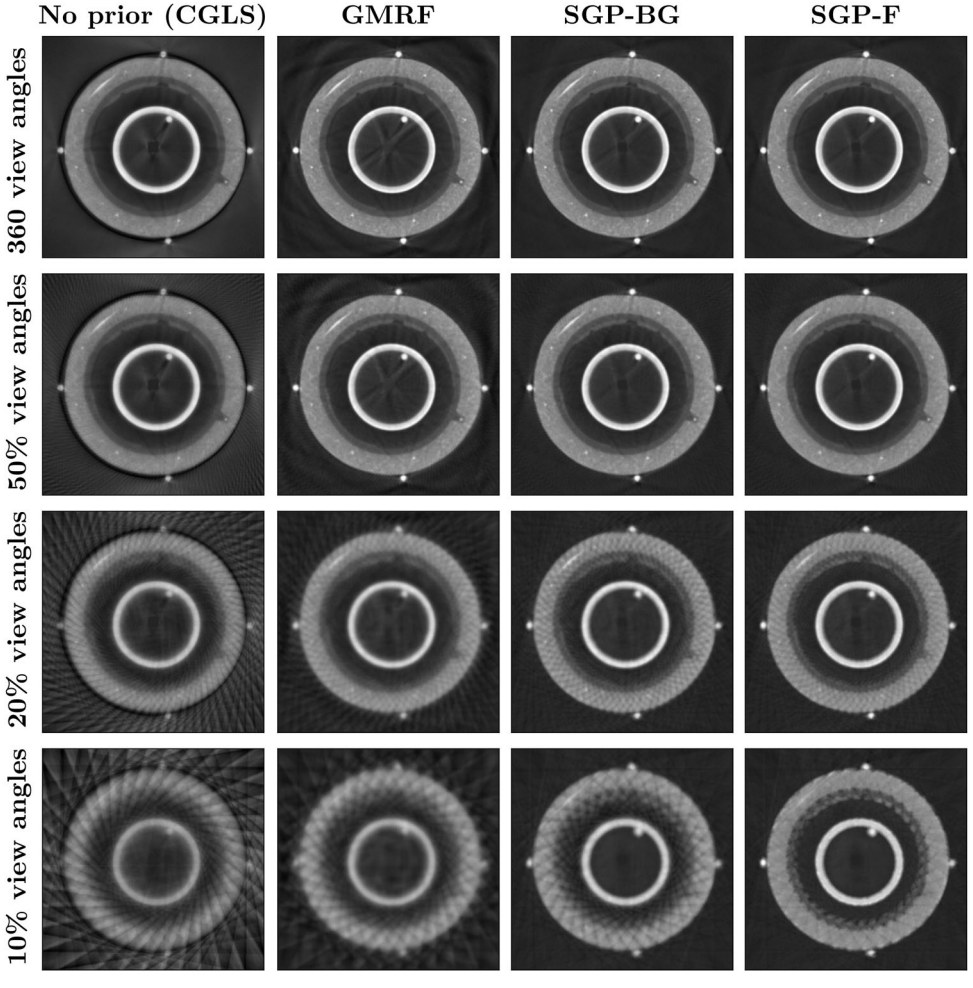



**Figure 9.** Posterior means using real data. The columns from left to right use increasingly informative priors: no prior (obtained as a CGLS reconstruction), Gaussian Markov Random Field (GMRF) prior, structural Gaussian prior on the background (SPG-BG), and structural Gaussian prior on the full image domain (SGP-F). The rows from top to bottom use decreasing numbers of view angles: 360 (or 100%), 50%, 20%, and 10% view angles. The colour range is $[-0.05, 0.20]$ for all images.

and thus it reflects higher uncertainty. The highest uncertainty is found at the boundary regions, where no particular attenuation coefficient is promoted. We note that the boundary regions seem to have relatively higher uncertainty for 10% and 20% view angles, than for 50% and 100%. This is caused by the lower GMRF precision $\delta_0$ used in the two former

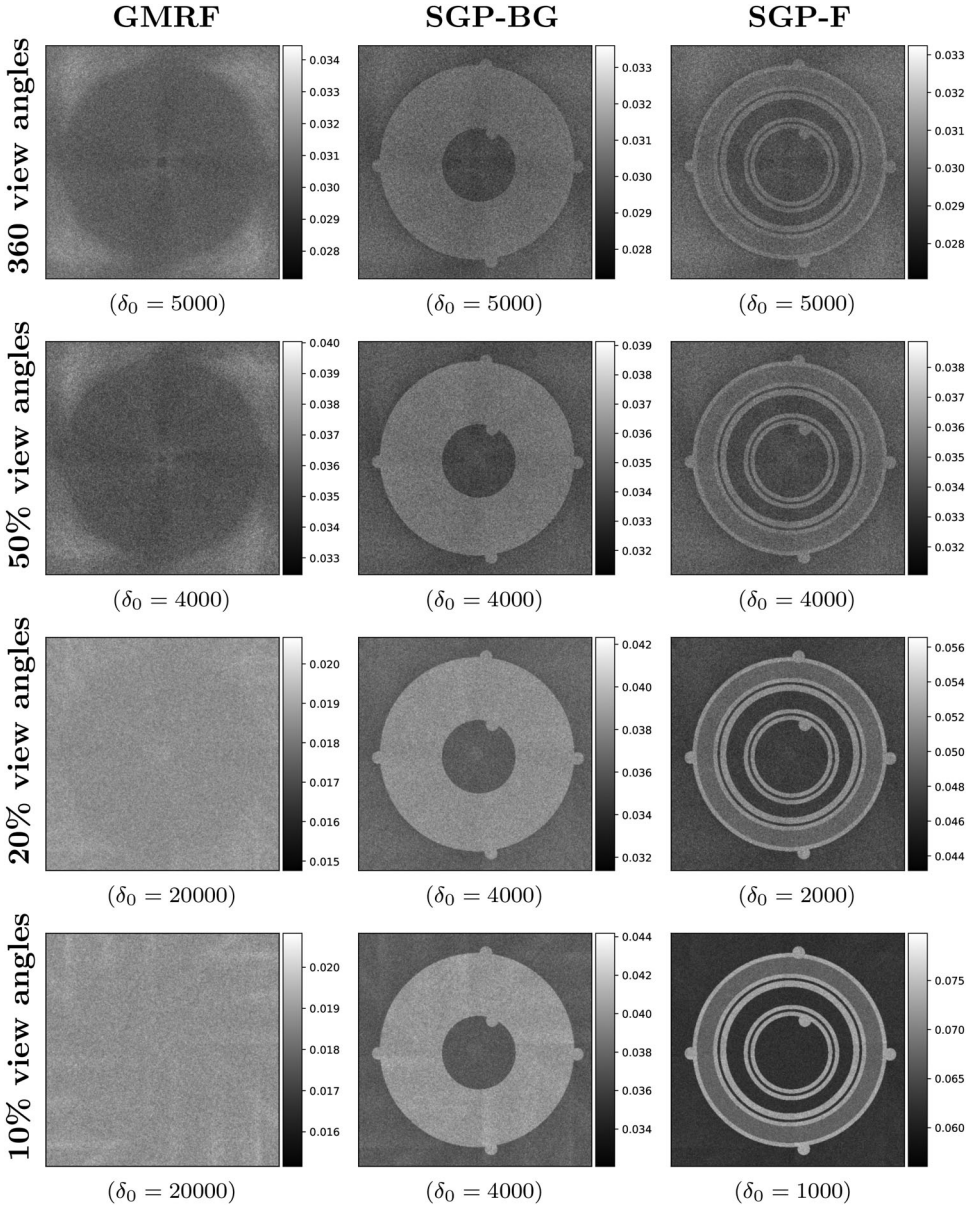



**Figure 10.** Posterior 95% interquantile ranges using real data. The columns from left to right use increasingly informative priors: Gaussian Markov Random Field (GMRF) prior, structural Gaussian prior on the background (SPG-BG), and structural Gaussian prior on the full image domain (SGP-F). The rows from top to bottom use decreasing numbers of view angles: 360 (or 100%), 50%, 20%, and 10% view angles. The colour range is $[-0.05, 0.20]$ for all images.

cases. Finally, an uncertainty pattern due to the likelihood is visible for the SGP-F prior cases with 50% and 100% view angles, but not with 10% and 20%. We suspect this is due to the prior having a relatively larger influence than the likelihood when data is sparse.

### 5.4. Results of real-data experiments

We now apply the SGPs to real data and obtain reconstructions with image qualities that appear comparable to the synthetic experiments. As for the synthetic data, we will study the effect of the SGP priors in combination with sparse view angles data. Figure 9 shows the posterior means from 16 experiments, where the priors and number of view angles are varied. Our observations in the synthetic-data experiments are in good agreement with the real-data case. Again, we see dominating artifacts when decreasing the number of view angles in the deterministic CGLS reconstruction with no prior information. We also see that they are reduced by introducing structural information. In the 20% view angles case, we observe that the SGPs not only reduce artifacts, but also help increase contrast between materials. In the most extreme case, with only 10% view angles, we lose some details in the reconstruction, such as the irregularities between the PU and PE layers. We do, however, still see traces of the large steel inclusion in the concrete layer.

In Figure 10, we see the 95% interquantile ranges associated with the posteriors. Again, we observe that the acquisition geometry is reflected when using the GMRF prior, and that the SGP structures are reflected (layers) when using these particular priors. Note that the steel rods used to fixate the pipe during measurements do not have any local prior promoting a particular attenuation coefficient.

## 6. Discussion and conclusions

We proposed a method of introducing prior structural information in the form of a so-called Structural Gaussian Prior (SGP) in Bayesian CT reconstruction when the internal structure and attenuation coefficients of an object are approximately known. We applied the proposed SGP in a case study of CT imaging of a subsea pipe with an offset fan-beam acquisition geometry. Numerical experiments with real and synthetic data showed that SGPs improve reconstructions in terms of suppressed artifacts and enhanced contrast between materials, compared to a conventional CGLS reconstruction with no prior information or with a standard smoothing GMRF prior. The SGPs were especially powerful in setups with sparse view angles. Information about all layers of the pipe is required to construct the SGP-F. Using this prior yielded the best results in the numerical experiments. In some cases, we might have access to less information, for instance only the pipe's exterior dimensions. Then, the SGP-BG prior is applicable, and also in such cases, the numerical results showed substantial improvements of the reconstructions compared to using a GMRF prior or no prior.

We also proposed a generalized linear RTO sampling method for efficient sampling of the Gaussian posterior with non-zero prior mean. Since the posterior is Gaussian and defined in terms of square-root precision matrices, it can be efficiently sampled by solving a system of linear equations, yielding approximately independent posterior samples. This is a useful property compared to conventional Markov chain Monte Carlo (MCMC) methods such as Metropolis-Hastings [44] or more modern Hamiltonian MCMC methods

[45], where independent samples cannot be guaranteed. To overcome this issue, MCMC methods often require a 'thinning' process where many of the samples are discarded. Therefore, MCMC methods require many more simulations, and thus, much longer computation times to obtain a sufficiently large quasi-independent set of samples.

If uncertainty estimates are not particularly required, our method based on SGP priors can also achieve reconstructions using MAP estimates that can be computed extremely efficiently. However, since one of our main goals is uncertainty quantification, we chose to sample the posterior.

Interpretation of the posterior uncertainty estimates is difficult and requires more study. We observed that the GMRF prior gives rise to uncertainty estimates that reflect the acquisition geometry, and this is in agreement with the results in [46]. When imposing the SGP, we noticed, that the uncertainty estimates directly reflected the prior. The prior's effect on the uncertainty estimates has also been seen in other studies. For instance, edge-preserving priors give rise to low uncertainty within regions and higher uncertainty on discontinuities in CT reconstructions [7,8]. In future work, we aim to further improve uncertainty estimates of the pipe and defect structure, for example to help distinguish artifacts from real features in the reconstruction.

## Acknowledgments

We would like to thank FORCE Technology for providing real data and information regarding subsea pipes.

## Data availability statement

The real CT data from the scan of the pipe is openly available in Zenodo at https://doi.org/10.5281/zenodo.6817690, reference number [47]. All Python code used to run the numerical experiments and produce figures is openly available in GitHub at https://doi.org/10.5281/zenodo.8039903, reference number [48].

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## Funding

This work was supported by The Villum Foundation [grant no. 25893].

## ORCID

*Silja L. Christensen* http://orcid.org/0000-0003-3995-3055
*Nicolai A. B. Riis* http://orcid.org/0000-0002-6883-9078
*Felipe Uribe* http://orcid.org/0000-0002-1010-8184
*Jakob S. Jørgensen* http://orcid.org/0000-0001-9114-754X

## References

[1] Briddon J, Metcalfe B. Inspecting pipelines with discovery$^{TM}$, the world's only subsea CT scanner. Pipeline Technol J. 2017;6:6–15. https://www.pipeline-journal.net/journal/pipeline-technology-journal-62017.

[2] Daaland A, Soldal R, Loetveit SA, et al. TomX: high-energy tomographic X-ray for in-service inspection of flexible pipeline systems. In: Proceedings for International Conference on Offshore Mechanics and Arctic Engineering; 2003. p. 739–746; Vol. 2. Safety and Reliability; Pipeline Technology.
[3] FORCE Technology, Effektiv og sikker kvalitetssikring af subsea-installationer [efficient and safe quality assurance of subsea intallations] [Internet]; 2018 [cited 2023-02-16]; https://forcetechnology.com/da/innovation/afsluttede-projekter/subsea-inspektion-konstruktion-levetid-reparation-planlaegning.
[4] Riis NAB, Frøsig J, Dong Y, et al. Limited-data X-ray CT for underwater pipeline inspection. Inverse Probl. 2018;34(3):034002. doi: 10.1088/1361-6420/aaa49c.
[5] Hansen PC. Discrete inverse problems: insight and algorithms. Philadelphia: Society for Industrial and Applied Mathematics (SIAM); 2010.
[6] Kaipio J, Somersalo E. Statistical and computational inverse problems. Dordrecht: Springer; 2005.
[7] Suuronen J, Emzir M, Lasanen S, et al. Enhancing industrial X-ray tomography by data-centric statistical methods. Data-Centric Eng. 2020;1(3):e10. doi: 10.1017/dce.2020.10.
[8] Uribe F, Bardsley JM, Dong Y, et al. A hybrid gibbs sampler for edge-preserving tomographic reconstruction with uncertain view angles. SIAM/ASA J Uncertain Quantif. 2022;10(3):1293–1320. doi: 10.1137/21M1412268.
[9] Hansen PC, Jørgensen JS, Lionheart WRB, editors. Computed Tomography: algorithms, insight, and just enough theory. Philadelphia: SIAM; 2021.
[10] Sidky EY, Pan X. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys Med Biol. 2008;53(17):4777–4807. doi: 10.1088/0031-9155/53/17/021.
[11] Bardsley JM. Uncertainty quantification for inverse problems. Philadelphia: SIAM; 2018.
[12] Bardsley JM. Gaussian Markov random field priors for inverse problems. Inverse Probl Imaging. 2013;7(2):397–416. doi: 10.3934/ipi.2013.7.397.
[13] Bardsley JM. Laplace-distributed increments, the Laplace prior, and edge-preserving regularization. J Inverse Ill-Posed Probl. 2012;20(3):271–285. doi: 10.1515/jip-2012-0017.
[14] Markkanen M, Roininen L, Huttunen JMJ, et al. Cauchy difference priors for edge-preserving Bayesian inversion. J Inverse Ill-posed Probl. 2019;27(2):225–240. doi: 10.1515/jiip-2017-0048.
[15] Suuronen J, Chada NK, Roininen L. Cauchy Markov random field priors for Bayesian inversion. Stat Comput. 2022;32(2):1573–1375. doi: 10.1007/s11222-022-10089-z.
[16] Zhang H, Han H, Ma J, et al. Deriving adaptive MRF coefficients from previous normal-dose CT scan for low-dose image reconstruction via penalized weighted least-squares minimization. Med Phys. 2014;41(4):041916. doi: 10.1118/1.4869160.
[17] Zhang H, Han H, Liang Z, et al. Extracting information from previous full-dose CT scan for knowledge-based Bayesian reconstruction of current low-dose CT images. IEEE Trans Med Imaging. 2016;35(3):860–870. doi: 10.1109/TMI.2015.2498148.
[18] Han H, Zhang H, Wei X, et al. Texture-preserved penalized weighted least-squares reconstruction of low-dose CT image via image segmentation and high-order MRF modeling. In: Medical Imaging 2016: Physics of Medical Imaging; 03; 2016. p. 97834F.
[19] Gao Y, Liang Z, Moore W, et al. A feasibility study of extracting tissue textures from a previous full-dose CT database as prior knowledge for Bayesian reconstruction of current low-dose CT images. IEEE Trans Med Imaging. 2019;38(8):1981–1992. doi: 10.1109/TMI.4201.
[20] Fukuda W, Maeda S, Kanemura A, et al. Bayesian x-ray computed tomography using material class knowledge. In: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing; 04; 2010. p. 2126–2129.
[21] Kaipio JP, Kolehmainen V, Vauhkonen M, et al. Inverse problems with structural prior information. Inverse Probl. 1999;15(3):713–729. doi: 10.1088/0266-5611/15/3/306.
[22] Heikkinen LM, Vauhkonen M, Savolainen T, et al. Modelling of internal structures and electrodes in electrical process tomography. Meas Sci Technol. 2001;12(8):1012–1019. doi: 10.1088/0957-0233/12/8/304.

[23] Karhunen K, Seppänen A, Lehikoinen A, et al. Electrical resistance tomography imaging of concrete. Cem Concr Res. 2010;40(1):137–145. doi: 10.1016/j.cemconres.2009.08.023.
[24] Arridge SR, Dorn O, Kolehmainen V, et al. Parameter and structure reconstruction in optical tomography. J Phys Conf Ser. 2008;135:012001. doi: 10.1088/1742-6596/135/1/012001.
[25] Ehrhardt MJ, Betcke MM. Multicontrast MRI reconstruction with structure-guided total variation. SIAM J Imaging Sci. 2016;9(3):1084–1106. doi: 10.1137/15M1047325.
[26] Rasch J, Kolehmainen V, Nivajarvi R, et al. Dynamic MRI reconstruction from undersampled data with an anatomical prescan. Inverse Probl. 2018;34(7):074001. doi: 10.1088/1361-6420/aac3af.
[27] Calvetti D, Sgallari F, Somersalo E. Image inpainting with structural bootstrap priors. Image Vis Comput. 2006;24(7):782–793. doi: 10.1016/j.imavis.2006.01.015.
[28] Heikkinen LM, Vauhkonen M, Savolainen T, et al. Electrical process tomography with known internal structures and resistivities. Inverse Probl Eng. 2001;9(5):431–454. doi: 10.1080/174159701088027775.
[29] Bardsley JM, Solonen A, Haario H, et al. Randomize-then-optimize: a method for sampling from posterior distributions in nonlinear inverse problems. SIAM J Sci Comput. 2014;36(4):A1895–A1910. doi: 10.1137/140964023.
[30] Björck A. Numerical methods for least squares problems. Philadelphia: SIAM; 1996.
[31] van Aarle W, Palenstijn WJ, De Beenhouwer J, et al. The ASTRA toolbox: a platform for advanced algorithm development in electron tomography. Ultramicroscopy. 2015;157:35–47. doi: 10.1016/j.ultramic.2015.05.002.
[32] van Aarle W, Palenstijn WJ, Cant J, et al. Fast and flexible X-ray tomography using the ASTRA toolbox. Opt Express. 2016;24(22):25129–25147. doi: 10.1364/OE.24.025129.
[33] Jørgensen JS, Ametova E, Burca G, et al. Core Imaging Library – Part I: a versatile python framework for tomographic imaging. Philos Trans Royal Soc A Math Phys Eng Sci. 2021;379(2204):20200192. doi: 10.1098/rsta.2020.0192.
[34] Papoutsellis E, Ametova E, Delplancke C, et al. Core Imaging Library – Part II: multichannel reconstruction for dynamic and spectral tomography. Philos Trans Royal Soc A Math Phys Eng Sci. 2021;379(2204):20200193. doi: 10.1098/rsta.2020.0193.
[35] Rafiei MM, Tavakoli-Anbaran H. Calculation of the exposure buildup factors for X-ray photons with continuous energy spectrum using Monte Carlo code. J Radiol Prot. 2018;38(1):207–217. doi: 10.1088/1361-6498/aa9cfa.
[36] Trubey DK. New gamma-ray buildup factor data for point kernel calculations: Ans-6. 4. 3 standard reference data. United States: International Nuclear Information System; 1988. Report ORNL/RSIC–49.
[37] Buzug T. Computed Tomography: from photon statistics to modern cone-beam CT. Berlin: Springer; 2008.
[38] Hubbell JH, Seltzer SM. Tables of X-ray mass attenuation coefficients and mass energy-absorption coefficients (version 1.4) [dataset] ; 2004 [cited 2021 Jan 19]. In: National Institute of Standards and Technology [Internet]. doi: 10.18434/T4D01F.
[39] Tarantola A. Inverse problem theory and methods for model parameter estimation. Philadelphia: Society for Industrial and Applied Mathematics; 2005.
[40] Robert CP. The Bayesian choice. 2nd ed. New York: Springer; 2007. (Springer Texts in Statistics).
[41] Rue H, Held L. Gaussian Markov random fields theory and applications. New York: Chapman & Hall/CRC; 2005.
[42] Coope ID. Circle fitting by linear and nonlinear least squares. J Optim Theory Appl. 1993;76:381–388. doi: 10.1007/BF00939613.
[43] Bickel P, Doksum K. Mathematical statistics: basic ideas and selected topics. Vol. 1, 2nd ed. New Jersey: Prentice Hall; 2001.
[44] Hastings WK. Monte carlo sampling methods using Markov chains and their applications. Biometrika. 1970;57:97–109. doi: 10.1093/biomet/57.1.97.
[45] Duane S, Kennedy A, Pendleton BJ, et al. Hybrid Monte Carlo. Phys Lett B. 1987;195(2):216–222. doi: 10.1016/0370-2693(87)91197-X.

[46] Bardsley JM. MCMC-based image reconstruction with uncertainty quantification. SIAM J Sci Comput. 2012;34(3):A1316–A1332. doi: 10.1137/11085760X.
[47] Jensen TH. X-ray CT of subsea pipe [dataset] ; 2023 May 05 [cited 2023 May 05]. In: Zenodo [Internet]. https://doi.org/10.5281/zenodo.6817690.
[48] Christensen SL, Riis NAB, Uribe F, et al. Code accompanying article "Structural Gaussian priors for Bayesian CT reconstruction of subsea pipes" [Code repository] 2023. In: Zenodo/GitHub [Internet]. Available from: https://doi.org/10.5281/zenodo.8039903.