Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2019-0060

Vol. **27**, No. 5, pp. 1485-1504
May 2020

Uncertainty Quantification of Density Reconstruction Using MCMC Method in High-Energy X-ray Radiography

Xinge Li, Haibo Xu*, Na Zheng, Qinggang Jia, Tongxiang Gu and Suhua Wei

Institute of Applied Physics and Computational Mathematics, Beijing 100094, P.R. China.

Received 12 April 2019; Accepted (in revised version) 25 September 2019

Abstract. High-energy X-ray radiography is a measuring technique for quantitative measurement and diagnosis of the object and its internal structure. Tomographic reconstruction determines the geometric and physical properties of the object according to the energy distribution on the imaging plane. Considering the noise and blur in the process of radiographing, we construct a general reconstruction model for the axisymmetric single image photographic system. This inverse problem is then cast within a statistical framework in order to compute volumetric object densities from X-ray radiographs and to quantify uncertainties in the reconstruction. A hierarchical Bayesian model is developed with a likelihood based on a Gaussian noise model and with priors placed on the unknown nonnegative density profile, the precision matrix, and two scale parameters. This results in a joint posterior distribution, which can be readily sampled using the Markov chain Monte Carlo (MCMC) method. To study the role of hyperparameters and their sensitivity analysis, a wide variety of tests were conducted which led to a number of definitive conclusions. Results of the density reconstructions and pointwise uncertainty estimates are presented for simulated signals with various physical factors in the imaging process included.

AMS subject classifications: 65K99, 65Z05, 62F15, 62P35

Key words: Inverse problem, density reconstruction, uncertainty quantification, Bayesian inference, MCMC method.

1 Introduction

With the aid of nondestructive characterization and high transmission ability of X-rays, X-ray tomography has been widely used in many areas, including radiation medicine,

*Corresponding author. *Email addresses:* lixinge@lsec.cc.ac.cn (X. Li), xu_haibo@iapcm.ac.cn (H. Xu), zheng_na@iapcm.ac.cn (N. Zheng), QGJIA_XJTU@126.com (Q. Jia), txgu@iapcm.ac.cn (T. Gu), wei_suhua@iapcm.ac.cn (S. Wei)

material science, earth science and nuclear physics. High-energy X-ray radiography measures the internal structure of the object according to the X-ray energy distribution on the imaging plane. The measured spatial density distribution is of great significance for studying the compression behavior of objects subjected to powerful shocks under the effect of explosives in hydro-test experiments. In this paper, we focus on the problem of uncertainty quantification of density reconstruction for high-energy X-ray radiography.

The essence of density reconstruction is to reconstruct the spatial density distribution of objects from the projection data obtained by transmission imaging. The existing reconstruction methods can be roughly classified into three categories, namely analytical methods [1–4], iterative methods [5–8] and statistical methods [9–13], which are extensively used in medical imaging diagnosis, industrial nondestructive detection and other fields. However, it is difficult to meet the reconstruction requirements of high-energy X-ray radiography under complex imaging engineering conditions. Uncertainty is a measure of the confidence in a calculated or measured value. A frequentist approach to uncertainty estimation would take a large number of experimental measurements and determine uncertainty from the spread of results obtained. This approach is not applicable in the case of hydro-test experiments which are costly and likely to be performed only once.

Given the challenging nature of hydro-test radiography, there are limitations in the quality of radiographs obtained. For the low photon flux that occurs with a highly attenuating object, the statistical variation of photons measured by the detector forms the output noise. Another significant issue is system blur, which results from the finite size of X-ray source spot and the energy spread within the detector. Both noise and blur make it difficult to extract information about edges and density distributions of materials, particularly near the interfaces. Therefore density reconstruction technology is highly demanded. Moreover, physical factors such as cone beam light source, scattering and energy spectrum effect must also be considered. In this paper, a Bayesian approach is used to reject those solutions that are unphysical, based on our experience and understanding of the physics involved. Examples of prior knowledge typically used are density nonnegativity and smoothness.

Bayesian formulations for inverse problems have gained considerable attention in the inverse problems community for their utility in uncertainty quantification [12,14,16,17]. It has been pointed out that many of the classical regularization methods for solving ill-posed problems can be viewed as constructing estimators based on the posterior distribution [12]. A large part of the literature discussing regularization techniques is devoted to the problem of selecting the regularization parameters. Moreover, the prior densities typically depend on parameters such as variance and expectation that are always assumed to be known in traditional Bayesian approaches. From the point of view of classical methods, this corresponds to knowing ahead of time the regularization parameters. In the Bayesian framework the answer to the question of how the parameters should be chosen is: If a parameter is not known, it is a part of the inference problem. This approach leads to hierarchical models, which allow the incorporation of empirical knowledge, not just about the reconstruction being sought but also about the structure of the prior [18–21].

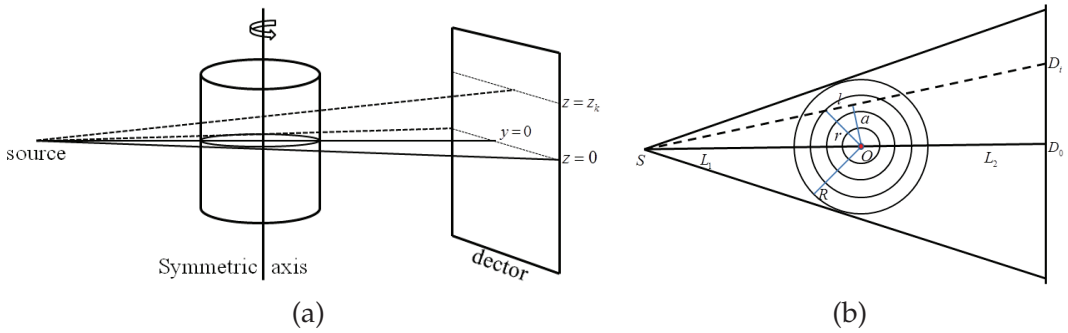


Figure 1: Illustration to single radiographic imaging system for radially symmetric object in 3D view (a) and 2D view (b).

In the application here, we seek an unknown that contains discontinuities but we do not precisely know the discontinuity locations, so we develop a hierarchical Bayesian model for localizing the discontinuities and computing object densities simultaneously.

The Bayesian approach combines the prior knowledge of the unknown parameters and the forward model to yield a posterior probability distribution of the model parameters. In this way, the unknown parameters can be characterized by their posterior distributions. The posterior distributions are typically not of analytical form or from a standard parametric family, and characterizing them exactly requires sampling approaches such as MCMC [22]. The major challenge of MCMC is the computational burden induced by the repeated evaluations of the forward model. When the forward model is computationally expensive, researchers have proposed various methods to accelerate inference. One way is to construct computationally inexpensive approximations of the forward model and to use these approximations as surrogates in the sampling procedure [23–25]. These approaches can enhance the efficiency compared to a standard MCMC approach when dealing with nonlinear reconstruction model.

In our experimental setting, a radially symmetric object with radius R and height H is situated so that its center-layer lies in the xy-plane and its axis of symmetry coincides with the z-axis (see Fig. 1(a)). Only a single radiograph is taken with a radiographic axis perpendicular to the symmetric axis of the object. The transmitted radiation is measured by a detector lying on a plane $x=x_0$. The X-ray source is placed sufficiently far from the object compared to its size, so that the X-rays can be assumed to be parallel on different layers. In each layer, we consider that the X-rays form a fan-beam shape, see Fig. 1(b). Each cross section of the object is projected onto a line of the detector plane.

The rest of the paper is organized as follows. In Section 2, the general density reconstruction model is described in detail. Section 3 discusses how to estimate the object density and its corresponding uncertainty quantitatively. In Section 4, we analyze the factors causing uncertainty in density reconstruction. Given that the French Test Object (FTO) is a standard experimental object to study the theory of high-energy X-ray radiography, uncertainty quantification of the FTO reconstruction are given in Section 5, as well

as the parameter selection method and sensitivity study. Finally, summary and outlook are contained in Section 6.

2 Density reconstruction model

Images captured by X-ray imaging systems are direct measures of the optical density. Attenuation is a function of both the photon energy spectrum and the materials' X-ray attenuation cross sections. The materials' photon attenuation properties can be encapsulated in a single quantity, areal density, also known as optical length. It combines both the volumetric densities of the materials and their thicknesses. The measured optical density, G, can be converted to areal density, b, using a measured transmission curve, which is given by

$$G = G_0 + k(X_D + X_S) = G_0 + k\left(\frac{X_0}{d^2}e^{-b} + X_S\right), \tag{2.1}$$

where G and G_0 are the optical density and background density at that point; k is the slope of the transmission curve; X_0 is the radiation exposure; d is the distance between source and the imaging plane; X_S is the scattered radiation; X_D is the direct radiation which is represented by Beer-Lambert Law:

$$X_D = \frac{X_0}{d^2}e^{-b}. \tag{2.2}$$

An example of the transmission curve is the film characteristic curve (H&D curve), which describes the relationship between different exposures and the corresponding densities after film development. Inversion of Eq. (2.1) yields the areal density in terms of measured optical density:

$$b = -\ln\frac{d^2}{kX_0} - \ln(G - G_0 - kX_s), \tag{2.3}$$

which is an accurate model for monochromatic X-rays and materials with similar X-ray cross sections.

The areal density is related to the object density ρ by the line integral along the X-ray track

$$b = \int \mu_\rho \rho \mathrm{d}l = \int \mu_l \mathrm{d}l, \tag{2.4}$$

where μ_ρ are materials' mass attenuation coefficients, and $\mu_l = \mu_\rho \rho$ are materials' linear attenuation coefficients. In high-energy X-ray radiography, the photon source has a non-monochromatic Bremsstrahlung spectrum. For simplicity, we consider the case where the effect of energy spectrum is approximated by the effective monoenergy, which means that μ_ρ are not functions of energy and are replaced by $\mu_\rho(\bar{E})$ at an average energy.

Considering the degradations that may occur during the imaging process, we discretize Eq. (2.4) and obtain the density reconstruction model as

$$\mathbf{b} = \mathbf{K}\mathbf{A}\boldsymbol{\rho} + \boldsymbol{\varepsilon}, \tag{2.5}$$

where $\mathbf{b}\in\mathbb{R}^m$ are the areal density values, which are computed from measured optical density values using Eq. (2.3); the unknown $\boldsymbol{\rho}\in\mathbb{R}^n$ here are actually the linear attenuation coefficients, which are also the input parameters for the model, and the density values to be reconstructed can be obtained by dividing the linear attenuation coefficients by materials' mass attenuation coefficients; $\boldsymbol{\varepsilon}\in\mathbb{R}^m$ is the unknown noise vector; $\mathbf{K}$ presents the blurring that may be produced in the process of radiographing, including the source blur and detector blur; $\mathbf{A}\in\mathbb{R}^{m\times n}$ denotes the forward projection matrix defined by the line integrals. Here, we assume that the system blur is Gaussian, and $\mathbf{K}$ denotes the Gaussian blur matrix. Also assume that the X-rays at different layers are parallel. Hence, the projection matrix $\mathbf{A}$ is always the same for each layer. In each layer, the projection matrix $\mathbf{A}$ is generated from fan-beam X-rays, as shown in Fig. 1(b).

Without loss of generality, we use the layer $z=0$ to illustrate how to formulate the projection operator $\mathbf{A}$, see Fig. 1(b). Let $\rho(r)$ be the radial density for the layer. Since the object has finite volume, we can assume that $\rho(r)=0$ when $r>R$. The distances from the source to the object and the object to the detector are denoted by L_1 and L_2 respectively, where $L_1=|SO|$ and $L_2=|OD_0|$. We consider Fig. 1(b) on the rectangular coordinate system with origin O. Thus, the detector lies at plane $x=L_2$ and the areal density $b(y)$ is measured at $y\in[-W,W]$, where $W>(L_1+L_2)\cdot R/L_1$ to ensure that the projection data for the whole object is covered. Separate $[-W,W]$ into $2m-1$ uniform partitions each with step size $\Delta w=2W/(2m-1)$. Considering the symmetry of the object, we only need half of the projection data. Take $\mathbf{b}$ to be a vector of m elements, with $b_i=b(y_i)$, for $i=1,\cdots,m$. For the cross section of radius R, we subdivide $[0,R]$ into n uniform partitions with step size $\Delta r=R/n$ and $r_i=i\Delta r$. Then $\boldsymbol{\rho}$ is a vector of n elements with its i-th entry being $\mu_\rho\rho(r_i)$. Hence, $\mathbf{A}$ is a matrix of size $m\times n$. The (i,j)-th element of $\mathbf{A}$ is equal to the length of the X-ray l_i lying between circles of $r=r_j$ and $r=r_{j+1}$.

3 Bayesian formulation of reconstruction problem

In this section, we introduce a hierarchical Bayesian model to compute object densities and quantify their corresponding uncertainties at the same time. By selecting conjugate prior distributions, the final posterior formulation can be easily and efficiently sampled using a Gibbs sampler, a specialized MCMC method.

3.1 Hierarchical Bayesian model

The noise in the intensity image is a composition of distributions. After the intensity is converted to areal density, the distribution of the noise in areal density space is difficult to estimate, so we begin with the standard additive Gaussian noise model [19, 26], i.e., $\boldsymbol{\varepsilon}\sim\text{Normal}(\mathbf{0},(\lambda\mathbf{I})^{-1})$ with precision λ. Then the conditional probability density of $\mathbf{b}$

given the information of $\boldsymbol{\rho}$ and λ, would be

$$p(\mathbf{b}|\boldsymbol{\rho},\lambda)\propto\lambda^{m/2}\exp\left(-\frac{\lambda}{2}\|\mathbf{K}\mathbf{A}\boldsymbol{\rho}-\mathbf{b}\|^2\right). \tag{3.1}$$

Assume that the prior model for $\boldsymbol{\rho}$ is also a Gaussian,

$$\boldsymbol{\rho}\sim\text{Normal}(\mathbf{0},(\delta\mathbf{L})^{-1}),$$

where $\mathbf{L}\in\mathbb{R}^{n\times n}$ is referred to as the precision matrix, and $\delta>0$ and $\mathbf{L}$ are poorly known. Hence, we write a conditional prior for $\boldsymbol{\rho}$, assuming that δ and $\mathbf{L}$ were known, as

$$p(\boldsymbol{\rho}|\delta,\mathbf{L})\propto\delta^{n/2}|\mathbf{L}|^{1/2}\exp\left(-\frac{\delta}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{L}\boldsymbol{\rho}\right), \tag{3.2}$$

that is, the prior density is conditioned on the knowledge of δ and $\mathbf{L}$.

In consideration of the nonnegativity of density value $\boldsymbol{\rho}$, we impose a nonnegativity constraint on the computed samples. Suppose $\mathcal{I}=\mathcal{I}(\boldsymbol{\rho})\overset{\text{def}}{=}\{i|\rho_i=0\}$ is the zero set and $p(\mathcal{I})$ a probability model for $\mathcal{I}$. Then the prior will depend upon the zero set. If we define $\mathbf{C}$ to be the diagonal matrix with diagonal entries $c_{ii}=1$ for $i\notin\mathcal{I}$, and $c_{ii}=0$ otherwise, the prior in the unconstrained case is modified as follows:

$$p(\boldsymbol{\rho}|\delta,\mathbf{L},\mathcal{I})\propto\delta^{n_p/2}|\mathbf{L}|^{1/2}\exp\left(-\frac{\delta}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{C}\mathbf{L}\mathbf{C}\boldsymbol{\rho}\right), \tag{3.3}$$

where $n_p=n-|\mathcal{I}|$, i.e. the number of positive elements in $\boldsymbol{\rho}$. Note that this prior is well defined for the indices $i\notin\mathcal{I}$, which is all we need.

Assume further that we have a hyperprior density $p(\mathbf{L})$ for the precision matrix $\mathbf{L}$. We suppose here that this density is a Wishart distribution, which is often used for inverse covariance matrices [30]. Thus $\mathbf{L}\sim\text{Wishart}(\Sigma,\nu)$, with probability density function

$$p(\mathbf{L})=\frac{1}{2^{\nu n/2}|\Sigma|^{\nu/2}\Gamma_n(\nu/2)}|\mathbf{L}|^{\frac{\nu-n-1}{2}}\exp\left(-\frac{1}{2}\text{tr}(\Sigma^{-1}\mathbf{L})\right), \tag{3.4}$$

where Σ is a positive definite scale matrix, and ν is the degrees of freedom parameter.

Last, a Gamma hyperprior distribution is chosen for both scale parameters λ and δ so that

$$p(\lambda)\propto\lambda^{\alpha_\lambda-1}\exp(-\beta_\lambda\lambda), \tag{3.5}$$

$$p(\delta)\propto\delta^{\alpha_\delta-1}\exp(-\beta_\delta\delta), \tag{3.6}$$

that is, $\lambda\sim\text{Gamma}(\alpha_\lambda,\beta_\lambda)$, $\delta\sim\text{Gamma}(\alpha_\delta,\beta_\delta)$, where α_λ and α_δ are Gamma shape parameters and β_λ and β_δ are Gamma rate parameters.

Considering now all $\boldsymbol{\rho}$, λ, δ, $\mathbf{L}$ and $\mathcal{I}$ as unknowns, we write Bayes' formula conditioned on $\mathbf{b}$ as

$$\begin{aligned} p(\boldsymbol{\rho},\lambda,\delta,\mathbf{L},\mathcal{I}|\mathbf{b}) &\propto p(\mathbf{b}|\boldsymbol{\rho},\lambda,\mathcal{I})p(\boldsymbol{\rho}|\delta,\mathbf{L},\mathcal{I})p(\mathbf{L})p(\delta)p(\mathcal{I})p(\lambda) \\ &= |\mathbf{L}|^{\frac{\nu-n}{2}}\lambda^{\frac{m}{2}+\alpha_\lambda-1}\delta^{\frac{np}{2}+\alpha_\delta-1}\exp(-\beta_\lambda\lambda-\beta_\delta\delta) \\ &\quad\times\exp\left(-\frac{\lambda}{2}\|\mathbf{KAC}\boldsymbol{\rho}-\mathbf{b}\|^2-\frac{\delta}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{CLC}\boldsymbol{\rho}-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}\mathbf{L})\right). \end{aligned} \tag{3.7}$$

This formula allows us to estimate $\boldsymbol{\rho}$, λ, δ and $\mathbf{L}$ simultaneously. That is,

$$p(\boldsymbol{\rho}|\lambda,\delta,\mathbf{L},\mathcal{I},\mathbf{b}) \propto \exp\left(-\frac{\lambda}{2}\|\mathbf{KAC}\boldsymbol{\rho}-\mathbf{b}\|^2-\frac{\delta}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{CLC}\boldsymbol{\rho}\right), \tag{3.8}$$

$$p(\lambda|\boldsymbol{\rho},\delta,\mathbf{L},\mathcal{I},\mathbf{b}) \propto \lambda^{\frac{m}{2}+\alpha_\lambda-1}\exp\left(-\left[\frac{1}{2}\|\mathbf{KAC}\boldsymbol{\rho}-\mathbf{b}\|^2+\beta_\lambda\right]\lambda\right), \tag{3.9}$$

$$p(\delta|\boldsymbol{\rho},\lambda,\mathbf{L},\mathcal{I},\mathbf{b}) \propto \delta^{\frac{np}{2}+\alpha_\delta-1}\exp\left(-\left[\frac{1}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{CLC}\boldsymbol{\rho}+\beta_\delta\right]\delta\right), \tag{3.10}$$

$$p(\mathbf{L}|\boldsymbol{\rho},\lambda,\delta,\mathcal{I},\mathbf{b}) \propto |\mathbf{L}|^{\frac{\nu-n}{2}}\exp\left(-\frac{1}{2}\left[\mathrm{tr}(\Sigma^{-1}\mathbf{L})+\delta\boldsymbol{\rho}^{\mathrm{T}}\mathbf{CLC}\boldsymbol{\rho}\right]\right), \tag{3.11}$$

which can be equivalently written as

$$\boldsymbol{\rho}|\lambda,\delta,\mathbf{L},\mathcal{I},\mathbf{b} \sim \mathrm{Normal}((\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{KA}+\delta\mathbf{L})^{\dagger}_{\mathcal{I}}\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{b},(\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{KA}+\delta\mathbf{L})^{\dagger}_{\mathcal{I}}), \tag{3.12}$$

$$\lambda|\boldsymbol{\rho},\delta,\mathbf{L},\mathcal{I},\mathbf{b} \sim \mathrm{Gamma}(m/2+\alpha_\lambda,\frac{1}{2}\|\mathbf{KAC}\boldsymbol{\rho}-\mathbf{b}\|^2+\beta_\lambda), \tag{3.13}$$

$$\delta|\boldsymbol{\rho},\lambda,\mathbf{L},\mathcal{I},\mathbf{b} \sim \mathrm{Gamma}(n_p/2+\alpha_\delta,\frac{1}{2}\boldsymbol{\rho}^{\mathrm{T}}\mathbf{L}_{\mathcal{I}}\boldsymbol{\rho}+\beta_\delta), \tag{3.14}$$

$$\mathbf{L}|\boldsymbol{\rho},\lambda,\delta,\mathcal{I},\mathbf{b} \sim \mathrm{Wishart}((\Sigma^{-1}+\delta(\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{T}})_{\mathcal{I}})^{-1},\nu+1), \tag{3.15}$$

where $\mathbf{B}_{\mathcal{I}} \overset{\mathrm{def}}{=} \mathbf{CBC}$ and $\dagger$ denotes pseudo-inverse. Note, again, that the distribution $p(\boldsymbol{\rho}|\lambda,\delta,\mathbf{L},\mathcal{I},\mathbf{b})$ is well-defined for indices $i\notin\mathcal{I}$, and that for $i\in\mathcal{I}$, $\rho_i=0$, as desired. Since $\mathbf{C}\boldsymbol{\rho}=\boldsymbol{\rho}$ (recall that $\rho_i=0$ for $i\in\mathcal{I}$), equivalent distributions result if $\mathbf{C}$ is removed in (3.13), $\mathbf{L}_{\mathcal{I}}$ is replaced by $\mathbf{L}$ in (3.14), and $(\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{T}})_{\mathcal{I}}$ is substituted as $\boldsymbol{\rho}\boldsymbol{\rho}^{\mathrm{T}}$ in (3.15). We do this in what follows.

It remains to define $p(\mathcal{I})$ and the conditional density $p(\mathcal{I}|\boldsymbol{\rho},\lambda,\delta,\mathbf{L},\mathbf{b})$. This can be accomplished by computing [14]

$$\hat{\boldsymbol{\rho}} = \underset{\boldsymbol{\rho}\geq\mathbf{0}}{\mathrm{argmin}}\left\{\frac{1}{2}\boldsymbol{\rho}^{\mathrm{T}}(\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{KA}+\delta\mathbf{L})\boldsymbol{\rho}-\boldsymbol{\rho}^{\mathrm{T}}(\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{b}+\mathbf{w})\right\}, \tag{3.16}$$

where $\mathbf{w}\sim\mathrm{Normal}(\mathbf{0},\lambda\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{KA}+\delta\mathbf{L})$. That is, solving (3.16) yields simultaneous samples of both $\boldsymbol{\rho}$ and, implicitly, $\mathcal{I}$ from $p(\mathcal{I}|\boldsymbol{\rho},\lambda,\delta,\mathbf{L},\mathbf{b})$. And, more remarkable, this optimization problem can be easily solved exploiting the Constrained Conjugate Gradient (CCG) method [15]. We still have not defined $p(\mathcal{I})$, but this is not necessary to define our MCMC method.

3.2 MCMC sampling of the posterior distribution

The power in (3.12)-(3.15) lies in the fact that samples from these four distributions can be easily computed using standard statistical software, though nonlinear optimization techniques will be needed for (3.12). A Gibbs sampler that results from sequential use of the conditional densities $p(\mathcal{I}|\boldsymbol{\rho},\lambda,\delta,\mathbf{L},\mathbf{b})$ and (3.12)-(3.15) can be written without an explicit sampling step for $\mathcal{I}$. The sampler begins with $\boldsymbol{\rho}$, and is initialized with λ_0, δ_0, and $\mathbf{L}_0$. A basic outline is listed in Algorithm 1.

Algorithm 1 (MCMC sampler)

1. Select λ_0, δ_0 and $\mathbf{L}_0$. Select a maximum number of samples, N, and set $k=0$.
2. Compute
$$\boldsymbol{\rho}^k=\underset{\boldsymbol{\rho}\geq\mathbf{0}}{\arg\min}\left\{\frac{1}{2}\boldsymbol{\rho}^{\mathrm{T}}(\lambda_k\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{A}+\delta_k\mathbf{L}_k)\boldsymbol{\rho}-\boldsymbol{\rho}^{\mathrm{T}}(\lambda_k\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{b}+\mathbf{w})\right\},$$
where $\mathbf{w}\sim\mathrm{Normal}(\mathbf{0},\lambda_k\mathbf{A}^{\mathrm{T}}\mathbf{K}^{\mathrm{T}}\mathbf{K}\mathbf{A}+\delta_k\mathbf{L}_k)$, using the CCG method.
3. Compute $\lambda_{k+1}\sim\mathrm{Gamma}(m/2+\alpha_\lambda,\frac{1}{2}\|\mathbf{K}\mathbf{A}\boldsymbol{\rho}^k-\mathbf{b}\|^2+\beta_\lambda)$.
4. Compute $\delta_{k+1}\sim\mathrm{Gamma}\left(n_p^k/2+\alpha_\delta,\frac{1}{2}(\boldsymbol{\rho}^k)^{\mathrm{T}}\mathbf{L}_k\boldsymbol{\rho}^k+\beta_\delta\right)$, where n_p^k is the number of positive entries in $\boldsymbol{\rho}^k$.
5. Compute $\mathbf{L}_{k+1}\sim\mathrm{Wishart}\left((\Sigma^{-1}+\delta_{k+1}\boldsymbol{\rho}^k(\boldsymbol{\rho}^k)^{\mathrm{T}})^{-1},\nu+1\right)$.
6. Set $k=k+1$. If $k<N$, return to step 2.

To simplify the computation of the inverse in step 5 of the sampler, the Sherman-Morrison formula [27] provides the relationship

$$(\Sigma^{-1}+\delta_{k+1}\boldsymbol{\rho}^k(\boldsymbol{\rho}^k)^{\mathrm{T}})^{-1}=\Sigma-\frac{\delta_{k+1}\Sigma\boldsymbol{\rho}^k(\boldsymbol{\rho}^k)^{\mathrm{T}}\Sigma}{1+\delta_{k+1}(\boldsymbol{\rho}^k)^{\mathrm{T}}\Sigma\boldsymbol{\rho}^k}. \tag{3.17}$$

Samples obtained via the above MCMC algorithm are used to compute the sample mean and 95% credibility intervals. The sample mean well characterizes the unknown true density, and the credibility intervals provide uncertainty quantification in the posterior estimate.

4 Analysis of physical factors causing uncertainty

Although empirical knowledge constraints have been added into the density reconstruction, it does not eliminate the uncertainty of reconstruction results (3.16), but only reduces

the uncertainty of reconstruction modeling (2.5), so it is necessary to analyze the factors causing uncertainty in density reconstruction. Generally, there are uncertainties caused by the conversion from optical density to areal density, the noise in areal density, the measurement error of areal density as well as the reconstruction method itself. Due to the effect of the energy spectrum, the conversion from intensity to areal density is the primary source of error in density reconstruction from real data. While in our simulations below, the direct radiation is monoenergetic, making the areal density transform in Eq. (2.3) appropriate. Thus errors in the reconstruction from the direct radiation data are mainly caused by the noise and measurement error in areal density and our approach to the density reconstruction problem (2.5).

It is well known that all experimental measurements have uncertainties. In high-energy X-ray radiography, scattered radiation is generally considered to be the greatest source of uncertainty. It arises from the interaction of X-rays with matter and needs to be subtracted, as these photons not only provide no useful information but also reduce the contrast and increase the noise level of the radiography. However, the scattering radiation cannot be measured directly in the experiment, especially for hydro-test experiments, so we have to count on theoretical and numerical methods, and this may result in errors. Besides, the incident radiation and the slope of the transmission curve also have errors in the measurement. Since the areal density is derived from measured optical density by using Eq. (2.3), these errors may bring about incorrect areal density transformation, and accordingly there are errors in the density reconstruction using the false areal density.

Under the existing experimental measurement conditions, it is applicable to approximate that the measurement errors of the optical density G and the distance d can be ignored. In this paper, we mainly consider the uncertainty of areal density measurement caused by the indeterminacy of measurements of physical quantities, including the scattering irradiation X_S, the incident exposure X_0 and the slope k of the transmission curve. From Eq. (2.3), we can obtain the measurement error of areal density induced by the determination error of scattering ΔX_S:

$$(\Delta b)_{X_S} = \left(\frac{\partial b}{\partial X_S}\right)\Delta X_S = \mathrm{SDR}\frac{\Delta X_S}{X_S}, \tag{4.1}$$

where $\mathrm{SDR} = \frac{X_S}{X_D}$ denotes the ratio of scattering to direct radiation. Similarly, the incident irradiation measurement error ΔX_0 brings

$$(\Delta b)_{X_0} = \left(\frac{\partial b}{\partial X_0}\right)\Delta X_0 = \frac{\Delta X_0}{X_0}, \tag{4.2}$$

to the uncertainty of areal density, and the measurement error of transmission curve Δk gives rise to

$$(\Delta b)_k = \left(\frac{\partial b}{\partial k}\right)\Delta k = (1+\mathrm{SDR})\frac{\Delta k}{k}, \tag{4.3}$$

then the uncertainty of areal density measurement is summarized as

$$\begin{aligned}\Delta b &= \sqrt{(\Delta b)^2_{X_S}+(\Delta b)^2_{X_0}+(\Delta b)^2_k} \\ &= \sqrt{\mathrm{SDR}^2\left(\frac{\Delta X_S}{X_S}\right)^2+\left(\frac{\Delta X_0}{X_0}\right)^2+(1+\mathrm{SDR})^2\left(\frac{\Delta k}{k}\right)^2}.\end{aligned} \tag{4.4}$$

5 Numerical experiments

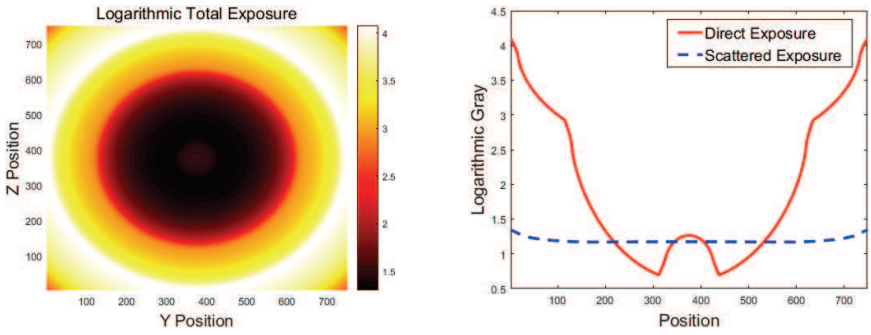
Instead of performing a real radiography experiment, we test on synthetic data generated by the Monte Carlo simulation, which is capable of simulating almost the entire physical process of radiography, to demonstrate the ability of MCMC method to reconstruct an axially symmetric object. Imperfect as this may sound, it is enough to test the principles of our method. We use the Monte Carlo N-Particle (MCNP) transport code [29] to produce simulated radiograph for the FTO, which was designed to allow French and American experimenters to collaborate on high-energy X-ray radiography methods, analysis, and their detection [28].

In this numerical simulation, the source is assumed as a monoenergetic photon beam of 4MeV, and the source blur is treated as Gaussian, whose full width at half maximum is 0.3cm. The FTO is placed 200cm from the source, and the detector is 250cm behind the object. The FTO consists of a set of concentric spheres with a void region at the center. The void has a radius of 1.0cm. The second layer is tungsten with radius of 4.5cm and density of $18.9\mathrm{g/cm}^3$. The third layer is copper with radius of 6.5cm and density of $8.9\mathrm{g/cm}^3$. The detector diagnostics can give separate tallies of the radial distributions of the direct and scattered exposures. Fig. 2(a) shows the synthetic radiograph of the total exposure, and Fig. 2(b) shows the profiles at the equator of the direct and scattered exposure respectively. We can calculate the radial distribution of areal density from the distribution of the direct exposure using Eq. (2.2).

Obviously, the introduction of the hyperprior distributions in (3.4)-(3.6) requires the choice of hyperparameters α_λ, β_λ, α_δ, β_δ, ν, and Σ. Setting the MCMC sampler with these hyperparameters and initializations λ_0, δ_0 and $\mathbf{L}_0$, we computed 100,000 samples and made the last 95,000 available, which demonstrated stationarity and had little correlation between samples (see Fig. 4). From these samples, we plot the sample mean as our reconstruction, which is known as conditional mean (CM) estimate of the unknown density, and 95% credibility intervals given by the 0.025 and 0.975 quantiles of the samples at each location, which were computed using empirical quantiles.

5.1 Parameter selection

We have first got to study the role of hyperparameters. General choices for the shape and rate parameters in each Gamma distribution are given in [31,32], that is, $\alpha_\lambda=\alpha_\delta=1$ and $\beta_\lambda=\beta_\delta=10^{-4}$. Then the hyperpriors can be deemed to be "uninformative", and their



(a) (b)

Figure 2: MCNP simulation results of the FTO. (a) Radiograph capturing both direct and scattered radiation; (b) the central cross sections of the direct and scattered exposure.

effect on the sampled values for λ and δ are negligible. More success has been found by adapting the magnitude of the parameters to the particular problem [30]. But we discovered a new way to tune parameters for λ and δ from the model.

For the hyperparameters $(\alpha_\lambda,\beta_\lambda)$, it is easy to see that λ characterises the noise level σ^2 of areal density. The smaller the value, the higher the noise. Since the mean and variance of the corresponding Gamma distribution are $\alpha_\lambda/\beta_\lambda$ and $\alpha_\lambda/\beta_\lambda^2$ respectively, we choose appropriate α_λ and β_λ such that $\alpha_\lambda/\beta_\lambda$ is close to σ^{-2} and $\alpha_\lambda/\beta_\lambda^2$ is smaller. In our numerical experiments, we just set $\alpha_\lambda=\sigma^{-2}$ and $\beta_\lambda=1$.

We now study the effect of the hyperparameters $(\alpha_\delta,\beta_\delta)$. To this end, we employ the variable-controlling approach by setting λ and $\mathbf{L}$ as known. We then fix $\alpha_\delta=1$, and the corresponding Gamma distribution degenerates to an exponential distribution. The maximum a posteriori estimate for the pair $(\boldsymbol{\rho},\delta)$ are calculated for various values of β_δ. We then pick a proper $\hat{\beta}_\delta$ with good performance of $\hat{\boldsymbol{\rho}}$. Similar to λ, we select suitable α_δ and β_δ such that $\alpha_\delta/\beta_\delta$ is equal to $1/\hat{\beta}_\delta$ and $\alpha_\delta/\beta_\delta^2$ is smaller. In our numerical experiments, we just set $\alpha_\delta=\hat{\beta}_\delta^{-1}$ and $\beta_\delta=1$.

Last but not least, setting the Wishart degrees of freedom parameter $\nu=n+1$ ensures that the distribution is well defined [33, 34]. Since the Wishart has mean $\nu\Sigma$, the quality of the final reconstruction is supposed to be greatly improved by choosing a Σ that incorporates edge information. We then set $\Sigma=\frac{1}{\nu}\mathbf{L}_{\mathrm{TV}}(\boldsymbol{\rho}_{\mathrm{TV}})$ to center the hyperprior for $\mathbf{L}$ around the initial edge estimate provided by total variation (TV) solution $\boldsymbol{\rho}_{\mathrm{TV}}$, where $\mathbf{L}_{\mathrm{TV}}(\boldsymbol{\rho}_{\mathrm{TV}})=\mathbf{D}^{\mathrm{T}}\psi(\boldsymbol{\rho}_{\mathrm{TV}})\mathbf{D}$, $\psi(\boldsymbol{\rho}_{\mathrm{TV}}):=\mathrm{diag}\big(1/\sqrt{(\mathbf{D}\boldsymbol{\rho}_{\mathrm{TV}})^2+\eta}\big)$, η is a small positive constant, and $\mathbf{D}$ is the forward differencing matrix [30]. Nonetheless, our computation in next subsection will show that the algorithm is relatively insensitive to this estimate.

5.2 Parameter sensitivity

We now reconstruct the density distribution of the FTO with the resulting areal density from the direct radiation shown in Fig. 2(b), and provide a parameter sensitivity analysis

Table 1: Parameters statistics and comparison of mean square errors in Fig. 3.

Figure	α_λ	β_λ	α_δ	β_δ	$\|\boldsymbol{\rho}_{\text{MCMC}}-\boldsymbol{\rho}_{\text{true}}\|/\sqrt{n}$
(b)	10^6	10^0	10^2	10^0	0.0637392
(c)	10^6	10^0	10^2	10^0	0.0637399
(d)	10^6	10^0	10^2	10^0	0.0637404
(e)	7.3×10^5	10^0	10^2	10^0	0.0636181
(f)	7.3×10^5	10^0	10^2	10^1	0.0632581

for Σ and β_δ. For the MCMC sampler, the degrees of freedom parameter ν is set to be $n+1$, the Wishart scale matrix Σ is computed to be $\frac{1}{\nu}\mathbf{L}_{\text{TV}}(\boldsymbol{\rho}_{\text{TV}})$, and $\mathbf{L}_{\text{TV}}(\boldsymbol{\rho}_{\text{TV}})$ is also used for $\mathbf{L}_0$. The initial λ and δ parameters are drawn from a Uniform$(0,1)$ distribution. All the remaining parameters for the Gibbs sampler are selected as noted above and are given in Table 1.

The sensitivity of the TV solution, which is used to inform the hyperparameter Σ, is tested by providing the sampler with a TV solution that contained no edge information and a TV solution that indicated incorrect edge locations and densities, shown by a dot-dashed line in Fig. 3(c) and (d), respectively. The mean reconstructions are computed and shown by the solid line in Fig. 3(c-d). For ease of comparison, we also calculate the CM estimate (see Fig. 3(b)), initializing the sampler with a proper TV solution, and it coincides well with the real one. The mean square error is computed using the mean reconstruction compared to the true density profile of the FTO. As shown in Table 1, the mean square error of the original reconstruction is 0.0637392 and each of these two reconstructions differs from that by no more than 0.000002. Both of these reconstructions demonstrate the Bayesian method's ability to overcome a poorly informed Σ parameter, relying on the data to determine edges and density scales.

To understand the sensitivity of the solution to the parameter β_δ, we carry out reconstruction experiments on the areal density (see the solid line in Fig. 3(a)), which is corrupted with Gaussian noise at level 1.5% of maximum of the noiseless projection data. Samples of the MCMC chain are drawn from the conditional distributions, holding all but β_δ constant and changing β_δ from 10^0 to 10^1. We calculate the sample mean and 95% credibility interval from the samples, as shown in Fig. 3(e) and (f), respectively. Combined with the mean square error in Table 1, we can observe that the MCMC sampler with different β_δ makes very little difference.

We can summarize from the data in Table 1 that $(\alpha_\delta,\beta_\delta)$ can be applicable to the projection data of the same object in different situations, while the selection of $(\alpha_\lambda,\beta_\lambda)$ is related to the noise level of the projected data, and different values are obtained for different data. It should be noted that this is consistent with our parameter selection rule in last subsection.

Discussion 5.2.1. To investigate the stationarity of the MCMC samples, we use the autocorrelation function of the chain to measure the correlation between samples. Ideally,

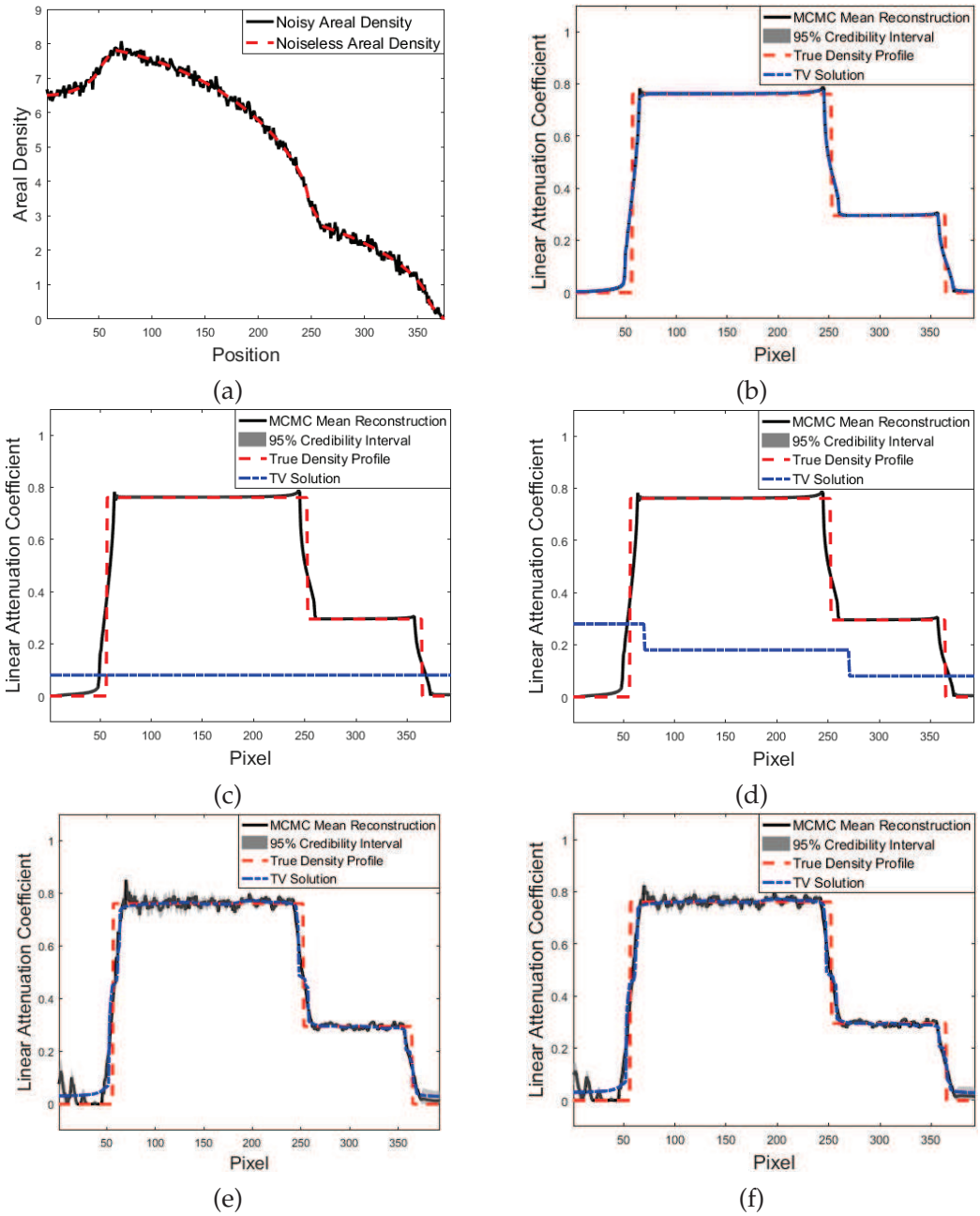


Figure 3: Density reconstructions for the data set which resulted from direct radiation only. (a) Noisy areal density is compared with noiseless areal density; (b) the original reconstruction from the noiseless projection data with the parameters chosen in Subsection 5.1; (c) a comparative reconstruction using a flat TV solution to create Σ; (d) a comparative reconstruction using a noisy TV solution to create Σ; (e) the original reconstruction from the noisy projection data with the parameters chosen in Subsection 5.1; (f) a comparative reconstruction with $\beta_\delta = 10^1$ as a test of the parameter sensitivity.

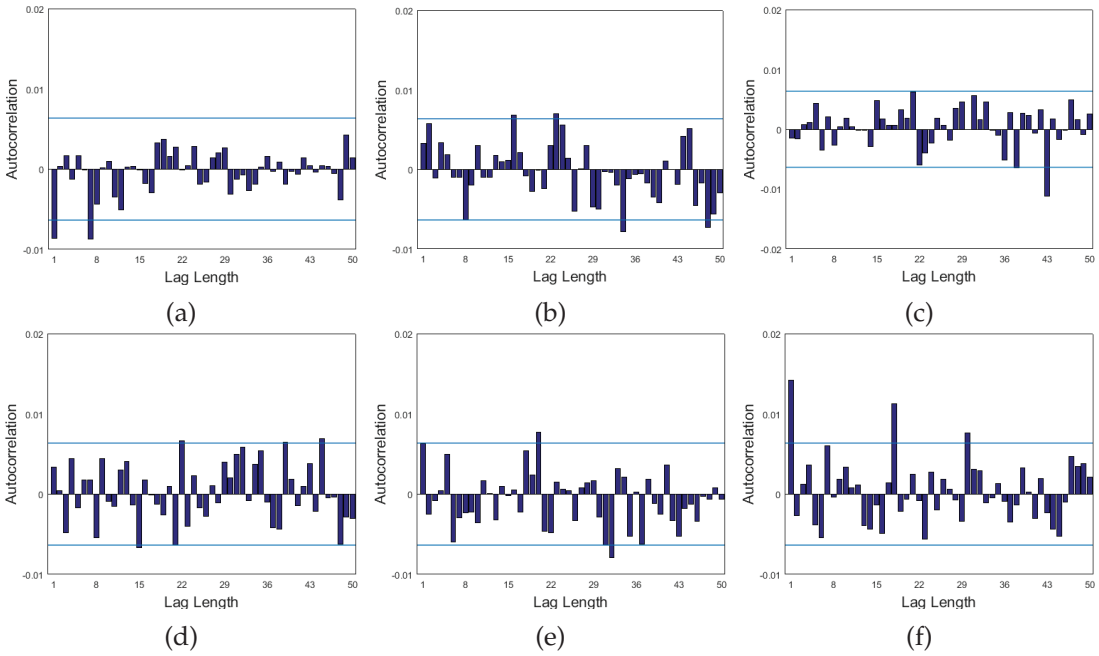


Figure 4: Autocorrelation plots for the MCMC samples of parameters ρ (a), λ (b) and δ (c), corresponding to the test in Fig. 3(b). The respective autocorrelation plots for the parameters samples corresponding to Fig. 3(e) are given in (d)-(f). The autocorrelations of the ρ samples are of the 197th element.

the autocorrelation function for a sequence of MCMC samples will decay quickly to zero. Autocorrelation plots for ρ, λ and δ are shown in Fig. 4 and demonstrate the samples are not significantly correlated. The rejection region lines for test of individual autocorrelations effectively indicate that the samples are stationary.

Discussion 5.2.2. For the convergence behavior of our MCMC approach, we have shown in Fig. 5 the one dimension marginal posterior distributions of the three parameters by the MCMC algorithm. We can observe that the marginal densities of the parameters samples are consistent with the posterior distributions (3.12)-(3.14). Specially, the probability peak in Fig. 5(a) is located at 0.7638, which is almost the true density value. The probability peak in Fig. 5(b) is at point 1.0002×10^6, and is in close agreement with the hyperprior distribution of λ chosen in Table 1.

5.3 Uncertainty caused by noise

To analyse the uncertainty introduced by noise in the imaging process, we fall back on the mean reconstruction and its uncertainty quantification shown in Fig. 3(e). Compared with the posterior estimation from noiseless data, as shown in Fig. 3(b), credibility interval at every location becomes obvious when noise added, and all of them are within 10% of the density value there. However, the uncertainty of density reconstruction from

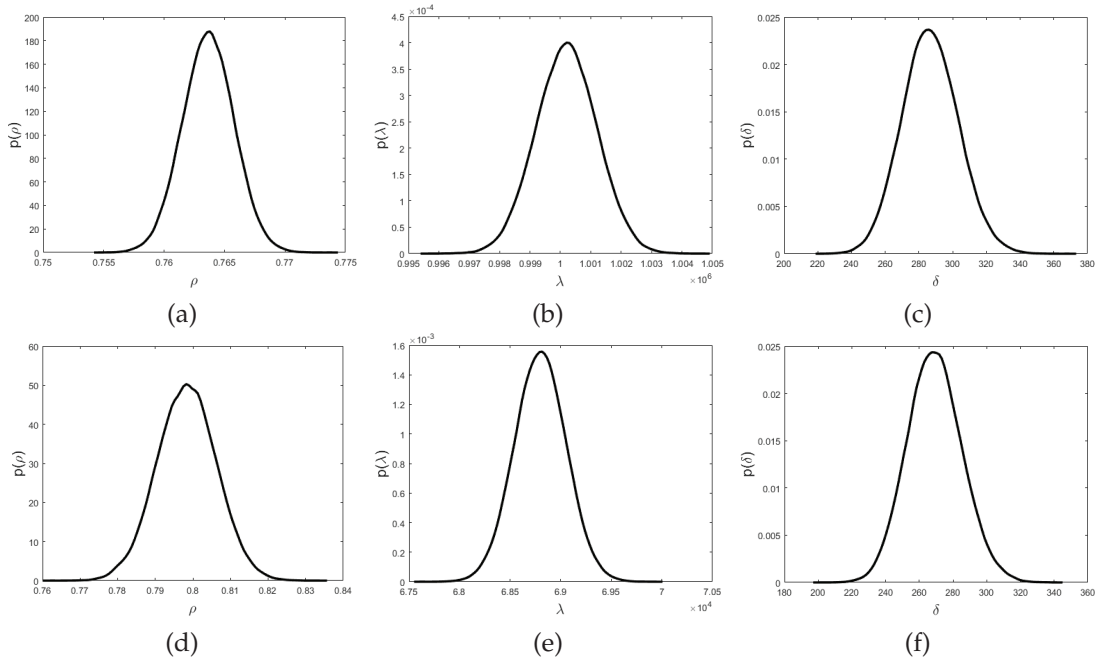


Figure 5: The one dimension marginal posterior distributions of parameters ρ (a), λ (b) and δ (c) obtained with our MCMC algorithm, corresponding to the sampling test in Fig. 3(b). The respective marginal densities for the parameters samples corresponding to Fig. 3(e) are given in (d)-(f). The marginal densities of the ρ samples are of the 197th element.

ideal radiograph mainly exists near the interface of materials. This is related to the ill-posedness of inverse problems, and reflects the uncertainty caused by numerical algorithm.

When comparing the CM estimate with TV estimate in Fig. 3(e), in the terminology of the classical regularization theory one is tempted to say that the former is underregularized compared to the latter. However, from the statistical point of view, the CM estimate is consistent with the prior. That is, if one uses a white noise prior with unknown variance, the CM estimate is expected to display the characteristics of a noise image.

5.4 Uncertainty caused by measurement error of physical quantities

According to the accuracy of experimental measurement, the error range of main physical quantities mentioned in Section 4 can be approximated by

$$\frac{\Delta X_S}{X_S}\leq 10\%,\quad \frac{\Delta X_0}{X_0}\leq 5\%,\quad \frac{\Delta k}{k}\leq 10\%,$$

then we can obtain the measurement error of areal density from Eq. (4.4), that is

$$\Delta b = 0.1\sqrt{\mathrm{SDR}^2+0.25+(1+\mathrm{SDR})^2}. \tag{5.1}$$

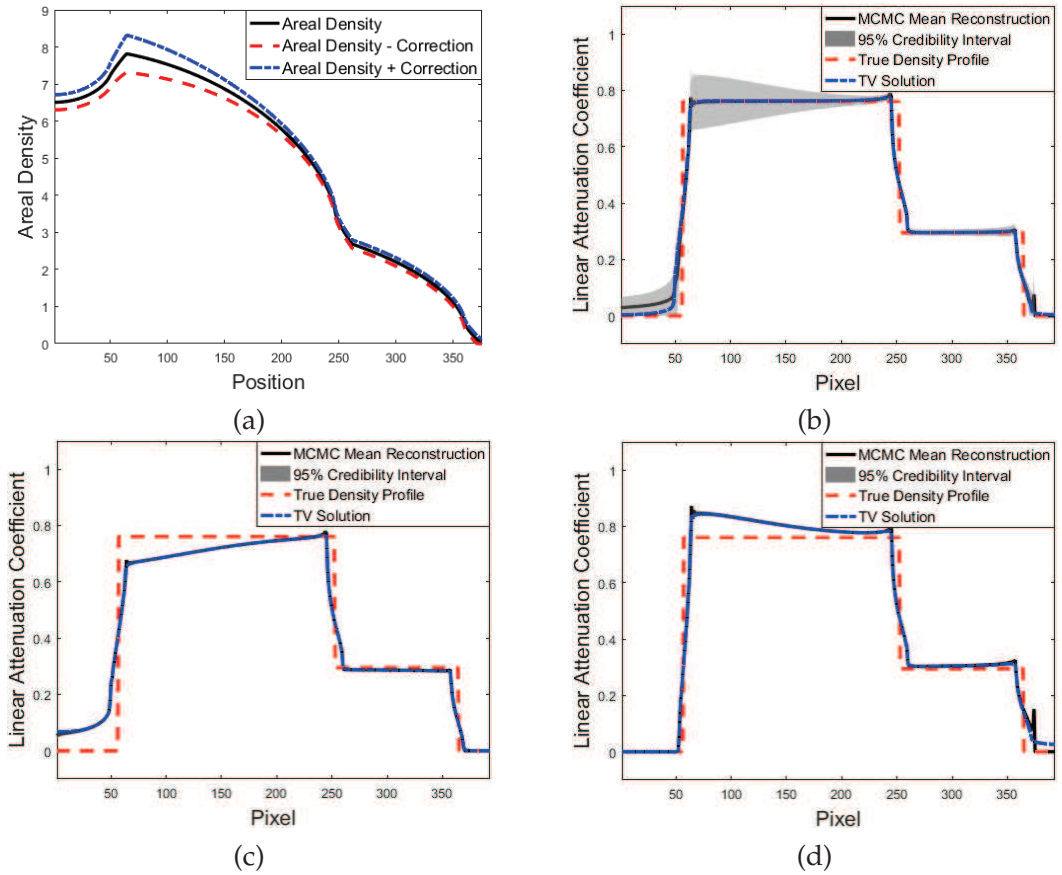


Figure 6: Density reconstruction for corrected projection data. (a) The corrected areal densities are compared to the original areal density; (b) the combination of (c) and (d); (c) reconstruction result from the areal density after subtracting correction, displayed as red dashed line in (a); (d) reconstruction result from the areal density after adding correction, displayed as blue dashdot line in (a).

In consideration of the uncertainties in the measurements of these physical quantities, we impose the correct value Δb on the areal density, see Fig. 6(a). Then we run MCMC for the corrected areal densities $b-\Delta b$ and $b+\Delta b$ separately. The mean reconstructions are given in Fig. 6(c) and (d), along with the 95% credibility intervals. To imitate the density uncertainty caused by measurement uncertainty, these two estimates are merged into single one, as shown in Fig. 6(b), whose density values are average of (c) and (d), and the pointwise credibility intervals are unions of two original intervals.

The mean square error between the density distributions in Fig. 6(b) and Fig. 3(b) is 0.0133972, which verifies the validity of our estimation method. The credible bands are fairly tight in (c) and (d) and significantly widen in (b), especially in the region near innermost material interface. That is, the uncertainty is low across the image except for the jump locations, as expected.

5.5 Uncertainty caused by noise and physical factors

Combining the above two Subsections 5.3 and 5.4, we get the noisy areal density in Fig. 3(a) to minus and plus the correction (5.1) separately, and obtain the corresponding noisy projection data with correction, as shown in Fig. 7(a). Based on the previously explained MCMC run, the CM estimates together with credible intervals are shown in Fig. 7(c) and (d), as well as the combination of them shown in Fig. 7(b).

It is easy to see that the MCMC reconstruction in Fig. 7(b) is very close to those in Fig. 3(e), and the mean square error between them is 0.0130941. However, the 95% credibility intervals become wider with the introduction of correction, particularly around the boundary of inner layers, and this is consistent with the case in Subsection 5.4.

In each example illustrated above, the uncertainty is seen to be lower near regions of constant density in the reconstruction and higher near edge locations. Moreover, the

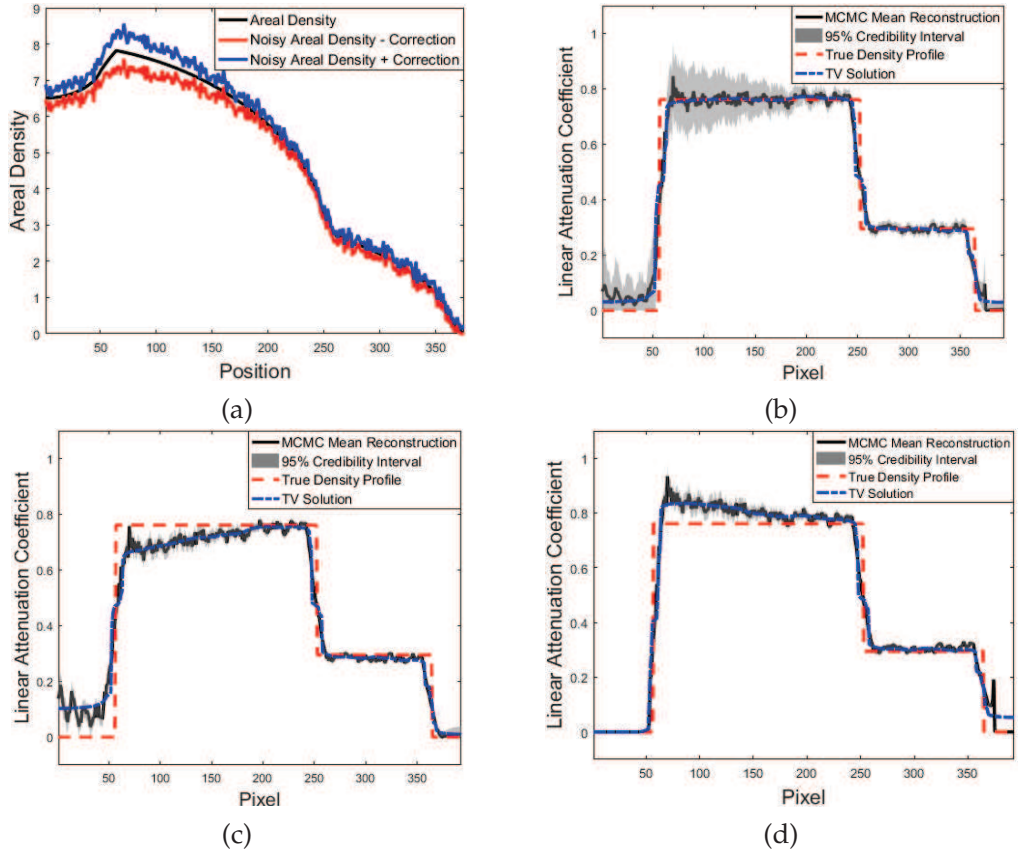


Figure 7: Density reconstruction for noisy projection data with correction. (a) The noisy areal densities with correction are compared to the original areal density; (b) the combination of (c) and (d); (c) reconstruction result from the noisy areal density after subtracting correction, displayed as red solid line in (a); (d) reconstruction result from the noisy areal density after adding correction, displayed as blue solid line in (a).

results show that some true density values do not lie within the 95% support of the estimated densities. This is because the blur in simulating the photography process leads to the non-unique boundary positions and density distributions around the boundaries, so the true density values of these areas can hardly be included in the 95% credibility intervals of the estimated densities.

6 Conclusion and future work

In this paper, a generalized density reconstruction model is presented based on the assumption that noise and blur will occur during the imaging process. Then, a hierarchical Bayesian model is proposed for computing object densities and estimating their uncertainties simultaneously. This approach includes a Gaussian noise model for X-ray radiographs in the construction of the likelihood function. A normal prior with nonnegativity restriction is assumed with both the scale parameter and precision matrix treated as hyperpriors. Combining the likelihood with the prior and hyperpriors, the resulting joint posterior distribution is appropriate for MCMC sampling. Density samples drawn from the conditional posterior distribution are insensitive to the choice in hyperparameter values selected based on the model. The sample mean and 95% credibility intervals are computed from the samples to characterize the unknown density and its corresponding uncertainty. Numerical experiments indicate that our MCMC method achieves quite effective reconstruction results and is comparable to TV regularization. Moreover, the uncertainty of density reconstruction is mainly introduced by noise and measurement error of physical quantities in the process of radiography.

In this paper, the X-rays are assumed to be parallel on different layers, and they form a fan-beam shape in each layer. In the future, we plan to estimate the uncertainty of cone beam reconstruction using the MCMC method. To this end, we have to develop parallel programs since the algorithm is computationally intensive and its memory overhead is large. Furthermore, we will consider other explicit noise models according to the characteristics of experimental data, such as poisson noise, impulse noise and a composition of them. We will also construct the prior density in more ways based on the nature of the prior information, for example, ℓ^1 prior, Cauchy density, entropy density, lognormal density, discontinuities prior and so on.

Acknowledgments

X. Li was supported by National Postdoctoral Program for Innovative Talents (grant No. BX201700038). H. Xu was supported by NSFC Funds (grant No. 11675021). N. Zheng was supported by NSFC Funds (grant No. 11505014). Q. Jia was supported by NSFC Funds (grant No. 11805018). T. Gu was supported by Science Challenge Project (grant No. TZ2016002) and NSFC Funds (grant No. 11671049). S. Wei was supported by NSFC Funds (grant No. 11571003).

References

[1] R. M. Lewitt, Reconstruction algorithms: Transform methods, Proceedings of the IEEE, 71 (1983), 390-408.
[2] F. Natterer, The Mathematics of Computerized Tomography, Wiley, New York, 1986.
[3] A. C. Kak and M. Slaney, Principles of Computerized Tomographic Imaging, IEEE Press, New York, 1987.
[4] X. Pan, E. Y. Sidky and M. Vannier, Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction, Inverse Problems, 25 (2008), 1230009-1230044.
[5] T. J. Asaki, R. Chartrand, K. R. Vixie and B. Wohlberg, Abel inversion using total-variation regularization, Inverse Problems, 21 (2005), 1895-1903.
[6] T. J. Asaki, P. R. Campbell, R. Chartrand, C. E. Powell, K. R. Vixie and B. E. Wohlberg, Abel inversion using total variation regularization: applications, Inverse Problems in Science and Engineering, 14 (2006), 873-885.
[7] S. Wei, S. Wang and H. Xu, Regularization method for axially symmetric objects tomography from a single X-ray projection data, Journal of Image and Graphics, 13 (2008), 2275-2280.
[8] R. H. Chan, H. Liang, S. Wei, M. Nikolova and X.-C. Tai, High-order total variation regularization approach for axially symmetric object tomography from a single radiograph, Inverse Problems and Imaging, 9 (2015), 55-77.
[9] L. A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography, IEEE Transactions on Medical Imaging, 1 (1982), 113-122.
[10] G. T. Herman, Application of Maximum Entropy and Bayesian Optimization Methods to Image Reconstruction from Projections, Springer Netherlands, 1985.
[11] S. Iwasaki, S. Odanaka, Y. Shintoku, M. Kitamura, M. Haruyama, M. Takase and K. Ara, New CT image reconstruction algorithm based on the Bayes estimation, Nucl. Instrum. Methods Phys. Res., Sect. A, 422 (1999), 683-687.
[12] J. P. Kaipio and E. Somersalo, Statistical and Computational Inverse Problems, Springer Science & Business Media, 2005.
[13] K. M. Hanson, A Bayesian Approach to Nonlinear Inversion: Abel Inversion from X-Ray Attenuation Data, Transport Theory, Invariant Imbedding, and Integral Equations, Lecture Notes in Pure and Applied Mathematics, 115 (1989), 363-378.
[14] J. M. Bardsley and C. Fox, An MCMC method for uncertainty quantification in nonnegativity constrained inverse problems, Inverse Problems in Science and Engineering, 20 (2012), 477-498.
[15] E. A. H. Vollebregt, The Bound-Constrained Conjugate Gradient Method for Non-negative Matrices, Journal of Optimization Theory and Applications, 162(2014), 931-953.
[16] A. M. Stuart, Inverse problems: A Bayesian perspective, Acta Numerica, 19 (2010), 451-559.
[17] L. Tenorio, Statistical Regularization of Inverse Problems, SIAM Review, 43(2001), 347-366.
[18] D. Calvetti and E. Somersalo, A Gaussian hypermodel to recover blocky objects, Inverse Problems, 23(2007), 733-754.
[19] D. Calvetti and E. Somersalo, Introduction to Bayesian Scientific Computing, Springer, New York, 2007.
[20] D. Calvetti and E. Somersalo, Hypermodels in the Bayesian imaging framework, Inverse Problems, 24(2008), 034013.
[21] B. Hosseini, C. Mougenot, S. Pichardo, E. Constanciel, J. M. Drake and J. Stockie, A bayesian approach for energy-based estimation of acoustic aberrations in high intensity focused ul-

trasound treatment, Commun. Comput. Phys., 25(2019), 1564-1590.
[22] S. Brooks, A. Gelman, G. L. Jones and X.-L. Meng, Handbook of Markov Chain Monte Carlo, Chapman & Hall/CRC, Boca Raton, 2011.
[23] J. Li and Y. M. Marzouk, Adaptive Construction of Surrogates for the Bayesian Solution of Inverse Problems, SIAM Journal on Scientific Computing, 36(2014), A1163-A1186.
[24] L. Yan and T. Zhou, Adaptive multi-fidelity polynomial chaos approach to Bayesian inference in inverse problems, Journal of Computational Physics, 381(2019), 110-128.
[25] L. M. M. V. D. Bos, B. Sanderse, W. A. A. M. Bierbooms and G. J. W. V. Bussel, Bayesian model calibration with interpolating polynomials based on adaptively weighted Leja nodes, Communications in Computational Physics, (2019).
[26] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier and F. Lenzen, Variational Methods in Imaging, Springer, New York, 2009.
[27] G. H. Golub and C. F. V. Loan, Matrix Computations, Johns Hopkins University Press, Baltimore, 1996.
[28] T. J. T. Kwan, A. R. Mathews, P. J. Christenson and C. M. Snell, Integrated system simulation in X-ray radiography, Computer Physics Communications, 142(2001), 263-269.
[29] X-5 Monte Carlo Team, MCNP – A General Monte Carlo N-Particle Transport Code, Version 5, LA-UR-03-1987, Los Alamos Nuclear Laboratory, 2003.
[30] M. Howard, M. Fowler, A. Luttman, S. E. Mitchell and M. C. Hock, Bayesian Abel inversion in quantitative X-Ray radiography, SIAM Journal on Scientific Computing, 38(2016), B396-B413.
[31] J. M. Bardsley, MCMC-based image reconstruction with uncertainty quantification, SIAM Journal on Scientific Computing, 34(2012), A1316-A1332.
[32] D. Higdon, A primer on space-time modeling from a Bayesian perspective, LA-UR-05-3097, Los Alamos Nuclear Laboratory, 2006.
[33] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.
[34] A. Gelman, J. B. Carlin, H. S. Stern, D. Dunson, A. Vehtari and D. B. Rubin, Bayesian Data Analysis, CRC Press, Boca Raton, 2013.