# THE BERNSTEIN-VON MISES THEOREM FOR THE INVERSE SCATTERING PROBLEM

TAKASHI FURUYA AND JENN-NAN WANG

Abstract. In this work, we study the inverse scattering problem of determining an unknown refractive index from the far-field measurements using the Bayesian approach. Our aim is to prove the Bernstein-von Mises theorem for this inverse scattering problem when the unknown coefficient is in a class of piecewise constant functions. We also provide numerical simulations based on the MCMC algorithm combining with a Gibbs-type sampling and MALA to verify the Bernstein-von Mises theorem.

## 1. Introduction

In this paper, we apply the Bayes method to the inverse scattering problem having a piecewise constant coefficient as the unknown refractive index. Our aim is to prove a Bernstein-von Mises theorem and uncertainty quantification of the estimation by the posterior mean. To describe the inverse problem, let $q \in L^\infty(\mathbb{R}^n)$ be a complex-valued function satisfying $q \equiv 0$ in $\mathbb{R}^n \setminus \bar{B}$, where $B$ is some open ball in $\mathbb{R}^n$ ($n = 2$ or 3). We consider the propagation of an acoustic wave in $\mathbb{R}^n$ scattered by the inhomogeneous medium $q$. Let $u_q = u^{\text{inc}} + u_q^{\text{sca}}$ satisfy

$$\Delta u_q + k^2(1+q)u_q = 0 \quad \text{in} \quad \mathbb{R}^n, \tag{1.1}$$

and the Sommerfeld radiation condition

$$\lim_{|x|\to\infty} |x|^{\frac{n-1}{2}} \left( \frac{\partial u_q^{\text{sca}}}{\partial |x|} - iku_q^{\text{sca}} \right) = 0. \tag{1.2}$$

Assume that $u^{\text{inc}}$ is the plane incident field, i.e. $u^{\text{inc}} = e^{ikx\cdot d}$ with $d \in \mathbb{S}^{n-1}$. Then the scattered field $u_q^{\text{sca}}$ satisfies

$$u_q^{\text{sca}}(x,d) = \frac{e^{ik|x|}}{|x|^{\frac{n-1}{2}}} u_q^\infty(\hat{x}, d) + O(|x|^{-\frac{n+1}{2}}) \quad \text{as} \quad |x| \to \infty, \tag{1.3}$$

where $\hat{x} = x/|x|$, see for example, [Ser17, Page 232]. The inverse scattering problem is to determine the perturbation of the refractive index $q$ from the knowledge of the *far-field pattern* or *scattering amplitude* $u_q^\infty(\hat{x}, d)$ for all $\hat{x}, d \in \mathbb{S}^{n-1}$ at one fixed energy $k^2$. The investigation of the inverse scattering problem of identifying the inhomogeneity or the obstacle has been well documented. We refer the reader to [CK19, CC14, CCH22, KG07, NP15] and references therein for the thorough development of the problem. On the other hand, the book [Che18] describes related results from the engineer's viewpoint.

It was known that the far-field pattern $u_q^\infty(\hat{x}, d)$ uniquely determines the near-field data of (1.1) on $\partial B$, which in turn, determines the Dirichlet-to-Neumann map of (1.1) provided $k^2$

2020 *Mathematics Subject Classification.* 35R30, 62G05.

*Key words and phrases.* Inverse scattering problem, Bayes method, Bernstein-von Mises theorem, MCMC.

Wang was partially supported by the National Science and Technology Council of Taiwan, NSTC 112-2115-M-002-010-MY3.

is not an Dirichlet eigenvalue of $\Delta + k^2(1+q)$ on $B$, for example, see [Nac88]. Combining this fact and the uniqueness results proved in [SU87, Buk08], one can show that the far-field pattern $u_q^\infty(\hat{x}, d)$ for all $\hat{x}, d \in \mathbb{S}^{n-1}$ uniquely determines $q$, at least when $q$ is essentially bounded. For the stability, a log-type estimate was derived in [HH01] reflecting the severe ill-posedness of the inverse problem. However, a Lipschitz stability estimate for the inverse scattering problem with a piecewise constant refractive index was proved in [Bou13] or in [AS22] by $u_q^\infty(\hat{x}, d)$, $\forall\, \hat{x}, d \in \mathbb{S}^{n-1}$ or by $u_q^\infty(\hat{x}, d)$ for finite number of $\hat{x}, d$, respectively.

A nonparametric Bayesian approach for estimation of the unknown inhomogeneity of the refractive index $q$ from noisy discrete measurements of the far-field pattern $\{u_q^\infty(\hat{x}_i, d_j)\}$ for a suitable set of $\hat{x}_i, d_j \in \mathbb{S}^{n-1}$ was recently studied in [FKW23], where the consistency of the posterior distribution with a contraction rate of $(\ln N)^{-\delta}$ for some $\delta > 0$ was established, where $N$ is the number of statistical experiments. We remark that the logarithmic-type contraction rate is due to the severe ill-posedness of the inverse scattering problem. There is a vast of literature investigating the consistency of the nonparametric Bayes method for both linear and nonlinear inverse problems. We refer the reader to [ALS13, ASZ14, GGVDV00, KVDVvZ11, Ray13, Vol13, AN20, GN20, Kek22, MNP21a] and references therein. Furthermore, two nice monographs [GVdV17, Nic23] contain more exhaustive references.

Inspired by the result in [Boh22], where the EIT problem was studied in the case of piecewise constant conductivities, we will study the statistical guarantees for the inverse scattering problem when the inhomogeneity $q$ lies in a *fixed finite dimensional space*. The goal is to prove the Bernstein-von Mises (BvM) theorem and to establish the estimation of $q$ by posterior means with optimal convergence guarantees. The optimality is understood in the sense of the asymptotic minimax optimal variance for estimators of $q$ with square loss. The BvM theorem implies that credible sets have valid frequentist coverage of true parameter asymptotically. Under the condition that the unknown parameter $q$ is piecewise constant, the main theorem can be regarded as a *parametric* BvM theorem. For more general description of the parametric BvM theorem, we refer to [VdV00] for further details.

In addition to the theoretical study of the BvM theorem, we also provide a numerical verification of the theorem in this paper. The key is to sample the posterior distribution obtained from the Bayes formula and to compute the posterior mean. Our numerical method is based on the MCMC algorithm combining with a Gibbs-type sampling and MALA (Metropolis-adjusted Langevin algorithm). The simulations in Section 6 clearly demonstrate that the posterior mean is an efficient estimator of the unknown piecewise constant inhomogeneity $q$.

This paper is organized as follows. In Section 2, we formulate the inverse scattering problem in the statistical setting and state the main results. We then describe the consequences of the BvM theorem, Theorem 2.4, in Section 3. In Section 4, we prove some analytic aspects of the inverse scattering problem in the case of piecewise constant inhomogeneity $q$. In Section 5, we present the proofs of theorems stated in Section 2. Finally, some numerical simulations are given in Section 6.

## 2. General framework and main results

In this section, we will follow the framework introduced in [MNP21b], [NP21], or [Boh22] to setup the statistical model. Let $(\mathcal{X}, \mathcal{A}, \lambda)$ be a probability space and $V$ be a finite-dimensional vector space of fixed dimension $D$ with inner product $\langle \cdot, \cdot \rangle_V$ and norm $\|\cdot\|_V$. We denote $L^\infty(\mathcal{X}) := L^\infty(\mathcal{X}, \mathcal{A}, \lambda)$ the bounded measurable functions and $L^2(\mathcal{X}) := L^2(\mathcal{X}, \mathcal{A}, \lambda)$

the space of $\lambda$-square integrable functions with corresponding inner product $\langle \cdot, \cdot \rangle_{L^2(\mathcal{X})}$. On the other hand, let $P$ be a probability measure on the product space $V \times \mathcal{X}$ and $L^2(P) = L^2(V \times \mathcal{X}, P)$ the square integrable functions on $V \times \mathcal{X}$ with respect to $P$.

Let $\Theta$ be the set of parameters which is a bounded subset of $K$-dimensional space. The inverse problem is encoded in a forward map

$$G : \Theta \to L^2(\mathcal{X}), \quad \Theta \ni \theta \to G_\theta.$$

We consider the measurement model with Gaussian noise

$$Y_i = G_\theta(X_i) + \varepsilon_i, \quad i = 1, \cdots, N, \tag{2.1}$$

where $X_i \overset{iid}{\sim} \lambda$ and $\varepsilon_i \overset{iid}{\sim} \mathcal{N}_D(0, I)$, a $D$-variant normal distribution with mean 0 and covariance $I$. Here $N$ denotes the number of experiments. We also assume that $\{X_i\}_{i=1}^N$ and $\{\varepsilon_i\}_{i=1}^N$ are drawn independently. Let $P_\theta$ be the probability distribution of $(Y_1, X_1)$ and $P_\theta^N = P_\theta \otimes \cdots \otimes P_\theta$ the law of the random vector $Z_N = \{(Y_i, X_i), \ i = 1, \cdots, N\}$. We use $E_\theta^N$ to represent the corresponding expectation operator. Finally, we denote the sequence of experiments $\mathcal{P}_N = \{P_\theta^N : \theta \in \Theta\}$. For $N = 1$, we write $\mathcal{P} = \mathcal{P}_1 = \{P_\theta : \theta \in \Theta\}$, which is sometimes called *model* or *experiment*.

Next we put the inverse scattering problem with a piecewise constant medium into the framework described above. Assume that $\{\hat{x}_j : j \in \mathbb{N}\} \subset \mathbb{S}^{n-1}$ and $\{d_l : l \in \mathbb{N}\} \subset \mathbb{S}^{n-1}$ are two dense subsets of $\mathbb{S}^{n-1}$ (possibly identical). Let $S \in \mathbb{N}$ be chosen (will be specified later) and set $(\mathcal{S}, U) = \{(1, \cdots, S), \mathrm{Uniform}\}$. We consider two kinds of observation data corresponding to (2.1). Let $q = q_\theta$ be determined by $\theta \in \Theta$ and define $G_\theta$ through the far-field pattern $u_q^\infty(\hat{x}, d)$.

(I) Consider the iid random samples $X_i \in (\mathcal{X}, \lambda) = (\mathcal{S}, U)$, $i = 1, 2, \cdots, N$. Let us denote

$$G_\theta(X_i) = \left(\mathrm{Re}(u_q^\infty(\hat{x}_{X_i}, d_1)), \mathrm{Im}(u_q^\infty(\hat{x}_{X_i}, d_1)), \cdots, \mathrm{Re}(u_q^\infty(\hat{x}_{X_i}, d_S)), \mathrm{Im}(u_q^\infty(\hat{x}_{X_i}, d_S))\right)^T.$$

Note that $G_\theta(X_i) \in \mathbb{R}^{2S}$ and thus $V = \mathbb{R}^{2S}$, i.e., $D = 2S$.

(II) Next, we consider $X_i = (X_i^{(1)}, X_i^{(2)}) \in (\mathcal{X}, \lambda) = (\mathcal{S} \otimes \mathcal{S}, U \otimes U)$, $i = 1, \cdots, N$, and set

$$G_\theta(X_i) = \left(\mathrm{Re}(u_q^\infty(\hat{x}_{X_i^{(1)}}, d_{X_i^{(2)}})), \mathrm{Im}(u_q^\infty(\hat{x}_{X_i^{(1)}}, d_{X_i^{(2)}}))\right)^T.$$

In this case, $G_\theta(X_i) \in \mathbb{R}^2$ and $V = \mathbb{R}^2$, i.e., $D = 2$.

Now we describe the parameter space $\Theta$. Let $\{B_i\}_{i=1}^J$ be non-empty open subsets of $B$ satisfying $B_i \cap B_j = \emptyset$ for $i \neq j$ and $\bar{B} = \cup_{i=1}^J \bar{B}_i$. Let $l, L, m, M$ be real numbers with $m > 0$ and $\{q_i\}_{i=1}^J$ be complex numbers satisfying

$$\ell \leq \mathrm{Re}(q_i) \leq L, \ \ m \leq \mathrm{Im}(q_i) \leq M, \ \text{ for } \ i = 1, \cdots, J. \tag{2.2}$$

We then assume the refractive index $q = \sum_{i=1}^J q_i \chi_i$, where $\chi_i$ is the characteristic function of $B_i$ and $q_i$ satisfies (2.2). In other words, the parameter space is

$$\Theta = ([\ell, L] \times [m, M])^J \subset \mathbb{R}^{2J}, \tag{2.3}$$

namely, $K = 2J$, and $\theta = \otimes_{i=1}^J (\mathrm{Re}(q_i), \mathrm{Im}(q_i)) \in \Theta$.

Let us consider a prior $\Pi$ on $\Theta$ with a continuous and positive Lebesgue density $\pi(\theta) \in C(\Theta, (0, \infty))$, i.e.,

$$d\Pi(\theta) = \pi(\theta) d\theta. \tag{2.4}$$

By Bayes' formula, the posterior distribution $\Pi(\cdot|Z_N)$ given the observation data $Z_N \in (V \times \mathcal{X})^N$ is

$$\Pi(A|Z_N) = \frac{\int_A e^{\ell_N(\theta)} d\Pi(\theta)}{\int_\Theta e^{\ell_N(\theta)} d\Pi(\theta)},$$

for any Borel set $A \subseteq \Theta$, where

$$\ell_N(\theta) = \ell_N(\theta|Z_N) = -\frac{1}{2}\sum_{i=1}^N \|Y_i - G_\theta(X_i)\|_V^2,$$

is the joint log-likelihood function.

To state main theorems, we would like to describe "information theoretic" features of the model $\mathcal{P}$.

**Definition 2.1.** The model $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is said to be *differentiable in quadratic mean* (DQM) if there exists a linear map (*score operator* of $\mathcal{P}$ at $\theta$)

$$\mathbb{A}_\theta : \mathbb{R}^K \to L^2(V \times \mathcal{X}, P_\theta),$$

such that

$$\int_{V\times\mathcal{X}} \left[ dP_{\theta+h}^{1/2} - dP_\theta^{1/2} - \frac{1}{2}\mathbb{A}_\theta[h] dP_\theta^{1/2} \right]^2 = o(\|h\|^2) \quad \text{as} \quad \|h\| \to 0, \tag{2.5}$$

where $\|h\|$ is any equivalent norm of $\mathbb{R}^K$. The *information matrix/operator* of $\mathcal{P}$ at $\theta$ is defined by

$$\mathbb{N}_\theta = E^{P_\theta}[\mathbb{A}_\theta^T \mathbb{A}_\theta] : \mathbb{R}^K \to \mathbb{R}^K, \text{ i.e., } \mathbb{N}_\theta \in \mathbb{R}^{K\times K}. \tag{2.6}$$

**Remark 2.2.** *As explained in* [Boh22, Remark 4.3], $\mathbb{A}_\theta^T(Y, X) = \dot{\ell}_\theta(Y, X)$ *and* $\mathbb{N}_\theta = E^{P_\theta}[\dot{\ell}_\theta(Y, X)\dot{\ell}_\theta^T(Y, X)]$*. We refer to* [VdV00, (5.38)] *for the definition of* $\dot{\ell}_\theta(Y, X)$*.*

The first theorem is to prove the invertibility of the information matrix $\mathbb{N}_\theta$ in our inverse scattering problem.

**Theorem 2.3.** *Let the model* $\mathcal{P}$ *arise from the statistical scattering experiment* (2.1) *with finite dimensional parameters* (*piecewise constant coefficient described in* (2.2))*. Then the information matrix* $\mathbb{N}_\theta$ *is invertible.*

Next theorem is the main result of the paper, a Bernstein-von Mises theorem for the inverse scattering problem with the piecewise constant coefficient. Denote $\mathcal{L}(\cdot)$ the law of a random variable in $\mathbb{R}^K$ and $\|\cdot\|_{\mathrm{TV}}$ the total variation of probability measures on $\mathbb{R}^K$. Furthermore, $|\cdot|$ denotes any Euclidean norm in $\mathbb{R}^K$.

**Theorem 2.4.** *Let the experiments* $\mathcal{P}_N, N \in \mathbb{N}$*, arise from the statistical scattering experiment* (2.1) *with the inhomogeneity* $q$ *defined by* (2.3)*. Assume that the prior* $\Pi$ *on* $\Theta$ *is given by* (2.4) *and* $\theta_0 \in \Theta \setminus \partial\Theta$*. Let the observation data* $Z_N = \{(Y_i, X_i)\}_{i=1}^N$ *be described in either* (I) *or* (II) *above with a sufficiently large* $S$ *given in* Lemma 4.1*. Then, if* $\theta \sim \Pi(\cdot|Z_N)$*, then*

$$\left\| \mathcal{L}\left(\sqrt{N}(\theta - \Psi_{\theta_0,N})|Z_N\right) - \mathcal{N}_K(0, \mathbb{N}_{\theta_0}^{-1}) \right\|_{\mathrm{TV}} \xrightarrow{P_{\theta_0}^N} 0, \text{ as } N \to \infty, \tag{2.7}$$

*where* $\mathcal{N}_K(0, \mathbb{N}_{\theta_0}^{-1})$ *is a* $K$*-variate normal distribution with mean* $0$ *and variance* $\mathbb{N}_{\theta_0}^{-1}$*, and the re-centering*

$$\Psi_{\theta_0,N} = \theta_0 + \frac{1}{N}\sum_{i=1}^N \mathbb{N}_{\theta_0}^{-1}\mathbb{A}_{\theta_0}(Y_i, X_i).$$

## 3. Consequences of Theorem 2.4

We will postpone the proofs of Theorem 2.3 and 2.4 to Section 5. Here we would like to discuss some interesting consequences of Theorem 2.4. In order to apply Theorem 2.4 to obtain an uncertainty quantification for estimation of $\theta$, we need to choose a more tractable centering instead of $\Psi_{\theta_0,N}$. A natural choice is the posterior mean $\bar{\theta}_N = E^{\Pi}[\theta|Z_N]$, which is computable, for example, via MCMC. In other words, the following holds true:

$$\left\|\mathcal{L}\left(\sqrt{N}(\theta-\bar{\theta}_N)|Z_N\right) - \mathcal{N}_K(0,\mathbb{N}_{\theta_0}^{-1})\right\|_{\mathrm{TV}} \xrightarrow{P_{\theta_0}^N} 0, \text{ as } N\to\infty, \tag{3.1}$$

where, as above, $\theta \sim \Pi(\cdot|Z_N)$. To prove (3.1), it suffices to show that

$$\sqrt{N}(E^{\Pi}[\theta|Z_N] - \Psi_{\theta_0,N}) \xrightarrow{P_{\theta_0}^N} 0, \text{ as } N\to\infty, \tag{3.2}$$

which can be achieved by following the argument in the proof of [MNP21b, Lemma 4.6] or in Step VII of the proof of [Nic20, Theorem 8]. In carrying out the proof, we need a consistency theorem with explicit contraction rate for the posterior distribution $\Pi(\cdot|Z_N)$. This consistency result can be established by using the method in [FKW23]. On the other hand, since $\sqrt{N}\Psi_{\theta_0,N} \xrightarrow{d} W$, where $W := \mathcal{N}_K(0,\mathbb{N}_{\theta_0}^{-1})$, it follows from (3.2) that

$$\sqrt{N}(\bar{\theta}_N - \theta_0) \xrightarrow{d} W. \tag{3.3}$$

On the other hand, from (2.7), we can also derive a consistency theorem for the posterior distribution $\Pi(\cdot|Z_N)$ with the optimal contraction rate $1/\sqrt{N}$. Precisely, for every sequence $\{M_N\}$ with $M_N\to\infty$ such that

$$\Pi(|\theta-\theta_0| > M_N/\sqrt{N}|Z_N) \xrightarrow{P_{\theta_0}^N} 0. \tag{3.4}$$

Indeed, (3.4) can be proved by using (2.7), Prohorov's theorem, and the CLT applied to $\frac{1}{\sqrt{N}}\sum_{i=1}^N \mathbb{N}_{\theta_0}^{-1}\mathbb{A}_{\theta_0}(Y_i,X_i)$.

Another useful application of Theorem 2.4 is the Bayesian uncertainty quantification for $\theta$ and its implication to the frequentist probability coverage of the true parameter. Let $0<\alpha<1$ be a given confidence level. Define Bayesian credible sets $C_N$ and the quantiles $R_N$ by

$$C_N = \{\theta\in\Theta : |\theta-\bar{\theta}_N| \le R_N/\sqrt{N}\}, \quad \Pi(C_N|Z_N) = 1-\alpha.$$

Observe that the determination of credible sets $C_N$ requires only the computation of the posterior mean $\bar{\theta}_N$ and the quantiles $R_N$, which can be calculated numerically by an MCMC method. Following the argument in [Nic23, Sec 4.1.3], let us compute the frequentist coverage of $C_N$ as $N\to\infty$. Denote $\Phi(t) = \Pr(|W|\le t)$. Consequently, $\Phi$ is continuous and strictly increasing. In view of (3.1), we have

$$\Phi(R_N) = \Phi(R_N) - \Pi(C_N|Z_N) + (1-\alpha) \xrightarrow{P_{\theta_0}^N} (1-\alpha).$$

The continuity of the inverse $\Phi^{-1}$ yields

$$R_N \xrightarrow{P_{\theta_0}^N} \Phi^{-1}(1-\alpha).$$

By Slutsky's lemma and (3.3), we obtain

$$\frac{\Phi^{-1}(1-\alpha)}{R_N}\sqrt{N}(\bar{\theta}_N-\theta_0) \xrightarrow{d} W. \tag{3.5}$$

Consequently, from (3.5) and the continuity of $|\cdot|$, it yields

$$\begin{aligned} P^N_{\theta_0}(\theta_0 \in C_N) =& P^N_{\theta_0}(\sqrt{N}|\theta_0 - \bar{\theta}_N| \leq R_N) \\ =& P^N_{\theta_0}\left(\frac{\Phi^{-1}(1-\alpha)}{R_N}\sqrt{N}|\theta_0 - \bar{\theta}_N| \leq \Phi^{-1}(1-\alpha)\right) \\ \to& \Phi(\phi^{-1}(1-\alpha)) = 1-\alpha. \end{aligned}$$

In other words, the posterior credible sets $C_N$ are efficient frequentist confidence sets for the parameter $\theta_0$.

## 4. Some analytic aspects of the inverse scattering problem

In this section, we will review some theoretical aspects of the inverse scattering problem before proving the theorems stated in Section 3. One of the main goal is to very the DQM property for the statistical experiment arising from the inverse scattering problem. Let $u_q(x,d)$ be the total field satisfying (1.1) and (1.2). Equivalently, $u_q$ can be expressed by the Lippmann-Schwinger integral equation

$$u_q(x,d) = u^{\mathrm{inc}}(x,d) + k^2 \int_B q(y) u_q(y,d) \Psi(x-y) dy, \tag{4.1}$$

where $u^{\mathrm{inc}}(x,d) = e^{ikx\cdot d}$ and $\Psi(x)$ is the fundamental solution of the Helmholtz operator in $\mathbb{R}^n$, i.e.,

$$\Psi(x) = \begin{cases} \dfrac{i}{4} H_0^{(1)}(k|x|) & (n=2) \\ \dfrac{e^{ik|x|}}{4\pi|x|} & (n=3) \end{cases},$$

where $H_0^{(1)}(z)$ is the Hankel function of order 0. The far-field pattern can be explicitly written as

$$u_q^\infty(\hat{x},d) = C_n \int_B e^{-ik\hat{x}\cdot y} q(y) u_q(y,d) dy, \tag{4.2}$$

where $u_q(x,d)$ is the solution to (4.1), and the constant $C_n$ depending on only $n$ is given by

$$C_n = \begin{cases} \dfrac{k^2 \mathrm{e}^{\frac{i\pi}{4}}}{\sqrt{8\pi k}} & (n=2) \\ \dfrac{k^2}{4\pi} & (n=3) \end{cases}.$$

It is not hard to show that for any $\hat{x}, d \in \mathbb{S}^{n-1}$

$$|u_q^\infty(\hat{x},d)| \leq C\|q\|_{L^\infty(B)}(1 + \|q\|_{L^\infty(B)}), \tag{4.3}$$

where $C > 0$ is an absolute constant. Define the forward map

$$T : L^\infty(B) \to L^2(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}),$$

by $Tq := u_q^\infty(\hat{x},d)$. As mentioned in the Introduction, from [Nac88] and [SU87], we see that

$$T \ \text{ is injective,} \tag{4.4}$$

i.e., if $q_1, q_2 \in L^\infty(B)$ and $Tq_1 = Tq_2$, then $q_1 = q_2$.

Define $L^\infty_+(B) = \{q \in L^\infty(B) : \mathrm{Im}(q) \geq m > 0 \text{ a.e.}\}$ and $E_J = \{q \in L^\infty(B) : q = q_\theta = \sum_{l=1}^J (\theta_l + i\tilde{\theta}_l)\chi_l, (\theta_1, \tilde{\theta}_1, \cdots, \theta_J, \tilde{\theta}_J) \in \mathbb{R}^{2J}\}$, where $\chi_l$ is the characteristic function of $B_l$ and $\{B_l\}_{l=1}^J$ are domains defined in Section 2. From (4.4), it follows that

$$T|_{E_J \cap L^\infty_+(B)} \text{ is injective.} \tag{4.5}$$

Recall from [Bou13, Lemma 3.3, Lemma 3.4] that the mapping $T : L^\infty_+(B) \to L^2(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ is differentiable and its Fréchet derivative at $q \in L^\infty_+(B)$, $T'(q) : L^\infty(B) \to L^2(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$, is bounded and injective. In fact, from [CK19, Theorem 11.6], the Fréchet derivative $T'(q)$ can be expressed explicitly by

$$T'(q)h = v^\infty_{q,h}(\hat{x}, d) \quad \text{for all} \quad h \in L^\infty(B),$$

where $v^\infty_{q,h}$ is the far field pattern of the radiating solution $v = v_{q,h}$ to

$$\Delta v + k^2(1+q)v = -k^2 h u_q(\cdot, d) \quad \text{in} \quad \mathbb{R}^n.$$

From [BL05, Lemmas 2.2, 2.3, 2.4, 2.6] (or some related results in [FP21]) , we can see that for all $\hat{x}, d \in \mathbb{S}^{n-1}$

$$|u^\infty_{q+h}(\hat{x}, d) - u^\infty_q(\hat{x}, d) - (T'(q)h)(\hat{x}, d)| \leq \quad C_0(1 + \|q\|_{L^\infty(B)})\|h\|^2_{L^\infty(B)}, \tag{4.6}$$

and

$$|u^\infty_{q+h}(\hat{x}, d) - u^\infty_q(\hat{x}, d)| \leq C'_0(1 + \|q\|_{L^\infty(B)})^2\|h\|_{L^\infty(B)}, \tag{4.7}$$

for some general constants $C_0, C'_0 > 0$.

We also prove a (Lipschitz) stability of determining a piecewise constant coefficient $q$ by some finite measurements of the far-field pattern.

**Lemma 4.1.** *Assume that $\{\hat{x}_i, i \in \mathbb{N}\}$ and $\{d_l : l \in \mathbb{N}\}$ are two dense subsets of $\mathbb{S}^{n-1}$ (possibly identical). Let $W$ be a finite-dimensional subspace of $L^\infty(B)$ and $\mathcal{K} \subset W \cap L^\infty_+(B)$ a convex and compact set. Then, there exists a large enough $S_0 \in \mathbb{N}$ such that if $S \geq S_0$*

$$\|q_1 - q_2\|_{L^\infty(B)} \leq C_1 \| \left(u^\infty_{q_1}(\hat{x}_i, d_l) - u^\infty_{q_2}(\hat{x}_i, d_l)\right)_{i,l=1}^S \|_2, \quad q_1, q_2 \in \mathcal{K}, \tag{4.8}$$

*for some constant $C_1 > 0$. Moreover, for any $q \in \mathcal{K} \subset E_J \cap L^\infty_+(B)$, then for all $S \geq S_0$, there exists a constant $C_2 = C_2(S) > 0$ such that*

$$\|h\|_{L^\infty(B)} \leq C_2 \| \left(v^\infty_{q,h}(\hat{x}_i, d_l)\right)_{i,l=1}^S \|_2, \quad \textit{for all} \quad h \in L^\infty(B). \tag{4.9}$$

*Proof.* Estimate (4.8) is proved in [AS22, Theorem 5]. We will prove (4.9) using the idea in [AS22]. Recall that the far-field pattern is analytic on $\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}$. Since $\dim(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}) = 2(n-1)$, and $3 > \frac{2(n-1)}{2}$ (if $n = 2, 3$), $H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ is continuously embedded into $C(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$, and so it is a reproducing kernel Hilbert space (RKHS) consisting of continuous functions. Let $Q_S : H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}) \to H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ be the projection onto $Z_S$ defined by

$$Z_S := \mathrm{span}\{k_{(\hat{x}_i, d_l)} : i, l = 1, ..., S\},$$

where $k_{(\hat{x}_i, d_l)} \in H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ is a reproducing kernel for RKHS $H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})$ satisfying

$$f(\hat{x}_i, d_l) = \langle f, k_{(\hat{x}_i, d_l)} \rangle_{H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})}, \quad f \in H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1}).$$

By the argument in [AS22, Example 2 and Theorem 7], we have

$$Q_S \to I_{H^3(\mathbb{S}^{n-1} \times \mathbb{S}^{n-1})} \text{ (in norm topology)}, \text{ as } S \to \infty, \tag{4.10}$$

and

$$\|Q_S f\|_{H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} \le C_S \left\| (f(x_i, d_l))_{i,l=1}^S \right\|_2, \quad f \in H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1}), \tag{4.11}$$

for some constant $C_S > 0$. In view of (4.10), we can choose a large enough $S_0 \in \mathbb{N}$ such that for all $S \ge S_0$, $(I - \|I - Q_S\|) > 0$. By the injectivity of $T'(q) : L^\infty(B) \to L^2(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})$ ([Bou13, Lemma 3.4]), there is a constant $C > 0$ such that

$$\|T'(q)h\|_{L^2(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} \ge C\|h\|_{L^\infty(B)}, \quad h \in L^\infty(B),$$

which combining with (4.11) implies

$$\begin{aligned} C_S\| \left(v^\infty_{q,h}(\hat{x}_i, d_l)\right)_{i,l=1}^S \|_2 &\ge \|Q_S T'(q)h\|_{H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} \\ &\ge \|T'(q)h\|_{H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} - \|(I - Q_S)T'(q)h\|_{H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} \\ &\ge (I - \|I - Q_S\|)\|T'(q)h\|_{H^3(\mathbb{S}^{n-1}\times\mathbb{S}^{n-1})} \\ &\ge C(I - \|I - Q_S\|)\|h\|_{L^\infty(B)}, \end{aligned}$$

and thus (4.9) with $C_2 = C_S[C(I - \|I - Q_S\|)]^{-1}$. □

**Remark 4.2.** *In the case of piecewise constant refractive index described above, we simply take $W = E_J$ in* Lemma 4.1*.*

## 5. Proofs of Theorems

Let the parameter space $\Theta$ be defined as in (2.3) and consider $q = q_\theta \in E_J$ with $\theta \in \Theta$. The statistical forward map $G_\theta(X)$ with $X \in (\mathcal{X}, \lambda)$ is defined in either (I) or (II) in Section 2. It follows directly from (4.3) that $\mathcal{G} = \sup_{\theta\in\Theta} \|G_\theta\|_\infty < \infty$. For $\theta$ in the interior of $\Theta$, the Fréchet derivative $T'(q)$ is written by

$$T'(q) = \mathbb{I}_\theta,$$

where $\mathbb{I}_\theta : \mathbb{R}^K \to L^2(\mathcal{X})$ is defined in Proposition A.1. In view of (4.6), we see that (A.1) holds true, i.e., for any $x \in \mathcal{X}$ and $h \in \mathbb{R}^K$,

$$\|G_{\theta+h}(x) - G_\theta(x) - \mathbb{I}_\theta[h](x)\|_V = o(\|h\|) \text{ as } \|h\| \to 0.$$

Moreover, it can be deduce from (4.7) that (A.2) is satisfied as well. Combining all discussions above, we can conclude that the statistical model $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ satisfies DQM with the help of Proposition A.1.

*Proof of* Theorem 2.3. It is readily seen from (4.9) that $\mathbb{I}_\theta$ is injective. Therefore, in view of (A.4), the information matrix $\mathbb{N}_\theta$ is invertible. □

*Proof of* Theorem 2.4. Theorem 2.4 can be seen as a parametric Bernstein-von Mises theorem. In view of Remark 2.2, Theorem 2.4 is equal to [VdV00, Theorem 10.1]. The invertibility of the information matrix $\mathbb{N}_\theta$ is already proved in Theorem 2.3. As explained above, the statistical experiment $\mathcal{P}$ has the DQM property. The prior given in (2.4) is clearly absolutely continuous at any interior point of $\Theta$. To verify the separation condition [VdV00, (10.2)], by noting the paragraph before [VdV00, Theorem 10.1], it suffices to use the fact that $\Theta$ is compact, $\theta \to P_\theta$ is continuous, and check the identifiability $P_\theta$, which is satisfied due to (4.8).

□

## 6. Numerical simulations

In this section, we provide numerical simulations to demonstrate the reconstruction of the piecewise constant inhomogeneous medium $q$ by the Bayes method and to justify the BvM theorem.

6.1. **Set up.** We simply consider the case when $n = 2$. The data in the inverse scattering problem is the far-field pattern $u_q^\infty$ which can be explicitly written as

$$u_q^\infty(\hat{x}, d) = \frac{k^2 \mathrm{e}^{\frac{i\pi}{4}}}{\sqrt{8\pi k}} \int_B e^{-ik\hat{x}\cdot y} q(y) u_q(y, d) dy, \quad (\hat{x}, d) \in \mathbb{S}^1 \times \mathbb{S}^1.$$

We approximately compute the above integral by the Monte Carlo method, where the total field $u_q$ in the integral is the solution to the Lippmann-Schwinger integral equation (4.1). This integral equation is numerically calculated based on Vainikko's method [SV01, Vai00, Kre13]. Let $S \in \mathbb{N}$ and set

$$\hat{x}_j := (\cos(2\pi j/S), \sin(2\pi j/S)), \quad j = 1, ..., S,$$

and

$$d_l := (\cos(2\pi l/S), \sin(2\pi l/S)), \quad l = 1, ..., S.$$

Then, our statistical experiments is given by

$$Y_i = G_{\theta^{(true)}}(X_i) + \epsilon_i,$$

and

$$X_i = (X_i^{(1)}, X_i^{(2)}) \overset{iid}{\sim} (\mathcal{S} \otimes \mathcal{S}, U \otimes U),$$

where $(\mathcal{S}, U) = \{(1, \cdots, S), \text{Uniform}\}$, and

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma I_2),$$

and $\theta^{(true)} \in \mathbb{R}^{2J}$ is the true parameter. Here, $G_\theta : \mathcal{S} \times \mathcal{S} \to \mathbb{R}^2$ is given by

$$G_\theta(X_i) = \left( \mathrm{Re}(u_\theta^\infty(\hat{x}_{X_i^{(1)}}, d_{X_i^{(2)}}), \mathrm{Im}(u_\theta^\infty(\hat{x}_{X_i^{(1)}}, d_{X_i^{(2)}}) \right).$$

In other words, in this simulation, we only consider the collection of observation data described in (II) in Section 2. We believe that the results for this case is enough to justify our method. Let the prior $\Pi$ be the uniform distribution on $\Theta$ and the observation data $Z_N = (X_i, Y_i)_{i=1,...,N}$, then the posterior distribution $\Pi(\theta|Z_N)$ is given by

$$\Pi(\theta|Z_N) \propto e^{\ell_N(\theta)}, \tag{6.1}$$

where $\ell_N(\theta)$ is the log-likelihood function given by

$$\ell_N(\theta) = -\frac{1}{2\sigma^2} \sum_{i=1}^N \|Y_i - G_\theta(X_i)\|_2^2.$$

We want to point out that the posterior (6.1) should be multiplied by the characteristic function of $\Theta$. Fortunately, in the numerical computation, ignoring this term is harmless.

**Algorithm 1** Metropolis–Hastings

**Require:** Initial guess $\theta^{(0)}$, Proposal ($2J$-dimensional) distribution $P$, Log-likelihood function $\ell_N$.

```
for t=0,...,T' do
    Sample θ^(*) from P(·|θ^(t)).
    Compute A = A(θ^(*)|θ^(t)).
    Generate a uniform random number u ∈ [0,1].
    if u < A then
        θ^(t+1) = θ^(*) (Accept θ^(*))
    else
        θ^(t+1) = θ^(t) (Reject θ^(*))
    end if
end for
Correct accepted states, re-named as {θ^(t)}_{t≤T}.
return A sequence {θ^(t)}_{t≤T}.
```

6.2. **Sampling algorithm.** Here, we adopt the Metropolis–Hastings algorithm [MRR+53, Has70, Tie98], which is a popular MCMC method to obtain samples of the posterior distribution $\Pi(\theta|Z_N)$. For completeness, we give the details of this scheme in Algorithm 1.

When parameter space is high-dimensional, it is known that finding a suitable proposal distribution is difficult as the individual dimension may behave differently. As an alternative approach, we also use the Gibbs sampling, in which one chooses the candidate for each dimension iteratively, rather than the one for all dimensions at once [GG84, GS90]. Denoting $\theta \in \mathbb{R}^{2J}$ by

$$\theta = (\theta_1, ..., \theta_{2J}), \quad \theta_j \in \mathbb{R},$$

we write the details of this scheme in Algorithm 2. The proposal distribution $P(\cdot|\theta^t)$ and the acceptance probability $A$ in Algorithm 1 and 2 will be specified below.

Here, we consider two cases of $P$ and $A$:

- Normal distribution, that is,

$$P(\cdot|\theta) = \mathcal{N}(\theta, hI).$$

  In this case, we employ the acceptance probability $A$ as

$$A(\cdot|\theta) = \min\{1, e^{\ell_N(\cdot) - \ell_N(\theta)}\}.$$

- MALA [Bes94, RT96, RR98], that is,

$$P(\cdot|\theta) = \mathcal{N}(\theta + \frac{h}{2}\nabla \ell_N(\theta), hI).$$

  In this case, we employ the acceptance probability $A$ as

$$A(\cdot|\theta) = \min\{1, e^{\ell_N(\cdot) - \ell_N(\theta) + \frac{1}{2h^2}\|\cdot - \theta + \frac{h^2}{2}\nabla_\theta \ell_N(\theta)\|_2^2 - \frac{1}{2h^2}\|\theta - \cdot + \frac{h^2}{2}\nabla_{\cdot} \ell_N(\cdot)\|_2^2}\}.$$

As above, $\mathcal{N}(\theta, \Sigma)$ denotes the multivariate normal distribution with mean $\theta$ and covariance matrix $\Sigma$.

**Algorithm 2** Gibbs sampling

**Require:** Initial guess $\theta^{(0)}$, Proposal (one-dimensional) distribution $P$, Log-likelihood function $\ell_N$.

$\theta^{(0,0)} = \theta^{(0)}$
**for** t=0,...,T' **do**
  **for** j=1,...,2J **do**
    Sample $\theta_j^{(*)}$ from $P(\cdot|\theta^{(t,j-1)})$.
    Denote by $\theta^{(*)} = (\theta_1^{(t,j-1)}, ..., \theta_{j-1}^{(t,j-1)}, \theta_j^{(*)}, \theta_{j+1}^{(t,j-1)}, ..., \theta_{2J}^{(t,j-1)})$.
    Compute $A = A(\theta^{(*)}|\theta^{(t,j-1)})$.
    Generate a uniform random number $u \in [0, 1]$.
    **if** $u < A$ **then**
      $\theta^{(t,j)} = \theta^{(*)}$ (Accept $\theta^{(*)}$)
    **else**
      $\theta^{(t,j)} = \theta^{(t,j-1)}$ (Reject $\theta^{(*)}$)
    **end if**
  **end for**
  $\theta^{(t+1,0)} = \theta^{(t,2J)}$
**end for**
Correct accepted states, re-named as $\{\theta^{(t)}\}_{t\le T}$.
**return** A sequence $\{\theta^{(t)}\}_{t\le T}$.

6.3. **Visualization.** In what follows, we always employ the following parameters:

- Wave number $k = 2.5$.
- Noise level $\sigma = 0.1$.
- Initial guess $\theta^{(0)} \equiv 0$ (both real and imaginary parts are 0).

By using sampling methods described in the previous subsection, we collect a sequence of samplings form the posterior distribution $\Pi(\theta|Z_N)$, denoting its sequence by $\{\theta^{(t)}\}_{0\le t\le T}$ where $T$ is the number of acceptances.

Figure 1 shows two inhomogeneous refractive indices with four and eight pieces, which are our ground truths. The dimensions of their parameter space are $2J = 2 \times 4 = 8$ and $2J = 2 \times 8 = 16$, respectively.

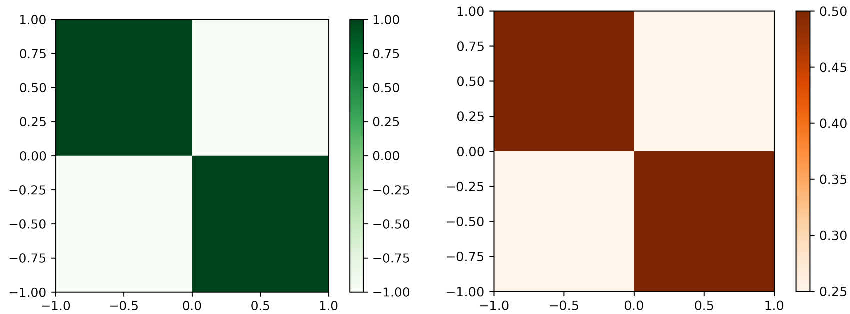
Figures 2, 3, 4, and 5 are the visualizations of sample means given by

$$\overline{\theta}^{(T)} = \frac{1}{T+1}\sum_{t=0}^{T}\theta^{(t)},$$

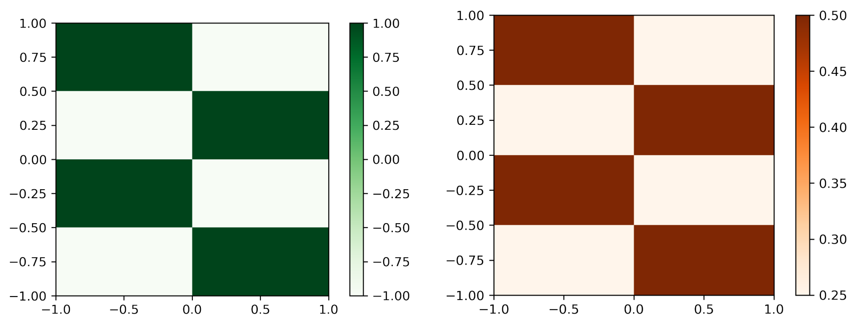
In our simulations, we use following four combinations of sampling algorithms and proposal distributions:

$$(\text{Algorithm, Proposal}) = \begin{cases} (\text{Metropolis–Hastings, Normal distribution}) \\ (\text{Metropolis–Hastings, MALA}) \\ (\text{Gibbs sampling, Normal distribution}) \\ (\text{Gibbs sampling, MALA}) \end{cases}.$$

For Metropolis–Hastings sampling and Gibbs sampling, we choose $h = 0.05$ and $h = 0.1$, respectively. In the case of four pieces, we take $S = 10$ and $N = 10$, while in the case of eight



(A) Four pieces



(B) Eight pieces

FIGURE 1. Two ground truth $q$'s. The first and second columns correspond to the real and imaginary parts, respectively.

pieces, we take $S = 30$ and $N = 30$. The first and second columns in Figures 2, 3, 4, and 5 correspond to the real and imaginary parts of sample means, respectively. The third column in Figures 2, 3, 4, and 5 represent the mean square error (MSE) between the true parameter $\theta^{(true)}$ and the sample mean up to the current state $\theta^{(t)}$ defined by

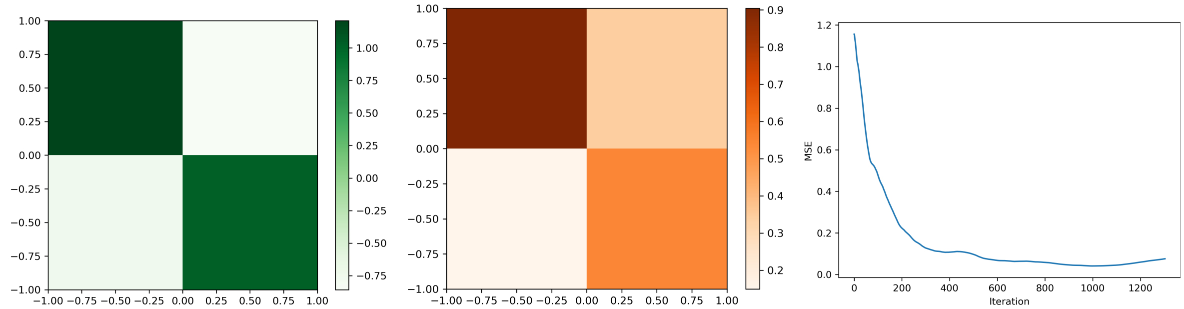
$$\overline{\theta}^{(t)} = \frac{1}{t+1}\sum_{i=0}^{t} \theta^{(t)},$$

i.e., MSE $e_t$ is given by

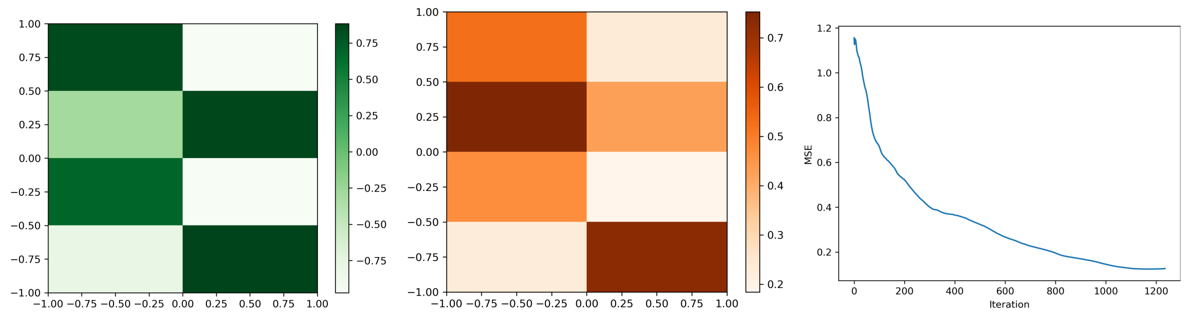
$$e_t := \|\theta^{(true)} - \overline{\theta}^{(t)}\|_2^2,$$

for $t = 0, ..., T$. From these figures, we can see that the combination of (Algorithm, Proposal) = (Gibbs sampling, MALA) provides the best reconstruction among other pairs.

The verification of the BvM theorem is shown in Figure 6. Precisely, Figure 6 shows the histogram of the samples obtained from the law $\mathcal{L}\left(\sqrt{N}(\theta - \bar{\theta}_N)|Z_N\right)$, where $\theta \sim \Pi(\cdot|Z_N)$, in the increasing order of $N = 5, 15, 25, 35$ and fixed $S = 100$. From the numerical results above, we demonstrate the BvM theorem using the combination of Gibbs sampling and MALA and the case of four pieces ($J = 4$) for the true parameter $q$. Each histogram is a one-dimensional plot showing the projection of the samples onto a certain eigenvector of $\mathbb{N}^{-1}_{\theta^{(true)}} \in \mathbb{R}^{8\times 8}$, which can be computed by (A.4). Here, we only plot three histograms corresponding to three largest eigenvalues. The blue curve in each figure is the one-dimensional normal with zero mean whose variance is the eigenvalue of the associated eigenvector for $\mathbb{N}^{-1}_{\theta^{(true)}}$.

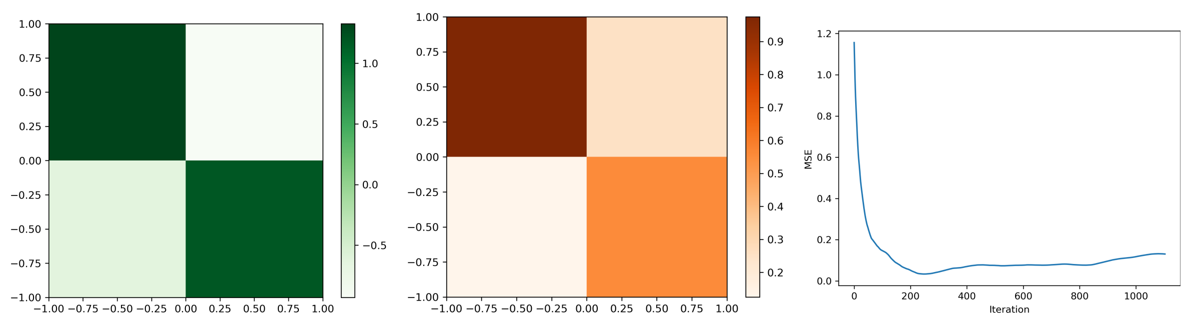



(A) Four pieces

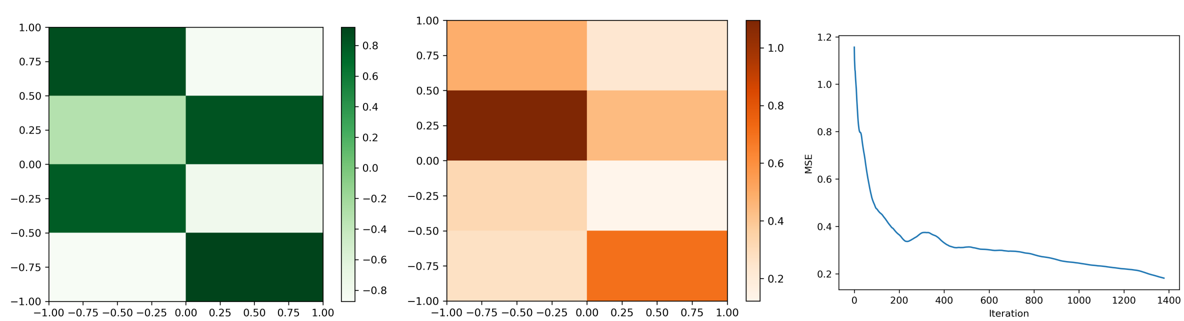



(B) Eight pieces

FIGURE 2. (Metropolis–Hastings, Normal distribution)

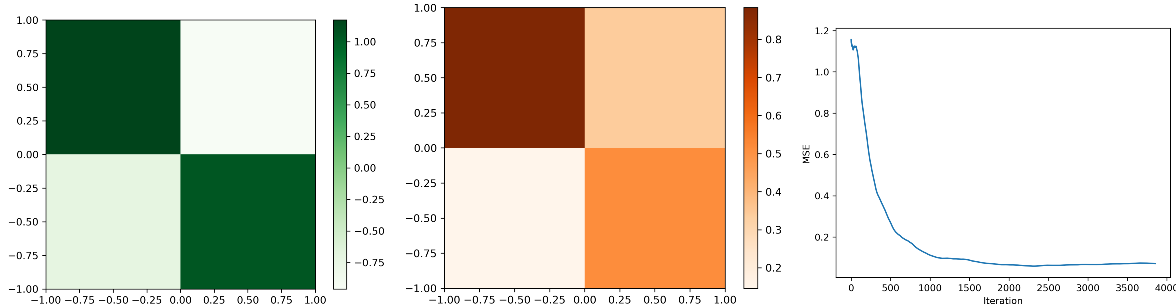



(A) Four pieces

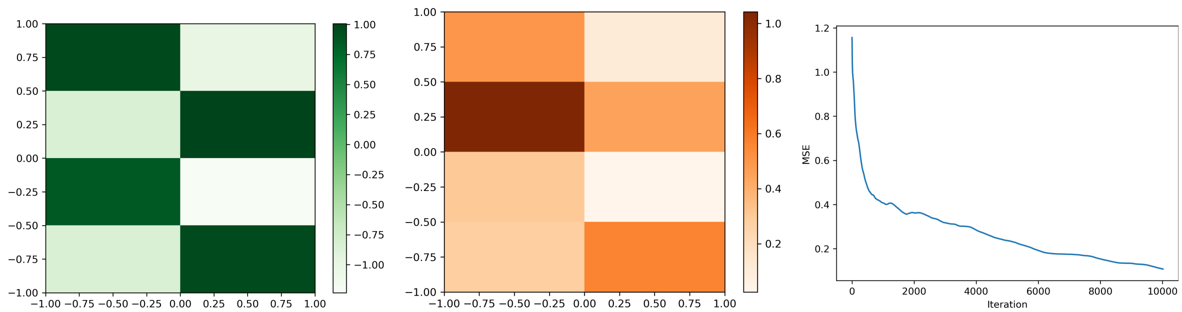



(B) Eight pieces

FIGURE 3. (Metropolis–Hastings, MALA)

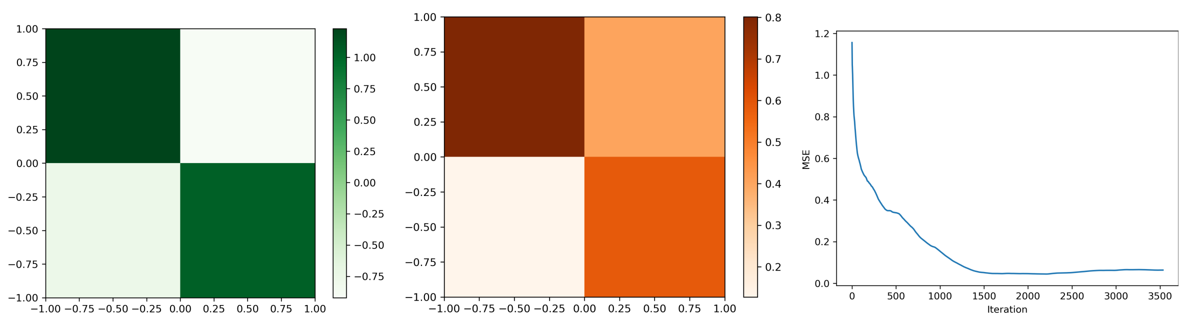



(A) Four pieces

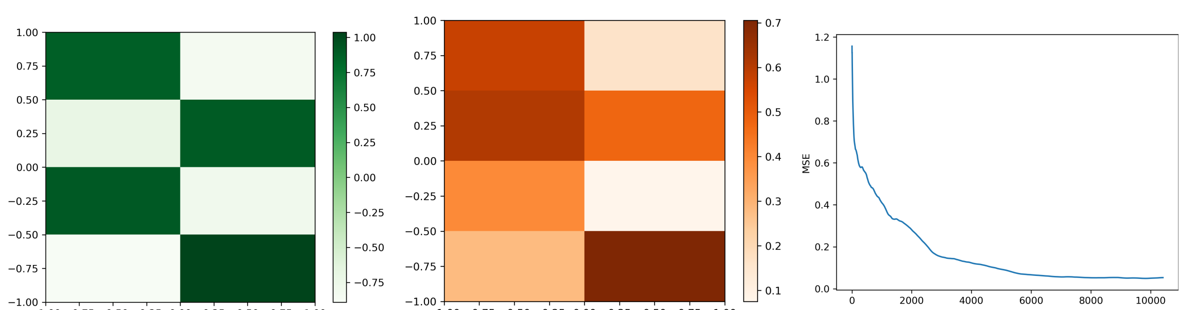



(B) Eight pieces

FIGURE 4. (Gibbs sampling, Normal distribution)

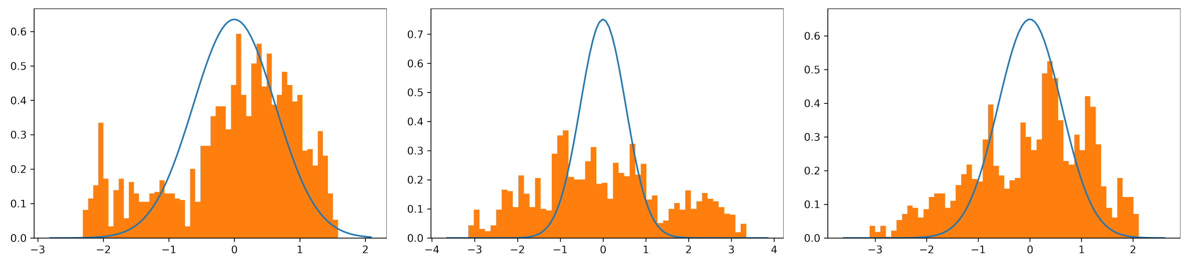



(A) Four pieces

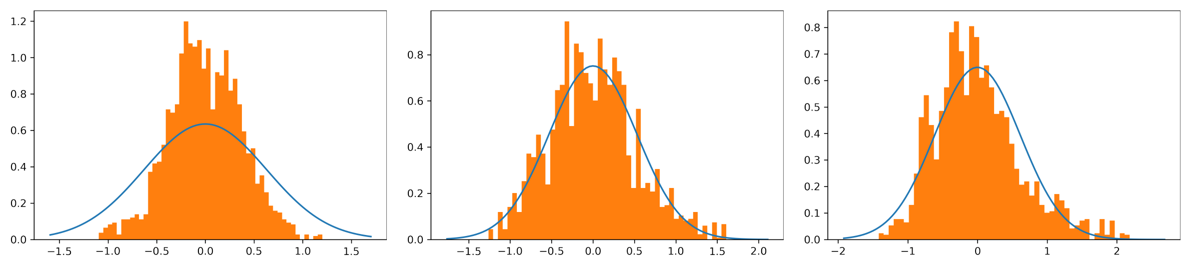



(B) Eight pieces

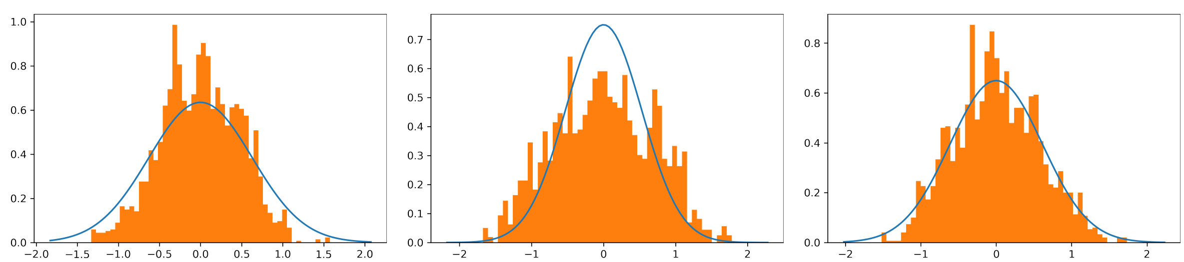
FIGURE 5. (Gibbs sampling, MALA)



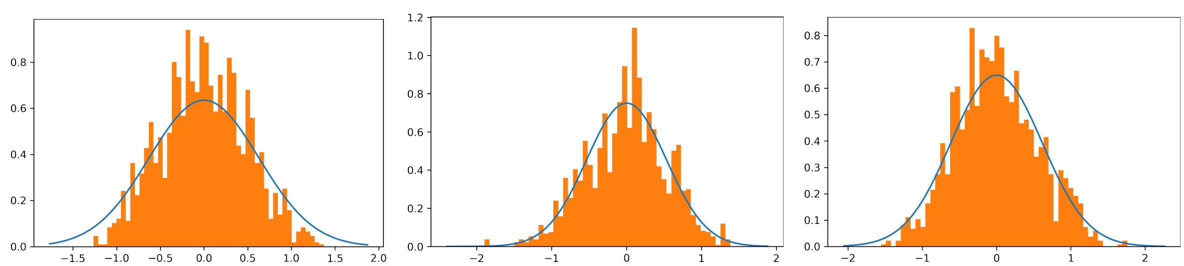
(A) $N = 5$



(B) $N = 15$



(C) $N = 25$



(D) $N = 35$

FIGURE 6. Demonstration of BvM theorem. Samples are obtained by using Gibbs sampling and MALA.

Figure 6 indicates that as $N$ increases, the shape of the histogram is getting closer to the blue curve, which clearly justifies the BvM theorem for the inverse scattering problem considered in this paper.

## Appendix A. DQM and Information operator

As described above, let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ be a statistical experiment polluted by Gaussian noise, arising from the forward map

$$G : \Theta \subset \mathbb{R}^K \to L^2(\mathcal{X}), \quad \theta \mapsto G_\theta.$$

The purpose of the section is to show that the differentiability of $G$ implies the DQM property of $\mathcal{P}$.

**Proposition A.1.** *Let $\mathcal{G} := \sup_{\theta\in\Theta} \|G_\theta\|_\infty < \infty$ and $\theta$ be in the interior of $\Theta$. Assume that there exists a bounded linear operator $\mathbb{I}_\theta : \mathbb{R}^K \to L^2(\mathcal{X})$ such that for all $x \in \mathcal{X}$*

$$\|G_{\theta+h}(x) - G_\theta(x) - \mathbb{I}_\theta[h](x)\|_V = o(\|h\|) \ \text{ as } \ \|h\| \to 0. \tag{A.1}$$

*Furthermore, suppose that there exist $\epsilon, B > 0$ such that*

$$\|G_{\theta+h} - G_\theta\|_\infty \le B\|h\| \ \text{ for } \ \|h\| < \epsilon. \tag{A.2}$$

*Then $\mathcal{P}$ satisfies the DQM property at $\theta$ defined in* Definition 2.1 *with the score operator $\mathbb{A}_\theta$ given by*

$$\mathbb{A}_\theta : \mathbb{R}^K \to L^2(V \times \mathcal{X}, P_\theta), \ \ \mathbb{A}_\theta[h](y, x) = \langle y - G_\theta(x), \mathbb{I}_\theta[h](x)\rangle_V. \tag{A.3}$$

The information operator $\mathbb{N}_\theta$ defined by (2.6) can be described explicitly as

$$\mathbb{N}_\theta = E[\mathbb{A}_\theta^T(Y, X)\mathbb{A}_\theta(Y, X)] \quad \text{where} \ \ (Y, X) \sim P_\theta.$$

On the other hand, according to [NP21, Proposition 1], we can see that the information operator can be expressed by

$$\mathbb{N}_\theta = \mathbb{I}_\theta^* \mathbb{I}_\theta. \tag{A.4}$$

*Proof of* Proposition A.1. Denote

$$\begin{aligned} f_h(y, x) &:= \log(dP_{\theta+h}^{1/2}/dP_\theta^{1/2}(y, x)) \\ &= \frac{1}{2}\langle y, G_{\theta+h}(x) - G_\theta(x)\rangle_V - \frac{1}{4}(\|G_{\theta+h}(x)\|_V^2 - \|G_\theta(x)\|_V^2). \end{aligned}$$

This proposition is an immediate consequence of the following limit

$$\int_{V\times\mathcal{X}} \left[\frac{e^{f_h} - 1 - \frac{1}{2}\mathbb{A}_\theta[h]}{\|h\|}\right]^2 dP_\theta \to 0 \ \text{ as } \ \|h\| \to 0. \tag{A.5}$$

For any $(y, x) \in V \times \mathcal{X}$, by (A.1) and the chain rule, the integrand in the above integral vanishes as $\|h\| \to 0$. Note that

$$\frac{|e^{f_h(y,x)} - 1|}{\|h\|} \le e^{|f_h(y,x)|}\frac{|f_h(y, x)|}{\|h\|}.$$

It is readily seen that

$$|f_h(y, x)| \le \frac{1}{2}\|y\|_V\|G_{\theta+h}(x) - G_\theta(x)\|_V + \frac{\mathcal{G}^2}{2}\|G_{\theta+h}(x) - G_\theta(x)\|_V,$$

and thus

$$\frac{|f_h(y, x)|}{\|h\|} \le \frac{1}{2}\|y\|_V\frac{\|G_{\theta+h}(x) - G_\theta(x)\|_V}{\|h\|} + \frac{\mathcal{G}^2}{2}\frac{\|G_{\theta+h}(x) - G_\theta(x)\|_V}{\|h\|}.$$

By (A.1), (A.2), (A.3), and Fernique's theorem, for any $\|h\|$ small,

$$\left[\frac{e^{f_h}-1-\frac{1}{2}\mathbb{A}_\theta[h]}{\|h\|}\right] \in L^2(P_\theta),$$

which implies (A.5) by the dominated convergence theorem. □

## References


[ALS13] Sergios Agapiou, Stig Larsson, and Andrew M Stuart. Posterior contraction rates for the bayesian approach to linear ill-posed inverse problems. *Stochastic Processes and their Applications*, 123(10):3828–3860, 2013.

[AN20] Kweku Abraham and Richard Nickl. On statistical calderón problems. *Mathematical Statistics and Learning*, 2(2):165–216, 2020.

[AS22] Giovanni S Alberti and Matteo Santacesaria. Infinite-dimensional inverse problems with finite measurements. *Archive for Rational Mechanics and Analysis*, 243:1–31, 2022.

[ASZ14] Sergios Agapiou, Andrew M Stuart, and Yuan-Xiang Zhang. Bayesian posterior contraction rates for linear severely ill-posed inverse problems. *Journal of Inverse and Ill-posed Problems*, 22(3):297–321, 2014.

[Bes94] JE Besag. Comments on "representation of knowledge in complex systems "by u. *Grenander and MI Miller: Journal of the Royal Statistical Society B*, 56:591592, 1994.

[BL05] Gang Bao and Peijun Li. Inverse medium scattering for the helmholtz equation at fixed frequency. *Inverse Problems*, 21(5):1621, 2005.

[Boh22] Jan Bohr. A bernstein–von-mises theorem for the calderón problem with piecewise constant conductivities. *Inverse Problems*, 39(1):015002, 2022.

[Bou13] Laurent Bourgeois. A remark on lipschitz stability for inverse problems. *Comptes Rendus Mathematique*, 351(5-6):187–190, 2013.

[Buk08] AL Bukhgeim. Recovering a potential from cauchy data in the two-dimensional case. *Journal of Inverse and Ill-Posed Problems*, 16(1):19–33, 2008.

[CC14] Fioralba Cakoni and David L Colton. *A qualitative approach to inverse scattering theory*, volume 767. Springer, 2014.

[CCH22] Fioralba Cakoni, David Colton, and Houssem Haddar. *Inverse scattering theory and transmission eigenvalues*. SIAM, 2022.

[Che18] Xudong Chen. *Computational methods for electromagnetic inverse scattering*, volume 244. Wiley Online Library, 2018.

[CK19] David Colton and Rainer Kress. *Inverse Acoustic and Electromagnetic Scattering Theory*, volume 93. Springer Nature, 2019.

[FKW23] Takashi Furuya, Pu-Zhao Kow, and Jenn-Nan Wang. Consistency of the bayes method for the inverse scattering problem. *Preprint*, 2023.

[FP21] Takashi Furuya and Roland Potthast. Inverse medium scattering problems with kalman filter techniques ii. nonlinear case. *arXiv preprint arXiv:2110.09518*, 2021.

[GG84] Stuart Geman and Donald Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6):721–741, 1984.

[GGVDV00] Subhashis Ghosal, Jayanta K Ghosh, and Aad W Van Der Vaart. Convergence rates of posterior distributions. *Annals of Statistics*, pages 500–531, 2000.

[GN20] Matteo Giordano and Richard Nickl. Consistency of bayesian inference with gaussian process priors in an elliptic inverse problem. *Inverse problems*, 36(8):085001, 2020.

[GS90] Alan E Gelfand and Adrian FM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.

[GVdV17] Subhashis Ghosal and Aad Van der Vaart. *Fundamentals of nonparametric Bayesian inference*, volume 44. Cambridge University Press, 2017.

[Has70] WK Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, pages 97–109, 1970.

[HH01] Peter Hähner and Thorsten Hohage. New stability estimates for the inverse acoustic inhomogeneous medium problem and applications. *SIAM journal on mathematical analysis*, 33(3):670–685, 2001.

[Kek22] Hanne Kekkonen. Consistency of bayesian inference with gaussian process priors for a parabolic inverse problem. *Inverse Problems*, 38(3):035002, 2022.

[KG07] Andreas Kirsch and Natalia Grinberg. *The factorization method for inverse problems*, volume 36. OUP Oxford, 2007.

[Kre13] R. Kress. *Linear Integral Equations*, volume 82. Springer Science & Business Media, 2013.

[KVDVvZ11] Bartek T Knapik, Aad W Van Der Vaart, and J Harry van Zanten. Bayesian inverse problems with gaussian priors. *The Annals of Statistics*, 39:2626–2657, 2011.

[MNP21a] François Monard, Richard Nickl, and Gabriel P Paternain. Consistent inversion of noisy non-abelian x-ray transforms. *Communications on Pure and Applied Mathematics*, 74(5):1045–1099, 2021.

[MNP21b] François Monard, Richard Nickl, and Gabriel P Paternain. Statistical guarantees for bayesian uncertainty quantification in nonlinear inverse problems with gaussian process priors. *The Annals of Statistics*, 49(6):3255–3298, 2021.

[MRR+53] Nicholas Metropolis, Arianna W Rosenbluth, Marshall N Rosenbluth, Augusta H Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 21(6):1087–1092, 1953.

[Nac88] Adrian I Nachman. Reconstructions from boundary measurements. *Annals of Mathematics*, 128(3):531–576, 1988.

[Nic20] Richard Nickl. Bernstein–von mises theorems for statistical inverse problems i: Schrödinger equation. *Journal of the European Mathematical Society*, 22(8):2697–2750, 2020.

[Nic23] Richard Nickl. *Bayesian Non-linear Statistical Inverse Problems*. Zurich Lectures in Advanced Mathematics. EMS Press, 2023.

[NP15] Gen Nakamura and Roland Potthast. *Inverse modeling*. IOP Publishing, 2015.

[NP21] Richard Nickl and Gabriel Paternain. On some information-theoretic aspects of non-linear statistical inverse problems. *arXiv preprint arXiv:2107.09488*, 2021.

[Ray13] Kolyan Ray. Bayesian inverse problems with non-conjugate priors. *Electronic Journal of Statistics*, 7:2516–2549, 2013.

[RR98] Gareth O Roberts and Jeffrey S Rosenthal. Optimal scaling of discrete approximations to langevin diffusions. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(1):255–268, 1998.

[RT96] Gareth O Roberts and Richard L Tweedie. Exponential convergence of langevin distributions and their discrete approximations. *Bernoulli*, pages 341–363, 1996.

[Ser17] Valery Serov. *Fourier series, Fourier transform and their applications to mathematical physics*, volume 197. Springer, 2017.

[SU87] John Sylvester and Gunther Uhlmann. A global uniqueness theorem for an inverse boundary value problem. *Annals of mathematics*, pages 153–169, 1987.

[SV01] J. Saranen and G. Vainikko. *Periodic integral and pseudodifferential equations with numerical approximation*. Springer Science & Business Media, 2001.

[Tie98] Luke Tierney. A note on metropolis-hastings kernels for general state spaces. *Annals of applied probability*, pages 1–9, 1998.

[Vai00] G. Vainikko. Fast solvers of the lippmann-schwinger equation. In *Direct and inverse problems of mathematical physics*, pages 423–440. Springer, 2000.

[VdV00] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

[Vol13] Sebastian J Vollmer. Posterior consistency for bayesian inverse problems through stability and regression results. *Inverse Problems*, 29(12):125011, 2013.

Education and Research Center for Mathematical and Data Science, Shimane University, Matsue, Japan
*Email address*: takashi.furuya0101@gmail.com

Institute of Applied Mathematical Sciences, National Taiwan University, Taipei 106, Taiwan.
*Email address*: jnwang@math.ntu.edu.tw