SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. 2, pp. 203–222

An Analysis of Infinite Dimensional Bayesian Inverse Shape Acoustic Scattering and Its Numerical Approximation*

Tan Bui-Thanh† and Omar Ghattas‡

Abstract. We present and analyze an infinite dimensional Bayesian inference formulation, and its numerical approximation, for the inverse problem of inferring the shape of an obstacle from scattered acoustic waves. Given a Gaussian prior measure on the shape space, whose covariance operator is the inverse of the Laplacian, the Bayesian solution of the inverse problem is the posterior measure given by the Radon–Nikodym derivative with respect to the Gaussian prior measure. The well-posedness of the Bayesian formulation in infinite dimensions is proved, including the justification of the Radon–Nikodym derivative and the continuous dependence of the posterior measure on the observation data via the Hellinger distance. The proof is made possible by a suitable shape parametrization in a Banach space setting and the regularity of the forward solution with respect to the smoothness of the shape. This also facilitates proving the Lipschitz continuity of the observation operator with respect to the scatterer shape via the shape derivative of the forward solution. Next, a finite dimensional approximation to the Bayesian posterior is proposed and the corresponding approximation error is quantified. The approximation strategy involves a Nyström scheme for approximating a boundary integral formulation of the forward Helmholtz problem and a Karhunen–Loève truncation for approximating the prior measure. Weak convergence of the resulting finite dimensional approximation, as well as convergence in the Hellinger distance, are investigated. In particular, we determine the convergence rate as a function of the number of Nyström quadrature points and the number of truncated terms in the Karhunen–Loève series. Finally, we estimate the error between the exact posterior moments, e.g., posterior mean and variance, and their finite dimensional approximate counterparts in terms of the errors due to forward equation approximation and prior approximation. The main result of this work is that the convergence rate for approximating the Bayesian inverse problem is spectral, and this directly inherits the spectral convergence rates of the approximations of both the prior and the forward problems.

Key words. shape derivative, Bayesian inversion, Radon–Nikodym derivative, Gaussian prior, posterior, convergence, error estimation, inverse shape scattering, well-posedness, acoustic wave propagation

AMS subject classifications. 65N12, 65N15, 49N45, 49K20, 49K40, 49J50, 49N60, 62G99

DOI. 10.1137/120894877

1. Introduction.

In this paper we adopt the Bayesian framework in [25] to the inverse shape scattering problem and extend it to include analysis of the contribution of the errors coming from both the likelihood discretization and the prior discretization. In particular, we

*Received by the editors October 11, 2012; accepted for publication (in revised form) January 21, 2014; published electronically May 15, 2014. This research was partially supported by AFOSR grant FA9550-09-1-0608; DOE grants DE-FC02-11ER26052, DE-FG02-09ER25914, DE-FG02-08ER25860, and DE-FC52-08NA28615; and NSF grants CMS-1028889 and OPP-0941678.

http://www.siam.org/journals/juq/2/89487.html

†Institute for Computational Engineering & Sciences, The University of Texas at Austin, Austin, TX 78712 (tanbui@ices.utexas.edu).

‡Institute for Computational Engineering & Sciences, Jackson School of Geosciences, and Department of Mechanical Engineering, The University of Texas at Austin, Austin, TX 78712 (omar@ices.utexas.edu).

are interested in inferring the shape of scatterers from reflected acoustic waves. For simplicity of the exposition, we shall exclusively work with the two dimensional setting. We further assume that the scatterer Ω_S (for convenience, one scatterer is considered, but all of the results in this paper hold for multiple scatterers) under consideration is *sound-soft*, namely, the total wave field must vanish on the boundary (see (1.1b)). If the incident wave is a plane wave, the acoustic scattering problem can be cast as the following exterior Helmholtz problem [10]:

$$\nabla^2 U + k^2 U = 0 \qquad \text{in } \Omega, \tag{1.1a}$$

$$U = -U^I \qquad \text{on } \Gamma_S, \tag{1.1b}$$

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial U}{\partial r} - ikU\right) = 0, \tag{1.1c}$$

where k is the wave number, $i^2 = -1$, U is the scattered field, U^I is the incident field which is assumed to be a solution of the Helmholtz equation (1.1a) in all of $\mathbb{R}^2$, Ω is the exterior domain given by $\Omega := \mathbb{R}^2 \setminus \overline{\Omega}_S$, Γ_S is the boundary of the scatterer Ω_S, and (1.1c) is the radiation condition, which is assumed to be valid uniformly in all directions $\frac{\mathbf{x}}{\|\mathbf{x}\|}$.

In the forward shape scattering problem, the task is to solve for the acoustic field U given a description of the scatterer shape Ω_S. In the inverse problem, the task is to reconstruct the scatterer shape given some available observations, e.g., scattered field data observed at some parts of the domain Ω. One can solve such an inverse problem using various methods, including the well-known least squares approach based on the farfield equation [10] and the linear sampling method [6], to name just a few (see [8, 24] for survey papers). Alternatively, one can cast the inverse problem in the framework of PDE-constrained optimization. To begin, let us consider the following additive noise-corrupted pointwise observation model:

$$z_j := U(\mathbf{x}_j) + \eta_j, \quad j = 1, \ldots, K, \tag{1.2}$$

where $\{\mathbf{x}_j\}_{j=1}^K$, with $\{\mathbf{x}_j\}_{j=1}^K \cap \Omega_S = \emptyset$, is the set of points at which the scattered field U is observed, η_j the additive noise, and z_j the actual noisy observations. Concatenating all the observations, one can rewrite (1.2) as

$$\mathbf{z} := \mathcal{G}(\Omega_S) + \boldsymbol{\eta}, \tag{1.3}$$

with $\mathcal{G} := [U(\mathbf{x}_1), \ldots, U(\mathbf{x}_K)]^T$ denoting the mapping from the shape Ω_S to the noise-free observables, $\boldsymbol{\eta}$ being normally distributed as $\mathcal{N}(0, \mathbf{L})$ with bounded covariance matrix $\mathbf{L}$, and $\mathbf{z} = [z_1, \ldots, z_K]^T$.

The inverse shape problem can now be formulated as

$$\min_{\Omega_S} \mathcal{J} = \frac{1}{2}\left|\mathbf{z} - \mathcal{G}(\Omega_S)\right|^2_{\mathbf{L}} \tag{1.4}$$

subject to the forward equation (1.1), where $|\cdot|^2_{\mathbf{L}} := \left(\cdot, \mathbf{L}^{-1}\cdot\right)$ denotes the weighted Euclidean norm. This optimization problem is, however, ill-posed. An intuitive reason is that observations are typically sparse, and hence they provide limited information about the shape of the scatterer. As a result, the Jacobian of the shape-to-observation map $\mathcal{G}$ contains limited spectral information. Indeed, we have shown that the Gauss–Newton approximation of the Hessian (which is the square of this Jacobian and equal to the full Hessian of the data misfit $\mathcal{J}$ evaluated at the optimal shape in the no-noise case) is a compact operator [5], and hence its range space is effectively finite dimensional.

One way to overcome the ill-posedness is to use *Tikhonov regularization*, which proposes solving the nearby problem

$$\min_{\Omega_S} \frac{1}{2} \left| \mathbf{z} - \mathcal{G}\left(\Omega_S\right) \right|_{\mathbf{L}}^2 + \frac{\kappa}{2} \left\| R^{1/2} \Omega_S \right\|^2, \tag{1.5}$$

where κ is a regularization parameter, R some regularization operator, and $\|\cdot\|$ some appropriate norm.

The above methods are representatives of deterministic inverse solution techniques that typically do not take into account the randomness due to measurements and other sources, though one can equip the deterministic solution with a confidence region by postprocessing (see, e.g., [26] and references therein). The question that needs to be addressed is how to incorporate randomness, including that coming from the measurement model (1.3), into the solution of the inverse problem. In this paper, we choose to tackle this question using a *Bayesian* framework proposed in [25]. Instead of seeking a shape that minimizes the Tikhonov-regularized functional (1.5), we seek a statistical description of all possible shapes that conform to some prior knowledge and at the same time are consistent with the observations. These shape are distributed according to the so-called posterior measure. The Bayesian approach does this by reformulating the inverse problem as a problem in *statistical inference*, incorporating uncertainties in the observations, the shape-to-observable map, and prior information on the shapes. This approach is appealing since it can incorporate most, if not all, kinds of randomness in a systematic manner. Our work is the first to adopt the Bayesian framework in [25] to the inverse shape scattering problem by providing a complete and rigorous analysis from the well-posedness of the Bayesian posterior measure in infinite dimensions to a priori convergence estimates of a finite dimensional approximation scheme.

In the following, we summarize the contributions of this article. In order to make inferences on the scatterer shape Ω_S from observation data $\mathbf{z}$, we propose using a suitable representation of Ω_S in a Banach space setting, which is the subject of section 2. The main result is the shape gradient of a general domain integral. Section 3 starts by recalling a regularity result of the scattered field with respect to the shape in [5]. It then studies the shape gradient of the scattered field, from which we can derive the Lipschitz continuity of the observation operator, which in turn is sufficient to prove the well-posedness of the infinite Bayesian formulation in section 4. In particular, in section 4 we first propose a Gaussian prior measure on the scatterer shape whose covariance operator is the inverse of the Laplacian. We then show that the infinite dimensional Bayesian posterior measure is well-defined and has a probability density with respect to the prior. Furthermore, it is Lipschitz continuous in the observation data with respect to the Hellinger distance. Section 5 proposes a finite dimensional approximation to the posterior measure. This involves a Nyström scheme for approximating the forward equation in section 5.1 and a Karhunen–Loève truncation strategy for approximating the prior measure in section 5.2. Weak convergence of the resulting finite dimensional approximation, as well as convergence in the Hellinger distance, are investigated in section 6. In particular, one of our main results is to provide the convergence rate as a function of the number of Nyström quadrature points and the number of truncated terms in the Karhunen–Loève series. Section 7 estimates the error between the exact posterior moments (e.g., posterior mean and variance) and their finite dimensional approximate counterparts, again in terms of the number of Nyström quadrature points and the number of truncated terms in the Karhunen–Loève series. Section 8 generalizes the results of the previous sections to the case where additional smoothness on the prior shape space is desirable. The main goal of section 8 is to show that the convergence rate is in fact spectral.

Our analysis shows how two sources of errors, one coming from approximating the prior and the other from approximating the forward model, translate into the error in approximating the posterior distribution. This distinguishes our work from the existing ones in the literature. One can find related work in [25, Example 4.7] in which inverting for the initial condition of the heat equation is considered; in this case, the two sources of error overlap, and hence the analysis needs to effectively take care only of the former. In [11], the convergence of the approximate posterior for Lagrangian data assimilation is analyzed for the case in which the forward model is exactly evaluated, and consequently only error due to approximating the prior is taken into account. Also in [11], a Galerkin approximation to the forward model is employed for Eulerian data assimilation, but Theorem A.3 is not applicable in this case since the condition on the accuracy for $G_{n,m}$ is not available; hence, the convergence is guaranteed but the convergence rate is unspecified. Another approach to the approximation of infinite dimensional Bayesian inverse problems is to first parametrize the prior and then pose the problem of finding the conditional expectation under the posterior measure (e.g., computing conditional moments) as one of evaluating deterministic infinite dimensional parametric integrals [23]. This can be done efficiently using a sparse deterministic approximation of the Radon–Nikodym derivative using generalized polynomial chaos (gPC). For finite dimensional Bayesian inverse problems, gPC has been used in approximating the forward map, which yields exponential convergence in approximating the posterior density in the Kullback–Leibler divergence if the error in approximating the forward map is exponentially small [18].

2. Shape derivatives in a Banach space setting.

In this section we adapt the shape parametrization and the shape derivative in [7, 13] to account for our exponential parametrization, which facilitates the analysis in the rest of the paper. Since this is not the main focus of the paper, we briefly present some results which are useful for the later analysis and refer the readers to [5, 7, 13] (and references therein) for the detailed development.

If we restrict our attention to a special shape space which is Banach, shape calculus becomes usual differential calculus on Banach spaces and many interesting conclusions can be drawn. As we shall see in what follows, only Fubini's theorem and Leibniz's rule are sufficient to derive the shape derivatives in this setting.

Following [7, 13], we represent the shape by its boundary. In particular we assume that the scatterer $\Omega_S \subsetneq \mathbb{R}^2$ is a simply connected domain and starlike with respect to the origin. Thus, its boundary Γ_S can be parametrized as

$$\Gamma_S := \{\mathbf{r} = r(\theta)\mathbf{e}_r : \theta \in [0, 2\pi]\}, \quad \mathbf{e}_r = [\cos\theta, \sin\theta]^T.$$

Since the radius $r(\theta)$ must be positive, the actual representation used in this paper is the following:

$$\Gamma_S := \{\mathbf{r} = \exp(p(\theta))\,\mathbf{e}_r : \theta \in [0, 2\pi]\}, \quad \mathbf{e}_r = [\cos\theta, \sin\theta]^T. \tag{2.1}$$

As a result, $p(\theta)$ can live in $L^2[0, 2\pi]$, where we have included the end points in the definition of the space of square integrable functions to signify the periodicity. We further assume that Γ_S belongs to Hölder continuous class $C^{m,\alpha}$ for $0 < \alpha \le 1$. That is, the "shape" parameter p lives in the Hölder continuous space $C^{m,\alpha}([0, 2\pi])$ with periodic condition

$$p^{(j)}(0) = p^{(j)}(2\pi), \quad j = 0, \dots, m, \tag{2.2}$$

where the superscript (j) denotes the jth derivative with respect to θ. Until the end of the paper, unless otherwise stated, functions defined on $[0, 2\pi]$ are periodic in the sense of (2.2).

The following lemma provides the first order shape derivative of a domain integral of the type

$$\mathcal{I} := \int_\Omega f(\mathbf{x})\, d\mathbf{x}.$$

Lemma 2.1. *Assume that $f \in C(\mathbb{R}^2)$; then $\mathcal{I}$ is continuously Fréchet differentiable at any $p \in C^1([0, 2\pi])$, and its shape derivative in direction $\hat{p}$ is given as*

$$D\mathcal{I}(p; \hat{p}) = -\int_0^{2\pi} f(\mathbf{r}) \exp(2p)\, \hat{p}\, d\theta = -\int_{\Gamma_s} f(\mathbf{x}) \frac{\exp(p)\, \hat{p}}{\sqrt{1 + \left(p^{(1)}\right)^2}}\, ds \tag{2.3}$$

for all $\hat{p} \in C^1([0, 2\pi])$.

Proof. Following the approach by [13] for interior problems, and later by [5] for exterior problems, we use Fubini's theorem to have

$$\mathcal{I} = \int_0^{2\pi} \int_{\exp(p(\theta))}^{\infty} f(r\cos\theta, r\sin\theta)\, r\, dr d\theta$$

and then the Leibniz rule to compute the first Gâteaux variation to obtain the shape derivative formula (2.3). Then it is obvious that $D\mathcal{I}(p; \hat{p})$ is linear and continuous with respect to $\hat{p}$. Now the continuity of $D\mathcal{I}(p; \hat{p})$ with respect to p is straightforward. Hence, a classical result on sufficiency for the Fréchet derivative [3] ends the proof. ■

3. Properties of the observation operator. In this section, we first study the shape gradient of the scattered field U that satisfies the forward equation (1.1) and then use it to derive the boundedness and Lipschitz continuity of the observation operator. To begin, let us introduce the standard single and double surface potentials [9, 10]

$$S\varphi(\mathbf{x}) := 2\int_{\Gamma_s} \Phi(\mathbf{x}, \mathbf{y})\varphi(\mathbf{y})\, ds(\mathbf{y}), \qquad \mathbf{x} \in \Gamma_S, \tag{3.1a}$$

$$D\varphi(\mathbf{x}) := 2\int_{\Gamma_s} \frac{\partial\Phi(\mathbf{x}, \mathbf{y})}{\partial\mathbf{n}(\mathbf{y})}\varphi(\mathbf{y})\, ds(\mathbf{y}), \qquad \mathbf{x} \in \Gamma_S, \tag{3.1b}$$

where Φ is the zero order Hankel function of the first kind, namely,

$$\Phi(\mathbf{x}, \mathbf{y}) := \frac{i}{4} H_0^1 (\mathbf{x} - \mathbf{y}) .$$

A standard approach for solving the forward equation using integral equation methods [10] is to look for a solution as a combination of single and double potentials

$$U(\mathbf{x}) := \tilde{D}\varphi(\mathbf{x}) - i\tilde{S}\varphi(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^2 \setminus \Gamma_S, \tag{3.2}$$

where

$$\tilde{S}\varphi(\mathbf{x}) := \int_{\Gamma_s} \Phi(\mathbf{x}, \mathbf{y})\varphi(\mathbf{y})\, ds(\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^2 \setminus \Gamma_S, \tag{3.3}$$

$$\tilde{D}\varphi(\mathbf{x}) := \int_{\Gamma_s} \frac{\partial \Phi(\mathbf{x}, \mathbf{y})}{\partial \mathbf{n}(\mathbf{y})}\varphi(\mathbf{y})\, ds(\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^2 \setminus \Gamma_S. \tag{3.4}$$

The ansatz (3.2) automatically satisfies the Helmholtz equation and the radiation condition, and hence v is analytic in $\mathbb{R}^2 \setminus \Gamma_S$. What remains is to determine φ, which satisfies the boundary condition on the scatterer surface. Thus, the boundary condition determines the space for φ, which in turn suggests the correct space for the solution v. Here is a regularity result for the forward solution, whose proof can be found in [5, Theorem 1].

Lemma 3.1. *Assume $\alpha \in (0, 1]$ and $\Gamma_S \in C^{2,\alpha}$. Then, the forward equation (1.1) is well-posed, in particular, $U \in C^{2,\alpha}\left(\mathbb{R}^2 \setminus \Omega_S\right)$, and the trace $U|_{\Gamma_S} \in C^{2,\alpha}(\Gamma_S)$. Both U and $U|_{\Gamma_S}$ depend continuously on $U^I\big|_{\Gamma_S}$ in the $C^{2,\alpha}$-norm. In particular, U lives in $C\left(\mathbb{R}^2 \setminus \Omega_S\right)$ and it depends continuously on $U^I\big|_{\Gamma_S}$ in the ∞-norm (uniform norm).*

The proof of Lemma 3.1 relies on the Riesz–Fredholm theory [9], and we repeat part of it here to find the correct space to which the density φ belongs. Using the standard limiting values on the boundary Γ_S for single and double potentials [9, 10], the trace of the solution on Γ_S can be written as

$$2U(\mathbf{x}) = (\varphi + D\varphi - iS\varphi)(\mathbf{x}) = -2U^I|_{\Gamma_S}(\mathbf{x}), \quad \mathbf{x} \in \Gamma_S. \tag{3.5}$$

Now $U^I|_{\Gamma_S} \in C^{2,\alpha}$ because the restriction of an analytic function on a curve is as smooth as the curve is. By the compactness of S and D in $C^{2,\alpha}$ and the injectivity of $I + D - iS$ [10], the Riesz–Fredholm theory [9] tells us that (3.5) is well-posed in the sense of Hadamard [15], namely, $I + D - iS : C^{2,\alpha}(\Gamma_S) \to C^{2,\alpha}(\Gamma_S)$ is bijective and its inverse $(I + D - iS)^{-1}$: $C^{2,\alpha}(\Gamma_S) \to C^{2,\alpha}(\Gamma_S)$ is bounded. As a result, $\varphi \in C^{2,\alpha}(\Gamma_S)$.

Remark 3.2. Note that it is sufficient to require $\Gamma_S \in C^2$ so that the single and double potentials S and D, respectively, are compact operators, and hence the Riesz–Fredholm theory can be applied. In this paper, it is more convenient to work with the smoother case, namely, $\Gamma_S \in C^{2,\alpha}$, as we shall show. On the other hand, it is possible to work with a boundary that is less regular than C^2 and still stay within the Riesz–Fredholm framework [19], but we avoid unnecessary technicalities here.

Remark 3.3. Unlike classical regularity results in which one needs to determine the smoothness of the solution from the smoothness of the forcing function U^I, the regularity result in Lemma 3.1 is with respect to the shape Γ_S.

For simplicity in writing, we find $C^{2,\alpha}([0,2\pi]) \equiv C^{2,\alpha}(\Gamma_S)$ and define $X := C^{2,\alpha}([0,2\pi])$. Throughout the paper, unless otherwise stated, the constant c at different places may have different value and depend on different variables. We are now in the position to bound the pointwise observation operator.

Proposition 3.4. *For every $\varepsilon > 0$, there exists $M = M(\varepsilon)$ such that*

$$|\mathcal{G}(\varphi, p)|_{\mathbf{L}} \leq c \exp(\|p\|_\infty) \quad \text{and} \quad |\mathcal{G}(\varphi, p)|_{\mathbf{L}} \leq \exp\left(\varepsilon \|p\|_X^2 + M\right)$$

hold for all $p \in X$.

Proof. It is sufficient to prove that for any $j \in \{1, \dots, K\}$

$$|U(\mathbf{x}_j)| \leq c \exp(\|p\|_\infty),$$

with some constant c independent of j and p. To this end, we shall show that

$$\left|\tilde{S}\varphi(\mathbf{x}_j)\right| \leq c \exp(\|p\|_\infty). \tag{3.6}$$

A similar estimate for $\left|\tilde{D}\varphi(\mathbf{x}_j)\right|$ follows similarly and hence is omitted. Since $\mathbf{x}_j \notin \Gamma_S$ for all $j = 1, \dots, K$, the kernel $\Phi(\mathbf{x}_j, \mathbf{y})$ is analytic in both arguments, and together with the fact that $\varphi \in C^{2,\alpha}(\Gamma_S)$ and $p \in C^{2,\alpha}([0,2\pi])$ we have

$$\left|\tilde{S}\varphi(\mathbf{x}_j)\right| \leq 2\pi \|\Phi(\mathbf{x}_j, \exp(p)\mathbf{e}_r)\|_\infty \|\varphi\|_{C^{2,\alpha}} \left\|\sqrt{1 + \left(p^{(1)}\right)^2}\right\|_\infty \exp(\|p\|_\infty),$$

which immediately yields (3.6). ■

Our next step is to show that scattered field U is Lipschitz continuous with respect to p. To this end, we first study the Gâteaux variation of U in the direction $\tilde{p}$, namely, $\tilde{U}$. Following [5], we can show that $\tilde{U}$ satisfies the so-called incremental forward equation

$$\nabla^2 \tilde{U} + k^2 \tilde{U} = 0 \quad \text{in } \Omega, \tag{3.7a}$$

$$\tilde{U} = -\frac{\partial\left(U + U^I\right)}{\partial \mathbf{e}_r} \exp(p)\,\tilde{p} \quad \text{on } \Gamma_S, \tag{3.7b}$$

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{\partial \tilde{U}}{\partial r} - ik\tilde{U}\right) = 0. \tag{3.7c}$$

Proposition 3.5. *The scattered field U is Fréchet differentiable in p. Consequently, U is Lipschitz continuous in p, i.e.,*

$$\|U(p_1) - U(p_2)\|_\infty \leq c \exp(\max\{\|p_1\|_\infty, \|p_2\|_\infty\}) \|p_1 - p_2\|_\infty$$

for all $p_1, p_2 \in X$.

Proof. Using the ansatz (3.2) together with the Riesz–Fredholm theory one can show that $\tilde{U}$ is a function in $C\left(\mathbb{R}^2 \setminus \Omega_S\right)$ (see [5, Theorem 1] for the details). Furthermore, $\tilde{U}$ depends linearly in $\tilde{p}$ and continuously on $\frac{\partial(U+U^I)}{\partial \mathbf{e}_r} \exp(p)\,\tilde{p}$, i.e.,

$$\left\|\tilde{U}\right\|_\infty \leq c \left\|\frac{\partial\left(U + U^I\right)}{\partial \mathbf{e}_r}\right\|_\infty \|\exp(p)\|_\infty \|\tilde{p}\|_\infty.$$

Hence, U is Fréchet differentiable in p and the Fréchet derivative is bounded by $c\exp\left(\|p\|_\infty\right)$. A mean value theorem [1] then concludes the proof. ■

Corollary 3.6. *The observation operator $\mathcal{G}$ is Lipschitz continuous in p, i.e.,*

$$\begin{aligned}|\mathcal{G}(\varphi,p_1)-\mathcal{G}(\varphi,p_2)|_{\mathbf{L}} &\le c\exp\left(\max\left\{\|p_1\|_\infty,\|p_2\|_\infty\right\}\right)\|p_1-p_2\|_\infty\\ &\le c\exp\left(\max\left\{\|p_1\|_X,\|p_2\|_X\right\}\right)\|p_1-p_2\|_X.\end{aligned}$$

4. A Bayesian inversion formulation. The preceding sections prepare all the necessary ingredients for an infinite Bayesian formulation of our shape inverse scattering problem. We shall show that the Bayesian formulation is well-posed and, in particular, the posterior measure is Lipschitz continuous with respect to the observation data. Let us introduce $\mathcal{A}$ as a Laplacian operator, namely, $\mathcal{A} := -\frac{d^2}{d\theta^2}$ with the following domain of definition:

$$D(\mathcal{A}) := \left\{v \in H^2[0,2\pi] : \int_0^{2\pi} v(\theta)\,d\theta = 0\right\},$$

where, again, we have used the square bracket in the definition of the domain $[0,2\pi]$ to signify the periodicity. Clearly, $m^2, m\in\mathbb{N}$, are eigenvalues of $\mathcal{A}$, each of which is repeated once with the corresponding eigenfunctions $\psi_{2m-1} := \sin(m\theta)/\sqrt{\pi}$ and $\psi_{2m} := \cos(m\theta)/\sqrt{\pi}$ (we also need $\psi_0 := 1/\sqrt{2\pi}$ in section 5.2). Furthermore, $\mathcal{A}^{-1}$ is symmetric positive definite (a direct consequence of the positiveness of $\mathcal{A}$) and a trace class operator since $\sum_{m=1}^\infty \frac{1}{m^2} < \infty$.

We now postulate a prior Gaussian measure (details on Gaussian measures on Banach and Hilbert spaces can be found in [4, 21], for example) μ on p such that the prior measure of $p^{(2)}(\theta)$ is a Gaussian measure on $L^2[0,2\pi]$ defined by

$$p^{(2)}(\theta) \sim \mathcal{N}\left(0,\mathcal{A}^{-1}\right)$$

for which the Karhunen–Loève expansion, which is in fact Fourier series in this case, reads as

$$p^{(2)} = \sum_{j=1}^\infty \frac{a_j}{j}\frac{\cos(j\theta)}{\sqrt{\pi}} + \frac{b_j}{j}\frac{\sin(j\theta)}{\sqrt{\pi}},$$

where $a_j, b_j \sim \mathcal{N}(0,1)$. We define

$$p(\theta) = \frac{p_0}{\sqrt{2\pi}} - \sum_{j=1}^\infty \frac{a_j}{j^3\sqrt{\pi}} + \frac{b_j}{j^2\sqrt{\pi}}\theta + \int_0^\theta\int_0^t p^{(2)}(s)\,ds dt,$$

where $p_0 \sim \mathcal{N}\left(u,\sigma^2\right)$ is a Gaussian random variable. The fact that integral operators are linear and continuous implies that the prior distribution μ of p is indeed a Gaussian (see, e.g., [16]).

We shall show in Lemma 4.2 that $p^{(2)} \in C^\alpha[0,2\pi]$, and hence we can integrate its Fourier expansion term by term to arrive at

$$p(\theta) = \frac{p_0}{\sqrt{2\pi}} - \sum_{j=1}^\infty \frac{a_j}{j^3}\frac{\cos(j\theta)}{\sqrt{\pi}} + \frac{b_j}{j^3}\frac{\sin(j\theta)}{\sqrt{\pi}},$$

which is the Karhunen–Loève expansion for p.

Remark 4.1. Note that we have followed [25] to define the Gaussian measure on p based on the Gaussian measure of its derivative, namely, $p^{(2)}$. This is clearly not a unique way to accomplish the task. Alternatively, one can postulate

$$p \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{2\pi}\right) \otimes \mathcal{N}\left(0, \mathcal{A}^{-3}\right),$$

but we omit the details here.

Here is a result on the support of μ.

Lemma 4.2. *There holds that* $\mu(X) = 1$ *for* $0 < \alpha < 1/2$.

Proof. First, note that $\mathcal{A}$, its eigenvalues, and its eigenfunctions satisfy Assumption 2.9 in [25]. Then, an application of the Kolmogorov theorem and the Gaussianity shows that $p^{(2)}(\theta)$ is almost surely in $C^{0,\alpha}$ for any $\alpha < 1/2$ [25, Lemma 6.25]. It immediately follows that p is $C^{2,\alpha}$-Hölder continuous almost surely, and hence $\mu(X) = 1$. ■

We are now in the position to consider the Bayesian inverse problem with observations given by (1.3), p as the inversion parameters, μ as the prior distribution of p, and U as the solution of the forward problem (1.1). Let us denote $\nu(p|\mathbf{z})$ as the posterior measure of p, given $\mathbf{z}$. We also need the following Hellinger distance between two probability measures μ_1 and μ_2, provided that they are both absolutely continuous with respect to the same reference measure μ_0:

$$d_{\text{Hellinger}}(\mu_1, \mu_2) := \sqrt{\frac{1}{2}\int\left(\sqrt{\frac{d\mu_1}{d\mu_0}} - \sqrt{\frac{d\mu_2}{d\mu_0}}\right)^2 d\mu_0}.$$

Theorem 4.3. *Let* $p \sim \mu$ *as in Lemma* 4.2*; then the following hold:*

(i) *The posterior measure* $\nu(p|\mathbf{z})$ *is absolutely continuous with respect to* μ*, i.e.,* $\nu \ll \mu$*, and the Radon–Nikodym derivative is given by*

$$\frac{d\nu}{d\mu}(p|\mathbf{z}) = \frac{1}{Z}\exp\left(-\frac{1}{2}\left|\mathbf{z} - \mathcal{G}(\varphi, p)\right|^2_{\mathbf{L}}\right), \tag{4.1}$$

where $Z = \int_X \exp\left(-\frac{1}{2}\left|\mathbf{z} - \mathcal{G}(\varphi, p)\right|^2_{\mathbf{L}}\right)\mu(dp)$.

(ii) *The posterior measure* $\nu(p|\mathbf{z})$ *is a well-defined probability measure on* $L^2[0, 2\pi]$.

(iii) *The posterior measure* $\nu(p|\mathbf{z})$ *is Lipschitz continuous with respect to the data* $\mathbf{z}$ *in the Hellinger distance, i.e.,*

$$d_{\text{Hellinger}}(\nu(p|\mathbf{z}_1), \nu(p|\mathbf{z}_2)) \le c(\gamma)\left|\mathbf{z}_1 - \mathbf{z}_2\right|,$$

with $\max\{|\mathbf{z}_1|, |\mathbf{z}_2|\} < \gamma$.

Proof. First, $\mu(X) = 1$ due to Lemma 4.2. Second, Proposition 3.4 and Corollary 3.6 show that Assumption A.1 holds. Finally, one readily applies Theorem A.2 to end the proof. ■

5. A finite dimensional approximation of the Bayesian posterior measure. We have formulated an infinite dimensional Bayesian formulation in section 4 and shown its well-posedness. In practice, one needs to approximate the Bayesian solution in some finite dimensional setting to make the inference feasible on computers. The goal of this section is to first devise such a finite dimensional approximation and then carry out its error analysis.

5.1. Finite dimensional approximation of the forward equation. We begin by approximating the observation operator $\mathcal{G}$, and this requires the solution of the integral equation (3.5). In this paper, boundary integrals for 2π-periodic functions are discretized using the trapezoidal rule (owing to its simplicity and high accuracy):

$$\int_0^{2\pi} f\, d\theta \approx \mathcal{I}_n(f) = \frac{\pi}{n}\sum_{j=0}^{2n-1} f(\theta_j),$$

where we have used $2n$ equidistant quadrature points. For k-times continuously differentiable 2π-periodic functions, the numerical integration error scales like

$$\left|\int_0^{2\pi} f\, d\theta - \frac{\pi}{n}\sum_{j=0}^{2n-1} f(\theta_j)\right| = \left|\int_0^{2\pi} f\, d\theta - \mathcal{I}_n(f)\right| \le c\frac{\left\|f^{(k)}\right\|_\infty}{(2n)^k} \tag{5.1}$$

for some constant c depending on k [17]. In particular, if f is analytic, the error decays exponentially in n.

Since the detailed presentation of the numerical method is already given in [10], let us now summarize the important points here. The boundary integral equation (3.5) can be written as

$$\varphi(t) - \underbrace{\int_0^{2\pi}\left[K_1(t,\tau)\ln\left(4\sin^2\frac{t-\tau}{2}\right) + K_2(t,\tau)\right]\varphi(\tau)\, d\tau}_{A\varphi} = -2\, U^I\big|_{\Gamma_S}(t) \tag{5.2}$$

for $0 \le t \le 2\pi$, where K_1 and K_2 are analytic in their arguments.

As suggested by [10], we use the Nyström method to discretize the boundary integral equation (5.2) to obtain

$$\varphi_n(t) - \underbrace{\sum_{j=0}^{2n-1}\left[R_n(j,t)\, K_1(t,t_j) + \frac{\pi}{n}K_2(t,t_j)\right]\varphi_n(t_j)}_{A_n\varphi_n} = -2\, U^I\big|_{\Gamma_S}(t),$$

where

$$R_n(j,t) = \frac{2\pi}{n}\sum_{\ell=1}^{n-1}\frac{1}{\ell}\cos(\ell(t-t_j)) - \frac{\pi}{n^2}\cos(n(t-t_j)), \quad j = 0,\dots,2n-1.$$

Lemma 5.1. $(I - A_n)^{-1}$ *is uniformly bounded.*

Proof. Following an argument similar to that in [17, section 12.3], one can show that the sequence $\{A_n\}_{n=1}^\infty$ is collectively compact and converges pointwise to A. By the Banach uniform boundedness principle, it follows that $\{A_n\}_{n=1}^\infty$ is uniformly bounded with respect to n. Furthermore,[1] there exists $n_0 \in \mathbb{N}$ such that for all $n \ge n_0$ we have

$$\left\|(I-A)^{-1}(A_n - A)A_n\right\| < 1/2,$$

[1] We refer the readers to [17] for the detailed proofs.

and the inverse operator $(I - A_n)^{-1}$ is bounded by

$$\left\|(I - A_n)^{-1}\right\| \leq \frac{1 + \left\|(I - A)^{-1} A_n\right\|}{1 - \left\|(I - A)^{-1} (A_n - A) A_n\right\|},$$

where the norms are understood as the operator norm induced by the uniform norm $\|\cdot\|_\infty$. As a direct consequence, we conclude that $(I - A_n)^{-1}$ is uniformly bounded. ■

The following result will be used in Proposition 5.6.

Lemma 5.2. *The interpolated density $\varphi_n(t) = A_n\varphi_n - 2\, U^I|_{\Gamma_S}(t)$ is uniformly bounded by the incident field, i.e.,*

$$\|\varphi_n\|_\infty \leq c \left\| U^I|_{\Gamma_S} \right\|_\infty,$$

with c independent of n.

Proof. Since $(I - A_n)^{-1}$ is uniformly bounded from above, we have

$$\|\varphi_n\|_\infty = 2\left\|(I - A_n)^{-1}\, U_n^I|_{\Gamma_S}\right\|_\infty \leq c\left\|U_n^I|_{\Gamma_S}\right\|_\infty \leq c\left\|U^I|_{\Gamma_S}\right\|_\infty,$$

where we have used the property of trigonometric interpolation [17, Lemma 11.5] in the last inequality. ■

Let us denote $\mathcal{G}_n(\varphi_n, p)$ ($U_n(\varphi_n, p, \mathbf{x})$, respectively) as the numerical approximation of the observation operator $\mathcal{G}(\varphi, p)$ ($U(\varphi, p, \mathbf{x})$, respectively). Here, $\varphi_n(\theta)$ is the density obtained by trigonometric interpolation of the solution of the boundary integral equation (3.5) via the Nyström method. Moreover, the subscript n of $\mathcal{G}_n$ signifies that fact that we also approximate the observation operator at $\varphi_n(\theta)$ using the trapezoidal rule. It should also be pointed out that $\varphi_n(\theta)$ is in fact a function of p, i.e., $\varphi_n(\theta) \equiv \varphi_n(p(\theta))$, but we often ignore this fact if it is not important in our error estimations. We have the following error bound for the observation operator due to approximating φ (and hence approximating the forward equation).

Lemma 5.3. *Suppose $p \sim \mu$ as in Lemma 4.2; then the following error estimations hold:*

$$|\mathcal{G}(\varphi, p) - \mathcal{G}_n(\varphi_n, p)|_{\mathbf{L}} \leq \frac{c}{2n}\exp\left(\|p\|_\infty\right)$$

and

$$|\mathcal{G}(\varphi, p) - \mathcal{G}_n(\varphi_n, p)|_{\mathbf{L}} \leq \frac{c}{2n}\exp\left(\varepsilon\,\|p\|_X^2\right)$$

for any $\varepsilon > 0$ and $c = c(\varepsilon)$.

Proof. Lemma 4.2 shows that p is $C^{2,\alpha}$-Hölder continuous almost surely. Using the Nyström method with trapezoidal quadrature rule (5.1), we have the following error estimation [10, 17] for the density:

$$\|\varphi_n(\theta) - \varphi(\theta)\|_\infty \leq \frac{c}{2n}, \tag{5.3}$$

owing to $p^{(1)} \in C^{1,\alpha}$. Now, it is sufficient to show that, for any $j \in \{1, \ldots, K\}$,

$$|U(\varphi, p, \mathbf{x}_j) - U_n(\varphi_n, p, \mathbf{x}_j)| \leq c\frac{\exp\left(\|p\|_\infty\right)}{2n}.$$

But this is straightforward by an application of the triangle inequality

$$
\begin{aligned}
&\left|U\left(\varphi, p, \mathbf{x}_j\right)-U_n\left(\varphi_n, p, \mathbf{x}_j\right)\right| \\
&\quad \leq \int_0^{2\pi}\left|\frac{\partial \Phi(\mathbf{x}_j, \mathbf{y}\left(\theta\right))}{\partial \mathbf{n}(\mathbf{y})}-i \Phi(\mathbf{x}_j, \mathbf{y}\left(\theta\right))\right|\left|\varphi-\varphi_n\right| \exp\left(p\left(\theta\right)\right) \sqrt{1+\left(p^{(1)}\left(\theta\right)\right)^2}\, d\theta \\
&+\left|\int_0^{2\pi} \underbrace{\left[\frac{\partial \Phi(\mathbf{x}_j, \mathbf{y}\left(\theta\right))}{\partial \mathbf{n}(\mathbf{y})}-i \Phi(\mathbf{x}_j, \mathbf{y}\left(\theta\right))\right] \varphi_n\left(\mathbf{y}\left(\theta\right)\right) \exp\left(p\left(\theta\right)\right) \sqrt{1+\left(p^{(1)}\left(\theta\right)\right)^2}}_{F}\, d\theta-\mathcal{I}_n\left(F\right)\right| \\
&\quad \leq c \frac{\exp\left(\|p\|_{\infty}\right)}{2n},
\end{aligned}
$$

where we have used (5.3) and (5.1) for the right-hand side of the first inequality, respectively, to obtain the second inequality. Note that this is possible due to the fact that $\frac{\partial \Phi(\mathbf{x}_j,\mathbf{y})}{\partial \mathbf{n}(\mathbf{y})} - i\Phi(\mathbf{x}_j,\mathbf{y})$ is analytic, and hence F is in $C^{1,\alpha}$. ■

5.2. Finite dimensional approximation of the prior. We have shown how to estimate the error in the observation operator due to the discretization of φ by the Nyström method, assuming that the shape parameter p is exactly represented. In practice, one has to approximate p as well, and accounting for the error in the observation operator induced from this approximation is necessary. We introduce the map

$$
Q : L^2\left[0, 2\pi\right] \ni p \mapsto Qp = \left\{\beta_j\right\}_{j=0}^{\infty} \in \ell^2\left(\mathbb{N}\right),
$$

with $\beta_j = (p, \psi_j)$, where $(\cdot, \cdot)$ denotes the $L^2\left[0, 2\pi\right]$ inner product. It follows that Q is an isometry. Consequently, we can identify

$$
\mu := \mathcal{N}\left(u, \sigma^2\right) \otimes \mathcal{N}\left(0, 1^{-6}\right) \otimes \mathcal{N}\left(0, 1^{-6}\right) \otimes \mathcal{N}\left(0, 2^{-6}\right) \otimes \mathcal{N}\left(0, 2^{-6}\right) \otimes \cdots
$$

as a probability measure on $\ell^2\left(\mathbb{N}_0\right)$ [20, 21], where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. This in turn implies that, on $\ell^2\left(\mathbb{N}_0\right)$, μ has the product structure.

We choose to construct a finite dimensional approximation to the parameter space $L^2\left[0, 2\pi\right]$ by projecting it into a subspace $V^m = \text{span}\left\{\psi_j\right\}_{j=0}^{2m}$ via the following projection operator Q_m:

$$
Q_m : L^2\left[0, 2\pi\right] \ni p \mapsto p_m := Q_m p := \frac{p_0}{\sqrt{2\pi}} - \sum_{j=1}^{m} p_j^c \frac{\cos\left(j\theta\right)}{\sqrt{\pi}} + p_j^s \frac{\sin\left(j\theta\right)}{\sqrt{\pi}} \in V^m,
$$

with $p_j^c, p_j^s \sim \mathcal{N}\left(0, j^{-6}\right)$. It follows that

$$
p_m^{(1)} = \sum_{j=1}^{m} j p_j^c \frac{\sin\left(j\theta\right)}{\sqrt{\pi}} - j p_j^s \frac{\cos\left(j\theta\right)}{\sqrt{\pi}} \quad \text{and} \quad p_m^{(2)} = \sum_{j=1}^{m} j^2 p_j^c \frac{\cos\left(j\theta\right)}{\sqrt{\pi}} + j^2 p_j^s \frac{\sin\left(j\theta\right)}{\sqrt{\pi}}.
$$

The probability measure μ_m on V^m, as a finite dimensional approximation of the prior measure μ, can therefore be identified with

$$
\mu_m := \mathcal{N}\left(u, \sigma^2\right) \otimes \mathcal{N}\left(0, 1^{-6}\right) \otimes \mathcal{N}\left(0, 1^{-6}\right) \otimes \cdots \otimes \mathcal{N}\left(0, m^{-6}\right) \otimes \mathcal{N}\left(0, m^{-6}\right).
$$

Let us denote $\mu_m^{\perp}$ as the complement measure of μ_m in μ such that $\mu = \mu_m \otimes \mu_m^{\perp}$. Next, we define

$$\mathcal{G}_{n,m}(\varphi_n, p_m) := \mathcal{G}_n(\varphi_n, p_m)$$

as an approximation to $\mathcal{G}(\varphi, p)$ taking into account the discretizations in both φ and p; similarly, $U_{n,m}(\varphi_n, p_m, \mathbf{x}) = U_n(\varphi_n, p_m, \mathbf{x})$. Let us now estimate the error between $\mathcal{G}_n(\varphi_n, p)$ and $\mathcal{G}_{n,m}(\varphi_n, p_m)$.

Lemma 5.4. *Suppose $p \sim \mu$ as in Lemma 4.2. There holds that*

$$|\mathcal{G}_n(\varphi_n, p) - \mathcal{G}_{n,m}(\varphi_n, p_m)|_{\mathbf{L}} \leq c\left(\frac{1}{n} + \frac{\log m}{m^{1+\alpha}}\right)\exp(\|p\|_\infty) \tag{5.4}$$

for some positive constant c independent of n and m.

Proof. Since $p \in C^{2,\alpha}$, a classical estimate for a truncated Fourier series [22] gives

$$\|p - p_m\|_\infty \leq c\frac{\log m}{m^{2+\alpha}} \quad \text{and} \quad \left\|p^{(1)} - p_m^{(1)}\right\|_\infty \leq c\frac{\log m}{m^{1+\alpha}}. \tag{5.5}$$

With F defined in the proof of Lemma 5.3, we have, by the triangle inequality,

$$\begin{aligned}
|U_{n,m}(\varphi_n, p_m, \mathbf{x}_j) - U_n(\varphi_n, p, \mathbf{x}_j)| &\leq \left|\mathcal{I}_n(F(\varphi_n, p_m)) - \int_0^{2\pi} F(\varphi_n, p_m)\, d\theta\right| \\
&+ \left|\int_0^{2\pi} F(\varphi_n, p_m)\, d\theta - \int_0^{2\pi} F(\varphi_n, p)\, d\theta\right| + \left|\int_0^{2\pi} F(\varphi_n, p)\, d\theta - \mathcal{I}_n(F(\varphi_n, p))\right| \\
&\leq c\left(\frac{1}{n} + \frac{\log m}{m^{1+\alpha}}\right)\exp(\|p\|_\infty),
\end{aligned}$$

where, in order to obtain the last inequality, we have used an argument similar to that in the proof of Lemma 5.3 for the first and the third terms of the right-hand side of the second inequality. In order to show the estimate for the second term, we first observe that

$$|F(\cdot, p) - F(\cdot, p_m)| \leq c\sum_{j=1}^{4} g_j |f_j(p) - f_j(p_m)|,$$

with

$$\begin{aligned}
&f_1(p) := \frac{\partial\Phi(\mathbf{x}_j, \mathbf{y}(p))}{\partial\mathbf{n}(\mathbf{y})} - i\Phi(\mathbf{x}_j, \mathbf{y}(p)), \quad f_2(p) := \sqrt{1 + \left(p^{(1)}\right)^2}, \\
&f_3(p) := \exp(p), \quad f_4(p) := \varphi_n(p), \quad g_j := \exp(\|p\|_\infty) \text{ for } j = 1, 2, 4, \quad \text{and } g_3 := 1.
\end{aligned}$$

Owing to the analyticity of f_1 and Lipschitz continuity of f_2, f_3, we have

$$\begin{aligned}
|f_j(p) - f_j(p_m)| &\leq c\left\|p^{(1)} - p_m^{(1)}\right\|_\infty \quad \text{for } j = 1, 2, \\
|f_3(p) - f_3(p_m)| &\leq c\|p - p_m\|_\infty \exp(\|p\|_\infty).
\end{aligned}$$

Then, an application of the triangle inequality gives

$$\begin{aligned}|\varphi_n(p) - \varphi_n(p_m)| &\le |\varphi_n(p) - \varphi(p)| + |\varphi(p) - \varphi(p_m)| + |\varphi(p_m) - \varphi_n(p_m)| \\ &\le \frac{c}{2n} + c\,\|p - r_m\|_\infty ,\end{aligned}$$

where we have used (5.3) for the first and the third terms of the right-hand side of the first inequality; for the second term, we use the fact that $\varphi \in C^{2,\alpha}(\Gamma_S)$ from the proof of Lemma 3.1, and hence it is Lipschitz continuous with respect to p. The proof is done by using the estimate (5.4) and the relation between $\mathcal{G}$ and U. ∎

We now have the error bound for the observation operator due to approximations in both the forward equation and the prior model.

Theorem 5.5. *Suppose $p \sim \mu$ as in Lemma 4.2; then the following error estimation holds:*

$$|\mathcal{G}(\varphi, p) - \mathcal{G}_{n,m}(\varphi_n, p_m)|_{\mathbf{L}} \le c\left(\frac{1}{n} + \frac{\log m}{m^{1+\alpha}}\right)\exp\left(\|p\|_\infty\right)$$

for some positive constant c.

Proof. With the help of Lemmas 5.3, 5.4 and the triangle inequality we have

$$\begin{aligned}|\mathcal{G}(\varphi, p) - \mathcal{G}_{n,m}(\varphi_n, p_m)|_{\mathbf{L}} &\le |\mathcal{G}(\varphi, p) - \mathcal{G}_n(\varphi_n, p)|_{\mathbf{L}} + |\mathcal{G}_n(\varphi_n, p) - \mathcal{G}_{n,m}(\varphi_n, p_m)|_{\mathbf{L}} \\ &\le c\left(\frac{1}{n} + \frac{\log m}{m^{1+\alpha}}\right)\exp\left(\|p\|_\infty\right). \qquad ∎\end{aligned}$$

5.3. Finite dimensional approximation of the posterior. We define an approximation to ν due to approximating the observation operator as

$$\frac{d\nu_{n,m}}{d\mu}(p|\mathbf{z}) := \frac{1}{Z_{n,m}}\exp\left(-\frac{1}{2}\left|\mathbf{z} - \mathcal{G}_{n,m}(\varphi_n, p_m)\right|_{\mathbf{L}}^2\right), \tag{5.6}$$

where $Z_{n,m} = \int_{V^m}\exp\left(-\frac{1}{2}\left|\mathbf{z} - \mathcal{G}_{n,m}(\varphi_n, p_m)\right|_{\mathbf{L}}^2\right)\mu_m(dp_m)$. The approximation (5.6) is well-defined, as we now show.

Proposition 5.6. *Let $p \sim \mu$ as in Lemma 4.2; then the following hold:*

(i) *The approximate measure $\nu_{n,m}$ is absolutely continuous with respect to μ, i.e., $\nu_{n,m} \ll \mu$, and the Radon–Nikodym derivative is given by (5.6).*

(ii) *The posterior measure $\nu_{n,m}$ is a well-defined probability measure on $L^2[0, 2\pi]$.*

(iii) *The posterior measure $\nu_{n,m}$ is Lipschitz continuous with respect to the data $\mathbf{z}$ in the Hellinger distance, i.e.,*

$$d_{\text{Hellinger}}(\nu_{n,m}(p|\mathbf{z}_1), \nu_{n,m}(p|\mathbf{z}_2)) \le C(\gamma)\,|\mathbf{z}_1 - \mathbf{z}_2|,$$

with $\max\{|\mathbf{z}_1|, |\mathbf{z}_2|\} < \gamma$.

Proof. As suggested by the proof of Theorem 4.3, it is sufficient to prove estimates similar to those in Proposition 3.4 and Corollary 3.6 for $\mathcal{G}_{n,m}(\varphi_n, p_m)$ (or $U_{n,m}(\varphi_n, p_m, \mathbf{x}_j)$). First, we have

$$\begin{aligned}|U_n(\varphi_n, p_m, \mathbf{x}_j)| &\le \frac{\pi}{n}\sum_{k=0}^{2n-1}|F(\varphi_n(p_m), p_m(\theta_k))| \le c\frac{\pi}{n}\sum_{k=0}^{2n-1}|\varphi_n|_\infty \exp\left(\|p_m\|_\infty\right) \\ &\le c\left\|U^I|_{\Gamma_S}\right\|_\infty \exp\left(\|p\|_\infty\right) \le \exp\left(\varepsilon\|p\|_\infty^2 + M\right),\end{aligned}$$

where we have used Lemma 5.2 in the second-to-last inequality. Second, using a technique similar to that in the proof of Lemma 5.4 we have

$$\begin{aligned}|F(\varphi_n(Q_m p_1), Q_m p_1(\theta_k)) - F(\varphi_n(Q_m p_2), Q_m p_2(\theta_k))| &\le c\,\|Q_m p_1 - Q_m p_2\|_{C^1[0,2\pi]}\\ &\le c\,\|p_1 - p_2\|_{C^1[0,2\pi]}\,.\end{aligned}$$

Consequently,

$$\begin{aligned}&|U_n(\varphi_n, Q_m p_1, \mathbf{x}_j) - U_n(\varphi_n, Q_m p_2, \mathbf{x}_j)|\\ &\le \frac{\pi}{n}\sum_{k=0}^{2n-1} |F(\varphi_n(\theta_k), Q_m p_1(\theta_k)) - F(\varphi_n(\theta_k), Q_m p_2(\theta_k))| \le c\,\|p_1 - p_2\|_{C^1[0,2\pi]}\,,\end{aligned}$$

and this ends the proof. ■

Owing to the fact that $\mu = \mu_m \otimes \mu_m^\perp$, $\nu_{n,m} \ll \mu$, and the right-hand side of (5.6) is a function only of p_m, we conclude that $\nu_{n,m} = \rho_{n,m} \otimes \mu_m^\perp$, where

$$\frac{d\rho_{n,m}}{d\mu_m}(p_m|\mathbf{z}) = \frac{1}{Z_{n,m}} \exp\left(-\frac{1}{2}|\mathbf{z} - \mathcal{G}_n(\varphi_n, p_m)|_{\mathbf{L}}^2\right). \tag{5.7}$$

Remark 5.7. It should be pointed out that $\nu_{n,m}$ can be viewed as an infinite dimensional approximation to the posterior ν by approximating the likelihood, while $\rho_{n,m}$, a part of $\nu_{n,m}$, is a finite dimensional approximation to ν by approximating both the likelihood and the prior. As a result, (5.7) is amenable to computer simulations.

6. Posterior convergence. The question we would like to address in this section is whether the approximate posterior measures $\nu_{n,m}$ in (5.6) converges to the infinite dimensional posterior measure ν in (4.1) in some sense. If it does converge, we would like to estimate the rate of convergence. To begin, let us denote $\mathbb{E}_\nu[\cdot]$ as the expectation under some measure ν. We have the following estimation in the Hellinger distance between $\nu_{n,m}$ and ν.

Theorem 6.1. *There exists a constant c, independent of n and m, such that the following holds:*

$$d_{\text{Hellinger}}(\nu, \nu_{n,m}) \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right).$$

Consequently, for any function $f : X \to H$ that has second moment with respect to both ν and $\nu_{n,m}$, where H is a Hilbert space, there holds that

$$\left\|\mathbb{E}_\nu[f] - \mathbb{E}_{\nu_{n,m}}[f]\right\|_H \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right).$$

Proof. We need only show that $\mathcal{G}_{n,m}(\varphi_n, p_m)$ satisfies Assumption A.1(i) uniformly in n and m, and then Lemma 5.3 allows us to use Theorem A.3 to conclude the proof. To that end, it is sufficient to show that

$$|U_{n,m}(\varphi_n, p_m, \mathbf{x}_j)| \le \exp\left(\varepsilon\,\|p\|_\infty^2 + M\right) \quad \forall p \in X,$$

for all $\varepsilon > 0$, and for $M = M(\varepsilon)$. But this is exactly the first inequality established in the proof of Proposition 5.6. ■

Corollary 6.2. *$\nu_{n,m}$ converges weakly to ν as n and m approach infinity.*

Proof. By the Portmanteau theorem [2], it is sufficient to show that

$$\mathbb{E}_{\nu_{n,m}}[f] \to \mathbb{E}_{\nu}[f] \quad \forall f : X \to \mathbb{R}, \text{ bounded and Lipschitz continuous.}$$

Since f is Lipschitz, we have

$$|f(p)| \le |f(0)| + c\,\|p\|_X\,,$$

which immediately shows that f has finite second moment with respect to both ν and $\nu_{n,m}$ by the Fernique theorem. Theorem 6.1 then gives

$$\left|\mathbb{E}_{\nu_{n,m}}[f] - \mathbb{E}_{\nu}[f]\right| \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right) \to 0 \quad \text{as } n, m \to \infty,$$

and this ends the proof. ■

Note that we cannot compare $\rho_{n,m}$ and $\nu_{n,m}$ (or ν) directly since $\rho_{n,m}$ is not a probability measure on the whole space X. Nevertheless, we can still quantify the uncertainty in approximating the posterior mean and the posterior covariance using $\rho_{n,m}$, as we shall show in section 7.

7. Error estimation for moments of finite dimensional posterior measures.

Section 5 presents a finite dimensional approximation to the Bayesian posterior measure in which $\rho_{n,m}$ is a computable approximation of ν. A question that immediately arises is how good $\rho_{n,m}$ is, in some sense. We address this question by estimating the errors of the posterior mean and posterior covariance under ν and $\rho_{n,m}$ in terms of n and m. We begin by defining the error in the mean and covariance as

$$\mathcal{E}_M = \mathbb{E}_{\nu}[p] - \mathbb{E}_{\rho_{n,m}}[p_m]\,,$$
$$\mathcal{E}_C = \mathbb{E}_{\nu}\left[(p - \mathbb{E}_{\nu}[p]) \otimes (p - \mathbb{E}_{\nu}[p])\right] - \mathbb{E}_{\rho_{n,m}}\left[\left(p_m - \mathbb{E}_{\rho_{n,m}}[p_m]\right) \otimes \left(p_m - \mathbb{E}_{\rho_{n,m}}[p_m]\right)\right].$$

Theorem 7.1. *There hold that*

$$\|\mathcal{E}_M\|_{L^2[0,2\pi]} \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right) \quad \textit{and} \quad \|\mathcal{E}_C\|_{L^2[0,2\pi]\otimes L^2[0,2\pi]} \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right).$$

Proof. We provide the proof for $\mathcal{E}_M$, and the proof for $\mathcal{E}_C$ follows similarly. Our proof technique is a variant of [12, Theorem 2.6]. For convenience in writing, we use $\|\cdot\|$ in place of $\|\cdot\|_{L^2[0,2\pi]}$. By the triangle inequality we have

$$\|\mathcal{E}_M\| \le \left\|\mathbb{E}_{\nu}[p] - \mathbb{E}_{\nu_{n,m}}[p]\right\| + \left\|\mathbb{E}_{\nu_{n,m}}[p] - \mathbb{E}_{\rho_{n,m}}[p_m]\right\|.$$

On the one hand, applying Theorem 6.1 we obtain

$$\left\|\mathbb{E}_{\nu}[p] - \mathbb{E}_{\nu_{n,m}}[p]\right\| \le c\left(\frac{1}{2n} + \frac{\log m}{m^{1+\alpha}}\right).$$

On the other hand, we observe that

$$\left\|\mathbb{E}_{\nu_{n,m}}[p] - \mathbb{E}_{\rho_{n,m}}[p_m]\right\| = \left\|\mathbb{E}_{\nu_{n,m}}[p] - \mathbb{E}_{\nu_{n,m}}[p_m]\right\| \le \mathbb{E}_{\nu_{n,m}}\left[\|p - p_m\|\right] \le c\frac{\log m}{m^{2+\alpha}}.$$

Combining these results concludes the proof. ■

8. Extension to general Hölder continuous priors. For the sake of concreteness and clarity, we have presented our results for $X = C^{2,\alpha}[0, 2\pi]$. However, if we believe a priori that p lives in a general smoother space, e.g., $C^{s,\alpha}[0, 2\pi]$, where $s \geq 2$, we can accommodate the above theory to have a better uncertainty quantification result consistent with the smoother prior, as we now show. We start by postulating a Gaussian measure on p such that the prior measure of $p^{(s)}$ is given by

$$p^{(s)}(\theta) \sim \mathcal{N}\left(0, \mathcal{A}^{-1}\right)$$

for which, similarly to section 4, the Karhunen–Loève expansion reads as

$$p^{(s)} = \sum_{j=1}^{\infty} \frac{a_j}{j} \frac{\cos(j\theta)}{\sqrt{\pi}} + \frac{b_j}{j} \frac{\sin(j\theta)}{\sqrt{\pi}}.$$

To generalize the results in section 4 we define p as

$$p(\theta) := \frac{p_0}{\sqrt{2\pi}} + \tilde{p}(\theta) + \tilde{p}^c(0),$$

where $\tilde{p}(\theta)$ is defined via the iteration

$$\widetilde{p^{(s-k-1)}} = \int_0^\theta p^{(s-k)} \quad \text{and} \quad p^{(s-k-1)} = \widetilde{p^{(s-k-1)}}(\theta) + \widetilde{p^{(s-k-1)}}^{c}(0), \quad k = 0, \ldots, s-1,$$

where the superscript c denotes the cosine part.

Within this setting, we now have $X := C^{s,\alpha}[0, 2\pi]$ and Lemma 4.2 holds again for $0 < \alpha < 1/2$, i.e., $\mu(X) = 1$. That is, a random function draw from μ is almost surely $C^{s,\alpha}$-Hölder continuous. In this case, a generalization of Lemma 3.1 (see [5, Theorem 1]) shows that $U \in C^{s,\alpha}\left(\mathbb{R}^2 \setminus \Omega_S\right)$, and the trace $U|_{\Gamma_S} \in C^{s,\alpha}(\Gamma_S)$. Both U and $U|_{\Gamma_S}$ depend continuously on $U^I\big|_{\Gamma_S}$ in the $C^{s,\alpha}$-norm. The numerical error in the density now becomes

$$\|\varphi_n(\theta) - \varphi(\theta)\|_\infty \leq \frac{c}{(2n)^{s-1}},$$

and the Fourier truncation error (5.5) reads as

$$\|p - p_m\|_\infty \leq c\frac{\log m}{m^{s+\alpha}} \quad \text{and} \quad \left\|p^{(1)} - p_m^{(1)}\right\|_\infty \leq c\frac{\log m}{m^{s-1+\alpha}}.$$

In addition, the general form for truncated prior measure μ_m is given by

$$\mu_m := \mathcal{N}\left(u, \sigma^2\right) \otimes \mathcal{N}\left(0, 1^{-(s+1)^2}\right) \otimes \mathcal{N}\left(0, 1^{-(s+1)^2}\right) \otimes \cdots \otimes \mathcal{N}\left(0, m^{-(s+1)^2}\right) \otimes \mathcal{N}\left(0, m^{-(s+1)^2}\right).$$

As a result, all the results in sections 4, 5, and 6 hold with $(2n)$ replaced by $(2n)^{s-1}$ and $m^{1+\alpha}$ by $m^{s-1+\alpha}$. In particular, we have the following general uncertainty quantification result.

Theorem 8.1. *For $s \geq 2$, there hold that*

$$d_{\text{Hellinger}}\left(\nu, \nu_{n,m}\right) \leq c\left(\frac{1}{(2n)^{s-1}} + \frac{\log m}{m^{s-1+\alpha}}\right),$$

$$\|\mathcal{E}_M\|_{L^2[0,2\pi]} \leq c\left(\frac{1}{(2n)^{s-1}} + \frac{\log m}{m^{s-1+\alpha}}\right),$$

$$\|\mathcal{E}_C\|_{L^2[0,2\pi]\otimes L^2[0,2\pi]} \leq c\left(\frac{1}{(2n)^{s-1}} + \frac{\log m}{m^{s-1+\alpha}}\right).$$

9. Conclusions. We have presented and analyzed an infinite dimensional Bayesian inference formulation, and its numerical approximation both in the likelihood and in the prior, for the inverse problem of inferring the shape of an obstacle from scattered acoustic waves. We derive the Lipschitz continuity of the observation operator, which is in turn sufficient to prove the well-posedness of the infinite dimensional Bayesian formulation. In particular, we first propose a Gaussian prior measure on the scatterer shape whose covariance operator is the inverse of the Laplacian. We then show that the Bayesian posterior measure is well-defined and it is Lipschitz continuous in the observation data with respect to the Hellinger distance. Next, a finite dimensional approximation to the posterior measure is proposed that involves a Nyström scheme for approximating the forward problem and a Karhunen–Loève truncation strategy for approximating the prior measure. Weak convergence of the resulting finite dimensional approximation, as well as convergence in the Hellinger distance, are investigated. In particular, we determine the convergence rate as a function of the number of Nyström quadrature points and the number of truncated terms in the Karhunen–Loève series. Finally, we estimate the error between the exact posterior moments, e.g., posterior mean and variance, and their finite dimensional approximate counterparts in terms of the errors due to forward equation approximation and prior approximation. The main result of this work is that the approximation of the Bayesian inverse problem exhibits spectral convergence, directly inheriting the spectral convergence rates of approximations of both the prior and the forward problem.

We have restricted ourselves to two dimensional acoustic scattering problems. To extend the infinite dimensional Bayesian results in this paper to three dimensions, one can use the regularity results of the scattered field with respect to the scatterer shape in [5] to obtain the boundedness and Lipschitz continuity of the observation operator similarly to Proposition 3.4 and Corollary 3.6. As a result, a well-posedness result for the infinite dimensional Bayesian posterior in three dimensions similar to Theorem 4.3 can be established in a straightforward manner. Finite dimensional approximations are more delicate, as we now discuss. Since the scatterer shape is assumed to be star-like with respect to the origin, it can be mapped to a sphere for which an efficient Nyström-type numerical method for three dimensional scattering problems is possible [14]. One can, in principle, go through exercises similar to those in section 5.1 to derive an error estimate for the observation operator when the forward equation is discretized. One can still postulate a Gaussian measure on the scatterer shape, but the inverse Laplacian is no longer a valid covariance operator for the L^2-shape. Indeed, $\Delta^{-\beta}$ for $\beta > 1$ is necessary [25]. In this case, one needs to replace the Fourier eigenfunctions by spherical harmonics; the Karhunen–Loève truncation approach is then still possible for the prior approximation. The details of carrying out the error analysis for the finite dimensional approximate posterior taking into account both the forward and the prior approximation errors is the subject of our future work.

Numerical results supporting our theoretical findings are still under investigation. The main difficulty is that evaluating the Hellinger distance in Theorem 6.1 numerically is not tractable due to the infinite dimensional nature of the prior measure on the shape space. Similarly, the exact mean $\mathbb{E}_{\nu}[f]$ is not practically available. Even computing the approximate mean $\mathbb{E}_{\nu_{n,m}}[f]$ accurately is a challenging problem since it involves evaluating an integral in (potentially) high dimensions. Monte Carlo integration seems to be a nature choice, but one needs to incorporate its error into the final estimate. A promising solution is to use the sparse deterministic approximation developed in [23].

Appendix. Let us collect some important results that are used in the paper. We begin with the key assumptions.

Assumption A.1 (see [25, Assumption 2.7]). The function $\mathcal{G}(\varphi,\cdot) : X \to \mathbb{R}^K$ satisfies the following:

(i) For each $\varepsilon > 0$, there exists an $M(\varepsilon) \in \mathbb{R}$ such that

$$|\mathcal{G}(\varphi,p)|_{\mathbf{L}} \leq \exp\left(\varepsilon \|p\|_X^2 + M\right) \quad \forall p \in X.$$

(ii) For each $r > 0$, there exists an $K(\varepsilon) > 0$ such that, with $\max\{\|p_1\|_X, \|p_2\|_X\} < r$,

$$|\mathcal{G}(\varphi,p_1) - \mathcal{G}(\varphi,p_2)| \leq K\|p_1 - p_2\| \quad \forall p_1, p_2 \in X.$$

Then, the following facts can be extracted from [25].

Theorem A.2 (see [25, Theorems 3.1, 4.1, and 4.2]). *Suppose that $\mathcal{G}(\varphi,\cdot)$ satisfies Assumption A.1 and that $\mu(X) = 1$.*

(i) *The posterior measure $\nu(p|\mathbf{z})$ is absolutely continuous with respect to μ, i.e., $\nu \ll \mu$, and the Radon–Nikodym derivative is given by*

$$\frac{d\nu}{d\mu}(p|\mathbf{z}) = \frac{1}{Z}\exp\left(-\frac{1}{2}|\mathbf{z} - \mathcal{G}(\varphi,p)|_{\mathbf{L}}^2\right),$$

where $Z = \int_X \exp\left(-\frac{1}{2}|\mathbf{z} - \mathcal{G}(\varphi,p)|_{\mathbf{L}}^2\right)\mu(dp)$.

(ii) *The posterior measure $\nu(p|\mathbf{z})$ is a well-defined probability measure on $L^2[0,2\pi]$.*

(iii) *The posterior measure $\nu(p|\mathbf{z})$ is Lipschitz continuous with respect to the data $\mathbf{z}$ in the Hellinger distance, i.e.,*

$$d_{\text{Hellinger}}(\nu(p|\mathbf{z}_1), \nu(p|\mathbf{z}_2)) \leq c(\gamma)|\mathbf{z}_1 - \mathbf{z}_2|,$$

with $\max\{|\mathbf{z}_1|, |\mathbf{z}_2|\} < \gamma$.

Theorem A.3 (see [25, Corollary 4.9]). *Suppose that ν and $\nu_{n,m}$ both have the Radon–Nikodym derivative with respect to μ given by (4.1) and (5.6), respectively. Assume that, for any $\varepsilon > 0$,*

$$|\mathcal{G}(\varphi,p) - \mathcal{G}_{n,m}(\varphi_n,p_m)| \leq c(\varepsilon)\exp\left(\varepsilon\|p\|_X^2\right)\zeta(n,m),$$

where $\zeta(n,m) \to 0$ as $n,m \to \infty$. If $\mathcal{G}_{n,m}$ satisfies Assumption A.1(i) uniformly in n and m, then there is a constant c, independent of n,m, such that

$$d_{\text{Hellinger}}(\nu, \nu_{n,m}) \leq c\zeta(n,m).$$

Consequently, the difference between expectation under ν *and* $\nu_{n,m}$ *of any polynomial bounded function* $f : X \to H$*, where* H *is a Hilbert space, is of order* $\zeta(n, m)$*.*

Acknowledgment. The authors would like to thank the anonymous referees for their critical and useful comments that significantly improved the paper.

REFERENCES

[1] T. Arbogast and J. L. Bona, *Methods of Applied Mathematics*, lecture notes in applied mathematics, University of Texas at Austin, Austin, TX, 2008.

[2] P. Billingsley, *Probability and Measure*, 2nd ed., Wiley, New York, 1986.

[3] P. Blanchard and E. Brüning, *Mathematical Methods in Physics*, Birhäuser Verlag, Basel, 2003.

[4] V. I. Bogachev, *Gaussian Measures*, Math. Surveys Monogr. 62, American Mathematical Society, Providence, RI, 1998.

[5] T. Bui-Thanh and O. Ghattas, *Analysis of the Hessian for inverse scattering problems. Part* I: *Inverse shape scattering of acoustic waves*, Inverse Problems, 28 (2012), 055001.

[6] F. Cakoni, D. Colton, and P. Monk, *The Linear Sampling Method in Inverse Electromagnetic Scattering*, SIAM, Philadelphia, 2011.

[7] J. Cea, *Problems of shape optimal design*, in Optimization of Distributed Parameter Structures, E. J. Haug and J. Cea, eds., Sijthoff & Noordhoff, Alphen aan den Rijn, The Netherlands, 1980, pp. 1005–1048.

[8] D. Colton, *The inverse scattering problem for time-harmonic acoustic waves*, SIAM Rev., 26 (1984), pp. 323–350.

[9] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*, John Wiley & Sons, New York, 1983.

[10] D. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering*, 2nd ed., Appl. Math. Sci. 93, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1998.

[11] S. L. Cotter, M. Dashti, and A. M. Stuart, *Approximation of Bayesian inverse problems for PDEs*, SIAM J. Numer. Anal., 48 (2010), pp. 322–345.

[12] M. Dashti and A. M. Stuart, *Uncertainty quantification and weak approximation of an elliptic inverse problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2524–2542.

[13] K. Eppler, *Second derivatives and sufficient optimality conditions for shape functionals*, Control Cybernet., 29 (2000), pp. 485–512.

[14] M. Ganesh and I. G. Graham, *A high-order algorithm for obstacle scattering in three dimensions*, J. Comput. Phys., 198 (2004), pp. 211–242.

[15] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Yale University Press, New Haven, CT, 1923.

[16] M. Hairer, *Introduction to Stochastic PDEs*, lecture notes, University of Warwick, Coventry, UK, Courant Institute, New York, NY, 2009.

[17] R. Kress, *Linear Integral Equations*, Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1989.

[18] Y. Marzouk and D. Xiu, *A stochastic collocation approach to Bayesian inference in inverse problems*, Commun. Comput. Phys., 6 (2009), pp. 826–847.

[19] S. G. Mikhlin, *Mathematical Physics: An Advanced Course*, North–Holland, Amsterdam, 1970.

[20] G. Da Prato, *An Introduction to Infinite-Dimensional Analysis*, Universitext, Springer, Berlin, 2006.

[21] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, UK, 1992.

[22] T. J. Rivlin, *An Introduction to the Approximation of Functions*, Dover, New York, 1981.

[23] C. Schwab and A. M. Stuart, *Sparse deterministic approximation of Bayesian inverse problems*, Inverse Problems, 28 (2012), 045003.

[24] B. D. Sleeman, *The inverse problem of acoustic scattering*, IMA J. Appl. Math., 29 (1980), pp. 113–142.

[25] A. M. Stuart, *Inverse problems: A Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559.

[26] B. Vexler, *Adaptive Finite Element Methods for Parameter Identification Problems*, Ph.D. thesis, University of Heidelberg, Heidelberg, Germany, 2004.