Available online at www.sciencedirect.com
ScienceDirect
Comput. Methods Appl. Mech. Engrg. 386 (2021) 114073
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Bayesian approach to inverse time-harmonic acoustic obstacle scattering with phaseless data generated by point source waves

Zhipeng Yang[a], Xinping Gui[a], Ju Ming[b], Guanghui Hu[c,*]

[a] *Department of Applied Mathematics, Beijing Computational Science Research Center, Beijing 100193, PR China*
[b] *School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, PR China*
[c] *School of Mathematical Sciences and LPMC, Nankai University, Tianjin 300071, PR China*

Received 1 September 2020; received in revised form 26 May 2021; accepted 20 July 2021
Available online 14 August 2021

Abstract

This paper concerns the Bayesian approach to inverse acoustic scattering problems of inferring the position and shape of a sound-soft obstacle from phaseless far-field data generated by two-dimensional point source waves. Given the total number of obstacle parameters, the Markov chain Monte Carlo (MCMC) method is employed to reconstruct the boundary of the obstacle in a high-dimensional space, which usually leads to slow convergence and prohibitively high computational cost. We use the Gibbs sampling and preconditioned Crank–Nicolson (pCN) algorithm with random proposal variance to improve the convergence rate, and design an effective strategy for the surrogate model constructed by the generalized polynomial chaos (gPC) method to reduce the computational cost of MCMC. Numerical examples are provided to illustrate the effectiveness of the proposed method.

MSC: 35R30; 35P25; 62F15; 78A46

Keywords: Inverse scattering; Helmholtz equation; Phaseless far-field data; Bayesian inference; MCMC

---

1. Introduction

Inverse acoustic, elastic and electromagnetic scattering problems occur in many applications such as geophysics, seismology, nondestructive testing, and radar imaging. In this paper, we consider the inverse problem of reconstructing an acoustically sound-soft obstacle from phaseless far-field data in two dimensions. Let $D \subset \mathbb{R}^2$ be a sound-soft scatterer, which occupies a bounded domain with C^2-smooth boundary ∂D such that its exterior $\mathbb{R}^2 \backslash \overline{D}$ is connected. Suppose that D is embedded in a homogeneous isotropic background medium and that it is illuminated by a given time-harmonic incident field u^{in}, which satisfies the Helmholtz equation

$$\Delta u^{\mathrm{in}} + k^2 u^{\mathrm{in}} = 0$$

at least in a neighboring area of D. Here $k > 0$ denotes the wavenumber of the background medium. The total field u is defined as $u := u^{\mathrm{in}} + u^{\mathrm{sc}}$ in $\mathbb{R}^2 \backslash \overline{D}$, where u^{sc} is the corresponding scattered (perturbed) field. Since D is

---

[*] Corresponding author.

E-mail addresses: yangzhp@csrc.ac.cn (Z. Yang), gui@csrc.ac.cn (X. Gui), jming@hust.edu.cn (J. Ming), ghhu@nankai.edu.cn (G. Hu).

https://doi.org/10.1016/j.cma.2021.114073
0045-7825/© 2021 Elsevier B.V. All rights reserved.

physically sound-soft, the total field u satisfies the Dirichlet boundary condition $u = 0$ on the boundary ∂D together with the Helmholtz equation $\Delta u + k^2 u = 0$ in $\mathbb{R}^2 \backslash \overline{D}$. Furthermore, the scattered field u^{sc} is required to fulfill the Sommerfeld radiation condition (see (2.4)) at infinity, leading to the asymptotic behavior

$$u^{\mathrm{sc}}(x) = \frac{e^{\mathrm{i}k|x|}}{\sqrt{|x|}} \left\{ u^{\infty}(\hat{\mathbf{x}}) + \mathcal{O}\left(\frac{1}{\sqrt{|x|}} \right) \right\}, \quad |x| \to \infty, \tag{1.1}$$

uniformly in all directions $\hat{\mathbf{x}} := x/|x| \in \mathbb{S}$. Here, $i = \sqrt{-1}$ is the imaginary unit, $\mathbb{S} := \{x \in \mathbb{R}^2 : |x| = 1\}$ and $|\cdot|$ denotes the norm of a vector or the modules of a complex number. The far-field pattern $u^{\infty} : \mathbb{S} \to \mathbb{C}$ is known as a real-analytic function defined over the unit circle $\mathbb{S}$. The above model also arises from the TE and TM polarizations of time-harmonic electromagnetic scattering from infinitely long cylinders. It follows from [1, Chapter 3.2] that the forward scattering problem always admits a unique solution $u \in C^2(\mathbb{R}^2 \backslash \overline{D}) \cap C(\overline{D})$.

Inverse obstacle scattering problems are well-known to be non-linear and severely ill-posed. We refer to monographs [1–5] for uniqueness, stability and inversion algorithms using phased far-field patterns, which have been extensively studied over the last three decades. In many applications, the accurate phase of the far-field pattern is usually difficult to obtain. Instead, the intensity of the data can be easily measured. This motivates us to consider inverse scattering problems using the phaseless far-field pattern $|u^{\infty}|$. Without the phase information, the inverse problems become more ill-posed. A key challenge lies in the translation invariance property (see e.g. [6]) of the phaseless far-field pattern caused by a single plane wave. To break the translation invariance property, one approach was recently proposed in [7,8] by using infinitely many sets of superposition of two plane waves with different directions as incoming waves, in which both uniqueness and algorithms were investigated. Analogously, the authors in [9] made use of different types of incident waves generated by superposition of a fixed plane wave and some point sources. Based on the idea of [7], in our previous work [10] we adapt the Bayesian approach to the recovery of a sound-soft disk, a line crack and a kite-shaped obstacle with less parameters from phaseless far-field patterns generated by plane waves.

The purpose of this paper is to recover obstacles with complex geometric shapes which can be parameterized in high dimensions. To break the translation invariance property caused by a single plane wave incidence, we suppose in this paper that the incoming waves are excited by the two-dimensional point source waves

$$u_{\ell}^{\mathrm{in}}(x) := \frac{i}{4} H_0^{(1)}(k|x - x_{\ell}^{\mathrm{in}}|), \quad x \in \mathbb{R}^2 \backslash \{x_{\ell}^{\mathrm{in}}\}, \quad \ell = 1, 2, \ldots, L. \tag{1.2}$$

Here, $x_{\ell}^{\mathrm{in}} \in \mathbb{R}^2 \backslash \overline{D}$ is the position of the ℓth point source, $H_0^{(1)}(\cdot)$ is the Hankel function of the first kind of order zero, and $L \in \mathbb{N}_0$ is the number of the incident waves. All source positions $x_{\ell}^{\mathrm{in}}, \ell = 1, 2, \ldots, L$ are assumed to lie on a large circle $|x| = R$ which contains the underlying obstacle inside. Emphasis of this paper will be placed upon how to reduce the computational cost of the Bayesian approach for recovering complex obstacles. An a-priori convergence estimate of the Bayesian approach is of great significance in practice, but will be explored in our future works. It should be remarked that the translation invariance property for phaseless far-field pattern generated by plane waves does not apply to point source waves. To the best of the authors' knowledge, the unique determination of a general sound-soft obstacle from phaseless far-field patterns corresponding to the aforementioned incoming waves still remains open. Using multi-frequency near-field data, Klibanov proved unique determination of a compactly supported potential of the stationary Schrödinger equation in 3D [11]. This was later extended in [12] for the reconstruction of a smooth wave speed in the three-dimensional Helmholtz equation. In a deterministic setting, we refer to [6,13,14] for inversion algorithms based on the Newton-type iterative scheme.

Recently, the Bayesian approach has attracted extensive attention for inverse problems; see e.g., [15–18] for an incomplete list. In [15,16], the authors built up a framework for the well-posedness of the posterior distribution, which was later used in [19] to determine the permeability of the subsurface from hydraulic head measurements. The Bayesian approach has also been used with great success to solve inverse scattering problems with phased far-field data in [20–22]. For simplification, we assume that the number of parameters (e.g., Fourier coefficients) for characterizing the geometrical shape of the sound-soft obstacle is known. This is based on the fact that one can always approximate the shape/boundary in a finite dimensional space. Such an assumption is also used in other deterministic optimization schemes for inverse scattering. The numerical reconstruction by our algorithm can be regarded as an approximation of the boundary, if the total number of obstacle parameters is larger than the given upper bound. Hence, our algorithm is also applicable to inverse scattering from general obstacles, if we increase the given upper bound of the dimension of obstacle parameters. In contract with the few unknown parameters in [10], in this paper we will consider acceleration strategies for recovering obstacles with many parameters. An obstacle with 11 unknown parameters will be recovered in Section 5.

The Bayesian method could be an alternative approach to inverse scattering, when we come up against challenges from deterministic inversion schemes, such as a good initial guess required in the optimization-based iterative schemes and a large number of observation data in non-iterative sampling methods. On the other hand, as a disadvantage, it always requires expensive computational cost due to the following reasons. (i) Since it is quite difficult or even impossible to gain an analytical form of the posterior distributions, we always choose sampling methods such as the Markov chain Monte Carlo (MCMC) method [23–25] to perform numerical approximation. However, an accurate estimation of the posterior distribution often requires a sufficient number of samples, especially for high-dimensional unknown parameters. (ii) In the iteration process of the Markov chain, accepting or rejecting a candidate state usually requires one or more forward solutions to calculate the associated Hastings ratio. Hence, the Bayesian method involves a quite large number of repeated solutions of the forward problem. Consequently, the computational cost of the Bayesian inference is prohibitively expensive, especially when the forward problem is computationally intensive. Unfortunately, in most applications, the forward model is always a nonlinear operator associated with partial differential equations. Therefore, how to reduce the computational cost of the MCMC method is a key point to implement the Bayesian method.

Roughly speaking, the total computational cost of MCMC is the product of the number of iteration steps and the computational cost of one forward solution, which gives rise to a criterion on how to save computational efforts. In this paper, the preconditioned Crank–Nicolson (pCN) algorithm [26] is adopted to reduce the number of iteration steps. Since the unknown parameters are high dimensional, we also adopt the Gibbs sampling [5,27,28] to accelerate the convergence of MCMC method. For the kite-shaped obstacle with six unknowns, 200 000 iterations are needed in our former work [10]. In this paper, we only use 20 000 iterations through the Gibbs sampling in determining the same obstacle. Furthermore, by the Gibbs sampling, 50 000 iterations are proven enough to get accurate reconstructions for an obstacle with 11 unknowns. However, even with these advanced methods the number of iteration steps is still high, which turns out to be the order of magnitude of tens of thousands or even hundreds of thousands. Then we have to resort to the idea of reducing the computational cost of the forward scattering model. Recently, substantial attempts have been made to accelerate the Bayesian method in inverse problems associated with a computationally intensive forward model. Using piecewise linear interpolation, Ma and Zabaras [29] adopted the adaptive hierarchical sparse grid collocation (ASGC) method to construct an approximation of the stochastic forward model. Similarly, Marzouk and Xiu [30] proposed sparse grid stochastic collocation methods to improve the efficiency of Bayesian inference. The latter are based on the generalized polynomial chaos to construct a stochastic surrogate model of the forward model over the support of the prior distribution. Yan and Guo [31] develop the same idea by combining the sparse grid stochastic collocation method with the compressive sensing (CS) method. They employ the ℓ_1-minimization to construct the stochastic surrogate model. In [32], the adaptive sparse-grid high-order stochastic collocation (aSG-hSC) method is used to construct the surrogate system of a nonlinear groundwater reactive transport model. Based on the truncated Karhunen–Loève (KL) expansions of the prior distribution, a reduced model [33] is constructed by the Galerkin projection onto a polynomial chaos basis. A data-driven strategy was employed in [34] to construct the reduced-order model by projecting the full forward model onto a reduced subspace. Liao and Li [35] proposed the Analysis of Variance (ANOVA) method to reduce the forward model both in the statistical and physical spaces.

The key of the aforementioned methods is to derive a reduced-order and computationally efficient surrogate for high-fidelity, large-scale, computationally costly forward models. This surrogate model is then utilized in place of the original forward model to reduce the computational cost in the MCMC iterations. In this paper, we employ the generalized polynomial chaos (gPC) method to construct a stochastic surrogate model. The forward scattering model is projected onto a limited number of basis functions over the support of the prior distribution. Such a projection approach will be used in our MCMC iterations. As concluded in [36], the surrogate model constructed by the gPC method will give rise to inaccurate results, especially when the prior distribution is significantly different from the posterior distribution. Several efforts with adaptive construction of the surrogate model [34,35,37,38] have been made to overcome this challenge. The snapshots of the reduced-order model are adaptively calculated in [34] during the iterations of MCMC method. Liao and Li [35] also derive an adaptive scheme to combine the reduced basis ANOVA model with the MCMC iterations. Regarding the surrogates constructed by the gPC method, Li and

Marzouk [37] consider a stochastic optimization problem in the sense of the Kullback–Leibler divergence and the authors of [38] propose the adaptive multi-fidelity surrogate for the MCMC method.

In this paper, we also develop a special strategy to combine the MCMC method with a surrogate model constructed by the gPC method. In the iteration of the MCMC method, the model-data misfit function determines whether we accept the candidate state or not. The probability of accepting the candidate state will be larger, if value of the model-data misfit function is smaller. In other words, the MCMC method will more rapidly converge to the posterior distribution, if it is easier to determine the new proposal and if value of the model-data misfit function is smaller. This motivates us with the following strategy to use the surrogate model constructed by the gPC method. To produce candidate states, we generate a large number of samples with the selected proposal algorithm, and then calculate their model-data misfit functions with the surrogate model. This enables us to evaluate these new proposal obstacle parameters through the model-data misfit function in low precision but with cheap computational cost. Furthermore, we can correct the evaluation of a part of these samples with the exact forward model at low cost. Then we can evaluate these samples in high precision and with cheap computational cost. Therefore, it will be more effective to generate appropriate candidate states, which will accelerate the convergence of the MCMC method and save computational cost. Compared with the results in [10], we obtain accurate reconstructions with lower computational cost for recovering the kite-shaped obstacle.

This paper is organized as follows. In Section 2, we introduce the deterministic forward scattering problem. Section 3 is devoted to the Bayesian framework to inverse scattering problems with phaseless data. In Section 4, we construct the surrogate model for the forward scattering problem. The generalized polynomial chaos method will be adopted to reduce the computational cost of MCMC method. Numerical examples will be reported in Section 5 and conclusions are finally made in Section 6.

2. Deterministic forward scattering problem

In this paper we want to recover the position and shape of an unknown sound-soft obstacle from phaseless far-field patterns corresponding to a set of incident point source waves. Before dealing with the inverse problem, we need to formulate the abstract nonlinear operator which maps the obstacle parameters to the far-field observation data in the deterministic setting.

Since the boundary $\partial D \subset \mathbb{R}^2$ is a closed C^2-smooth curve, we can represent or approximate ∂D by a finite set $\mathbf{Z}$ of variables

$$\mathbf{Z} := (z_1, z_2, \ldots, z_N)^\top \in \mathbb{R}^N, \quad N \in \mathbb{N}_0. \tag{2.1}$$

For example, we can use $\mathbf{Z} := (a_1, b_1, a_2, b_2, \ldots, a_N, b_N)^\top$ to approximate a star-shaped closed curve where $\{(a_n, b_n) : n = 1, \ldots, N\}$ stands for the Fourier coefficients in the truncated Fourier expansion. Let $u_\ell^{\mathrm{in}}(x)$, $\ell = 1, 2, \ldots, L$ be incident waves given by the formula (1.2). The forward scattering problem is to find the scattered field u^{sc} to the Helmholtz equation

$$\Delta u^{\mathrm{sc}} + k^2 u^{\mathrm{sc}} = 0 \ \text{ in } \ \mathbb{R}^2 \backslash \overline{D}, \tag{2.2}$$

which satisfies the inhomogeneous Dirichlet boundary condition

$$u^{\mathrm{sc}} = -u^{\mathrm{in}} \ \text{ on } \ \partial D, \tag{2.3}$$

and the Sommerfeld radiation condition

$$\lim_{r\to\infty} \sqrt{r} \left(\frac{\partial u^{\mathrm{sc}}}{\partial r} - iku^{\mathrm{sc}} \right) = 0, \quad r = |x|, \tag{2.4}$$

uniformly in all directions $\hat{\mathbf{x}} \in \mathbb{S}$. The far-field pattern of the scattering models (2.2)–(2.4) with and without phase information can be expressed in terms the obstacle parameters $\mathbf{Z}$ by

$$u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, x_\ell^{\mathrm{in}}, k), \quad |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, x_\ell^{\mathrm{in}}, k)|, \quad \ell = 1, 2, \ldots, L, \quad \hat{\mathbf{x}} \in \mathbb{S}. \tag{2.5}$$

We rewrite the forward scattering problem by the operator $F^\ell : \mathbb{R}^N \to C(\mathbb{S})$ as

$$F^\ell(\mathbf{Z}) := |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, x_\ell^{\mathrm{in}}, k)|, \quad \hat{\mathbf{x}} \in \mathbb{S}, \quad \ell = 1, 2, \ldots, L. \tag{2.6}$$

Here $C(\mathbb{S})$ denotes the space of continuous functions on $\mathbb{S}$. In other words, each operator $F^\ell (\ell = 1, 2, \ldots, L)$ can be regarded as an abstract map from the space of obstacle parameters to the space of phaseless far-field patterns in the continuous sense. From the well-posedness of forward scattering, the operators $F^\ell, \ell = 1, 2, \ldots, L$, are continuous but highly nonlinear.

Let $G = (g_1, g_2, \ldots, g_M)^\top : C(\mathbb{S}) \to \mathbb{R}^M$ be a bounded linear observation operator with $g_m : C(\mathbb{S}) \to \mathbb{R}_+$ defined as

$$g_m(|u^\infty(\hat{\mathbf{x}})|) := |u^\infty(\hat{\mathbf{x}}_m)|, \quad m = 1, 2, \ldots, M, \tag{2.7}$$

where $\{\hat{\mathbf{x}}_m \in \mathbb{S}\}_{m=1}^M$ denotes the set of discrete observation directions, and $M \in \mathbb{N}_0$ is the number of observation directions. Corresponding to the incident wave $u_\ell^{\text{in}}(x)$ and the obstacle parameters $\mathbf{Z} \in \mathbb{R}^N$, we denote the map $\mathcal{G}^\ell : \mathbb{R}^N \to \mathbb{R}^M$ from the obstacle parameter space to observation space as

$$\begin{aligned} \mathcal{G}^\ell(\mathbf{Z}) &:= G \circ F^\ell(\mathbf{Z}) \\ &= \big(|u^\infty(\hat{\mathbf{x}}_1; \mathbf{Z}, x_\ell^{\text{in}}, k)|, |u^\infty(\hat{\mathbf{x}}_2; \mathbf{Z}, x_\ell^{\text{in}}, k)|, \ldots, |u^\infty(\hat{\mathbf{x}}_M; \mathbf{Z}, x_\ell^{\text{in}}, k)|\big)^\top. \end{aligned} \tag{2.8}$$

Let $\mathbf{Y}^\ell = (y_1^\ell, y_2^\ell, \ldots, y_M^\ell)^\top$ be the measurement data of the phaseless far-field pattern with the observation noise $\eta^\ell = (\eta_1^\ell, \eta_2^\ell, \ldots, \eta_M^\ell)^\top \in \mathbb{R}^M$. Then we can express the observation data as

$$\mathbf{Y}^\ell = \mathcal{G}^\ell(\mathbf{Z}) + \eta^\ell, \quad \ell = 1, 2, \ldots, L, \tag{2.9}$$

or equivalently,

$$y_m^\ell = g_m(F^\ell(\mathbf{Z})) + \eta_m^\ell = |u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, x_\ell^{\text{in}}, k)| + \eta_m^\ell, \quad m = 1, 2, \ldots, M. \tag{2.10}$$

Now our inverse problem can be stated as: determine the obstacle parameters $\mathbf{Z} \in \mathbb{R}^N$ from the observation data $\mathbf{Y}^\ell \in \mathbb{R}^M$ with the noise pollution $\eta^\ell \in \mathbb{R}^M$, $\ell = 1, 2, \ldots, L$.

3. Bayesian framework

Within the Bayesian framework, all parameters are random variables and the key issue is to estimate the posterior distribution of the obstacle parameters based on the Bayes formula [16] and the given assumptions of the prior distribution and the observation pollution. Since an explicit expression of the posterior distribution is not available, we adopt the Markov chain Monte Carlo method (MCMC) to get an approximation of the posterior distribution.

3.1. Posterior distribution

The prior distribution of the obstacle parameters $\mathbf{Z}$ depends on the distribution of $z_n, n = 1, 2, \ldots, N$. Let $\{z_n\}_{n=1}^N$ be independent Gaussian variables as $z_n \sim \mathcal{N}(m_n, \sigma_n), n = 1, 2, \ldots, N$, where m_n, σ_n are the mean and variance of the distribution of z_n. For simplicity, we assume that $\sigma_1 = \cdots = \sigma_N = \sigma_{\text{pr}}$, implying that $\mathbf{Z} \sim \mathcal{N}(\mathbf{m}_{\text{pr}}, \sigma_{\text{pr}}\mathbf{I})$, where $\mathbf{m}_{\text{pr}} = (m_1, m_2, \ldots, m_N)^\top$ and $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix. Then the prior distribution $P_{\text{pr}}(\mathbf{Z})$ is given by

$$P_{\text{pr}}(\mathbf{Z}) = (2\pi\sigma_{\text{pr}})^{-\frac{N}{2}} \exp\Big(-\frac{1}{2\sigma_{\text{pr}}}|\mathbf{Z} - \mathbf{m}_{\text{pr}}|^2\Big). \tag{3.1}$$

We assume that the observation pollution η^ℓ is independent of u^∞, and drawn from the Gaussian distribution $\mathcal{N}(\mathbf{0}, \Sigma_\eta^\ell)$ with density ρ^ℓ, where $\Sigma_\eta^\ell \in \mathbb{R}^{M\times M}$ is a self-adjoint positive matrix, $\ell = 1, 2, \ldots, L$. From the observation data with noise (2.9), it follows the relationship $\mathbf{Y}^\ell|\mathbf{Z} \sim \mathcal{N}(\mathcal{G}^\ell(\mathbf{Z}), \Sigma_\eta^\ell), \ell = 1, 2, \ldots, L$. Define the model-data misfit function $\Phi^\ell(\mathbf{Z}; \mathbf{Y}^\ell) : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$ as

$$\Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big) = \frac{1}{2}|\mathbf{Y}^\ell - \mathcal{G}^\ell(\mathbf{Z})|_{\Sigma_\eta^\ell}^2, \quad \ell = 1, 2, \ldots, L. \tag{3.2}$$

Here the norm $|\cdot|_{\Sigma_\eta^\ell}$ is defined as

$$|x|_{\Sigma_\eta^\ell}^2 := x^\top\big(\Sigma_\eta^\ell\big)^{-1}x, \quad x \in \mathbb{R}^M, \quad \ell = 1, 2, \ldots, L. \tag{3.3}$$

Hence, the likelihood function $P_{\rm lhd}\big(\mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^L; \mathbf{Z}\big)$ is given by

$$\begin{aligned} P_{\rm lhd}\big(\mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^L; \mathbf{Z}\big) &= \prod_{\ell=1}^{L} \rho^\ell\big(\mathbf{Y}^\ell - \mathcal{G}^\ell(\mathbf{Z})\big) \\ &= (2\pi)^{-\frac{LM}{2}} \prod_{\ell=1}^{L}\big(\det(\Sigma_\eta^\ell)\big)^{-\frac{1}{2}} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big)\Big). \end{aligned} \tag{3.4}$$

Furthermore, by Bayes' theorem [16,18], the posterior distribution $P_{\rm post}(\mathbf{Z}; \mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^L)$ is given as

$$P_{\rm post}(\mathbf{Z}; \mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^L) = C_z^{-1} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big) - \frac{1}{2\sigma_{\rm pr}}|\mathbf{Z} - \mathbf{m}_{\rm pr}|^2\Big). \tag{3.5}$$

Here C_z is the normalization constant

$$C_z = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big) - \frac{1}{2\sigma_{\rm pr}}|\mathbf{Z} - \mathbf{m}_{\rm pr}|^2\Big) d\mathbf{Z}. \tag{3.6}$$

Let $\mu_{\rm pr}$ and $\mu_{\rm post}$ be the probability measure of the prior distribution $P_{\rm pr}$ and the posterior distribution $P_{\rm post}$, respectively. From (3.5), (3.6) and the Radon–Nikodym derivative [39], we have

$$\frac{d\mu_{\rm post}}{d\mu_{\rm pr}}(\mathbf{Z}) = C_{\mathbf{Y}}^{-1} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big)\Big). \tag{3.7}$$

Here, $C_{\mathbf{Y}}$ is a modification of the normalization constant C_z depending on the observation data $\{\mathbf{Y}^\ell\}_{\ell=1}^L$, given by

$$C_{\mathbf{Y}} = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big)\Big) d\mu_{\rm pr}(\mathbf{Z}). \tag{3.8}$$

Before proving well-posedness of the posterior distribution, we shall discuss properties of the matrices $\{\Sigma_\eta^\ell\}_{\ell=1}^L$, forward maps $\{\mathcal{G}^\ell\}_{\ell=1}^L$, and model-data misfit functions $\{\Phi^\ell\}_{\ell=1}^L$ in the following lemmas.

Lemma 3.1. *Denote by* λ_{min}^ℓ $(\ell = 1, 2, \ldots, L)$ *the minimum eigenvalue of the matrix* Σ_η^ℓ *and set* $\lambda_{min} := \min_{\ell=1,2,\ldots,L}\{\lambda_{min}^\ell\}$. *Then we have* $\lambda_{min}^\ell > 0$ *for* $\ell = 1, 2, \ldots, L$ *and* $\lambda_{min} > 0$.

Proof. Since Σ_η^ℓ is a self-adjoint positive matrix, it is easy to see $\lambda_{\rm min}^\ell > 0$ and $\lambda_{\rm min} > 0$. □

Lemma 3.2. *For* $\ell = 1, 2, \ldots, L$, *there is a constant* C_F *depending on the operators* $\{F^\ell\}_{\ell=1}^L$ *such that the forward map* $\mathcal{G}^\ell : \mathbb{R}^N \to \mathbb{R}^M$ *satisfies*

$$|\mathcal{G}^\ell(\mathbf{Z})| \le C_F, \tag{3.9}$$

$$|\mathcal{G}^\ell(\mathbf{Z_1}) - \mathcal{G}^\ell(\mathbf{Z}_2)| \le C_F|\mathbf{Z_1} - \mathbf{Z_2}|. \tag{3.10}$$

Proof. The continuity and differentiability of the far-field mapping have been sufficiently studied in [1, Chapter 5.3]. Consequently, the phaseless far-field pattern $\{|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, x_\ell^{\rm in}, k)|\}$ is bounded and Lipschitz continuous with respect to $\mathbf{Z} \in \mathbb{R}^N$ for each $\ell = 1, 2, \ldots, L$. This gives rise to the conclusions in (3.9) and (3.10). □

Using the definition (3.2) and (3.3) together with Lemmata 3.1 and 3.2, one can prove

Lemma 3.3. *Given* $\big\{\mathbf{Y}^\ell\big\}_{\ell=1}^L$, $\big\{\mathbf{Y}_1^\ell\big\}_{\ell=1}^L$, $\big\{\mathbf{Y}_2^\ell\big\}_{\ell=1}^L$ *such that* $\max_{\ell=1,2,\ldots,L}\big\{|\mathbf{Y}^\ell|, |\mathbf{Y}_1^\ell|, |\mathbf{Y}_2^\ell|\big\} < r$ *and* $\mathbf{Z}$, $\mathbf{Z}_1$, $\mathbf{Z}_2 \in \mathbb{R}^N$, *there is a constant* $C_\Phi = C_\Phi(C_F, r, L, M, \lambda_{min})$ *such that the model-data misfit functions* $\Phi^\ell(\mathbf{Z}; \mathbf{Y}^\ell) : \mathbb{R}^N \times \mathbb{R}^M \to \mathbb{R}$, $\ell = 1, 2, \ldots, L$ *have the following properties:*

$$\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}; \mathbf{Y}^\ell\big) \le C_\Phi, \tag{3.11}$$

$$\left|\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_1^{\ell}\big)-\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_2^{\ell}\big)\right|\le C_{\Phi}\sum_{\ell=1}^{L}\left|\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right|, \tag{3.12}$$

$$\left|\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z}_1;\mathbf{Y}^{\ell}\big)-\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z}_2;\mathbf{Y}^{\ell}\big)\right|\le C_{\Phi}\left|\mathbf{Z}_1-\mathbf{Z}_2\right|. \tag{3.13}$$

Proof. For all $\left\{\mathbf{Y}^{\ell}\right\}_{\ell=1}^{L}$ with $\max_{\ell=1,2,\dots,L}\left\{|\mathbf{Y}^{\ell}|\right\}<r$ and $\mathbf{Z}\in\mathbb{R}^N$, we have

$$\begin{aligned}\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}^{\ell}\big)&=\sum_{\ell=1}^{L}\frac{1}{2}|\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})|^2_{\Sigma_\eta^{\ell}}=\sum_{\ell=1}^{L}\frac{1}{2}\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)\\&\le\sum_{\ell=1}^{L}\frac{1}{2\lambda_{\min}^{\ell}}\left|\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right|^2\le\sum_{\ell=1}^{L}\frac{1}{\lambda_{\min}^{\ell}}\left(\left|\mathbf{Y}^{\ell}\right|^2+\left|\mathcal{G}^{\ell}(\mathbf{Z})\right|^2\right)\\&\le\sum_{\ell=1}^{L}\frac{1}{\lambda_{\min}}\left(r^2+C_F^2\right)=\frac{L\left(r^2+C_F^2\right)}{\lambda_{\min}}=C_{\Phi}.\end{aligned}$$

For all $\left\{\mathbf{Y}_1^{\ell}\right\}_{\ell=1}^{L}$, $\left\{\mathbf{Y}_2^{\ell}\right\}_{\ell=1}^{L}$ with $\max_{\ell=1,2,\dots,L}\left\{|\mathbf{Y}_1^{\ell}|,|\mathbf{Y}_1^{\ell}|\right\}<r$ and $\mathbf{Z}\in\mathbb{R}^N$, we have

$$\begin{aligned}&\left|\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_1^{\ell}\big)-\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_2^{\ell}\big)\right|\le\sum_{\ell=1}^{L}\left|\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_1^{\ell}\big)-\Phi^{\ell}\big(\mathbf{Z};\mathbf{Y}_2^{\ell}\big)\right|\\&=\frac{1}{2}\sum_{\ell=1}^{L}\left(\left(\mathbf{Y}_1^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}_1^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)-\left(\mathbf{Y}_2^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}_2^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z})\right)\right)\\&=\frac{1}{2}\sum_{\ell=1}^{L}\left(\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right)\\&\le\sum_{\ell=1}^{L}\frac{1}{2\lambda_{\min}^{\ell}}\left|\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right|^2\le\sum_{\ell=1}^{L}\frac{1}{2\lambda_{\min}}\left(\left|\mathbf{Y}_1^{\ell}\right|+\left|\mathbf{Y}_2^{\ell}\right|\right)\left|\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right|\\&\le\frac{r}{\lambda_{\min}}\sum_{\ell=1}^{L}\left|\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right|=C_{\Phi}\sum_{\ell=1}^{L}\left|\mathbf{Y}_1^{\ell}-\mathbf{Y}_2^{\ell}\right|.\end{aligned}$$

For all $\left\{\mathbf{Y}^{\ell}\right\}_{\ell=1}^{L}$ with $\max_{\ell=1,2,\dots,L}\left\{|\mathbf{Y}^{\ell}|\right\}<r$ and $\mathbf{Z}_1$, $\mathbf{Z}_2\in\mathbb{R}^N$, we have

$$\begin{aligned}&\left|\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z}_1;\mathbf{Y}^{\ell}\big)-\sum_{\ell=1}^{L}\Phi^{\ell}\big(\mathbf{Z}_2;\mathbf{Y}^{\ell}\big)\right|\le\sum_{\ell=1}^{L}\left|\Phi^{\ell}\big(\mathbf{Z}_1;\mathbf{Y}^{\ell}\big)-\Phi^{\ell}\big(\mathbf{Z}_2;\mathbf{Y}^{\ell}\big)\right|\\&=\frac{1}{2}\sum_{\ell=1}^{L}\left(\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z}_1)\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z}_1)\right)-\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathbf{Y}^{\ell}-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right)\right)\\&=\frac{1}{2}\sum_{\ell=1}^{L}\left(\mathcal{G}^{\ell}(\mathbf{Z}_1)-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right)^{\top}\left(\Sigma_\eta^{\ell}\right)^{-1}\left(\mathcal{G}^{\ell}(\mathbf{Z}_1)-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right)\\&\le\sum_{\ell=1}^{L}\frac{1}{2\lambda_{\min}^{\ell}}\left|\mathcal{G}^{\ell}(\mathbf{Z}_1)-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right|^2\\&\le\sum_{\ell=1}^{L}\frac{1}{2\lambda_{\min}}\left(\left|\mathcal{G}^{\ell}(\mathbf{Z}_1)\right|+\left|\mathcal{G}^{\ell}(\mathbf{Z}_2)\right|\right)\left|\mathcal{G}^{\ell}(\mathbf{Z}_1)-\mathcal{G}^{\ell}(\mathbf{Z}_2)\right|\\&\le\frac{LC_F^2}{\lambda_{\min}}\left|\mathbf{Z}_1^{\ell}-\mathbf{Z}_2^{\ell}\right|=C_{\Phi}\left|\mathbf{Z}_1^{\ell}-\mathbf{Z}_2^{\ell}\right|.\quad\square\end{aligned}$$

Inspired by the ideas in [15,16,20–22], we prove well-posedness of the posterior distribution as follows.

Theorem 3.4. *For all* $\left\{\mathbf{Y}^{\ell}\right\}_{\ell=1}^{L}$ *with* $\max_{\ell=1,2,\ldots,L}\left\{|\mathbf{Y}^{\ell}|\right\} < r$*, the posterior measure* μ_{post} *is a well-defined probability measure on* $\mathbb{R}^{N}$ *and is absolutely continuous with respect to prior measure* μ_{pr}*. Furthermore, the posterior measure* μ_{post} *is Lipschitz continuous in the observation data* $\left\{\mathbf{Y}^{\ell}\right\}_{\ell=1}^{L}$*, with respect to the Hellinger distance: if* μ_{post}^{1} *and* μ_{post}^{2} *are two posterior measures corresponding to data* $\left\{\mathbf{Y}_1^{\ell}\right\}_{\ell=1}^{L}$ *and* $\left\{\mathbf{Y}_2^{\ell}\right\}_{\ell=1}^{L}$*, then there exists a constant* $C_H = C_H(C_\Phi) > 0$ *such that,*

$$d_{Hell}(\mu_{post}^{1}, \mu_{post}^{2}) \le C_H \sum_{\ell=1}^{L} |\mathbf{Y}_1^{\ell} - \mathbf{Y}_2^{\ell}|, \tag{3.14}$$

for all $\left\{\mathbf{Y}_1^{\ell}\right\}_{\ell=1}^{L}$, $\left\{\mathbf{Y}_2^{\ell}\right\}_{\ell=1}^{L}$ *such that* $\max_{\ell=1,2,\ldots,L}\left\{|\mathbf{Y}_1^{\ell}|, |\mathbf{Y}_2^{\ell}|\right\} < r$*. Here the Hellinger distance is defined by*

$$d_{Hell}\left(\mu_1, \mu_2\right) := \sqrt{\frac{1}{2}\int\left(\sqrt{\frac{d\mu_1}{d\mu_0}} - \sqrt{\frac{d\mu_2}{d\mu_0}}\right)^2 d\mu_0}, \tag{3.15}$$

where μ_1, μ_2 *are two measures that are absolutely continuous with respect to* μ_0*.*

Proof. From (3.11), the constant $C_{\mathbf{Y}}$ in the Radon–Nikodym derivative (3.7) is bounded from below and above by

$$C_{\mathbf{Y}} = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}^{\ell}\big)\Big) d\mu_{\text{pr}}(\mathbf{Z}) \le \int_{\mathbb{R}^N} 1 d\mu_{\text{pr}}(\mathbf{Z}) = 1, \tag{3.16}$$

$$C_{\mathbf{Y}} = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}^{\ell}\big)\Big) d\mu_{\text{pr}}(\mathbf{Z}) \ge \int_{\mathbb{R}^N} e^{-C_\Phi} d\mu_{\text{pr}}(\mathbf{Z}) = e^{-C_\Phi} > 0. \tag{3.17}$$

On the other hand, one can conclude from (3.12) and (3.13) that $\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}^{\ell}\big)$ is μ_{pr}-measurable. Hence, the posterior measure μ_{post} denoted in (3.8) is well-defined.

Let $C_{\mathbf{Y}_1}, C_{\mathbf{Y}_2}$ denote the normalization constants for $\left\{\mathbf{Y}_1^{\ell}\right\}_{\ell=1}^{L}$, $\left\{\mathbf{Y}_2^{\ell}\right\}_{\ell=1}^{L}$, given by

$$C_{\mathbf{Y}_1} = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_1^{\ell}\big)\Big) d\mu_{\text{pr}}(\mathbf{Z}),$$

$$C_{\mathbf{Y}_2} = \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_2^{\ell}\big)\Big) d\mu_{\text{pr}}(\mathbf{Z}).$$

From (3.12), we have

$$\begin{aligned} |C_{\mathbf{Y}_1} - C_{\mathbf{Y}_2}| &\le \int_{\mathbb{R}^N} \left|\exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_1^{\ell}\big)\Big) - \exp\Big(-\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_2^{\ell}\big)\Big)\right| d\mu_{\text{pr}}(\mathbf{Z}) \\ &\le \int_{\mathbb{R}^N} \left|\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_1^{\ell}\big) - \sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_2^{\ell}\big)\right| d\mu_{\text{pr}}(\mathbf{Z}) \\ &\le C_\Phi \sum_{\ell=1}^{L} \left|\mathbf{Y}_1^{\ell} - \mathbf{Y}_2^{\ell}\right|. \end{aligned} \tag{3.18}$$

Recalling the definition of Hellinger distance, we have

$$\begin{aligned} &2d_{\text{Hell}}\left(\mu_1, \mu_2\right) \\ &= \int_{\mathbb{R}^N} \left(C_{\mathbf{Y}_1}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_1^{\ell}\big)\Big) - C_{\mathbf{Y}_2}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2}\sum_{\ell=1}^{L} \Phi^{\ell}\big(\mathbf{Z}; \mathbf{Y}_2^{\ell}\big)\Big)\right)^2 d\mu_{\text{pr}}(\mathbf{Z}) \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^N} \Bigg(C_{\mathbf{Y}_1}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_1^\ell\big)\Big) - C_{\mathbf{Y}_1}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Big) \\
&\qquad + C_{\mathbf{Y}_1}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Big) - C_{\mathbf{Y}_2}^{-\frac{1}{2}} \exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Big)\Bigg)^2 d\mu_{\mathrm{pr}}(\mathbf{Z}) \\
&\le H_1 + H_2,
\end{aligned} \tag{3.19}$$

where

$$H_1 = \frac{2}{C_{\mathbf{Y}_1}} \int_{\mathbb{R}^N} \Bigg(\exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_1^\ell\big)\Big) - \exp\Big(-\frac{1}{2} \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Big)\Bigg)^2 d\mu_{\mathrm{pr}}(\mathbf{Z}), \tag{3.20}$$

$$H_2 = 2\left|C_{\mathbf{Y}_1}^{-\frac{1}{2}} - C_{\mathbf{Y}_2}^{-\frac{1}{2}}\right|^2 \int_{\mathbb{R}^N} \exp\Big(-\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Big) d\mu_{\mathrm{pr}}(\mathbf{Z}). \tag{3.21}$$

From (3.12) and (3.17), we have

$$\begin{aligned}
H_1 &\le \frac{1}{2\exp(-C_\Phi)} \int_{\mathbb{R}^N} \Bigg(\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_1^\ell\big) - \sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z};\mathbf{Y}_2^\ell\big)\Bigg)^2 d\mu_{\mathrm{pr}}(\mathbf{Z}) \\
&\le \frac{1}{2\exp(-C_\Phi)} \int_{\mathbb{R}^N} \Bigg(C_\Phi \sum_{\ell=1}^{L} \left|\mathbf{Y}_1^\ell - \mathbf{Y}_2^\ell\right|\Bigg)^2 d\mu_{\mathrm{pr}}(\mathbf{Z}) \\
&= \frac{C_\Phi^2}{2\exp(-C_\Phi)} \Bigg(\sum_{\ell=1}^{L} \left|\mathbf{Y}_1^\ell - \mathbf{Y}_2^\ell\right|\Bigg)^2 .
\end{aligned} \tag{3.22}$$

From (3.11), (3.17) and (3.18), we have

$$H_2 \le \frac{2\left|C_{\mathbf{Y}_1} - C_{\mathbf{Y}_2}\right|^2}{\exp(-3C_\Phi)} \int_{\mathbb{R}^N} 1 d\mu_{\mathrm{pr}}(\mathbf{Z}) \le \frac{2C_\Phi^2}{\exp(-3C_\Phi)} \Bigg(\sum_{\ell=1}^{L} \left|\mathbf{Y}_1^\ell - \mathbf{Y}_2^\ell\right|\Bigg)^2 . \tag{3.23}$$

Combining (3.19), (3.22) and (3.23) gives the desired result (3.14). □

3.2. Markov chain Monte Carlo method

In the posterior distribution (3.5), it is challenging to give an explicit expression of the normalization constant C_z by the integration (3.6). Hence, a suitable numerical method is needed to calculate the posterior distribution. For this purpose we adopt the Markov chain Monte Carlo method (MCMC) [23–25] to generate a large number of samples subject to the posterior distribution. The numerical approximation of the posterior distribution of unknown obstacle parameters can be obtained by statistical analysis on these samples.

In this section we use the Metropolis–Hastings [40,41] algorithm to construct MCMC samples. Since the dimension N of the space of obstacle parameters is large, the Metropolis–Hastings algorithm may stay at one state for a quite long time with a huge number of iterations. In each iteration of the Metropolis–Hastings algorithm, we have to choose a proper candidate multi-dimensional sample, which however relies heavily on the dimension of obstacle parameters. To overcome this challenge we adopt the Gibbs sampling [5,27,28], especially when the dimension is conditioned only on a small number of other dimensions. In the iteration of the Gibbs sampling, samples can be chosen to be dependent on partial dimensions of obstacle parameters and are not necessarily uniform in all dimensions. Noticing that z_n, $n = 1, 2, \ldots, N$ are independent variables, the Gibbs sampling will significantly improve the convergence rate.

As done in our previous work [10], the preconditioned Crank–Nicolson (pCN) algorithm with a random proposal variance [26] can be applied to generate an ergodic Markov chain and to improve the convergence rate of the MCMC method. Using this scheme, the candidate state of the obstacle parameters $\tilde{\mathbf{Z}}$ can be iteratively updated from the current state (initial guess) $\mathbf{Z}$ through the formula

$$\tilde{\mathbf{Z}} = \mathbf{m}_{\mathrm{pr}} + (1-\beta^2)^{1/2}(\mathbf{Z} - \mathbf{m}_{\mathrm{pr}}) + \beta\omega, \tag{3.24}$$

where $\beta \in [0,1]$ is the random proposal variance coefficient and $\omega \sim \mathcal{N}(\mathbf{0}, \Sigma_{pcn})$ is a zero-mean normal random vector with the covariance matrix $\Sigma_{pcn} = \sigma_{\text{pr}}\mathbf{I} \in \mathbb{R}^{N\times N}$. The numerical algorithm is described as follows.

Algorithm 3.5 (*Gibbs Sampling with Random Proposal Variance*).

- Initialize $\mathbf{Z}_0 \in \mathbb{R}^N$ from the prior distribution $P_{\text{pr}}(\mathbf{Z})$ and $\beta_{1_0} = \beta_{2_0} = \cdots = \beta_{N_0} \in [0,1]$.
- Repeat iteration from $\mathbf{Z}_j = (z_{1_j}, z_{2_j}, \ldots, z_{N_j})^\top$ to $\mathbf{Z}_{j+1} = (z_{1_{j+1}}, z_{2_{j+1}}, \ldots, z_{N_{j+1}})^\top$, $j = 0, 1, \ldots, J_0 - 1$. For $n = 1, 2, \ldots, N$:

  1. Generate new proposal variance coefficient β_{new} from β_{n_j}. First we set

     $$\tilde{\beta} = (1-\gamma^2)^{1/2}\beta_{n_j} + \gamma(\omega_\beta - 0.5), \quad \omega_\beta \sim \mathcal{U}(0,1), \tag{3.25}$$

     with $\gamma \in [0,1]$. In our case we choose $\gamma = 0.1$. Then β_{new} can be updated by

     $$\beta_{\text{new}} = \begin{cases} \tilde{\beta}, & \text{if } \ \tilde{\beta} \in [0,1], \\ -\tilde{\beta}, & \text{if } \ \tilde{\beta} < 0, \\ \tilde{\beta} - 1, & \text{if } \ \tilde{\beta} > 1. \end{cases} \tag{3.26}$$

  2. Draw the candidate state $\tilde{\mathbf{Z}}_n = (z_{1_{j+1}}, \ldots, z_{(n-1)_{j+1}}, \tilde{z}_{n_j}, z_{(n+1)_j}, \ldots, z_{N_j})^\top$ by modifying the nth component of the current state $\mathbf{Z}_{j,n} = (z_{1_{j+1}}, \ldots, z_{(n-1)_{j+1}}, z_{n_j}, z_{(n+1)_j}, \ldots, z_{N_j})^\top$ using the pCN algorithm (3.24) with the proposal variance coefficient β_{new} as:

     $$\tilde{z}_{n_j} = m_n + (1-\beta_{\text{new}}^2)^{1/2}(z_{n_j} - m_n) + \beta_{\text{new}}\,\omega_n, \quad \omega_n \sim \mathcal{N}(0, \sigma_{\text{pr}}); \tag{3.27}$$

  3. Compute Hastings ratio $\alpha(\cdot,\cdot) : \mathbb{R}^N \times \mathbb{R}^N \to [1,\infty)$ as:

     $$\alpha\big(\mathbf{Z}_{j,n}, \tilde{\mathbf{Z}}_n\big) = \min\left\{1, \exp\left(\sum_{\ell=1}^{L} \Phi^\ell\big(\mathbf{Z}_{j,n}; \mathbf{Y}^\ell\big) - \sum_{\ell=1}^{L} \Phi^\ell\big(\tilde{\mathbf{Z}}_n; \mathbf{Y}^\ell\big)\right)\right\}; \tag{3.28}$$

  4. Accept or Reject $\tilde{\mathbf{Z}}_n$: draw $U \sim \mathcal{U}(0,1)$ and then update $\mathbf{Z}_{j,n}$ by the criterion

     $$\mathbf{Z}_{j,n+1} = \begin{cases} \tilde{\mathbf{Z}}_n, & \text{if } \ U \le \alpha\big(\tilde{\mathbf{Z}}_n, \mathbf{Z}_{j,n}\big) \\ \mathbf{Z}_{j,n}, & \text{if } \ \text{otherwise}. \end{cases} \tag{3.29}$$

  5. Update the proposal variance coefficient $\beta_{n_{j+1}}$:

     $$\beta_{n_{j+1}} = \begin{cases} \beta_{\text{new}}, & \text{if } \ \mathbf{Z}_{j,n+1} = \tilde{\mathbf{Z}}_n \\ \beta_{n_j}, & \text{if } \ \mathbf{Z}_{j,n+1} = \mathbf{Z}_{j,n}. \end{cases} \tag{3.30}$$

  6. Set $\tilde{\mathbf{Z}}_1 = (\tilde{z}_{1_j}, z_{2_j}, \ldots, z_{N_j})^\top$, $\tilde{\mathbf{Z}}_N = (z_{1_{j+1}}, z_{2_j}, \ldots, z_{(N-1)_{j+1}}, \tilde{z}_{N_j})^\top$, $\mathbf{Z}_{j,1} = \mathbf{Z}_j$, $\mathbf{Z}_{j,N} = (z_{1_{j+1}}, z_{2_{j+1}}, \ldots, z_{(N-1)_{j+1}}, z_{N_j})^\top$, $\mathbf{Z}_{j+1} = \mathbf{Z}_{j,N+1}$.

- Select $\mathbf{Z}_{\tilde{j}}$ with the indices $\tilde{j} = J_1 + (\hat{j} - 1)J_2$, $\hat{j} = 1, 2, \ldots, J_3$.

In the first step of Algorithm 3.5, the parameter γ appearing in (3.25) is used to update the proposal variance coefficient β. On the one hand, the parameter γ should be chosen large enough for converging to a suitable β with few iterations. But on the other hand, the parameter γ should also be small to ensure a small variation of β when the MCMC sequence converges. Notice that the new β updated by (3.25), (3.26), (3.30) depends on the parameter γ and the new candidate state $\tilde{\mathbf{Z}}_n$. Then we can get a suitable β, when the candidate state $\tilde{\mathbf{Z}}_n$ is accepted. Therefore, after a larger number of iterations, the proposal variance coefficient β will finally converge and vary slightly, even if we do not use an optimal γ. The choice $\gamma = 0.1$ is not the best one, but it is sufficient to give rise to accurate numerical reconstructions.

In Algorithm 3.5, the numbers $J_0, J_1, J_2, J_3 \in \mathbb{N}_0$ are four positive integers. The integer J_0 is the number of total iterations and J_1 the number of initial states which will be discarded to ensure that the Markov chain converges to the posterior distribution. By the number J_2 we mean that J_2 states are taken to guarantee the independence of the selected states. The integer J_3 denotes the number of totally selected states to approximate the posterior distribution. These four numbers (J_0, J_1, J_2, J_3) should be as large as possible, especially when the number of unknown parameters is large. But this will lead to prohibitive computational cost for large J_0, J_1, J_2, J_3. Hence, we have to balance the computational cost and the accuracy. In our numerical examples (see Section 5), we firstly set $J_2 = 100$, $J_3 = 101$, and then choose integers J_0, J_1 large enough when the percent relative error f_{PRE} (see (5.8)) between the phaseless far-field data of the reconstructed boundary and the measurement data is smaller than 1%.

4. Stochastic Surrogate model

The computational cost of each iteration in Algorithm 3.5 is dominated by the computational cost of the forward map $\mathcal{G}^\ell(\mathbf{Z})$, $\ell = 1, 2, \ldots, L$. The total computational cost T of Algorithm 3.5 is given by

$$T = T_0 N J_0, \tag{4.1}$$

where T_0 denotes the computational cost of the forward maps $\mathcal{G}^\ell(\mathbf{Z})$, $\ell = 1, 2, \ldots, L$, for L incident waves. In the previous section, Gibbs sampling and the pCN algorithm with a random proposal variance have been adopted to reduce the number of total iterations J_0. Below we discuss how to reduce the computational cost T_0. Our idea is to adopt the generalized polynomial chaos method (gPC) to construct a surrogate model for the forward map $\mathcal{G}^\ell(\mathbf{Z})$, $\ell = 1, 2, \ldots, L$.

Let $x \sim \mathcal{N}(0, 1)$ be the one-dimensional standard Gaussian variable with the distribution $P_0(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. The corresponding Gaussian space $\mathcal{L}^2_{P_0}(\mathbb{R})$ is defined as

$$\mathcal{L}^2_{P_0}(\mathbb{R}) := \left\{ f(x) : \int_{\mathbb{R}} P_0(x) |f(x)|^2 dx < \infty \right\}. \tag{4.2}$$

The normalized one-dimensional Hermite polynomials of order m are defined as

$$h_m(x) := (-1)^m e^{x^2/2} \frac{d^m}{dx^m} e^{-x^2/2}, \quad m = 0, 1, 2, \ldots, \quad x \in \mathbb{R}. \tag{4.3}$$

It is well known that the set $\{h_m(x)\}_{m=0}^\infty$ is a complete orthonormal basis of $\mathcal{L}^2_{P_0}(\mathbb{R})$ with respect to the Gaussian distribution $P_0(x)$, that is,

$$\mathbb{E}[h_n h_m] = \int_{\mathbb{R}} P_0(x) h_n(x) h_m(x) dx = \delta_{nm}, \quad n, m = 0, 1, 2, \ldots. \tag{4.4}$$

Here δ_{nm} is the Kronecker delta function.

Recall that the components of the obstacle parameters $\mathbf{Z}$ are Gaussian variables with the prior distribution $P_{pr,n}(z_n)$ given by

$$z_n \sim \mathcal{N}(m_n, \sigma_{\mathrm{pr}}), \quad P_{pr,n}(z_n) = \frac{1}{\sqrt{2\pi \sigma_{\mathrm{pr}}}} \exp\left(-\frac{1}{2\sigma_{\mathrm{pr}}} (z_n - m_n)^2 \right), \quad n = 1, 2, \ldots, N. \tag{4.5}$$

Then for each z_n, the Gaussian space $\mathcal{L}^2_{P_{pr,n}}(\mathbb{R})$ has a complete orthonormal basis $\{h_{n,m}(x)\}_{m=0}^\infty$, which can be obtained by modifying the set $\{h_m(x)\}_{m=0}^\infty$ given by (4.3). Clearly, the tensor product of the elements of $\{h_{n,m}(x)\}_{m=0}^\infty$, $n = 1, 2, \ldots, N$ forms a complete basis of the corresponding N-dimensional Gaussian probability space $\mathcal{L}^2_{P_{\mathrm{pr}}}(\mathbb{R}^N)$ with respect to the prior distribution $P_{\mathrm{pr}}(\mathbf{Z})$.

Let $\mathcal{I}$ denote the N dimensional multi-indexes:

$$\mathcal{I} := \{ \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N) : \ \alpha_n \in \mathbb{N}_0, \ |\alpha| = \sum_{n=1}^N \alpha_n < \infty \}. \tag{4.6}$$

Then a complete basis of the space $\mathcal{L}^2_{P_{\mathrm{pr}}}(\mathbb{R}^N)$ is given by the set $\{\mathcal{H}_\alpha(\mathbf{Z})\}_{\alpha \in \mathcal{I}}$, defined by

$$\mathcal{H}_\alpha(\mathbf{Z}) := \prod_{n=1}^N h_{\alpha_n}(z_n), \quad \alpha \in \mathcal{I}. \tag{4.7}$$

By properties of the operator $\mathcal{G}^\ell$, we claim that $|u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, x_\ell^{\mathrm{in}}, k)| \in \mathcal{L}^2_{P_{\mathrm{pr}}}(\mathbb{R}^N)$, $\ell = 1, 2, \ldots, L$, $m = 1, 2, \ldots, M$. By the Cameron–Martin theorem [42], the phaseless data $|u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, x_\ell^{\mathrm{in}}, k)|$ can be expanded into the series

$$|u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, x_\ell^{\mathrm{in}}, k)| := \sum_{\alpha \in \mathcal{I}} u_\alpha^{\ell,m} \ \mathcal{H}_\alpha(\mathbf{Z}), \quad \ell = 1, 2, \ldots, L, \quad m = 1, 2, \ldots, M, \tag{4.8}$$

where $u_\alpha^{\ell,m} \in \mathbb{C}$ are referred to as the chaos coefficients. From the orthogonality of the one dimensional basis $\{h_m(x)\}_{m=0}^{\infty}$, the N dimensional basis functions $\{\mathcal{H}_\alpha(\mathbf{Z})\}_{\alpha\in\mathcal{I}}$ are also orthogonal. Then we have

$$
\begin{aligned}
&\mathbb{E}\big[|u^\infty(\hat{\mathbf{x}}_m;\mathbf{Z},x_\ell^{\text{in}},k)|\mathcal{H}_{\alpha_1}(\mathbf{Z})\big] = \int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})|u^\infty(\hat{\mathbf{x}}_m;\mathbf{Z},x_\ell^{\text{in}},k)|\mathcal{H}_{\alpha_1}(\mathbf{Z})d\mathbf{Z}\\
&= \int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})\sum_{\alpha\in\mathcal{I}}\left(u_\alpha^{\ell,m}\ \mathcal{H}_\alpha(\mathbf{Z})\right)\mathcal{H}_{\alpha_1}(\mathbf{Z})d\mathbf{Z} = \sum_{\alpha\in\mathcal{I}}\left(u_\alpha^{\ell,m}\int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})\mathcal{H}_\alpha(\mathbf{Z})\mathcal{H}_{\alpha_1}(\mathbf{Z})d\mathbf{Z}\right)\\
&= u_{\alpha_1}^{\ell,m}\int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})\mathcal{H}_{\alpha_1}(\mathbf{Z})\mathcal{H}_{\alpha_1}(\mathbf{Z})d\mathbf{Z} + \sum_{\alpha\in\mathcal{I},\alpha\neq\alpha_1}\left(u_\alpha^{\ell,m}\int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})\mathcal{H}_\alpha(\mathbf{Z})\mathcal{H}_{\alpha_1}(\mathbf{Z})d\mathbf{Z}\right)\\
&= u_{\alpha_1}^{\ell,m}, \quad \ell = 1,2,\ldots,L, \quad m = 1,2,\ldots,M, \quad \forall\alpha_1\in\mathcal{I}.
\end{aligned}
$$

In other words, the chaos coefficients $u_\alpha^{\ell,m} \in \mathbb{C}$ ($\alpha\in\mathcal{I}, \ell = 1,2,\ldots,L, m = 1,2,\ldots,M$) are given by

$$
u_\alpha^{\ell,m} = \mathbb{E}\big[|u^\infty(\hat{\mathbf{x}}_m;\mathbf{Z},x_\ell^{\text{in}},k)|\mathcal{H}_\alpha(\mathbf{Z})\big] = \int_{\mathbb{R}^N} P_{\text{pr}}(\mathbf{Z})|u^\infty(\hat{\mathbf{x}}_m;\mathbf{Z},x_\ell^{\text{in}},k)|\mathcal{H}_\alpha(\mathbf{Z})d\mathbf{Z}. \tag{4.9}
$$

In this paper, we define the surrogate $\tilde{\mathcal{G}}^\ell$ of the forward operator $\mathcal{G}^\ell$, $\ell = 1,2,\ldots,L$ through the gPC approximation of order $\tilde{N}\in\mathbb{N}_0$ as

$$
\tilde{\mathcal{G}}^\ell(\mathbf{Z}) := \Big(|\tilde{u}^\infty(\hat{\mathbf{x}}_1;\mathbf{Z},x_\ell^{\text{in}},k)|,\ |\tilde{u}^\infty(\hat{\mathbf{x}}_2;\mathbf{Z},x_\ell^{\text{in}},k)|,\ \cdots,\ |\tilde{u}^\infty(\hat{\mathbf{x}}_M;\mathbf{Z},x_\ell^{\text{in}},k)|\Big)^\top, \tag{4.10}
$$

where

$$
|\tilde{u}^\infty(\hat{\mathbf{x}}_m;\mathbf{Z},x_\ell^{\text{in}},k)| = \sum_{\alpha\in\mathcal{I},|\alpha|=0}^{|\alpha|=\tilde{N}} u_\alpha^{\ell,m}\mathcal{H}_\alpha(\mathbf{Z}), \quad \ell = 1,2,\ldots,L, \quad m = 1,2,\ldots,M. \tag{4.11}
$$

We adopt the Monte Carlo method [43] to calculate the chaos coefficients $u_\alpha^{\ell,m}$ in (4.9). More details will be presented in Example 4 of the subsequent section. The surrogate model $\{\tilde{\mathcal{G}}^\ell\}_{\ell=1}^L$ given by (4.10) will be used in place of the original operator $\{\mathcal{G}^\ell\}_{\ell=1}^L$ to generate the candidate state $\tilde{\mathbf{Z}}_n$, $n = 1,2,\ldots,N$ (see (3.27)) in Algorithm 3.5. We summarize the algorithm as follows.

Algorithm 4.1 (*Gibbs Sampling with Surrogate Model*)**.** Let $j = 0,1,\ldots,J_0-1$, $n = 1,2,\ldots,N$ be given in the second step of Algorithm 3.5. We replace the formula (3.27) for generating a candidate state $\tilde{z}_{n_j}$ by the following steps:

1. Drew $\hat{J}_1$ candidate states $\tilde{\mathbf{Z}}_n^{\hat{j}}$, $\hat{j} = 1,2,\ldots,\hat{J}_1$, by the pCN algorithm (3.24) with the proposal variance coefficient β_{new}. Here the nth component of $\tilde{\mathbf{Z}}_n^{\hat{j}}$ is given by:
$$
\tilde{z}_{n_j}^{\hat{j}} = (1-\beta_{\text{new}}^2)^{1/2}z_{n_j} + \beta_{\text{new}}\ \omega_n^{\hat{j}}, \quad \omega_n^{\hat{j}} \sim \mathcal{N}(0,\sigma_{\text{pr}}), \quad \hat{j} = 1,2,\ldots,\hat{J}_1; \tag{4.12}
$$
2. Compute $\tilde{\phi}^{\hat{j}} = \sum_{\ell=1}^L \tilde{\Phi}^\ell\big(\tilde{\mathbf{Z}}_n^{\hat{j}};\mathbf{Y}^\ell\big)$, $\hat{j} = 1,2,\ldots,\hat{J}_1$. Here the function $\tilde{\Phi}^\ell$ is defined by (cf. the model-data function (3.2))
$$
\tilde{\Phi}^\ell\big(\mathbf{Z};\mathbf{Y}^\ell\big) = \frac{1}{2}|\mathbf{Y}^\ell - \tilde{\mathcal{G}}^\ell(\mathbf{Z})|_{\Sigma_\eta^\ell}^2, \quad \ell = 1,2,\ldots,L. \tag{4.13}
$$
Noting that the forward operator $\mathcal{G}^\ell$ in (3.2) has been substituted by the surrogate $\tilde{\mathcal{G}}^\ell$ in (4.13);
3. Sort the candidate state $\{\tilde{\mathbf{Z}}_n^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_1}$ that corresponds to the sequence $\{\tilde{\phi}^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_1}$ from the smallest to the largest;
4. Choose the first $\hat{J}_2$ states of the sorted candidate states $\{\tilde{\mathbf{Z}}_n^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_1}$. Then compute $\phi^{\hat{j}} = \sum_{\ell=1}^L \Phi^\ell\big(\tilde{\mathbf{Z}}_n^{\hat{j}};\mathbf{Y}^\ell\big)$, $\hat{j} = 1,2,\ldots,\hat{J}_2$, with the model-data function Φ^ℓ;
5. Sort the candidate state $\{\tilde{\mathbf{Z}}_n^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_2}$ that corresponds to the sequence $\{\phi^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_2}$ from smallest to largest;
6. Set the candidate state $\tilde{\mathbf{Z}}_n$ as the first state of the sorted candidate states $\{\tilde{\mathbf{Z}}_n^{\hat{j}}\}_{\hat{j}=1}^{\hat{J}_2}$.

We remark that, with the surrogate model, the new Algorithm 4.1 explores $\hat{J}_1$ states for each iteration in the MCMC method, whereas Algorithm 3.5 explores one state only. Let T_1 be the computational cost in calculating $\left\{\mathcal{H}_\alpha(\mathbf{Z}^{\hat{j}})\right\}_{\alpha\in\mathcal{I},|\alpha|=0}^{|\alpha|=\tilde{N}}$, $\hat{j}=1,2,\ldots,\hat{J}_1$. The computational cost of Algorithm 4.1 is

$$\hat{T}=(T_1+T_0\hat{J}_2)NJ_0. \tag{4.14}$$

Assume that there are also $\hat{J}_1$ states at each iteration in Algorithm 3.5. Without using the surrogate model $\left\{\tilde{\mathcal{G}}^\ell\right\}_{\ell=1}^L$, these $\hat{J}_1$ states can be evaluated by the model-data function Φ^ℓ given by (3.2). To further compare the previous two schemes, we rewrite Algorithm 3.5 in the form of Algorithm 4.1 as follows.

Algorithm 4.2 (*Gibbs Sampling with Multi Candidate*)**.**

1. Drew $\hat{J}_1$ candidate states $\tilde{\mathbf{Z}}_n^{\hat{j}}$, $\hat{j}=1,2,\ldots,\hat{J}_1$, by the formula (4.12);
2. Compute $\phi^{\hat{j}}=\sum\limits_{\ell=1}^{L}\Phi^\ell\left(\tilde{\mathbf{Z}}_n^{\hat{j}};\mathbf{Y}^\ell\right)$, $\hat{j}=1,2,\ldots,\hat{J}_1$, with the model-data function Φ^ℓ defined by the formula (3.2);
3. Sort the candidate state $\left\{\tilde{\mathbf{Z}}_n^{\hat{j}}\right\}_{\hat{j}=1}^{\hat{J}_1}$ corresponding to the sequence $\left\{\phi^{\hat{j}}\right\}_{\hat{j}=1}^{\hat{J}_1}$ from smallest to largest;
4. Set the candidate state $\tilde{\mathbf{Z}}_n$ by the first state of the sorted candidate states $\left\{\tilde{\mathbf{Z}}_n^{\hat{j}}\right\}_{\hat{j}=1}^{\hat{J}_1}$.

It is easy to find that Algorithm 4.2 has improved Algorithm 3.5 by using smaller integers J_0, J_1 and J_2. In particular, the number J_2 could be taken as $J_2=1$, since states $\{\mathbf{Z}_j\}_{j=J_1}^{J_0}$ are independent of each other. However, the computation cost of Algorithm 4.2, given by

$$T^{\text{multi}}=T_0\hat{J}_1NJ_0, \tag{4.15}$$

turns out to be greatly expensive. The ratio between the computational cost of Algorithms 4.1 and 4.2 is

$$R_T=\frac{\hat{T}}{T^{\text{multi}}}=\frac{(T_1+T_0\hat{J}_2)NJ_0}{T_0\hat{J}_1NJ_0}=\frac{T_1+T_0\hat{J}_2}{T_0\hat{J}_1}. \tag{4.16}$$

Compared with Algorithms 4.2, 4.1 also explores $\hat{J}_1$ states in each iteration. However, using the surrogate model $\left\{\tilde{\mathcal{G}}^\ell\right\}_{\ell=1}^L$, Algorithm 4.1 has the advantage of giving a rough estimate of these $\hat{J}_1$ states with cheaper computational cost. Noting that $\hat{J}_2$ could be smaller if the surrogate model is more accurate, and the number T_1 could decrease if the total number of basis functions of the set $\left\{\mathcal{H}_\alpha(\mathbf{Z})\right\}_{\alpha\in\mathcal{I},|\alpha|=0}^{|\alpha|=\tilde{N}}$ is smaller. In other words, we can reduce the computational cost with a smaller ratio R_T, if the surrogate model is more accurate and can be evaluated at cheaper computational cost. Furthermore, the computational cost T_0 will linearly increase with respect to L (the number of the incident waves) and M (the number of the observation directions), while the computational cost T_1 is insensitive to these parameters. This implies that, in the case of a large number of incident waves and observation directions, the total computational cost can be significantly reduced by using the surrogate model.

5. Numerical examples

In this section we exhibit numerical examples to demonstrate the effectiveness of the Bayesian method. In our former work [10], we consider inverse scattering of plane waves from a sound-soft disk with three unknown parameters, a line crack with four unknown parameters and a kite-shaped obstacle with six unknown parameters. In this paper, we shall take point source waves as incoming waves and extend the scenario to the kite-shaped obstacle with six unknowns in Example 1 and to other sound-soft obstacles with five (resp. eleven) unknowns in Example 2 (resp. Example 4). Furthermore, the gPC method will be used to reduce the computational cost for recovering an obstacle with many parameters; see Examples 3 and 4. In comparison with our previous work [10], we perform in this section additional numerics to test the sensitivity of the proposed Bayesian approach to the involved parameters. In each example, we investigate the accuracy of numerical reconstructions when the parameters L (the number of incident waves) and M (the number of observation directions) become large. In Example 1, we show the influence of the choice of the radius R of the circle where the source positions and initial guesses are located, and verify that our scheme is not sensitive to the initial guess. In Example 2, we show how to capture the localized features and micro-structures of the scattering interface by using the data excited at different wavenumbers.

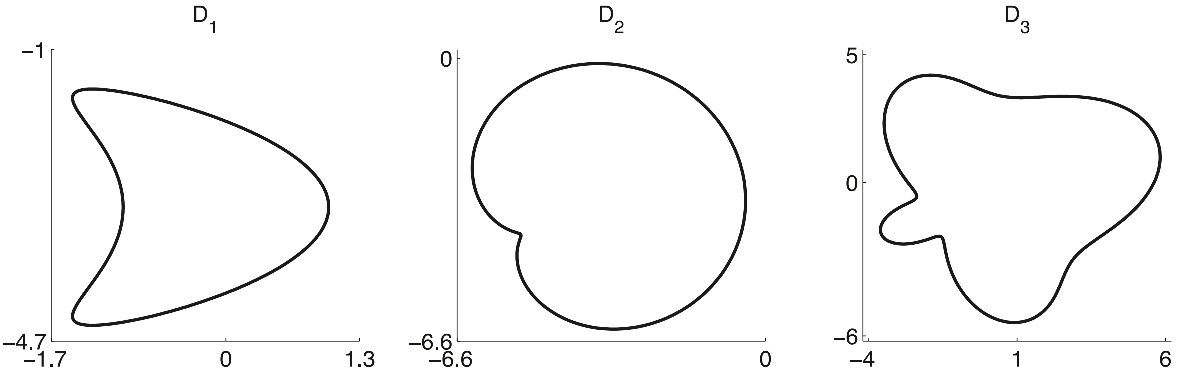


Fig. 5.1. The boundaries of three acoustically sound-soft scatterers.

Below we represent the boundaries of three acoustically sound-soft scatterers in two dimensions (see Fig. 5.1):

- The first obstacle D_1,

$$\partial D_1 = \left\{ x(t) = \left(-0.65 + \cos t + 0.65\cos(2t),\ -3 + 1.5\sin t\right)^\top, \quad 0 \le t \le 2\pi \right\}, \tag{5.1}$$

  which is the classical kite-shaped domain in inverse scattering problems.

- The second obstacle D_2,

$$\partial D_2 = \left\{ x(t) = \left(a + r(t)\cos t,\ b + r(t)\sin t\right)^\top, \quad 0 \le t \le 2\pi \right\}, \tag{5.2}$$

$$r(t) = a_0 + \sum_{n_r=1}^{N_r} a_{n_r}\cos(n_r t) + b_{n_r}\sin(n_r t), \quad 0 \le t \le 2\pi, \tag{5.3}$$

  with $N_r = 1$, $(a, b, a_0, a_1, b_1)^\top = (-5, -4, 2.5, 2, 1)^\top$. This obstacle can be regarded as a disk with a notch.

- The third obstacle D_3,

$$\partial D_3 = \left\{ x(t) = \left(a + r(t)\cos t,\ b + r(t)\sin t\right)^\top, \quad 0 \le t \le 2\pi \right\}, \tag{5.4}$$

$$r(t) = a_0 + \sum_{n_r=1}^{N_r} a_{n_r}\cos(n_r t) + b_{n_r}\sin(n_r t), \quad 0 \le t \le 2\pi. \tag{5.5}$$

  Here $N_r = 4$, $(a, b, a_0, a_1, b_1, \ldots, a_4, b_4)^\top = (-1, -1, 4, 2, 1, 0, 0, 0, 0, 0, 1)^\top$. There are 11 parameters in this scatterer. It will be used to investigate the effectiveness of the numerical method for recovering complex scatterers with high dimensional parameters.

We choose the Hausdorff distance (HD) to quantify the numerical error between the reconstructed and exact boundaries. Such a metric is frequently used in the stability estimate of various inverse acoustic, elastic and electromagnetic obstacle scattering problems. The Hausdorff distance between two obstacles ∂D_1 and ∂D_2 is defined by

$$d_H(\partial D_1, \partial D_2) := \max\left\{ \sup_{x_1\in\partial D_1} \inf_{x_2\in\partial D_2} |x_1 - x_2|, \quad \sup_{x_2\in\partial D_2} \inf_{x_1\in\partial D_1} |x_2 - x_1| \right\}. \tag{5.6}$$

The percent relative error (PRE) can be used to measure to what extent we have made use of the information of the observation data. In this paper, we denote by f_{PRE} the percent relative error (PRE) between the phaseless far-field data of the reconstructed boundary and the polluted observation. Notice that the polluted observation (2.9) is of the form of a $M \times L$ matrix, that is, $\mathbb{Y} = (\mathbf{Y}^1, \mathbf{Y}^2, \ldots, \mathbf{Y}^L) \in \mathbb{R}^M \times \mathbb{R}^L$. One can calculate the phaseless far-field data $\tilde{\mathbb{Y}} = (\tilde{\mathbf{Y}}^1, \tilde{\mathbf{Y}}^2, \ldots, \tilde{\mathbf{Y}}^L)$ of the numerical reconstruction $\tilde{\mathbf{Z}}$ by

$$\tilde{\mathbf{Y}}^\ell = \mathcal{G}^\ell(\tilde{\mathbf{Z}}), \quad \ell = 1, 2, \ldots, L, \tag{5.7}$$

where $\mathcal{G}^\ell$ denotes the forward operator mapping the space of obstacle parameters to the space of phaseless far-field pattern in the discrete sense (see (2.8)). Then the percent relative error f_{PRE} is given by

$$f_{\text{PRE}} := \|\mathbb{Y} - \tilde{\mathbb{Y}}\| / \|\mathbb{Y}\|. \tag{5.8}$$

If f_{PRE} is small enough, we can claim that we have made full use of the observation data.

Recalling the definition (2.9) of the polluted observation, we can construct two types of observations. Firstly, we consider an ideal setting where the observation data is given by the synthetic data polluted by a special sample of noise. The forward numerical solutions are obtained via the integral equation method (see e.g., [1, Chapter 3.6]) in the resonance case. The scattered field is represented as a combined double-layer and single-layer operator, and the Nyström method is used for solving the one-dimensional integral equation of second kind over $[0, 2\pi]$. Compared with the collocation and Galerkin methods, the Nyström method requires the least computational cost. In our examples, the Nyström method leads to an approximation error which decreases exponentially, because the underlying obstacles all have analytical boundaries. Therefore, the observation data in the ideal setting would contain enough information to derive an accurate numerical reconstruction. Then the accuracy of our numerical method can be examined in such an ideal setting, that is, the Hausdorff distance between the reconstructed and exact boundaries indicates whether the numerical method is accurate or not. Secondly, we consider a more practical setting where the the observation of the phaseless far-field data of the exact boundary is artificially polluted by a general sample of noise. In this case we can test the robustness of the numerical method in practical applications.

Unless otherwise stated, we always perform our numerical examples with the following assumptions.

- The wavenumber is taken as $k = 2$. Equivalently, we have the wavelength $\lambda = 2\pi/k = \pi$, which is comparable with the diameter of the underlying obstacles. We are mostly interested in this resonance case, because the forward map is extremely non-linear.
- The incident waves are excited at source positions located at

  $$x_\ell^{\text{in}} = (R\cos\theta_\ell,\ R\sin\theta_\ell)^\top, \quad \theta_\ell = 2\pi(\ell - 1)/L, \quad \ell = 1, 2, \ldots, L; \tag{5.9}$$

- The observation directions are

  $$\hat{\mathbf{x}}_m = \big(\cos\theta_m,\ \sin\theta_m\big)^\top, \quad \theta_m = -\pi + 2\pi(m-1)/M, \quad m = 1, 2, \ldots, M; \tag{5.10}$$

- For simplification, we only consider $L = M$;
- In the setting of the prior distribution P_{pr}, the initial guess is supposed to be a unit circle centered at the origin with the variance variable $\sigma_{\text{pr}} = 1$;
- Corresponding to the ℓth incident wave $u_\ell^{\text{in}}(x)$, the observation pollution $\eta^\ell = (\eta_1^\ell, \eta_2^\ell, \ldots, \eta_M^\ell)^\top$ is supposed to be a M-dimensional Gaussian variable, given by

  $$\eta_m^\ell = \sigma_\eta \times \left|u^\infty(\hat{\mathbf{x}}_m; \hat{\mathbf{Z}}, x_\ell^{\text{in}}, k)\right| \omega_m^\ell, \quad m = 1, 2, \ldots, M, \tag{5.11}$$

  where $\hat{\mathbf{Z}}$ represents the exact obstacle parameters, $\omega_m^\ell \sim \mathcal{N}(0, 1)$, $m = 1, 2, \ldots, M$, and σ_η is the noise coefficient. In other words, for $\ell = 1, 2, \ldots, L$, we take $\eta^\ell \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{\eta^\ell})$ and the diagonal matrix $\mathbf{\Sigma}_{\eta^\ell} = \text{diag}(\sigma_1^\ell, \sigma_2^\ell, \ldots, \sigma_M^\ell)$ with $\sigma_m^\ell = \left(\sigma_\eta \times \left|u^\infty(\hat{\mathbf{x}}_m; \hat{\mathbf{Z}}, x_\ell^{\text{in}}, k)\right|\right)^2$, $m = 1, 2, \ldots, M$. In our numerical tests, we choose $\sigma_\eta = 3\%, 6\%, 9\%$;
- The observation $\mathbf{Y}^\ell = (y_1^\ell, y_2^\ell, \ldots, y_M^\ell)^\top$, $\ell = 1, 2, \ldots, L$, is constructed as

  $$y_m^\ell = \left|u^\infty(\hat{\mathbf{x}}_m; \hat{\mathbf{Z}}, x_\ell^{\text{in}}, k)\right| + \eta_m^\ell, \quad m = 1, 2, \ldots, M. \tag{5.12}$$

  Here $\hat{\mathbf{Z}}$ is the exact obstacle parameters, and η_m^ℓ, $\ell = 1, 2, \ldots, L$, $m = 1, 2, \ldots, M$ is the observation noise specified in (5.11);
- In the ideal setting, the special sample of the observation noise is supposed to be $\omega_m^\ell \sim \mathcal{N}(0, 1)$ with $\omega_m^\ell = 0$, $\ell = 1, 2, \ldots, L$, $m = 1, 2, \ldots, M$, and $\sigma_\eta = 3\%$;

Table 5.1
Reconstruction of parameters and Hausdorff distance vs. $L(M = L)$.

L	Reconstructed parameters z_j	HD
10	−0.4673, −2.7488, 0.8614, 0.7405, 1.3993, 0.1154	0.1660
15	−0.6073, −2.9121, 0.9478, 0.6881, 1.4661, 0.0431	0.0431
20	−0.6179, −2.9475, 1.0013, 0.6423, 1.5169, −0.0028	0.0489
25	−0.6526, −2.9720, 0.9983, 0.6462, 1.4966, 0.0038	0.0216
40	−0.6479, −2.9935, 1.0123, 0.6473, 1.4966, −0.0028	0.0048
100	−0.6507, −2.9975, 1.0010, 0.6520, 1.4988, 0.0018	0.0007
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0		

- In the practical settings of the observations, we generate 1000 samples of the observation noise $\omega_m^\ell \sim \mathcal{N}(0, 1)$, $\ell = 1, 2, \dots, L$, $m = 1, 2, \dots, M$. This leads to 1000 samples of the noise-polluted observation data with the noise coefficient σ_η and 1000 numerical reconstructions for these samples. We perform statistical analysis of these reconstructions to demonstrate the robustness of our numerical method. For simplicity of presentation, we use E(HD) and std(HD) in the following tables to represent mean and standard deviation of a Hausdorff distance.
- The integral equation method is used for getting forward solutions. For this purpose we adopt the MATLAB code given by [5, Chapter 8];
- All calculations are performed using MATLAB R2014a on a personal laptop with a 2.29 GHz CPU and 7.90 GB RAM.

5.1. Example 1: Kite-shaped domain

In this subsection, the kite-shaped sound-soft obstacle D_1 expressed by (5.1) is supposed to be illuminated by incident waves (1.2). Obviously, the boundary of D_1 can be parameterized by

$$\mathbf{Z} := (z_1, z_2, \dots, z_6)^\top. \tag{5.13}$$

The exact parameters are $\hat{\mathbf{Z}} = (\hat{z}_1, \hat{z}_2, \dots, \hat{z}_6)^\top = (-0.65, -3, 1, 0.65, 1.5, 0)^\top$. To implement Algorithm 3.5, we choose $J_0 = 20000$, $J_1 = 10000$, $J_2 = 100$ and $J_3 = 101$.

In our former work [10], we recover this kite-shaped obstacle using the superposition of two incident plane waves, while point source waves are considered in this paper. Besides, the Gibbs sampling in Algorithm 4.2 is used to reduce the total number of iterations in the MCMC sequence. In [10], we choose $J_0 = 200000$, $J_1 = 180000$, $J_2 = 20$ and $J_3 = 1001$. Similar to the results in [10], we can obtain more accurate reconstructions with larger number of incident waves and observation directions, and get reliable reconstructions using the noise polluted observation data. Being different from [10], the surrogate model constructed by the gPC method will be used in Example 3 for improving the convergence. We use ideal observations to investigate the accuracy of our method. In this part, we will examine the effect of the choice of the parameters L, M, R and the initial guess.

Firstly, we fix R and the initial guess. Recall that the source positions are given by (5.9) with $R = 6$, and the mean of the prior distribution P_{pr} is $\mathbf{m}_{\mathrm{pr}} = (0, 0, 1, 0, 1, 0)^\top$. Note that the initial guess of the obstacle is assumed to be a unit circle centered at the origin. For different choices of L $(M = L)$, we exhibit in Table 5.1 and Fig. 5.2 the numerical reconstructions and the Hausdorff distance between the numerical and exact boundaries. We find that the reconstructed parameters are getting more accurate as the number of incident and observation directions becomes larger. The Hausdorff distances are less than 0.01 if we choose L and M large enough such as $L = M = 40{,}100$. The numerical solutions with $L = M = 15, 20$ are less accurate, since the Hausdorff distances are larger than 0.04 in these cases. The numerical solution with $L = M = 10$ turns out to be unreliable; see Fig. 5.2.

To verify that the scheme is independent of initial guess, we consider other four initial guesses by fixing $R = 6$, $L = M = 25$. We display the numerical reconstructions and the Hausdorff distances between the numerical and exact boundaries in Table 5.2. In Fig. 5.3, we plot the initial guesses, positions of the incoming point sources together with the exact and reconstructed boundaries. Observing that the Hausdorff distances are smaller than 0.03, one may conclude that the reconstructed parameters with different initial guesses are all accurate. Thus, the inversion scheme indeed does not depend on the choice of initial guess.

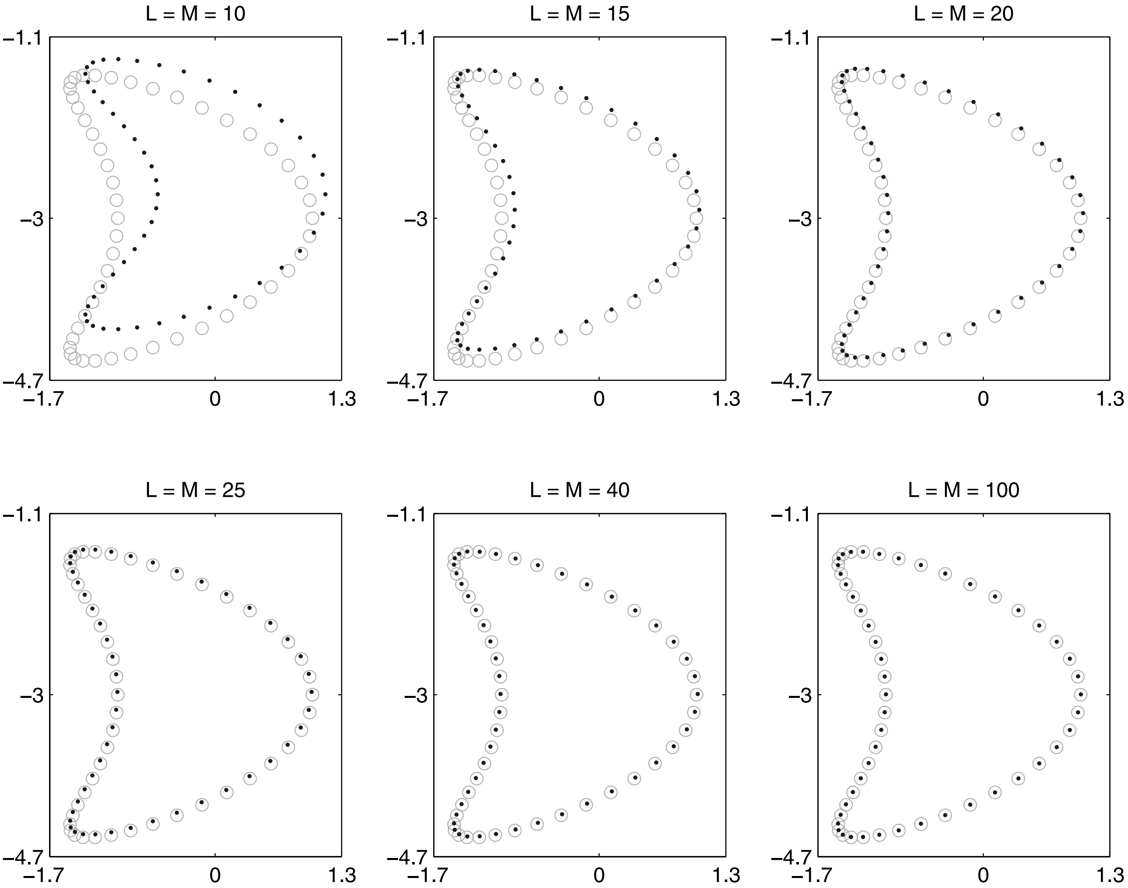


Fig. 5.2. Reconstructions of the kite-shaped obstacle at the ideal setting with different $L(M = L)$. The circle ∘ and dot · denote the exact and reconstructed boundaries, respectively.

Table 5.2
Reconstruction of parameters and Hausdorff distance vs. different initial guesses with $L = M = 25$.

Initial guess	Reconstructed parameters z_j	HD
0, 0, 1, 0, 1, 0	−0.6526, −2.9720, 0.9983, 0.6462, 1.4966, 0.0038	0.0216
1, 2, 0.5, 0, 0.5, 0	−0.6541, −2.9940, 1.0125, 0.6384, 1.5005, 0.0025	0.0287
2, −1, 1, 0, 1, 0	−0.6345, −2.9874, 1.0045, 0.6457, 1.4978, 0.0009	0.0249
−1, 3, 0.5, 0, 0.5, 0	−0.6501, −2.9880, 1.0062, 0.6450, 1.5052, −0.0009	0.0178
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0		

Secondly, with the fixed initial guess parameters $(0, 0, 0, 1, 1, 0)^\top$, we test the influence of the choice of R, $L(M = L)$, and exhibit the results in Table 5.3 and Fig. 5.4. For each fixed L, the Hausdorff distance between the numerical and exact boundaries becomes larger when we increase R. Specially, Algorithm 3.5 does not converge, as the percent relative error f_{PRE} is larger than 1% when we take $R = 20$, $L = 50$ or $R = 20$, $L = 100$. For larger R, accurate reconstructions can still be achieved if we increase the parameter L. In conclusion, more incident waves and observation directions are needed to guarantee the accuracy of reconstructions for larger R. This may be caused by the fast decay of the point source wave, as the distance between the location of source positions and the obstacle becomes larger.

Thirdly, we consider the noise-polluted phaseless far-field data. The mean and standard deviation (SD) are used to perform statistical analysis over 1000 reconstructions. The mean of these reconstructions is reported in Table 5.4,

Table 5.3
The percent relative error f_{PRE} and numerical reconstructions vs. R and $L(M = L)$.

$R, L(M = L)$	f_{PRE}	Reconstructed parameters z_j	HD
6, 25	0.37%	−0.6526, −2.9720, 0.9983, 0.6462, 1.4966, 0.0038	0.0216
8, 25	0.70%	−0.6284, −2.9585, 1.0125, 0.6309, 1.4952, −0.0066	0.0645
8, 50	0.20%	−0.6524, −2.9893, 1.0084, 0.6484, 1.5037, 0.0029	0.0167
10, 50	0.33%	−0.6209, −2.9930, 1.0007, 0.6426, 1.4988, 0.0030	0.0374
10, 100	0.12%	−0.6516, −2.9931, 0.9964, 0.6505, 1.5003, −0.0014	0.0091
20, 50	4.61%	−0.7352, −2.6130, 0.9948, 0.6466, 1.4939, 0.0017	0.4014
20, 100	1.76%	−0.6349, −2.8584, 0.9996, 0.6554, 1.4955, 0.0055	0.1484
20, 200	0.22%	−0.6520, −2.9836, 1.0002, 0.6516, 1.5051, −0.0013	0.0223
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0			

Table 5.4
Numerical solutions vs. σ_η with $L = M = 25$.

σ_η	Mean of reconstructed parameters z_j	E(HD)	std(HD)
3%	−0.6372, −2.9724, 1.0089, 0.6436, 1.4988, 0.0024	0.0229	0.0128
6%	−0.6215, −2.9442, 1.0040, 0.6463, 1.4902, 0.0116	0.0467	0.0233
9%	−0.6038, −2.9150, 0.9925, 0.6559, 1.4822, 0.0276	0.0659	0.0356
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0			

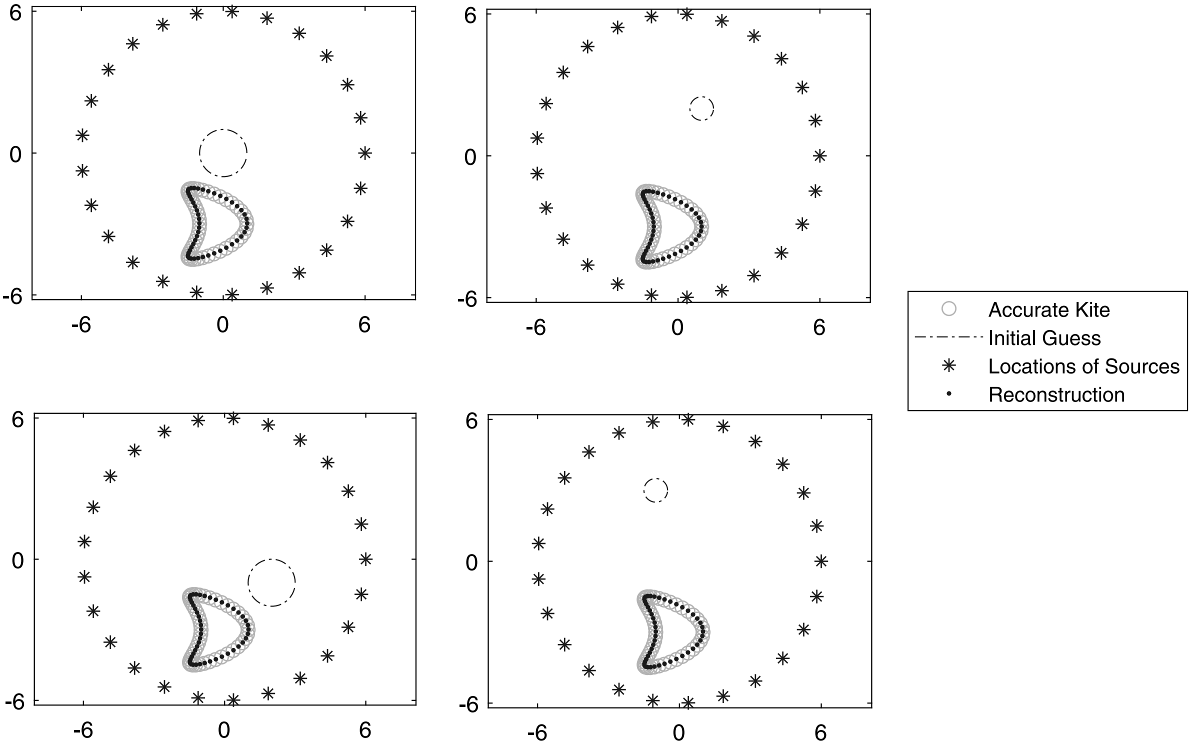


Fig. 5.3. Reconstruction of a kite-shaped obstacle at the ideal setting with $L = M = 25$ for four different initial guesses.

and the mean and standard deviation (SD) of the Hausdorff distances at different noise levels σ_η with $L = M = 25$ are described in Fig. 5.5. From Fig. 5.5, we find that the numerical reconstructions are distorted for large observation noise. However, the results in Table 5.4 show that, our approach is robust against the noise pollution, because the mean and standard deviation of the Hausdorff distances are relatively small even at the noise level of 9%. Further,

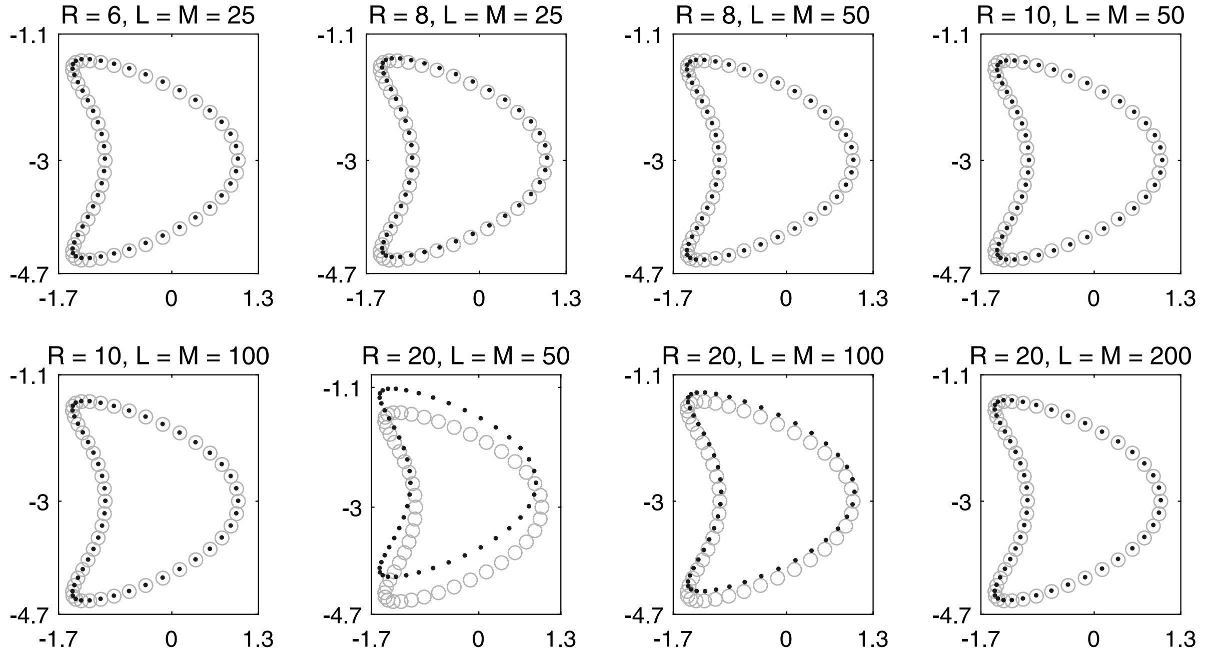


Fig. 5.4. Reconstructions of the kite-shaped obstacle at the ideal setting with different $R, L(M = L)$ and R. The circle ∘ and dot · denote the exact and reconstructed boundaries, respectively.

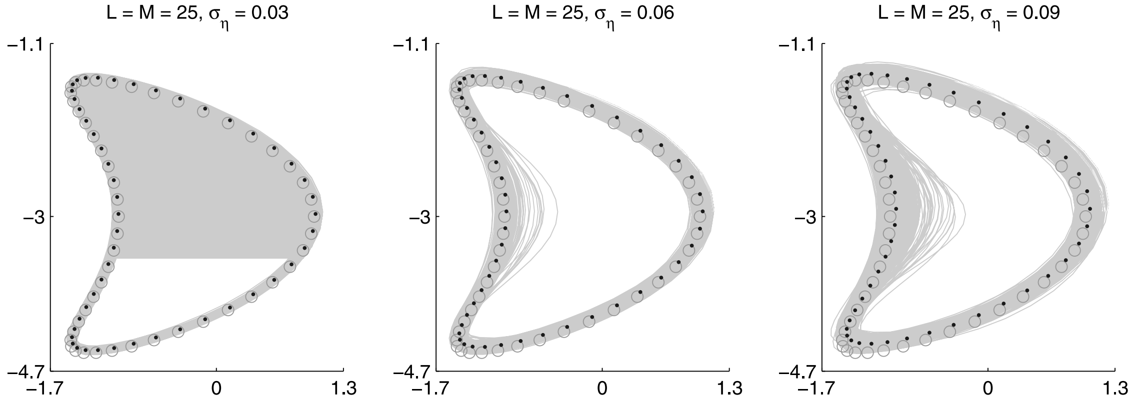


Fig. 5.5. Reconstructions of the kite-shaped obstacle with $L = M = 25$ at different noise levels $\sigma_\eta = 3\%$ (left), 6% (center), 9% (right). The circle ∘ denotes the exact boundary, the closed curves — are numerical reconstructions and the dot · is the mean of 1000 numerical reconstructions from noisy observations.

one can observe that less noise could give rise to more reliable reconstructions. We also show the numerical results with $L = M = 100$ in Fig. 5.6. From there one can claim that more reliable results can be obtained if we increase the number of observation directions in practical applications. Notice that, the numerical reconstruction is still accurate even if at the noise level of $\sigma_\eta = 0.09$.

Using (5.8), we can compute the percent relative error f_{PRE} between the noise polluted observation and the observation of the phaseless far-field data of the exact boundary. Recall from (5.11) that the phaseless far-field data are polluted by observation noise at the level σ_η. Let $\hat{\mathbb{Y}} = (\hat{\mathbf{Y}}^1, \hat{\mathbf{Y}}^2, \dots, \hat{\mathbf{Y}}^L)$ be the observation of the phaseless far-field data of the obstacle $\hat{\mathbf{Z}}$, and let $\mathbb{Y}_j = (\mathbf{Y}_j^1, \mathbf{Y}_j^2, \dots, \mathbf{Y}_j^L)$ be the 1000 samples of the noise polluted observation. Given $M = L$ and the noise level σ_η, we calculate the percent relative error $f_{\text{PRE}}^j = \|\hat{\mathbb{Y}} - \mathbb{Y}_j\| / \|\hat{\mathbb{Y}}\|$ for $j = 1, 2, \dots,$ 1000, and show their histograms in Fig. 5.7. These 1000 samples of f_{PRE} are divided into 10 intervals. In Fig. 5.7,

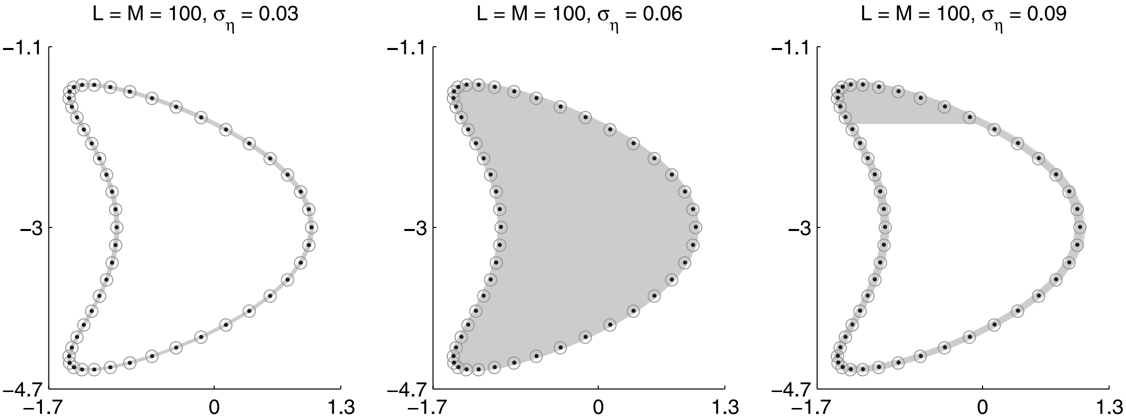


Fig. 5.6. Reconstructions of the kite-shaped obstacle with $L = M = 100$ at different noise levels $\sigma_\eta = 3\%$ (left), 6% (center), 9% (right). The circle ∘ is the exact boundary, the closed curves — are the numerical reconstructions and the dot · is the mean of 1000 numerical reconstructions from polluted observations.

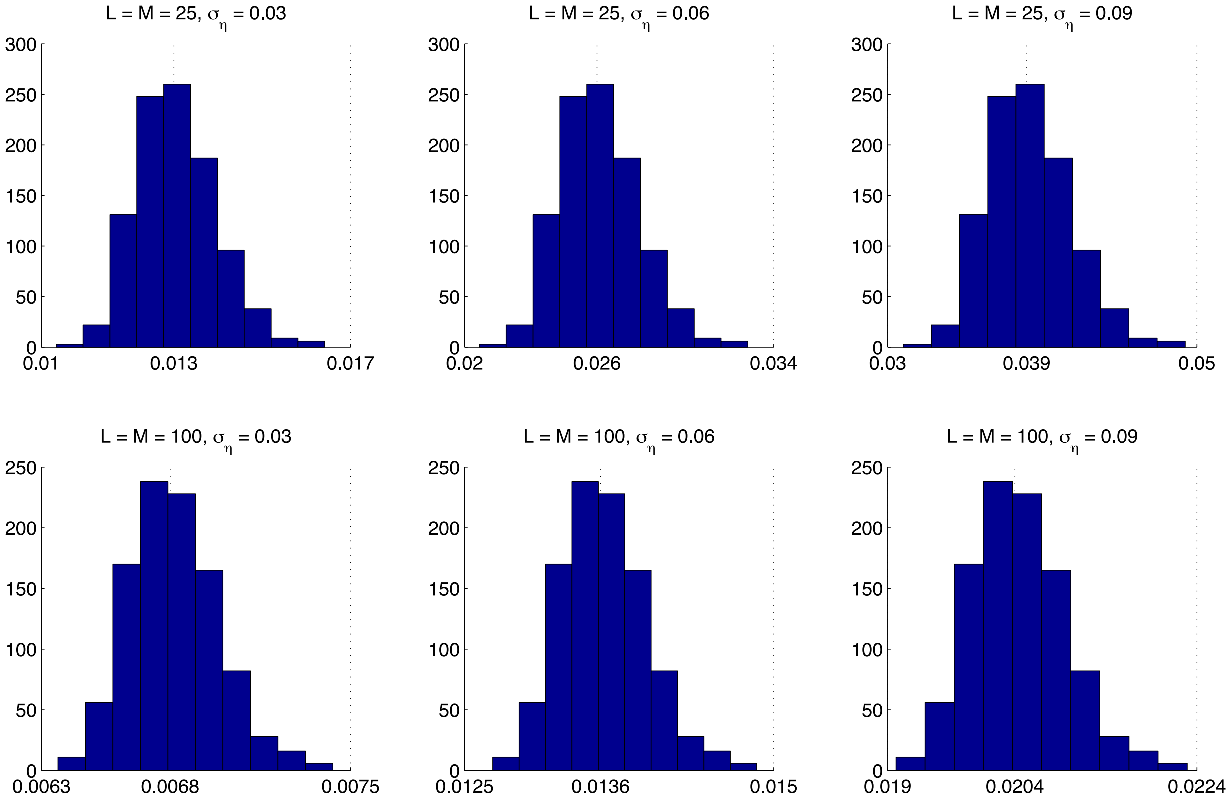


Fig. 5.7. Histograms of f_{PRE} between the noise polluted observation and the observation of the phaseless far-field data of the obstacle $\hat{\mathbf{Z}}$ with $L = M = 25$ (top), $L = M = 100$ (bottom) at $\sigma_\eta = 3\%$ (left), 6% (middle), 9% (right).

the abscissa denotes the value of f_{PRE}, and the ordinate the number of samples in each interval. We observe that $\left\{f_{\mathrm{PRE}}^{j}\right\}_{j=1}^{1000}$ decrease when the number of observations and incident directions increases, and they are extraordinarily small in comparison with the noise level σ_η. Then we consider a second type of observation noise defined by

$$\eta_m^\ell = \sigma_\eta \times \, \omega_m^\ell, \quad \omega_m^\ell \sim \mathcal{N}(0, 1), \quad \ell = 1, 2, \ldots, L, \quad m = 1, 2, \ldots, M. \tag{5.14}$$

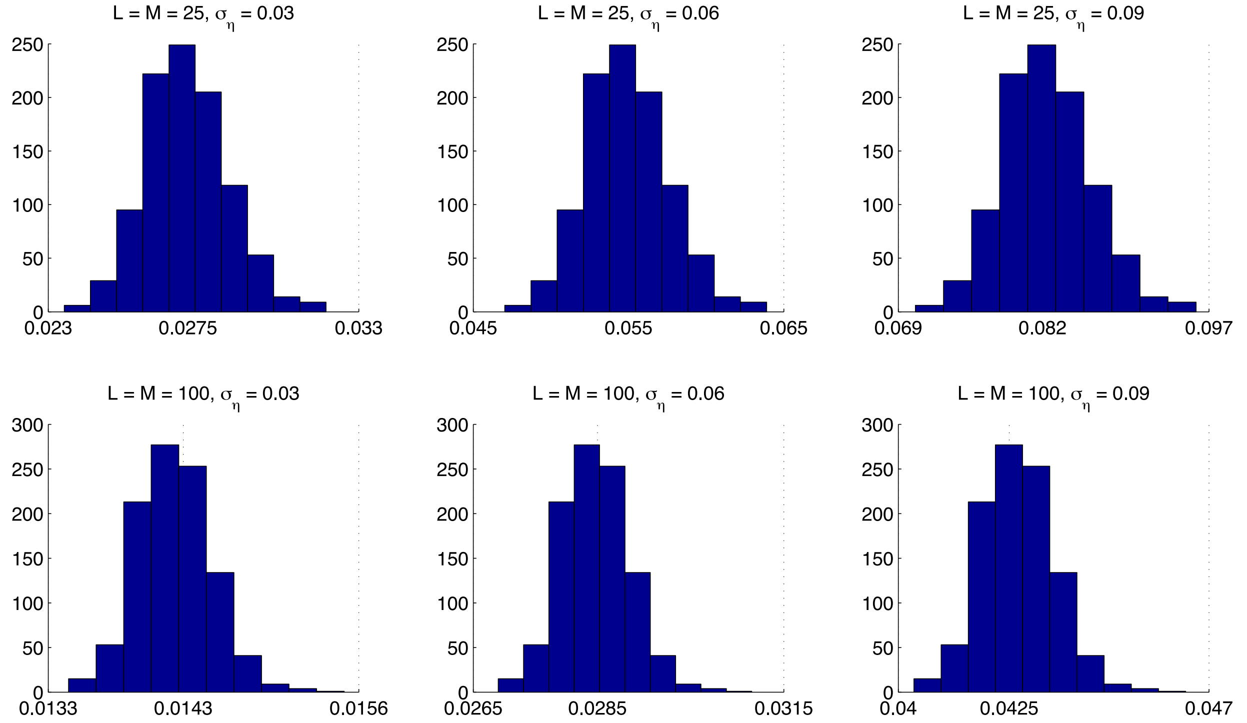


Fig. 5.8. Histograms of f_{PRE} between the noise polluted observation defined by (5.14) and the observation of the phaseless far-field data of the obstacle $\hat{\mathbf{Z}}$ with $L = M = 25$ (top), $L = M = 100$(bottom) at $\sigma_\eta = 3\%$ (left), 6% (middle), 9% (right).

The corresponding histograms of $\left\{f^{j}_{\mathrm{PRE}}\right\}_{j=1}^{1000}$ are plotted in Fig. 5.8, which are proven larger than those corresponding to the observation noise (5.11). In the case $L = M = 25$, the percent relative error is closer to the given noise level σ_η.

5.2. Example 2

The boundary of the second sound-soft obstacles (5.2)–(5.3) can be parameterized by five parameters

$$\mathbf{Z} := (z_1, z_2, \ldots, z_5)^\top = (a, b, a_0, a_1, b_1)^\top, \tag{5.15}$$

where the vector $(a, b)^\top$ indicates the position/location of this obstacle. The exact obstacle parameters are $\hat{\mathbf{Z}} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_5)^\top = (-5, -4, 2.5, 2, 1)^\top$. In this example we shall test the capability of our approach for recovering localized features (that is, the notch of ∂D_2) of a sound-soft obstacle.

We generate incident waves (1.2) by setting $R = 9$ in (5.9). Let the mean of the prior distribution P_{pr} be $\mathbf{m}_{\mathrm{pr}} = (0, 0, 1, 0, 0)^\top$. In Algorithm 3.5, we choose $J_0 = 20000$, $J_1 = 10000$, $J_2 = 100$ and $J_3 = 101$.

In the case of ideal observations, we discuss the accuracy of the numerical solutions for different choices of L ($M = L$) in Table 5.5 and Fig. 5.9. Obviously, the Hausdorff distance between the reconstructed and exact boundaries decreases as the number of incident and observation directions becomes larger. A satisfactory reconstruction of the location/position of the obstacle can be achieved even if L, M are small such as $L = M = 20$. However, an accurate recovery of the notch requires large L and M such as $L = M = 50, 70, 100$. Fig. 5.10 shows that our approach is not sensitive to the initial guess.

Now we discuss the accuracy of the numerical solutions at different wavenumbers. Since the wavenumber k is inversely proportional to the wave length, it is more difficult to extract the information of the notch at smaller wavenumbers. Hence, numerical reconstructions cannot be expected to be accurate when the wavenumber k becomes smaller; see the results shown in Table 5.6 and Fig. 5.11 with $L = M = 50$ in the ideal setting. The Hausdorff distance decreases as the wavenumber k increases. We also calculate the percent relative error f_{PRE} between the

Table 5.5
Numerical Reconstructions vs. $L(M = L)$.

L	Reconstructed parameters z_j	HD
20	−4.6808, −3.8144, 2.6498, 1.6899, 0.8117	0.1434
30	−4.7925, −3.8674, 2.6087, 1.8064, 0.9058	0.0763
40	−4.8550, −3.9437, 2.5645, 1.8497, 0.9180	0.0500
50	−4.9163, −3.9550, 2.5437, 1.8923, 0.9479	0.0194
70	−4.9581, −3.9744, 2.5227, 1.9565, 0.9842	0.0187
100	−4.9603, −3.9846, 2.5254, 1.9466, 0.9696	0.0071
Exact parameters: −5, −4, 2.5, 2, 1		

Table 5.6
The percent relative error f_{PRE} and numerical reconstructions vs. wavenumber k with $L = M = 50$.

k	f_{PRE}	Reconstructed parameters z_j	HD
0.02	0.37%	−2.1874, −2.1006, 2.6857, −1.1903, −1.1858	0.3109
0.2	0.67%	−2.4166, −1.9278, 2.6838, −0.9258, −1.3725	0.3126
0.5	1.36%	−2.7712, −2.5510, 2.8827, −0.4930, −0.5941	0.1612
1	0.43%	−4.7948, −3.8822, 2.6234, 1.7193, 0.8458	0.0399
2	0.25%	−4.9163, −3.9550, 2.5437, 1.8923, 0.9479	0.0194
5	0.16%	−4.9984, −3.9879, 2.4955, 2.0064, 1.0079	0.0105
Exact parameters: −5, −4, 2.5, 2, 1			

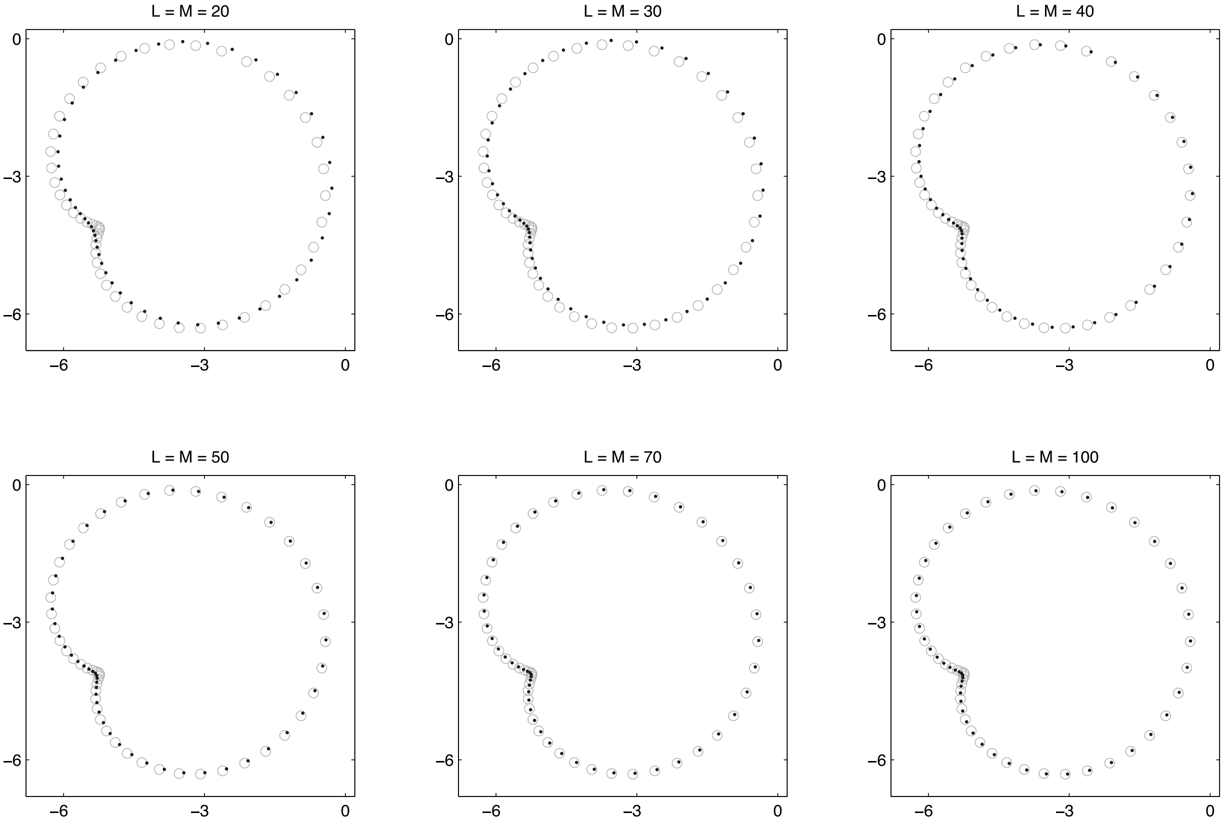


Fig. 5.9. Recovery of the second obstacle at the ideal setting with different L ($M = L$). The circle $\circ$ denotes the exact boundary and the dot $\cdot$ the reconstructions.

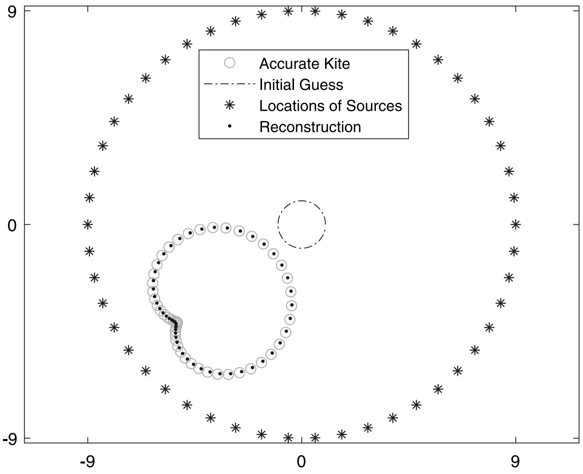


Fig. 5.10. Reconstructions of the second obstacle at the ideal setting with $L = M = 50$.

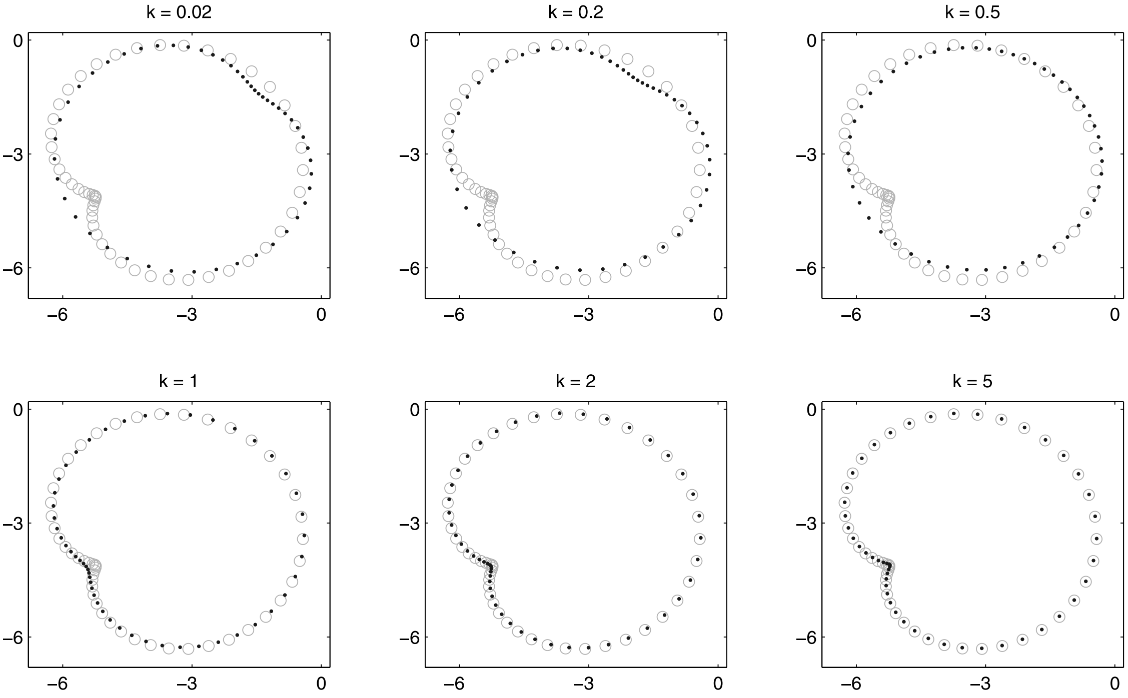


Fig. 5.11. Reconstructions of the second obstacle with different wavenumbers at the ideal setting. We set $L = M = 50$. The circle $\circ$ and dot $\cdot$ denote respectively the exact and reconstructed boundaries.

phaseless far-field data of the reconstructed and polluted observations in Table 5.6. The decreasing f_{PRE} for larger wavenumbers is consistent with the reconstructed Hausdorff distance. For the chosen wavenumbers, see Fig. 5.11, one can always get a good approximation of obstacle location, while an accurate reconstruction of the notch can be achieved only for large k. To conclude, a large k would be better for recovering the micro-structures of the obstacle.

Using polluted data, we get numerical reconstructions for 1000 samples of the observation noise. In Table 5.7 we calculate the mean of each reconstructed parameter, the mean and standard deviation (SD) of the Hausdorff distances at different noise levels σ_η with $L = M = 50$. It can be observed from Fig. 5.12 that the reconstructed boundaries may be inaccurate if the noise level is large, because the reconstruction of the notch is blurred when

Table 5.7
Numerical solutions vs. σ_η with $L = M = 50$.

σ_η	Mean of reconstructed parameters	E(HD)	std(HD)
3%	−4.9116, −3.9528, 2.5479, 1.8949, 0.9482	0.0301	0.0135
6%	−4.8488, −3.9173, 2.5739, 1.8322, 0.9177	0.0541	0.0199
9%	−4.7936, −3.8844, 2.5926, 1.7835, 0.8938	0.0702	0.0249
Exact parameters: −5, −4, 2.5, 2, 1			

Table 5.8
Numerical reconstructions vs. L ($M = L$) by the Algorithm 4.1.

L	Reconstruction of parameters z_j	HD
25	−0.6471, −2.9991, 1.0039, 0.6515, 1.4983, 0.0004	0.0084
40	−0.6514, −3.0032, 0.9935, 0.6515, 1.4981, 0.0053	0.0016
100	−0.6506, −3.0005, 1.0006, 0.6495, 1.4970, −0.0005	0.0006
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0		

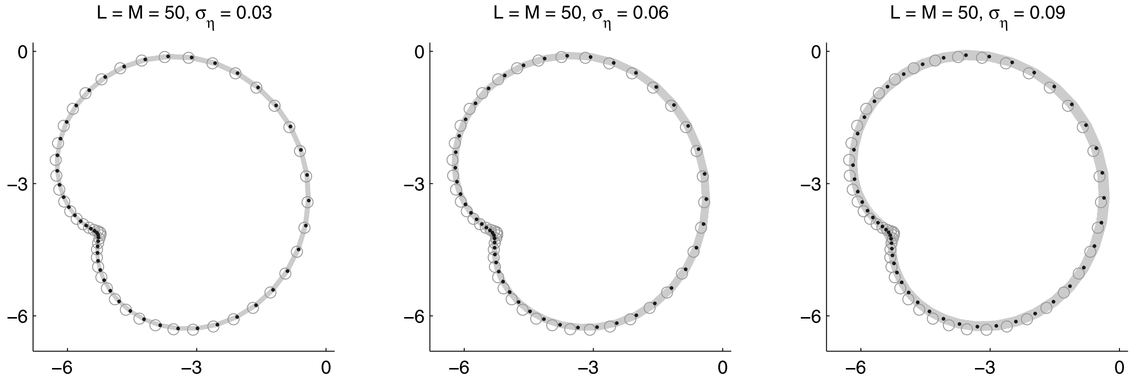


Fig. 5.12. Reconstructions of the second obstacle with $L = M = 50$ at different noise levels $\sigma_\eta = 3\%$ (left), 6% (center), 9% (right). The circle ∘ is the exact boundary, the closed curves — are numerical reconstructions and the dot · is the mean of 1000 numerical reconstructions from polluted observation data.

the noise level increases. But the numerical method is robust against the noise pollution, as the mean and standard deviation of the Hausdorff distance are small; see Table 5.7.

5.3. Example 3: gPC method for the kite-shaped domain

To demonstrate the efficiency of the surrogate model constructed by the gPC method, we apply Algorithm 4.1 to reconstruct the kite-shaped domain D_1, and compare the results with those obtained in Example 1.

Recall the exact obstacle parameters $\hat{\mathbf{Z}} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_6)^\top = (-0.65, 1, 0.65, -3, 1.5, 0)^\top$. As in previous settings we set $R = 6$ and let the mean of the prior distribution P_{pr} be $\mathbf{m}_{\mathrm{pr}} = (0, 1, 0, 0, 1, 0)^\top$. For the surrogate model, we take $\tilde{N} = 9$. To apply Algorithm 4.1, we choose $J_0 = 100$, $J_1 = 50$, $J_2 = 1$, $J_3 = 51$, $\hat{J}_1 = 1000$ and $\hat{J}_2 = 100$.

First, we need to calculate the chaos coefficients $u_\alpha^{\ell,m}$, $\alpha \in \mathcal{I}, |\alpha| = 0, 1, \ldots, \tilde{N}$, $\ell = 1, 2, \ldots, L$, $m = 1, 2, \ldots, M$. For this purpose, 6000 samples of the obstacle parameters $\mathbf{Z}$ are generated from the prior distribution $P_{\mathrm{pr}}(\mathbf{Z})$. Then, for every given L and M, the integration in Eq. (4.9) is calculated through the Monte Carlo method [43]. The tic-toc time spent on calculating the chaos coefficients with $L = M = 25$ was 1.27 min.

The numerical results with $L = M = 25, 40, 100$ are shown in Table 5.8 and Fig. 5.13 in the noise-free case. We observe that the reconstructions using the surrogate model are more accurate than those reported in Example 1; cf. Table 5.1. This is due to the fact that Algorithm 4.1 in the current Example 4 has explored 100 000 states in the iteration of the MCMC method, while Algorithm 3.5 in Example 1 only explores 20 000 states.

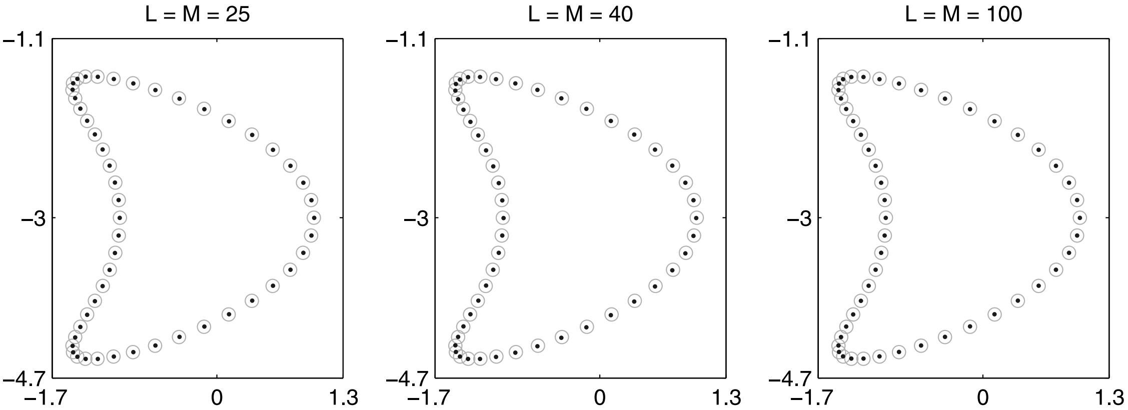


Fig. 5.13. Reconstructions of the kite-shaped obstacle by Algorithm 4.1 at the ideal setting with different L ($M = L$). The circle ∘ and dot · denote respectively the exact and reconstructed boundaries.

Table 5.9
Comparison of numerical reconstructions using Algorithms 3.5 and 4.1 with $L = M = 25$.

Method	Reconstructed parameters z_j	HD	Tic-toc time
Algorithm 3.5	−0.6445, −2.9904, 1.0079, 0.6487, 1.4965, 0.0066	0.0188	51.08 min
Algorithm 4.1	−0.6471, −2.9991, 1.0039, 0.6515, 1.4983, 0.0004	0.0084	14.34 min
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0			

Table 5.10
Tic-toc time of Algorithms 3.5 and 4.1 with $L = M = 25$.

	Algorithm 3.5	Algorithm 4.1
Total time	51.08 min	14.34 min
Calculate forward solutions online	50.91 min	2.59 min
Calculate surrogate model online	0	10.50 min
Generate surrogate model offline	0	1.21 min
Others	0.17 min	0.04 min

Then we implement Algorithm 3.5 at $L = M = 25$ by exploring 100 000 states, with the results shown in Table 5.9. Comparing the results in Tables 5.9 and 5.1, we find that more accurate reconstructions can be achieved by exploring more states, since the Hausdorff distance decreases from 0.0216 (see Table 5.1 with $L = M = 25$) to 0.0188 (see Table 5.9). But the resulting Hausdorff distance 0.0188 is still much larger than the reconstructed distance 0.0084 using the surrogate model. On the other hand, the computational cost with the surrogate model (14.34 min) is much cheaper than that in Algorithm 3.5 (51.088 min). It is worth noting that the tic-toc time of 14.34 min also includes the computational cost (1.21 min) for calculating the chaos coefficients. The detailed computational cost of each part of Algorithms 3.5 and 4.1 is shown in Table 5.10. From the second column in Table 5.10, we observe that the calculation of forward solutions takes up almost the whole computational cost of Algorithm 3.5. Algorithm 4.1 is able to reduce computation cost, because it spends much less time in calculating forward solutions; cf. the second and third columns in Table 5.10. In addition, we observe that computing the surrogate model takes up most of the computational cost in Algorithm 4.1. This suggests that we could further improve this method by using a more efficient surrogate model. In summary, using the surrogate model constructed by the gPC method, we have indeed reduced the computational cost.

Finally, the numerical results from the noisy data of the level σ_η are shown in Table 5.11 and Fig. 5.14. As in Example 1, the numerical method is robust against the noise pollution, and the phaseless data with less noise give rise to a more reliable reconstruction result. Furthermore, with much cheaper computational cost, the mean and standard deviation (SD) of Hausdorff distances in Table 5.11 are smaller than the corresponding results in Table 5.4. Hence, the surrogate model has improved the numerical results reported in the first example.

Table 5.11
Numerical solutions vs. σ_η with $L = M = 25$ by Algorithm 4.1.

σ_η	Mean of reconstructed parameters z_j	E(HD)	std(HD)
3%	−0.6499, −3.0000, 0.9999, 0.6497, 1.4999, −0.0001	0.0076	0.0062
6%	−0.6500, −3.0003, 0.9991, 0.6497, 1.4997, −0.0001	0.0158	0.0130
9%	−0.6499, −3.0003, 0.9988, 0.6496, 1.4997, −0.0001	0.0238	0.0197
Exact parameters: −0.65, −3, 1, 0.65, 1.5, 0			

Table 5.12
Numerical reconstructions vs. $L(M = L)$.

L	Reconstructed parameters z_j	HD
20	−0.790, −1.543, 3.923, 1.821, 1.317, −0.079, 0.148, −0.089, 0.094, −0.033, 0.961	0.1495
30	−0.986, −1.329, 3.941, 2.003, 1.365, 0.027, −0.063, −0.166, −0.009, −0.047, 0.997	0.0688
40	−0.966, −1.251, 3.904, 2.014, 1.224, 0.020, −0.039, −0.090, −0.035, −0.028, 1.003	0.0453
60	−1.075, −1.131, 3.962, 2.070, 1.099, −0.021, 0.037, −0.046, −0.044, −0.002, 1.031	0.0100
80	−1.047, −1.051, 3.963, 2.084, 1.042, −0.007, 0.014, −0.036, −0.039, 0.005, 1.028	0.0074
100	−0.980, −1.058, 3.988, 1.981, 1.080, 0.022, −0.023, −0.027, 0.005, −0.010, 1.011	0.0039
Exact parameters: −1, −1, 4, 2, 1, 0, 0, 0, 0, 0, 1		

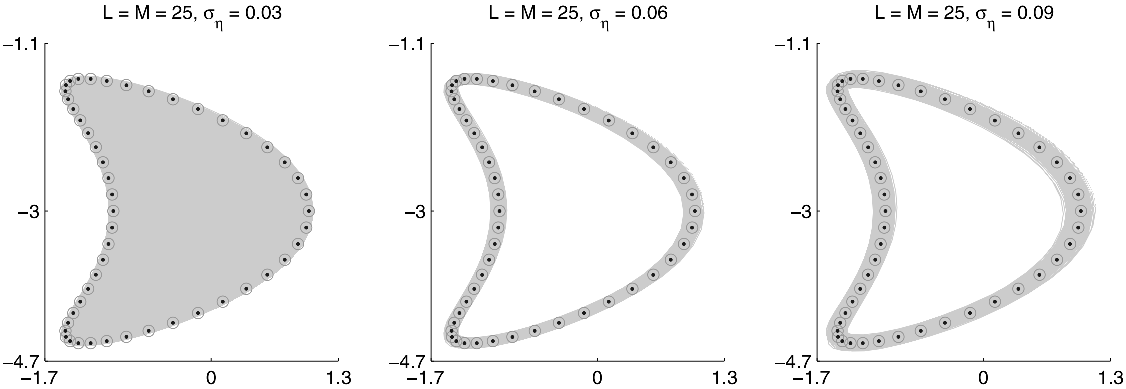


Fig. 5.14. Reconstructions of the kite-shaped obstacle by Algorithm 4.1 with $L = M = 25$ at different noise levels $\sigma_\eta = 3\%$ (left), 6% (center), 9% (right). The circle ∘ is the exact boundary, the closed curves — are the numerical reconstructions from 1000 samples of the observation noise and the dot · is the mean of these 1000 reconstructions.

5.4. Example 4

At last, we apply Algorithms 3.5 and 4.1 to the third obstacles (5.4)–(5.5), whose boundary consists of more parameters than the previous ones. The boundary of the third obstacles (5.4)–(5.5) can be parameterized by the following 11 parameters

$$\mathbf{Z} := (z_1, z_2, \ldots, z)^\top = (a, b, a_0, a_1, b_1, \ldots, a_4, b_4)^\top, \tag{5.16}$$

and the exact obstacle parameters are $\hat{\mathbf{Z}} = (\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_{11})^\top = (-1, -1, 4, 2, 1, 0, 0, 0, 0, 0, 1)^\top$.

We generate incident point source wave (1.2) with $R = 8$ in (5.9). The mean of the prior distribution P_{pr} is set to be $\mathbf{m}_{\mathrm{pr}} = (0, 0, 1, 0, \ldots, 0)^\top$.

To implement Algorithm 3.5, we choose $J_0 = 50000$, $J_1 = 40000$, $J_2 = 100$ and $J_3 = 101$. In the noise-free case, the accuracy of the numerical solutions for different choice of $L(M = L)$ is shown in Table 5.12 and Fig. 5.15. Although there are 11 unknown parameters, an accurate reconstruction can be obtained if we have enough input and output information such as $L = M = 60, 80, 100$. In the noisy case, as we have done in the previous examples, the mean and standard deviation are again used to statistically analyze 1000 reconstructions, which correspond to 1000 samples of the observation noise; see Table 5.13 and Fig. 5.16. As observed in the previous examples, the

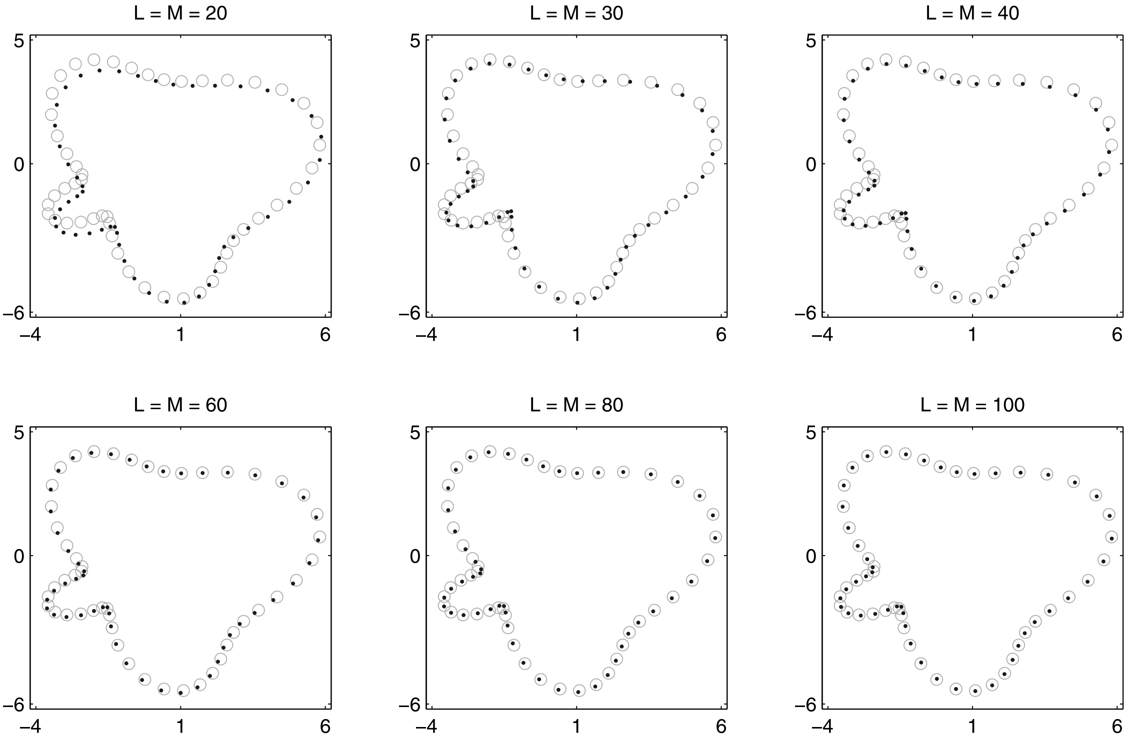


Fig. 5.15. Reconstructions of the third obstacle at the ideal setting with different L $(M = L)$. The circle ∘ denotes the exact boundary and the dot · denotes the reconstructed boundary.

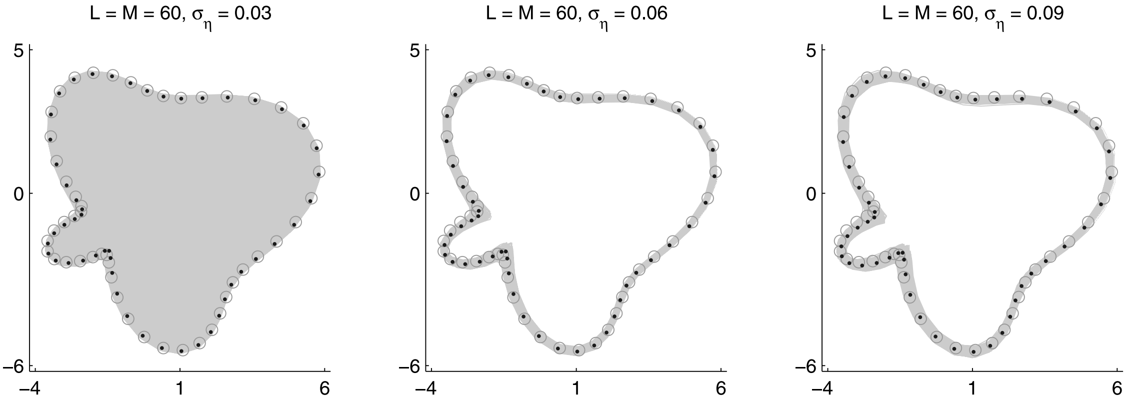
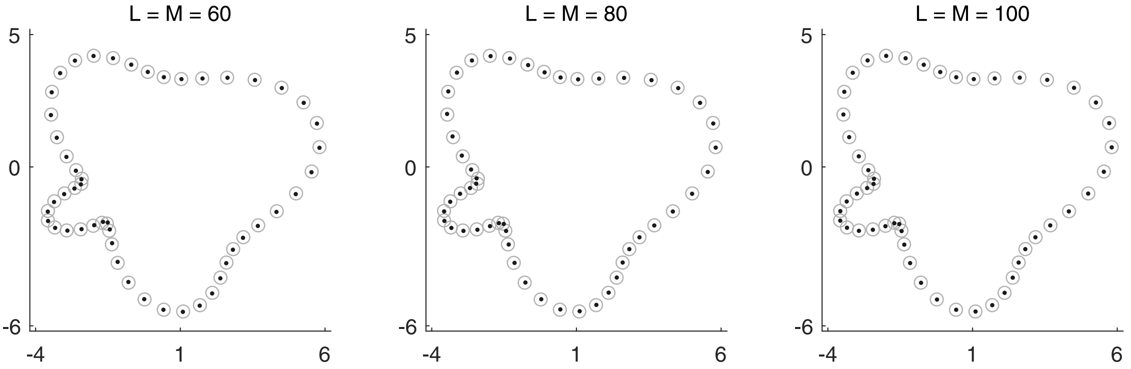


Fig. 5.16. Reconstructions of the third obstacle with $L = M = 60$ at different noise levels $\sigma_\eta = 3\%$ (left), 6% (center), 9% (right). The circle ∘ is the exact boundary, the closed curves — are numerical reconstructions and the dot · is the mean of 100 numerical reconstructions from polluted data.

numerical reconstruction may be inaccurate if the observation noise is large in Fig. 5.16. However, the numerical method is still robust against the noise pollution, because the mean and standard deviation of the Hausdorff distance are small.

To apply Algorithm 4.1, we choose $J_0 = 150$, $J_1 = 100$, $J_2 = 1$, $J_3 = 51$, $\hat{J}_1 = 1000$ and $\hat{J}_2 = 100$. For the surrogate model, we take $\tilde{N} = 3$. To calculate the chaos coefficients $u_\alpha^{\ell,m}$, $\alpha \in \mathcal{I}$, $|\alpha| = 0, 1, \ldots, \tilde{N}$, $\ell = 1, 2, \ldots, L$, $m = 1, 2, \ldots, M$, 7000 samples of the obstacle parameters $\mathbf{Z}$ are generated from the prior distribution $P_{\text{pr}}(\mathbf{Z})$.

In the noise-free case, the numerical results with $L = M = 60, 80, 100$ are shown in Table 5.14 and Fig. 5.17. Comparing the results in Tables 5.12 and 5.14, we also find that the reconstructions using the surrogate model are


L = M = 60

L = M = 80

L = M = 100

Fig. 5.17. Reconstructions of the kite-shaped obstacle by Algorithm 4.1 at the ideal setting with different L ($M = L$). The circle $\circ$ and dot $\cdot$ denote respectively the exact and reconstructed boundaries.

Table 5.13
Numerical solutions vs. σ_η with $L = M = 60$.

σ_η	Mean of reconstructed parameters z_j	E(HD)	std(HD)
3%	−1.02, −1.10, 3.96, 2.04, 1.11, 0.02, −0.01, −0.05, −0.03, −0.01, 1.03	0.0233	0.0161
6%	−1.02, −1.14, 3.94, 2.04, 1.14, 0.01, 0.001, −0.06, −0.03, −0.01, 1.03	0.0319	0.0213
9%	−1.00, −1.20, 3.94, 2.01, 1.17, −0.001, 0.01, −0.06, −0.02, −0.02, 1.02	0.0401	0.0256
Exact parameters: −1, −1, 4, 2, 1, 0, 0, 0, 0, 0, 1			

Table 5.14
Numerical reconstructions vs. L ($M = L$) by Algorithm 4.1.

L	Reconstruction of parameters z_j	HD
60	−1.025, −1.025, 3.987, 2.037, 1.032, 0.002, −0.002, −0.020, −0.016, 0.003, 1.018	0.0028
80	−0.975, −1.014, 4.004, 1.966, 1.018, 0.010, −0.008, −0.003, 0.013, −0.005, 0.992	0.0019
100	−0.992, −1.008, 4.002, 1.990, 1.004, −0.003, 0.003, 0.001, 0.003, −0.004, 0.999	0.0002
Exact parameters: −1, −1, 4, 2, 1, 0, 0, 0, 0, 0, 1		

Table 5.15
Comparison of numerical reconstructions using Algorithms 3.5 and 4.1 with $L = M = 60$.

Algorithm	Reconstructed parameters z_j	HD
Algorithm 3.5	−1.055, −1.076, 3.954, 2.097, 1.081, 0.002, 0.014, −0.043, −0.050, −0.002, 1.030	0.0067
Algorithm 4.1	−1.025, −1.025, 3.987, 2.037, 1.032, 0.002, −0.002, −0.020, −0.016, 0.003, 1.018	0.0028
Exact parameters: −1, −1, 4, 2, 1, 0, 0, 0, 0, 0, 1		

more accurate, since Algorithms 4.1 and 3.5 have explored 150 000 and 50 000 states in the iteration of the MCMC method, respectively.

Then we implement Algorithm 3.5 at $L = M = 60$ by exploring 150 000 states, compare the results with those of Algorithm 4.1 in Table 5.15, and exhibit the computational costs of Algorithms 3.5 and 4.1 in Table 5.16. As the kite-shaped obstacle case, owing to fewer times spent on calculating forward solutions online, more accurate reconstructions can be obtained with lower computational cost by Algorithm 4.1. The results using the polluted observation data are omitted, since they are quite similar to those given earlier for the kite-shaped obstacle.

6. Conclusion

In this paper, we apply the Bayesian approach to inverse time-harmonic scattering problems of recovering sound-soft obstacles from the phaseless far-field data excited by point source waves. The well-posedness of the posterior measure was discussed. Special attention has been paid to complex obstacles with high-dimensional parameters.

Table 5.16
Tic-toc time of Algorithms 3.5 and 4.1 with $L = M = 60$.

	Algorithm 3.5	Algorithm 4.1
Total time	383.77 min	36.20 min
Calculate forward solutions online	380.84 min	24.97 min
Calculate surrogate model online	0	5.31 min
Generate surrogate model offline	0	2.66 min
Others	2.93 min	3.26 min

When the dimension of unknown parameters becomes larger, both the computational cost of the forward model and the number of iteration steps in the MCMC method would increase sharply, giving rise to prohibitively high cost in the MCMC simulation. We adopt the Gibbs method and the stochastic surrogate model based on the generalized polynomial chaos method to overcome this challenge. We also develop a strategy to combine the stochastic surrogate model with the MCMC method. In our numerical examples, the efficiency of these schemes is demonstrated without sacrificing too much accuracy. Our future efforts will be devoted to recovering physical properties (e.g., refractive index) of a penetrable acoustically scatterer, which contains more parameters than its geometrical shape as discussed here. In this paper, the total number of boundary parameters is assumed to be known in advance. Removing or relaxing this assumption would also lead to a high-dimensional inverse scattering problem, and more efficient schemes are needed in designing the MCMC method and the reduced model. It would also be interesting and important to investigate the convergence of the proposed algorithm. Research outcomes along these directions will be reported in our forthcoming publications.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

G. Hu is supported by the National Natural Science Foundation of China (No. 12071236) and the Fundamental Research Funds for Central Universities in China.

References

[1] D. Colton, R. Kress, Inverse Acoustic and Electromagnetic Scattering Theory, Vol. 93, third ed., Springer, 2013.
[2] F. Cakoni, D. Colton, A Qualitative Approach to Inverse Scattering Theory, Vol. 188, Springer Science & Business Media, 2014.
[3] A. Kirsch, N. Grinberg, The Factorization Method for Inverse Problems.
[4] A. Kirsch, An Introduction to the Mathematical Theory of Inverse Problems, Vol. 120, Springer Science & Business Media, 2011.
[5] G. Nakamura, R. Potthast, Inverse Modeling: An Introduction to the Theory and Methods of Inverse Problems and Data Assimilation, IOP Publishing, Bristol UK, 2015.
[6] R. Kress, W. Rundell, Inverse obstacle scattering with modulus of the far field pattern as data, in: H.W. Engl (Ed.), Inverse Problems in Medical Imaging and Nondestructive Testing, Springer, Wien, 1997, pp. 75–92.
[7] X. Xu, B. Zhang, H. Zhang, Uniqueness in inverse scattering problems with phaseless far-field data at a fixed frequency, SIAM J. Appl. Math. 78 (3) (2018) 1737–1753.
[8] B. Zhang, H. Zhang, Fast imaging of scattering obstacles from phaseless far-field measurements at a fixed frequency, Inverse Problems 34 (10) (2018) 104005.
[9] D. Zhang, Y. Guo, Uniqueness results on phaseless inverse acoustic scattering with a reference ball, Inverse Problems 34 (8) (2018) 085002.
[10] Z. Yang, X. Gui, J. Ming, G. Hu, Bayesian approach to inverse time-harmonic acoustic scattering with phaseless far-field data, Inverse Problems 36 (6) (2020) 065012.
[11] M.V. Klibanov, Phaseless inverse scattering problems in three dimensions, SIAM J. Appl. Math. 74 (2) (2014) 392–410.
[12] M.V. Klibanov, A phaseless inverse scattering problem for the 3-d Helmholtz equation, Inverse Probl. Imaging 11 (2) (2017) 263–276.
[13] O. Ivanyshyn, Shape reconstruction of acoustic obstacles from the modulus of the far field pattern, Inverse Probl. Imaging 1 (4) (2007) 609–622.
[14] O. Ivanyshyn, R. Kress, Identification of sound-soft 3D obstacles from phaseless data, Inverse Probl. Imaging 4 (1) (2010) 131–149.
[15] T. Bui-Thanh, O. Ghattas, An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation, SIAM/ASA J. Uncertain. Quantif. 2 (1) (2014) 203–222.

[16] A.M. Stuart, Inverse problems: a Bayesian perspective, Acta Numer. 19 (2010) 451–559.
[17] J. Kaipio, E. Somersalo, Statistical and Computational Inverse Problems, Vol. 160, Springer Science & Business Media, 2006.
[18] K. Law, A.M. Stuart, K. Zygalakis, Data Assimilation: A Mathematical Introduction, in: Texts in Applied Mathematics, vol. 62, Springer International Publishing, 2015.
[19] M.A. Iglesias, K. Lin, A.M. Stuart, Well-posed Bayesian geometric inverse problems arising in subsurface flow, Inverse Problems 30 (11) (2014) 114001.
[20] A. Baussard, D. Prémel, O. Venard, A Bayesian approach for solving inverse scattering from microwave laboratory-controlled data, Inverse Problems 17 (6) (2001) 1659.
[21] J. Liu, Y. Liu, J. Sun, An inverse medium problem using Stekloff eigenvalues and a Bayesian approach, Inverse Problems 35 (9) (2019) 094004.
[22] Y. Wang, F. Ma, E. Zheng, Bayesian method for shape reconstruction in the inverse interior scattering problem, Math. Probl. Eng. (2015).
[23] S. Brooks, A. Gelman, G. Jones, X. Meng, Handbook of Markov Chain Monte Carlo, CRC press, 2011.
[24] D. Gamerman, H.F. Lopes, Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, second ed., Chapman and Hall/CRC, 2006.
[25] C.J. Geyer, Practical Markov chain Monte Carlo, Statist. Sci. 7 (4) (1992) 473–483.
[26] S.L. Cotter, G.O. Roberts, A.M. Stuart, D. White, MCMC methods for functions: Modifying old algorithms to make them faster, Statist. Sci. 28 (3) (2013) 424–446.
[27] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, IEEE Trans. Pattern Anal. Mach. Intell. (6) (1984) 721–741.
[28] J.S. Liu, Monte Carlo Strategies in Scientific Computing, Springer Science & Business Media, 2008.
[29] X. Ma, N. Zabaras, An efficient Bayesian inference approach to inverse problems based on an adaptive sparse grid collocation method, Inverse Problems 25 (3) (2009) 035013.
[30] Y. Marzouk, D. Xiu, A stochastic collocation approach to Bayesian inference in inverse problems, 2009.
[31] L. Yan, L. Guo, Stochastic collocation algorithms using l_1-minimization for bayesian solution of inverse problems, SIAM J. Sci. Comput. 37 (3) (2015) A1410–A1435.
[32] G. Zhang, D. Lu, M. Ye, M. Gunzburger, C. Webster, An adaptive sparse-grid high-order stochastic collocation method for Bayesian inference in groundwater reactive transport modeling, Water Resour. Res. 49 (10) (2013) 6871–6892.
[33] Y.M. Marzouk, H.N. Najm, Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, J. Comput. Phys. 228 (6) (2009) 1862–1902.
[34] T. Cui, Y.M. Marzouk, K.E. Willcox, Data-driven model reduction for the Bayesian solution of inverse problems, Internat. J. Numer. Methods Engrg. 102 (5) (2015) 966–990.
[35] Q. Liao, J. Li, An adaptive reduced basis ANOVA method for high-dimensional Bayesian inverse problems, J. Comput. Phys. 396 (2019) 364–380.
[36] F. Lu, M. Morzfeld, X. Tu, A.J. Chorin, Limitations of polynomial chaos expansions in the Bayesian solution of inverse problems, J. Comput. Phys. 282 (2015) 138–147.
[37] J. Li, Y.M. Marzouk, Adaptive construction of surrogates for the Bayesian solution of inverse problems, SIAM J. Sci. Comput. 36 (3) (2014) A1163–A1186.
[38] L. Yan, T. Zhou, Adaptive multi-fidelity polynomial chaos approach to Bayesian inference in inverse problems, J. Comput. Phys. 381 (2019) 110–128.
[39] D. Rick, Probability: Theory and Examples, fourth ed., Cambridge university press, 2010.
[40] W.K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika 57 (1) (1970) 97–109.
[41] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. 21 (6) (1953) 1087–1092.
[42] R.H. Cameron, W.T. Martin, The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals, Ann. of Math. 48 (2) (1947) 385–392.
[43] C.P. Robert, G. Casella, Monte Carlo Statistical Methods, second ed., Springer, New York, 2004.