

PAPER

## Bayesian approach to inverse time-harmonic acoustic scattering with phaseless far-field data

To cite this article: Zhipeng Yang *et al* 2020 *Inverse Problems* **36** 065012

View the [article online](#) for updates and enhancements.



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Bayesian approach to inverse time-harmonic acoustic scattering with phaseless far-field data

Zhipeng Yang<sup>1</sup> , Xinping Gui<sup>1</sup>, Ju Ming<sup>2</sup> and Guanghui Hu<sup>3,4</sup> 

<sup>1</sup> Beijing Computational Science Research Center, Beijing 100193, People's Republic of China

<sup>2</sup> School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China

<sup>3</sup> School of Mathematical Sciences, Nankai University, Tianjin 300071, People's Republic of China

E-mail: [yangzhp@csrc.ac.cn](mailto:yangzhp@csrc.ac.cn), [gui@csrc.ac.cn](mailto:gui@csrc.ac.cn), [jming@hust.edu.cn](mailto:jming@hust.edu.cn) and [ghhu@nankai.edu.cn](mailto:ghhu@nankai.edu.cn)

Received 30 July 2019, revised 18 February 2020

Accepted for publication 24 March 2020

Published 25 May 2020



## Abstract

This paper is concerned with inverse acoustic scattering problem of inferring the position and shape of a sound-soft obstacle from phaseless far-field data. We propose the Bayesian approach to recover sound-soft disks, line cracks and kite-shaped obstacles through properly chosen incoming waves in two dimensions. Given the Gaussian prior measure, the well-posedness of the posterior measure in the Bayesian approach is discussed. The Markov Chain Monte Carlo (MCMC) method is adopted in the numerical approximation and the preconditioned Crank–Nicolson (pCN) algorithm with random proposal variance is utilized to improve the convergence rate. Numerical examples are provided to illustrate effectiveness of the proposed method.

Keywords: inverse scattering problem, phaseless far-field pattern, Bayesian inference, MCMC

(Some figures may appear in colour only in the online journal)

## 1. Introduction

Time-harmonic inverse scattering problems have attracted extensive attention due to their numerous applications in many areas such as radar and sonar detection, geophysical prospection, medical imaging, nondestructive testing and so on. In this paper, we are interested in the

<sup>4</sup> Author to whom any correspondence should be addressed.

inverse problem of reconstructing the location and shape of an acoustically sound-soft obstacle using phaseless far-field data.

The propagation of a time-harmonic incident field  $u^{\text{in}}$  in a homogeneous and isotropic medium is governed by the Helmholtz equation

$$\Delta u^{\text{in}} + k^2 u^{\text{in}} = 0 \text{ in } \mathbb{R}^2, \quad (1.1)$$

where  $k > 0$  is the wavenumber. Let  $D \subset \mathbb{R}^2$  be a sound-soft scatterer, which occupies a bounded subset with  $C^2$ -smooth boundary  $\partial D$  such that the exterior  $\mathbb{R}^2 \setminus \bar{D}$  of  $D$  is connected. In this paper  $D$  maybe a domain or a curve, which represents an extended obstacle or a crack in acoustics. The forward scattering problem is to find the scattered (perturbed) field  $u^{\text{sc}}$  to the Helmholtz equation

$$\Delta u^{\text{sc}} + k^2 u^{\text{sc}} = 0 \text{ in } \mathbb{R}^2 \setminus \bar{D}, \quad (1.2)$$

which satisfies the Dirichlet boundary condition

$$u^{\text{sc}} = -u^{\text{in}} \text{ on } \partial D, \quad (1.3)$$

and the Sommerfeld radiation condition

$$\lim_{r \rightarrow \infty} \sqrt{r} \left( \frac{\partial u^{\text{sc}}}{\partial r} - iku^{\text{sc}} \right) = 0, \quad r = |x|, \quad (1.4)$$

uniformly in all directions  $\hat{x} = x/|x| \in \mathbb{S} := \{x : |x| = 1\}$ ,  $x \in \mathbb{R}^2 \setminus \bar{D}$ , and  $i = \sqrt{-1}$  is the imaginary unit. The total field  $u$  is defined as  $u = u^{\text{in}} + u^{\text{sc}}$  in  $\mathbb{R}^2 \setminus \bar{D}$ . The Sommerfeld radiating solution  $u^{\text{sc}}$  has an asymptotic behavior of the form

$$u^{\text{sc}}(x) = \frac{e^{ik|x|}}{\sqrt{|x|}} \left\{ u^\infty(\hat{x}) + O\left(\frac{1}{\sqrt{|x|}}\right) \right\}, \quad |x| \rightarrow \infty, \quad (1.5)$$

where  $u^\infty(\hat{x})$  is called the far-field pattern at the observation direction  $\hat{x} \in \mathbb{S}$ . Note that  $u^\infty : \mathbb{S} \rightarrow \mathbb{C}$  is an analytic function with phase information. The above model also appears in the TE polarization of time-harmonic electromagnetic scattering from infinitely long and perfectly conducting cylinders.

The uniqueness, stability and inversion algorithms for recovering  $\partial D$  from phased far-field patterns have been extensively studied with one or many incoming plane and point source waves. We refer to the monographs [6, 8, 23, 24, 37] for historical remarks, an overview of recent progresses and the comparison between different approaches. In many practical applications, the phase information of the far-field pattern cannot be measured accurately compared with its modulus or intensity. For instance, in optics it is not trivial to measure the phase of electromagnetic waves incited at high frequencies. One of the essential difficulties in using phaseless far-field data lies in the translation invariance property for plane wave incidence, which we state as follows. Let  $u^\infty(\hat{x}; D, \mathbf{d})$  be the far-field pattern corresponding to the incident plane wave  $e^{ikx \cdot \mathbf{d}}$  ( $\mathbf{d} \in \mathbb{S}$  is the incident direction) and the sound-soft obstacle  $D$ . For the shifted obstacle  $D_z := \{x + z : x \in D\}$ , the corresponding far-field pattern is given by (see [29])

$$u^\infty(\hat{x}; D_z, \mathbf{d}) = e^{ikz \cdot (d - \hat{x})} u^\infty(\hat{x}; D, \mathbf{d}) \quad \text{for all } \hat{x} \in \mathbb{S}. \quad (1.6)$$

Hence, we get

$$|u^\infty(\hat{x}; D_z, \mathbf{d})| = |u^\infty(\hat{x}; D, \mathbf{d})| \quad \text{for all } \mathbf{d}, \hat{x} \in \mathbb{S}.$$

This implies that it is impossible to recover the location of  $D$  from the phaseless far-field pattern of a plane wave.

There has been tremendous interest in inverse scattering with phaseless data or in phase retrieval problems in optics and other physical and engineering areas (see, e.g. [1, 2, 12, 17, 20, 25, 26, 34, 35] and the references therein). In a deterministic setting where randomness are not taken into account, Kress & Rundell and Ivanyshyn & Kress proposed a Newton-type iterative approach to reconstruct the shape of sound-soft obstacles from only the modulus of the far-field pattern in [17, 20, 29]. The approach of [17, 20] was based on a pair of nonlinear and ill-posed integral equations motivated by an inverse boundary value problem for the Laplace equation with phase information [18, 30]; see also [12, 19, 28] for inverse scattering from sound-soft cracks using a single far-field pattern with phase or phaseless information. Klibanov proved unique determination of a compactly supported potential of the stationary three-dimensional Schrödinger equation from the phaseless near-field data incited by an interval of frequencies [25]. This was later extended in [26] to the reconstruction of a smooth wave speed in the three-dimensional Helmholtz equation. To broke the translation invariance property, it was recently proposed in [41] that, phaseless far-field patterns generated by infinitely many sets of superpositions of two plane waves with different directions can be used to uniquely determine a penetrable or impenetrable scatterer; see [42] for a fast imaging algorithm based on this idea. Similar uniqueness results were derived by Zhang & Guo [43] where the superposition of a fixed plane wave and some point sources was taken as incident waves and a reference ball technique was proposed. Uniqueness and direct sampling algorithms using the superposition of plane waves and fixed source location point sources were considered in [21]. In this paper we propose to generate the phaseless data using the following superposition of two plane waves (see [41])

$$u_\ell^{\text{in}}(x) := e^{ikx \cdot d_0} + e^{ikx \cdot d_\ell}, \quad \ell = 0, 1, 2, \dots, L, \quad (1.7)$$

and then to recover a sound-soft disk, a line crack or a kite-shaped obstacle through the Bayesian approach. In (1.1), we fix  $d_0 \in \mathbb{S}$  and change  $d_\ell \in \mathbb{S}$  as incident directions, due to the *a priori* information of the obstacle; see theorem 3.1 for a uniqueness proof for sound-soft disks.

In recent years, the Bayesian method has received increasing attention for inverse problems [10, 16, 22, 31, 39], which also has been applied to the inverse scattering problems [3, 5, 14, 32, 33, 40] with phase far-field data. In particular, the authors of [5] adopt the Bayesian framework of [39] to shape identification problems in inverse scattering and establish a framework for proving well-posedness of the Bayesian formulation using a suitable shape parametrization and the regularity of shape derivatives. The aim of this paper is to propose the Bayesian method using a single far-field pattern without phase information. The Bayesian method provides us a new perspective to view the inverse scattering problem in the form of statistical inferences. In this statistical approach, all parameters are random variables and the key issue is to estimate the posterior distribution of the unknown quantities based on the Bayes' formula [39] and the known prior distribution. The Bayesian method could be an alternative method to overcome the challenges in deterministic inverse problems, although it usually leads to expensive computational cost. The advantageous over deterministic inversion schemes in inverse scattering (for example, optimization-based iterative schemes and non-iterative sampling methods) are summarized as follows. (i) Instead of deterministic reconstructions, the Bayesian approach gives rise to statistic characteristics of the posterior distribution of unknown parameters and provides a quantification of the uncertainties arising from the corresponding model predictions. (ii) The Bayesian method could lead to all possible solutions of the inverse problem. For example, using one plan wave rather than the superposition of two plane waves, the Bayesian method

could reconstruct the shape of an unknown obstacle which is located at every possible position due to the translation invariance (1.6). (iii) For inverse scattering problems, the theoretical analysis and numerical methods in the Bayesian framework are only based on the deterministic forward model. Hence, it is easy to perform theoretical analysis and numerical examples. (iv) The Bayesian method needs less measurement data without phase information and does not require a good initial guess. In this paper, the Markov chain Monte Carlo (MCMC) method [4, 11, 13] is proposed to accomplish the characterization of the posterior distribution, while the preconditioned Crank–Nicolson (pCN) algorithm [9] is adopted to improve the convergence rate in the iteration of MCMC method. Since the MCMC method and the pCN algorithm are adopted to calculate the numerical approximation, the numerical method in this paper is insensitive to the initial guess of the obstacle shape. In our numerical examples, we exhibit that accurate reconstructions can be achieved when the number of incident waves and observation directions is small.

This paper is organized as follows. In section 2, we adapt the Bayesian framework to inverse scattering problems with phaseless data. In section 3, we exhibit numerical results for recovering a disk, a line crack and a kite-shaped obstacle. Conclusions are given in section 4.

## 2. Bayesian framework

In this paper we want to recover an unknown sound-soft obstacle from phaseless far-field patterns corresponding to a set of superposition of two plane waves. We propose the Bayesian approach to solve this inverse scattering problem. First of all, we set a suitable parameterization of the position and the shape of an obstacle. Then, in the Bayesian framework, we estimate the posterior distribution of the unknown obstacle parameters. By the Bayes' theorem [31, 39], the posterior distribution of these parameters can be obtained from the prior distribution and the likelihood function to be specified in this section. Numerically, we will adopt the Markov chain Monte Carlo method (MCMC) to get an approximation of the posterior distribution.

### 2.1. Parameterization of the obstacle

Since the boundary of the underlying obstacle is a  $C^2$ -smooth curve, we can represent or approximate its geometrical shape by a finite set  $\mathbf{Z}$  of variables

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

For example, we can use four parameters  $\mathbf{Z} := (z_1, z_2, z_3, z_4)^\top$  to represent a line segment, where  $(z_1, z_2)^\top$  and  $(z_3, z_4)^\top$  denote respectively the two ending points, or we can use  $\mathbf{Z} := (a_1, b_1, a_2, b_2, \dots, a_N, b_N)^\top$  to approximate a star-shaped closed curve where  $\{(a_j, b_j) : j = 1, \dots, N\}$  stand for the Fourier coefficients in the truncated Fourier expansion.

Recalling the incident waves  $u_\ell^{\text{in}}(x)$ ,  $\ell = 1, 2, \dots, L$  in the form of a set of superpositions of two plane waves (1.7), we express the (phased) far-field patterns of the scattering model (1.2)–(1.4) by

$$u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k), \quad \ell = 1, 2, \dots, L, \quad \hat{\mathbf{x}} \in \mathbb{S}. \quad (2.2)$$

Correspondingly, the phaseless far-field patterns are denoted by

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k)|, \quad \ell = 1, 2, \dots, L, \quad \hat{\mathbf{x}} \in \mathbb{S}, \quad (2.3)$$

where  $|\cdot|$  is the modulus of a complex number.

## 2.2. Prior distribution

By (2.1), the prior distribution of the obstacle parameters  $\mathbf{Z}$  depends on the distribution of  $z_n, n = 1, 2, \dots, N$ . Let  $\{z_n\}_{n=1}^N$  be independent variables with the prior density  $\pi_{\text{pr}}^n$  and prior measure  $\mu_{\text{pr}}^n$ . Then the prior density  $\pi_{\text{pr}}$  and prior measure  $\mu_{\text{pr}}$  of  $\mathbf{Z}$  are respectively given by

$$\pi_{\text{pr}}(\mathbf{Z}) = \prod_{n=1}^N \pi_{\text{pr}}^n(z_n), \quad (2.4)$$

$$\mu_{\text{pr}}(d\mathbf{Z}) = \prod_{n=1}^N \mu_{\text{pr}}^n(dz_n). \quad (2.5)$$

In this paper we assume that  $z_n$  are random variables with the Gaussian distribution, that is,

$$\mu_{\text{pr}}^n = \mathcal{N}(m_n, \sigma_n), \quad n = 1, 2, \dots, N.$$

For simplicity, we assume that  $\sigma_1 = \dots = \sigma_N = \sigma_{\text{pr}}$ , implying that  $\mu_{\text{pr}} = N(\mathbf{m}_{\text{pr}}, \sigma_{\text{pr}}\mathbf{I})$ , where  $\mathbf{m}_{\text{pr}} = (m_1, m_2, \dots, m_N)^\top$  and  $\mathbf{I} \in \mathbb{R}^{N \times N}$  is the identity matrix.

## 2.3. Observation of far-field pattern

To bridge the parameterization (2.1) of the obstacle and the associated phaseless far-field data (2.3), we define an operator  $F : \mathbb{R}^N \rightarrow C^\infty(\mathbb{S})$  as

$$|u^\infty(\hat{\mathbf{x}})| = F(\mathbf{Z}), \quad \hat{\mathbf{x}} \in \mathbb{S}, \quad (2.6)$$

which can be regarded an abstract map from the space of obstacle parameters to the space of observation data in the continuous sense. From the well-posedness of forward scattering,  $F$  is continuous but highly non-linear.

Let  $G = (g_1, g_2, \dots, g_M)^\top : C^\infty(\mathbb{S}) \rightarrow \mathbb{R}^M$  be a bounded linear observation operator with  $g_m : C^\infty(\mathbb{S}) \rightarrow \mathbb{R}_+$  given by

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

where  $\{\hat{\mathbf{x}}_m\}_{m=1}^M \subset \mathbb{S}$  is the set of discrete observation directions. Then the observation at the observation direction  $\hat{\mathbf{x}}_m$  can be rephrased as

$$y_m = g_m(|u^\infty(\hat{\mathbf{x}})|) + \eta_m = |u^\infty(\hat{\mathbf{x}}_m)| + \eta_m, \quad (2.7)$$

where  $\eta_m$  represents the noise polluting the observation data at the direction  $\hat{\mathbf{x}}_m$ .

Set  $\mathbf{Y} = (y_1, y_2, \dots, y_M)^\top \in \mathbb{R}^M$  and  $\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_M)^\top \in \mathbb{R}^M$ . Denote by  $G = G \circ F$  the map from the obstacle parameter space  $\mathbb{R}^N$  to observation space  $\mathbb{R}^M$ , that is,

$$\mathbf{Y} = G(\mathbf{Z}) + \boldsymbol{\eta}, \quad \mathbf{Y}, \boldsymbol{\eta} \in \mathbb{R}^M, \quad \mathbf{Z} \in \mathbb{R}^N. \quad (2.8)$$

Our inverse problem in this paper is to determine the obstacle parameters  $\mathbf{Z} \in \mathbb{R}^N$  from the observation data  $\mathbf{Y} \in \mathbb{R}^M$  with the noise pollution  $\boldsymbol{\eta} \in \mathbb{R}^M$ .

## 2.4. Likelihood

We assume the observation pollution  $\boldsymbol{\eta}$  is independent of  $u^\infty$  and drawn from the Gaussian distribution  $N(\mathbf{0}, \Sigma_\eta)$  with the density  $\rho$ , where  $\Sigma_\eta \in \mathbb{R}^{M \times M}$  is a self-adjoint positive matrix.

By the observation of the phaseless data with noise (2.8), we can get the relationship  $\mathbf{Y}|\mathbf{Z} \sim N(G(\mathbf{Z}), \Sigma_\eta)$ . Define the model-data misfit function  $\Phi(\mathbf{Z}; \mathbf{Y}) : \mathbb{R}^N \times \mathbb{R}^M \rightarrow \mathbb{R}$  as

$$\Phi(\mathbf{Z}; \mathbf{Y}) = \frac{1}{2} |\mathbf{Y} - G(\mathbf{Z})|_{\Sigma_\eta}^2, \quad (2.9)$$

where  $|\cdot|_{\Sigma_\eta} = |\Sigma_\eta^{-\frac{1}{2}} \cdot|$ . Hence, the likelihood function is given by

$$\rho(\mathbf{Y} - G(\mathbf{Z})) = \frac{1}{((2\pi)^M \det(\Sigma_\eta))^{1/2}} e^{-\Phi(\mathbf{Z}; \mathbf{Y})}.$$

Furthermore, the posterior density  $\pi_{\text{post}}$  and the posterior measure  $\mu_{\text{post}}$  are connected to the prior measure  $\mu_{\text{pr}}$  through the Radon–Nikodym derivative [38], given by

$$\frac{d\mu_{\text{post}}}{d\mu_{\text{pr}}}(\mathbf{Z}) \propto e^{-\Phi(\mathbf{Z}; \mathbf{Y})}. \quad (2.10)$$

### 2.5. Well-posedness of Bayesian framework

The well-posedness arguments of [5, 39] can be applied to deal with our inverse scattering problem with the Bayesian approach. In our phaseless case, we are required to justify the following assumption 2.1, relying on regularity properties of the forward operator  $G$ .

**Assumption 2.1.** The map  $G : \mathbb{R}^N \rightarrow \mathbb{R}^M$  satisfies

- (a) For every  $\varepsilon > 0$ , there is an  $\hat{M} = \hat{M}(\varepsilon) \in \mathbb{R}$  such that, for all  $\mathbf{Z} \in \mathbb{R}^N$ ,

$$|\mathcal{G}(\mathbf{Z})|_{\Sigma_\eta} \leq e^{\|\mathbf{Z}\|_2^2 + \hat{M}},$$

where  $\|\cdot\|_2$  is the Euclidean norm.

- (b) For every  $r > 0$ , there is a  $K = K(r) > 0$  such that, for all  $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathbb{R}^N$  with  $\max\{\|\mathbf{Z}_1\|_2, \|\mathbf{Z}_2\|_2\} < r$ , it holds that

$$|\mathcal{G}(\mathbf{Z}_1) - \mathcal{G}(\mathbf{Z}_2)|_{\Sigma_\eta} \leq K\|\mathbf{Z}_1 - \mathbf{Z}_2\|_2.$$

We remark that there is no essential difference in proving assumption 2.1 (a) between the phased and phaseless inverse scattering problems. The assumption 2.1 (b) follows directly from the triangle inequality  $\|a\| - \|b\| \leq \|a - b\|$  for complex numbers  $a, b \in \mathbb{C}$  and the corresponding assumption for phased inverse scattering problems. Hence, when  $D$  is sound-soft scatterer, assumption 2.1 can be proved following the phased arguments of [5]; see also [32, 40] for the proofs in the case of limited aperture data and for interior scattering problems. If  $D$  is sound-soft crack, the same results can be verified by applying the Fréchet differentiability with respect to the boundary of the far field operator ([27, 28]). The above assumptions together with the choice of the Gaussian prior measure (which satisfies  $\mu_{\text{pr}}(\mathbb{R}^N) = 1$ ) lead to well-posedness of the Bayesian inverse problem, which is a result of application of lemma 2.8, theorems 4.1, 4.2 and 6.31 in [39]. Before stating the well-posedness (see theorem 2.1 below), we recall the Hellinger distance defined by

$$d_{\text{Hell}}(\mu_1, \mu_2) := \sqrt{\frac{1}{2} \int \left( \sqrt{\frac{d\mu_1}{d\mu_0}} - \sqrt{\frac{d\mu_2}{d\mu_0}} \right)^2 d\mu_0}, \quad (2.11)$$



where  $\mu_1, \mu_2$  are two measures that are absolutely continuous with respect to  $\mu_0$ .

**Theorem 2.1.** *If the operator  $G$  satisfies the assumption 2.1 and the prior measure  $\mu_{\text{pr}}$  satisfies  $\mu_{\text{pr}}(\mathbb{R}^N) = 1$ , then the posterior measure  $\mu_{\text{post}}$  is a well-defined probability measure on  $\mathbb{R}^N$  and absolutely continuous with respect to prior measure  $\mu_{\text{pr}}$ . What's more, the posterior measure  $\mu_{\text{post}}$  is Lipschitz in the data  $\mathbf{Y}$ , with respect to the Hellinger distance: if  $\mu_{\text{post}}^1$  and  $\mu_{\text{post}}^2$  are two posterior measures corresponding to data  $\mathbf{Y}_1$  and  $\mathbf{Y}_2$ , then there exists  $C = C(r) > 0$  such that,*

$$d_{\text{Hell}}(\mu_{\text{post}}^1, \mu_{\text{post}}^2) \leq C \|\mathbf{Y}_1 - \mathbf{Y}_2\|_2,$$

for all  $\mathbf{Y}_1, \mathbf{Y}_2$  with  $\max \{\|\mathbf{Y}_1\|_2, \|\mathbf{Y}_2\|_2\} < r$ .

## 2.6. Preconditioned Crank–Nicolson (pCN) algorithm with random proposal variance

This subsection is devoted to the numerical approximation of the posterior distribution. We adopt the Markov chain Monte Carlo method (MCMC) [4, 11, 13] 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 statistic analysis on these samples. The Metropolis–Hastings [15, 36] algorithm will be used to construct MCMC samples. To improve the convergence rate of the MCMC method, we apply the preconditioned Crank–Nicolson algorithm [9].

According to the pCN algorithm, the new obstacle parameter  $\mathbf{X}$  can be iteratively updated by the old parameter (initial guess)  $\mathbf{Z}$  through the formula

$$\mathbf{X} = \mathbf{m}_{\text{pr}} + (1 - \beta^2)^{1/2}(\mathbf{Z} - \mathbf{m}_{\text{pr}}) + \beta\omega, \quad (2.12)$$

where  $\beta \in [0, 1]$  is the proposal variance coefficient and  $\omega \sim N(\mathbf{0}, \Sigma_{\text{pcn}})$  is a zero-mean normal random vector with covariance matrix  $\Sigma_{\text{pcn}} \in \mathbb{R}^{N \times N}$ . We remark that it is important and very tricky to select a suitable  $\beta$ , because the value of  $\beta$  dominates the proposal variance in the pCN algorithm. If  $\beta \ll 1$  is small, the parameter  $\mathbf{Z}$  will be updated slightly in the MCMC sequence, leading to a time-consuming iteration process to get the ergodic in the space of obstacle parameters. If  $\beta$  is big, the parameter  $\mathbf{Z}$  may stay at one state for quite a long time with a huge number of iterations in the MCMC method. Consequently, one cannot get enough number of samples to approximate the posterior distribution, due to the computational cost prohibition. To overcome this difficulty, we recommend the pCN algorithm with a random proposal variance [9] to obtain good MCMC sequences (see below for the description).

## 3. Numerical examples

In this section we exhibit numerical examples to demonstrate the effectiveness of the method described in the previous section. To save computational costs, we consider three types of acoustically sound-soft scatterers in two dimensions:

- Sound-soft disks with unknown centers and radii;
- Line cracks with unknown starting and ending points;
- Kite-shaped obstacles with unknown position and shape.

There are totally three unknown parameters for disks, four parameters for cracks and six parameters for kite-shaped obstacle in 2D, implying that the unknown parameters always lie in a finite space with low dimensions.



**Algorithm 1.** pCN algorithm with random proposal variance

- 
- Initialize  $\mathbf{Z}_0 \in \mathbb{R}^N$  and  $\beta_0 \in [0, 1]$ .
  - Repeat
    1. Draw new obstacle parameter  $\mathbf{X}$  from the old state  $\mathbf{Z}_j$  by the pCN algorithm (2.12) with the proposal variance coefficient  $\beta_j$  as:
 
$$\mathbf{X} = \mathbf{m}_{\text{pr}} + (1 - \beta_j^2)^{1/2}(\mathbf{Z}_j - \mathbf{m}_{\text{pr}}) + \beta_j \omega, \quad \omega \sim N(\mathbf{0}, \Sigma_{\text{pcn}}); \quad (2.13)$$
    2. Compute Hasting ratio  $\alpha(\cdot, \cdot) : \mathbb{R}^N \times \mathbb{R}^N \rightarrow [1, \infty)$  as:
 
$$\alpha(\mathbf{X}, \mathbf{Z}_j) = \min\{1, e^{\Phi(\mathbf{Z}_j; \mathbf{Y}) - \Phi(\mathbf{X}; \mathbf{Y})}\}; \quad (2.14)$$
    3. Accept or reject  $\mathbf{X}$ : draw  $U \sim U(0, 1)$  and then update  $\mathbf{Z}_j$  by the criterion
 
$$\mathbf{Z}_{j+1} = \begin{cases} \mathbf{X}, & \text{if } U \leq \alpha(\mathbf{X}, \mathbf{Z}_j), \\ \mathbf{Z}_j, & \text{if otherwise;} \end{cases} \quad (2.15)$$
    4. Generate new proposal variance coefficient  $\beta_{j+1}$  from  $\beta_j$ . First we set
 
$$\beta_{\text{new}} = (1 - \gamma^2)^{1/2} \beta_j + \gamma(\omega_\beta - 0.5), \quad \omega_\beta \sim U(0, 1), \quad (2.16)$$
 with  $\gamma \in [0, 1]$ . In our case we choose  $\gamma = 0.1$ . Then  $\beta_{j+1}$  can be updated by
 
$$\beta_{j+1} = \begin{cases} \beta_{\text{new}}, & \text{if } \beta_{\text{new}} \in [0, 1], \\ -\beta_{\text{new}}, & \text{if } \beta_{\text{new}} < 0, \\ \beta_{\text{new}} - 1, & \text{if } \beta_{\text{new}} > 1. \end{cases} \quad (2.17)$$
  - Select  $\mathbf{Z}_{\hat{j}}$ ,  $\hat{j} = J_1 + (\hat{j} - 1)J_2$ ,  $\hat{j} = 1, 2, \dots, J_3$ ,  $J_1, J_2, J_3 \in \mathbb{N}_0$ .
- 

In the algorithm 1, the randomness of the proposal variance has the potential advantage of including the possibility of large and small proposal variance coefficients  $\beta$ . The large proposal variance coefficient  $\beta$  helps the pCN algorithm explore the state space efficiently, and the small proposal variance coefficient  $\beta$  protects the iterations from dropping into a fixed state. Then the random proposal variance gives rise to ergodic Markov chains. The sequence  $\mathbf{Z}_j$  in the algorithm 1 is selected to approximate the posterior distribution. Here,  $J_1$  is the number of initial states; we take every  $J_2$  states to guarantee the selected states are independent;  $J_3$  is the number of the total selected states.

With the definition of the observation (2.8), we construct two types of observations. Since the observation noise  $\eta$  in (2.8) is assumed to be a Gaussian random vector, the zero vector is a special sample of the observation noise. Then in the first type observations, we consider an ideal model where the phaseless far-field data corresponding to the exact obstacle are polluted by this special sample ( $\eta = \mathbf{0}$ ) of observation noise. In the second type observations, we consider a practical model with noise-polluted far-field data, which will be used to discuss the robustness of the numerical method in practical applications. The Hausdorff difference between the reconstructed and exact scatterers indicates the accuracy of our numerical method.

### 3.1. Disk

In this subsection, we assume the underlying obstacle is a sound-soft disk. The inverse problem of recovering disks arises from, for instance, the polarization model of time-harmonic electromagnetic scattering from perfectly conducting cylinders whose cross-section is a disk. The parameterization of a disk is given by

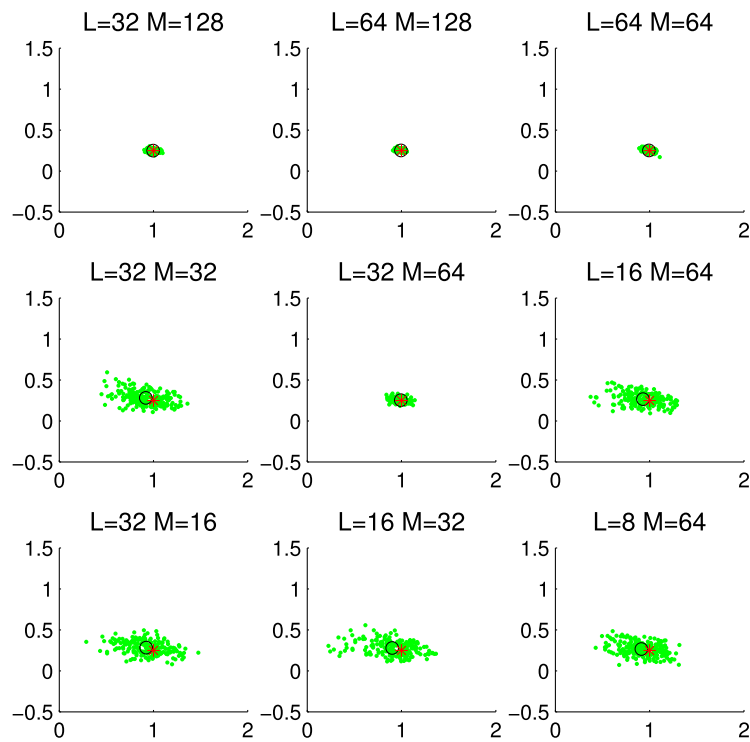
$$\mathbf{Z} := (z_1, z_2, z_3)^\top = (x_1, x_2, \log r)^\top, \quad (3.1)$$

where  $(x_1, x_2)^\top$  is the center and  $r$  is the radius of the disk. Since  $r > 0$ , we assume that  $r$  is a lognormal random variable, i.e.,  $z_3 = \log r$  is a Gaussian random variable.

Let the incident wave be given by the sum of two plane waves of the form (1.7). If the disk is located at the origin, it is well-known that the corresponding far-field pattern incited by the

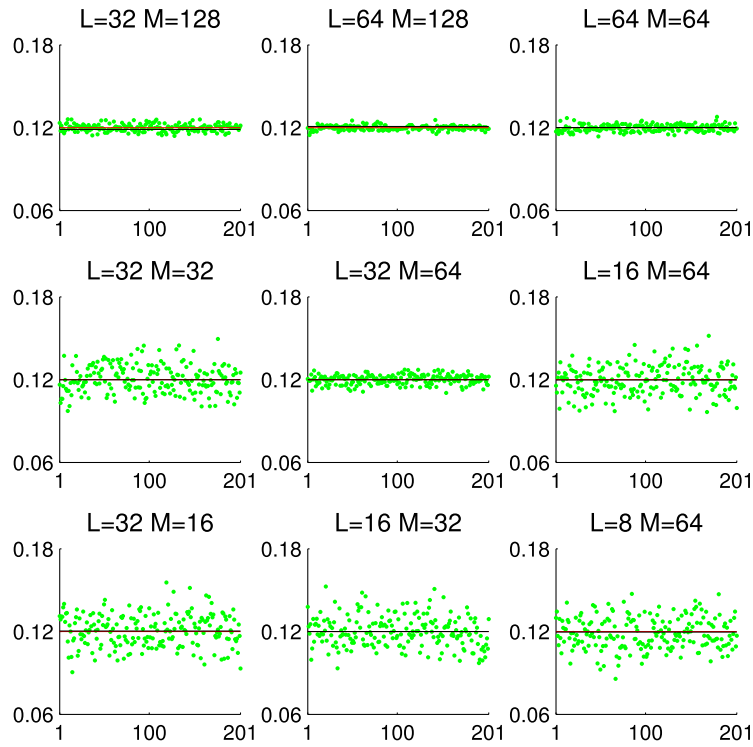
**Table 1.** Mean, standard deviation and relative error of  $(x_1, x_2)$ ,  $r$  vs  $L$  and  $M$ .

$L$	$M$	Mean	Standard deviation	Relative error
32	128	(0.9976, 0.2494), 0.1199	(0.0402, 0.0173), 0.0026	(0.24%, 0.24%), 0.07%
32	32	(0.9192, 0.2804), 0.1203	(0.1893, 0.0847), 0.0100	(8.08%, 12.15%), 0.28%
32	16	(0.9222, 0.2833), 0.1209	(0.1997, 0.0808), 0.0118	(7.78%, 13.30%), 0.77%
64	128	(0.9941, 0.2514), 0.1199	(0.0278, 0.0144), 0.0016	(0.59%, 0.56%), 0.10%
32	64	(0.9934, 0.2542), 0.1200	(0.0606, 0.0287), 0.0034	(0.66%, 1.67%), 0.03%
16	32	(0.9023, 0.2794), 0.1201	(0.2207, 0.0841), 0.0112	(9.77%, 11.74%), 0.11%
64	64	(0.9934, 0.2521), 0.1198	(0.0399, 0.0204), 0.0025	(0.66%, 0.83%), 0.18%
16	64	(0.9290, 0.2660), 0.1196	(0.1857, 0.0763), 0.0104	(7.10%, 6.41%), 0.36%
8	64	(0.9127, 0.2661), 0.1187	(0.1830, 0.0828), 0.0115	(8.73%, 6.43%), 1.07%

**Figure 1.** Reconstructions of the center of a sound-soft disk with different  $L$  and  $M$  at the idea setting. The red star  $*$  denotes the accurate center, the green dots  $\cdot$  are the numerical centers and the black  $\circ$  is the mean of the numerical centers.

plane wave  $u^{\text{in}}(x) = e^{ikx \cdot d_\ell}$  is given by the convergent series

$$u^\infty(\hat{x}; \mathbf{Z}, \mathbf{d}_\ell, k) = -e^{-i\frac{\pi}{4}} \sqrt{\frac{2}{\pi k}} \left[ \frac{J_0(kr)}{H_0^{(1)}(kr)} + 2 \sum_{n=1}^{\infty} \frac{J_n(kr)}{H_n^{(1)}(kr)} \cos(n\theta_\ell) \right], \quad \ell = 0, 1, \dots, L. \quad (3.2)$$



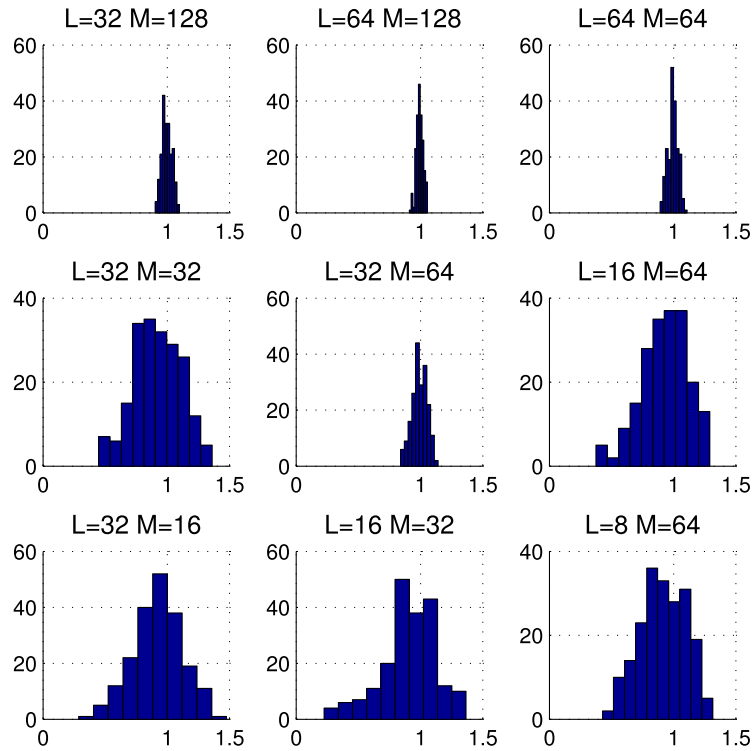
**Figure 2.** Reconstructions of the radius of a sound-soft disk with different  $L$  and  $M$  at the idea setting. The red line is the accurate radius, the black line is the mean of the numerical radii and the green dots are the numerical radii.

Here,  $\theta_\ell = \angle(\hat{\mathbf{x}}, \mathbf{d}_\ell)$  denotes the angle between the observation direction  $\hat{\mathbf{x}}$  and the incident direction  $\mathbf{d}_\ell$ ,  $J_n(\cdot)$  is the Bessel function of order  $n$  and  $H_n^{(1)}(\cdot)$  is the Hankel function of the first kind of order  $n$ . If the disk is located at  $(x_1, x_2)^\top \in \mathbb{R}^2$ , by the translational formula (1.6) and the linear superposition principle, the exact far-field pattern with phase information of the scattered waves can be expressed as

$$\begin{aligned}
 u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k) &= -e^{-i\frac{\pi}{4}} \sqrt{\frac{2}{\pi k}} \left[ \frac{J_0(kr)}{H_0^{(1)}(kr)} + 2 \sum_{n=1}^{\infty} \frac{J_n(kr)}{H_n^{(1)}(kr)} \cos(n\theta_0) \right] e^{ik(x_1, x_2)^\top \cdot (\mathbf{d}_0 - \hat{\mathbf{x}})} \\
 &\quad - e^{-i\frac{\pi}{4}} \sqrt{\frac{2}{\pi k}} \left[ \frac{J_0(kr)}{H_0^{(1)}(kr)} + 2 \sum_{n=1}^{\infty} \frac{J_n(kr)}{H_n^{(1)}(kr)} \cos(n\theta_\ell) \right] e^{ik(x_1, x_2)^\top \cdot (\mathbf{d}_\ell - \hat{\mathbf{x}})}, \\
 \ell &= 1, 2, \dots, L, \quad \hat{\mathbf{x}} \in \mathbb{S}.
 \end{aligned} \tag{3.3}$$

Note that the first line on the right hand side of (3.3) denotes the far-field pattern corresponding to the incoming plane wave  $e^{ik\mathbf{x} \cdot \mathbf{d}_0}$ , while the second line corresponding to  $e^{ik\mathbf{x} \cdot \mathbf{d}_\ell}$ .

The following theorem states that our phaseless data set is sufficient to uniquely identify a sound-soft disk.



**Figure 3.** Histograms of 201 reconstruction results of the  $x_1$ -component of the center with different  $L$  and  $M$  at the idea setting.

**Theorem 3.1.** Let  $k > 0$  and  $\mathbf{d}_0 \in \mathbb{S}$  be fixed. Then the data  $\{|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}, k) : \mathbf{d} \in \mathbb{S}\}$  uniquely determine a sound-soft disk (that is, the center and radius of a disk).

**Proof.** Suppose that  $D_j := \{z = (z_1, z_2)^\top \in \mathbb{R}^2 : \|z - x^{(j)}\|_2 < r_j\}$  ( $j = 1, 2$ ) are two sound-soft disks centered at  $x^{(j)} = (x_1^{(j)}, x_2^{(j)})^\top \in \mathbb{R}^2$  with the radius  $r_j > 0$  ( $j = 1, 2$ ). Set  $\mathbf{Z}^{(j)} = (x_1^{(j)}, x_2^{(j)}, r_j)^\top$  and denote by  $u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(j)}, \mathbf{d}_0, \mathbf{d}, k)$  the far-field data corresponding to  $D_j$  and the incident wave (1.1). Suppose that the phaseless far-field pattern are identical, i.e.,

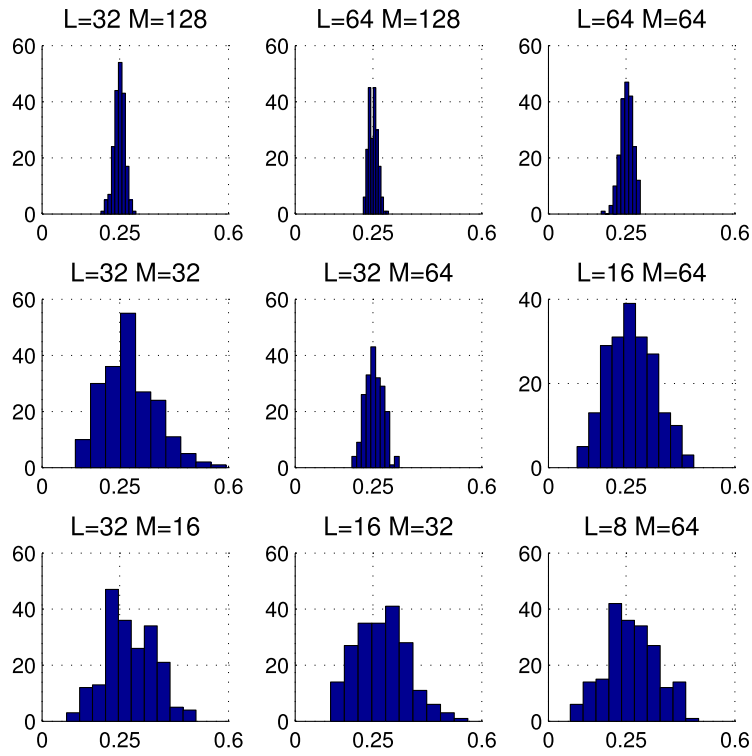
$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(1)}, \mathbf{d}_0, \mathbf{d}, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(2)}, \mathbf{d}_0, \mathbf{d}, k)| \quad \text{for all } \mathbf{d}, \hat{\mathbf{x}} \in \mathbb{S}. \quad (3.4)$$

In particular, choosing  $\mathbf{d} = \mathbf{d}_0$  in the previous relation yields

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(1)}, \mathbf{d}_0, \mathbf{d}_0, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(2)}, \mathbf{d}_0, \mathbf{d}_0, k)| = 2|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(j)}, \mathbf{d}_0, k)|, \quad j = 1, 2,$$

for all  $\hat{\mathbf{x}} \in \mathbb{S}$ , where  $u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(j)}, \mathbf{d}_0, k)$  stands for the far-field pattern corresponding to the plane wave  $e^{ikx \cdot \mathbf{d}_0}$  incident onto  $D_j$ . This implies that

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(1)}, \mathbf{d}_0, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(2)}, \mathbf{d}_0, k)| \quad \text{for all } \hat{\mathbf{x}} \in \mathbb{S}.$$



**Figure 4.** Histograms of 201 reconstruction results of the  $x_2$ -component of the center with different  $L$  and  $M$  at the idea setting.

Next, we shift the center of the disk  $D_j$  to the origin and set  $\mathbf{Z}_0^{(j)} := (0, 0, r_j)^\top$ . Recalling the translational formula (1.6), we obtain

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(1)}, \mathbf{d}_0, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(2)}, \mathbf{d}_0, k)| \quad \text{for all } \hat{\mathbf{x}} \in \mathbb{S}.$$

Since the shifted disks with the parameters  $\mathbf{Z}_0^{(j)}$  are rotationally invariant, the far-field pattern  $u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(j)}, \mathbf{d}, k)$  only depends on the angle between the incident direction  $\mathbf{d}$  and the observation direction  $\hat{\mathbf{x}}$ . For any  $\mathbf{d} \in \mathbb{S}$ , there exist an orthogonal matrix  $Q$  such that  $\mathbf{d} = Q\mathbf{d}_0$ . It then follows that (see e.g., [8, chapter 5.1])

$$u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(j)}, \mathbf{d}, k) = u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(j)}, Q\mathbf{d}_0, k) = u^\infty(Q\hat{\mathbf{x}}; \mathbf{Z}_0^{(j)}, \mathbf{d}_0, k), \quad \forall \hat{\mathbf{x}} \in \mathbb{S}.$$

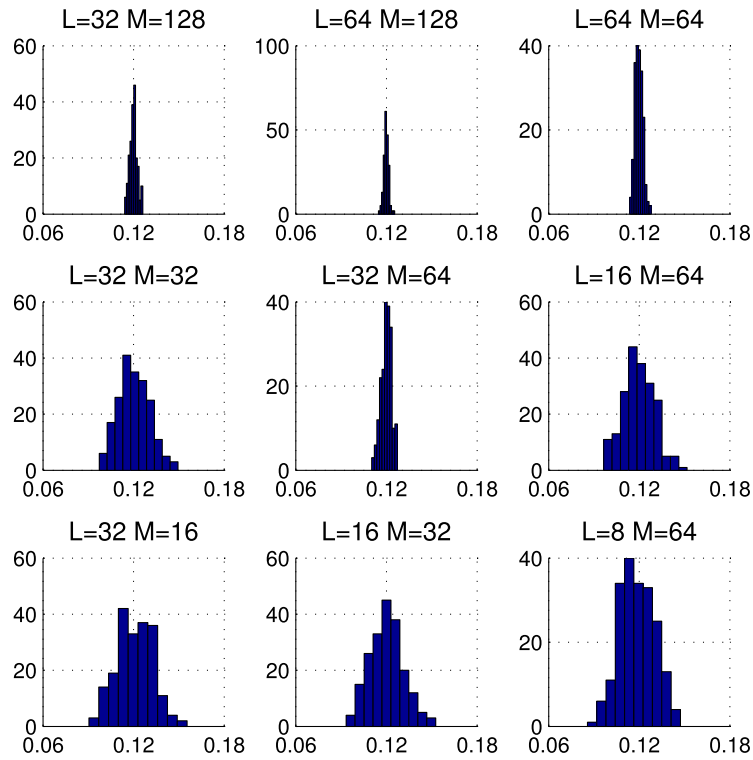
Combining the previous two identities yields

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(1)}, \mathbf{d}, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}_0^{(2)}, \mathbf{d}, k)| \quad \text{for all } \hat{\mathbf{x}}, \mathbf{d} \in \mathbb{S},$$

which together with the translational formula implies

$$|u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(1)}, \mathbf{d}, k)| = |u^\infty(\hat{\mathbf{x}}; \mathbf{Z}^{(2)}, \mathbf{d}, k)| \quad \text{for all } \hat{\mathbf{x}}, \mathbf{d} \in \mathbb{S}. \quad (3.5)$$

As a consequence of [41, theorem 2.2], the relations (3.4) and (3.5) lead to the coincidence of  $D_1$  and  $D_2$ , which proves theorem 3.1.  $\square$



**Figure 5.** Histograms of 201 reconstruction results of the radius of a sound-soft disk with different  $L$  and  $M$  at the idea setting.

To apply the pCN algorithm (2.12) or (2.13), we need to set the key parameters  $\beta$  and  $\Sigma_{\text{pcn}}$ . For sound-soft disks, we set  $\beta_j = \beta = 0.1$  and let  $\Sigma_{\text{pcn}} = \mathbf{I}$  be the  $N$ -by- $N$  identity matrix. Then the proposal is given by

$$\mathbf{X} = \mathbf{m}_{\text{pr}} + \sqrt{0.99}(\mathbf{Z}_j - \mathbf{m}_{\text{pr}}) + 0.1\omega, \quad \omega \sim N(\mathbf{0}, \mathbf{I}). \quad (3.6)$$

We describe the settings of our computational performance as follows:

- Unless otherwise specified, the wave number is always taken as  $k = 1$ ;
- The incident directions are

$$\mathbf{d}_\ell = (\cos \theta_\ell, \sin \theta_\ell), \quad \theta_\ell = -\frac{\pi}{2} + \frac{2\pi\ell}{L+1}, \quad \ell = 0, 1, \dots, L; \quad (3.7)$$

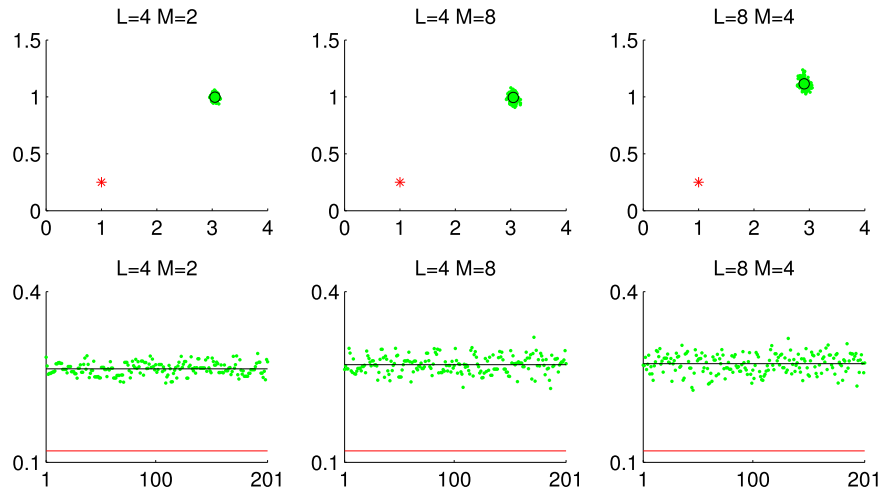
- The observation directions are

$$\hat{\mathbf{x}}_m = (\cos \theta_m, \sin \theta_m), \quad \theta_m = -\frac{\pi}{2} + \frac{2\pi m}{M}, \quad m = 1, 2, \dots, M; \quad (3.8)$$

- To compute the far-field pattern (3.3), we truncate the infinite series of (3.3) by using the Bessel and first-kind Hankel functions of order  $n = 0, 1, 2, \dots, 100$ ;
- In the algorithm 1, we choose  $J_1 = 9000, J_2 = 5, J_3 = 201$ ;
- In the setting of the prior distribution  $\pi_{\text{pr}}$ , we assume  $\sigma_{\text{pr}} = 1$ ;

**Table 2.** Mean and relative error of  $(x_1, x_2), r$  vs small  $L$  and  $M$ .

$L$	$M$	Mean	Relative error
4	2	(3.0494, 0.9982), 0.2643	(2.05%, 2.99%), 1.20%
4	8	(3.0457, 0.9949), 0.2717	(2.05%, 2.98%), 1.26%
8	4	(2.9069, 1.1143), 0.2732	(1.91%, 3.46%), 1.28%

**Figure 6.** Reconstructions of the center (top) and radius (bottom) of a sound-soft disk with a small number of incident waves ( $L$ ) and observation directions ( $M$ ) at the idea setting. The red stars \* (resp. lines) are the accurate center (resp. radius); the green dots  $\cdot$  are the numerical reconstructions; the black dots  $\circ$  (lines) are the mean values of the center (resp. radius).

- For the observation  $\mathbf{Y}_\ell$  corresponding to the incident wave  $u_\ell^{\text{in}}(x)$  in (1.7),  $\ell = 1, 2, \dots, L$ , we assume the observation pollution  $\eta^\ell = (\eta_1^\ell, \eta_2^\ell, \dots, \eta_M^\ell)^\top$  is an  $M$ -dimensional Gaussian variable, given by

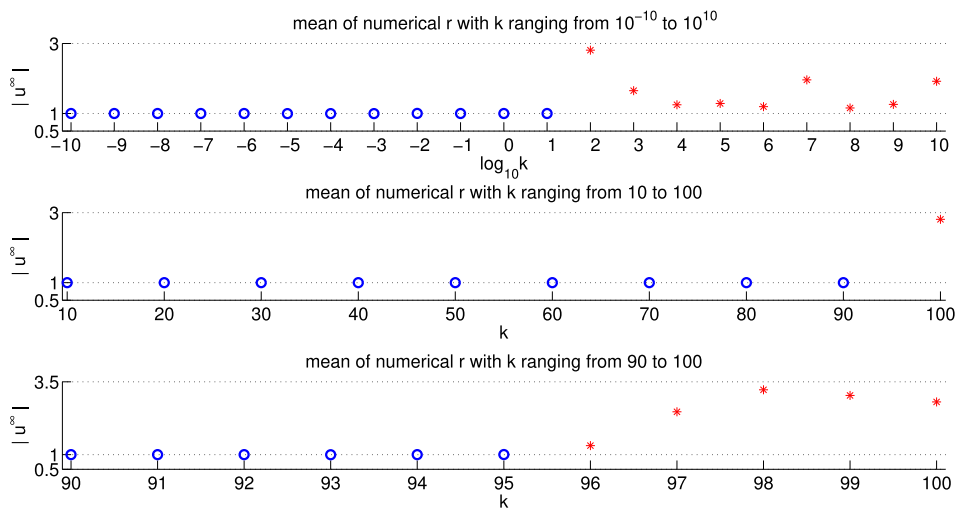
$$\eta_m^\ell = \sigma_\eta \times |u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k)| \omega_m^\ell, \quad (3.9)$$

where  $\omega_m^\ell \sim N(0, 1)$ ,  $m = 1, 2, \dots, M$  and  $\sigma_\eta$  is the noise coefficient. In our numerical tests we choose  $\sigma_\eta = 3\%, 6\%, 9\%$ . In other words, we take  $\eta^\ell \sim N(\mathbf{0}, \Sigma_{\eta^\ell})$  and the diagonal matrix  $\Sigma_{\eta^\ell} = \text{diag}(\sigma_1^\ell, \sigma_2^\ell, \dots, \sigma_M^\ell)$  with  $\sigma_m^\ell = (\sigma_\eta \times |u^\infty(\hat{\mathbf{x}}_m; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k)|)^2$ ,  $m = 1, 2, \dots, M$ ;

- Unless otherwise specified, the accurate obstacle parameters are set as  $\hat{\mathbf{Z}} = (\hat{x}_1, \hat{x}_2, \hat{r})^\top = (1, 0.25, 0.12)^\top$ , that is, a disk centered at  $(1, 0.25)^\top$  with the radius 0.12;
- We assume the initial guess is a disk centered at origin  $(0, 0)^\top$  with the radius 0.05. Then the mean of the prior distribution is  $\mathbf{m}_{\text{pr}} = (0, 0, \log 0.05)^\top$ .

In the first part, we adopt the ideal setting. Noting that in the formula (3.9)  $\omega_m^\ell \sim N(0, 1)$ , we can gain a special sample of the observation noise with  $\omega_m^\ell = 0$  and  $\sigma_\eta = 3\%$ ,  $\ell = 1, 2, \dots, L$ ,  $m = 1, 2, \dots, M$ . This would help us to investigate the accuracy of the numerical method and to verify the above uniqueness result in theorem 3.1.





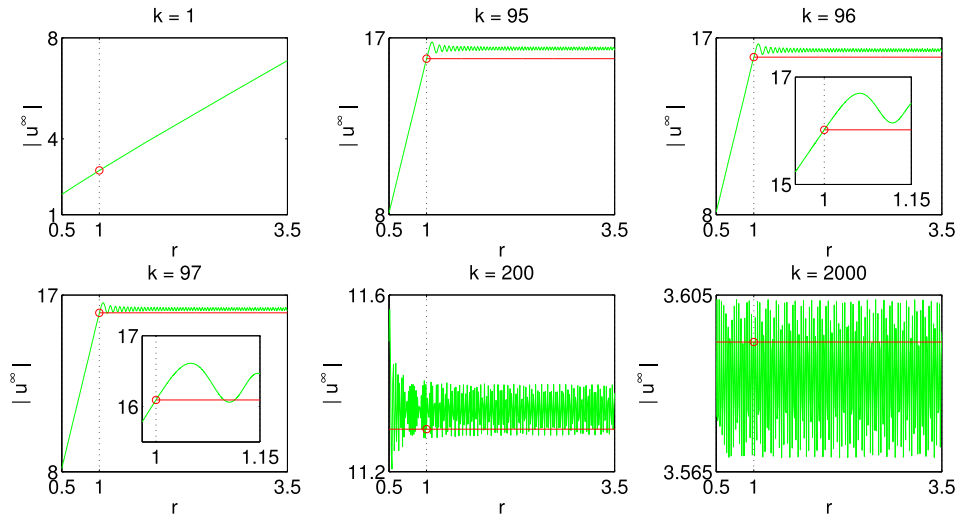
**Figure 7.** The reconstructed radii with different wavenumbers  $k$  ranging from  $10^{-10}$  to  $10^{10}$  (top), from 10 to 100 (middle) and those from 90 to 100 (bottom). Both the blue dots  $\circ$  and red stars  $*$  are the reconstructed radii. The blue dot  $\circ$  represents the accurate reconstruction, while the red star  $*$  represents the inaccurate reconstruction.

At first, we discuss the accuracy of the numerical solutions for different choice of  $L$  and  $M$ . Recall that the parameter  $L$  denotes the number of incident waves and  $M$  the number of observation directions. In table 1, we show the mean, standard deviation and the relative error of the reconstructed parameters with different choice of  $L, M$ . The numerical solutions of the recovered centers and radii are shown in figures 1 and 2. The histograms of numerical solutions of the centers and radii are shown in figures 3–5, respectively.

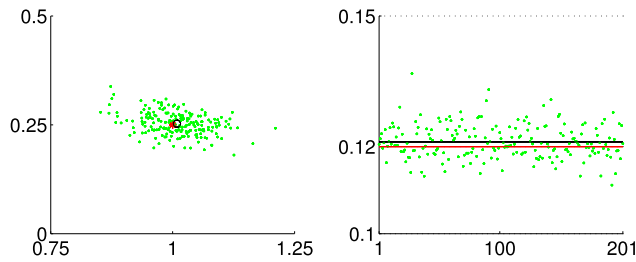
Based on these results, we find that the reconstructed parameters are getting more accurate as the number of incident or observation directions become larger. The numerical solutions with  $(L, M) = (32, 32), (32, 16), (16, 32), (16, 64), (8, 64)$  are relatively inaccurate, since the resulted relative errors are larger than 5% in these cases. In contrast, the relative error is less 1% if we choose  $L$  and  $M$  large enough such as  $(L, M) = (64, 128), (64, 64), (32, 128)$ . On the other hand, it can be observed from table 1 that the standard deviation decreases as  $L$  or  $M$  increases. Table 2 and figure 6 illustrate that small  $L$  and  $M$  may lead to unreliable reconstructions.

In the following we suppose that the location of the center  $(\hat{x}_1, \hat{x}_2)^\top = (2, 2)^\top$  is known and the knowledge of the radius needs to be recovered. Since only one parameter of the scatter remains unknown, we make use of minimal number of incident and observation directions by setting  $L = M = 1$ . In our tests we set incident directions  $\mathbf{d}_0 = (0, -1)^\top, \mathbf{d}_1 = (0, 1)^\top$ , observation direction  $\hat{\mathbf{x}}_1 = (0, -1)^\top$  and accurate radius  $\hat{r} = 1$ . The numerical approximations of radius  $r$  vs different wave numbers  $k$  are exhibited in figure 7. For each fixed  $k$ , we plot the phaseless far-field pattern  $|u^\infty(\hat{\mathbf{x}}_1)|$  against the radius  $r$  in figure 8.

From the numerical results we conclude that an accurate approximation of the radius can be obtained if the wave number  $k$  is less than a threshold. It is seen from figure 8 that the function  $r \rightarrow |u^\infty(\hat{\mathbf{x}}_1)|$  is monotonically increasing in  $(0, R(k))$  where  $R(k) \rightarrow 0^+$  as  $k \rightarrow +\infty$ . This suggests that for large  $k$  such as  $k = 97, 200, 2000$ , there are more than one radii corresponding to the measured phaseless far-field pattern at  $\hat{\mathbf{x}}_1$ . Hence, the reconstructed radii are inaccurate.



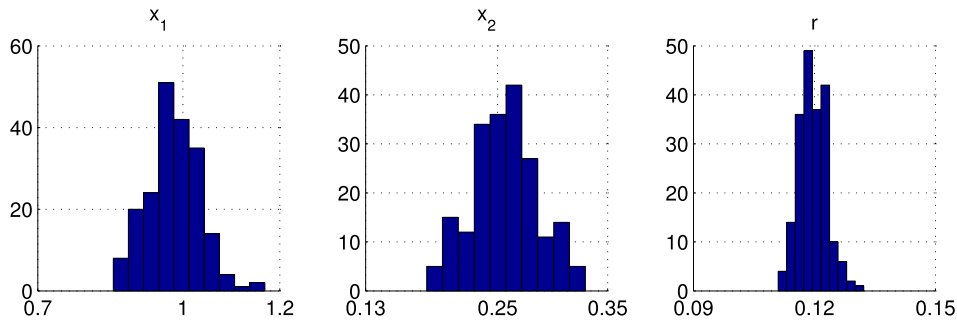
**Figure 8.** Phaseless far-field pattern  $|u^\infty(\hat{x}_1)|$  vs radius  $r$  for different wavenumbers  $k$ .



**Figure 9.** Reconstruction of the center (left) and radius (right) of a disk with one sample of observation noise at  $\sigma_\eta = 3\%$ . The red star\* (resp. line) is the accurate center (radius); the green dots  $\cdot$  are the numerical reconstructions; the black circle  $\circ$  (resp. line) is the mean of the centers and radii. We choose  $k = 1$ ,  $L = 32$  and  $M = 64$ .

These findings are consistent with the uniqueness result of [34], which states that a sound-soft disk can be uniquely determined from the phaseless far-field pattern at one observation direction, provided the radius is sufficiently small for a fixed wave number. The monotonicity property of the backscattered phaseless data with respect to the radius was rigorously justified in [34].

Having verified the accuracy of our inversion scheme at the ideal setting with a special sample of the observation noise, we now consider the inverse problem with a general sample of the observation noise at the noise coefficient  $\sigma_\eta = 3\%$ . In the second part, we estimate the obstacle parameters by setting  $k = 1$ ,  $L = 32$  and  $M = 64$ . We generate one sample of the observation noise, which is a matrix with  $M \times L$  elements constructed by the formula (3.9). The numerical approximations from the polluted observation data are exhibited in figures 9 and 10. The mean of the numerical center and radius are  $(1.0089, 0.2527)^\top$  and 0.1211, respectively. The standard deviations of these parameters are  $(0.0601, 0.0259)^\top$ , 0.0039 and the relative errors are  $(0.89\%, 1.08\%)^\top$ , 0.92%.



**Figure 10.** Histograms of the reconstructed parameters  $x_1, x_2, r$  with one sample of the noisy data polluted at the level  $\sigma_\eta = 3\%$ .

**Table 3.** The mean, standard deviation and relative error of  $(x_1, x_2), r$  vs noise coefficient  $\sigma_\eta$ .

$\sigma_\eta$	Mean	Standard deviation	Relative error
3%	(0.9917, 0.2520), 0.1199	(0.0077, 0.0039), 0.0005	(0.83%, 0.79%), 0.10%
6%	(0.9914, 0.2519), 0.1199	(0.0137, 0.0065), 0.0009	(0.86%, 0.77%), 0.10%
9%	(0.9918, 0.2517), 0.1199	(0.0199, 0.0095), 0.0013	(0.82%, 0.69%), 0.10%

To demonstrate the robustness of the numerical scheme, we generate 1000 samples of the observation noise. For each sample of the observation noise, one can gain a corresponding reconstruction of the parameters  $x_1, x_2, r$ . Hence, we can perform statistical analysis over totally 1000 reconstructions of  $x_1, x_2, r$ . In our tests, we pollute the phaseless data at different levels  $\sigma_\eta = 3\%, 6\%, 9\%$  and exhibit the numerics in table 3, figure 11 and 12. From these reconstructed parameters we conclude that the mean and relative error are robust against the noise pollution, but the standard deviation is very sensitive to the noisy level. Further, the phaseless data with less noise give rise to a more reliable reconstruction result.

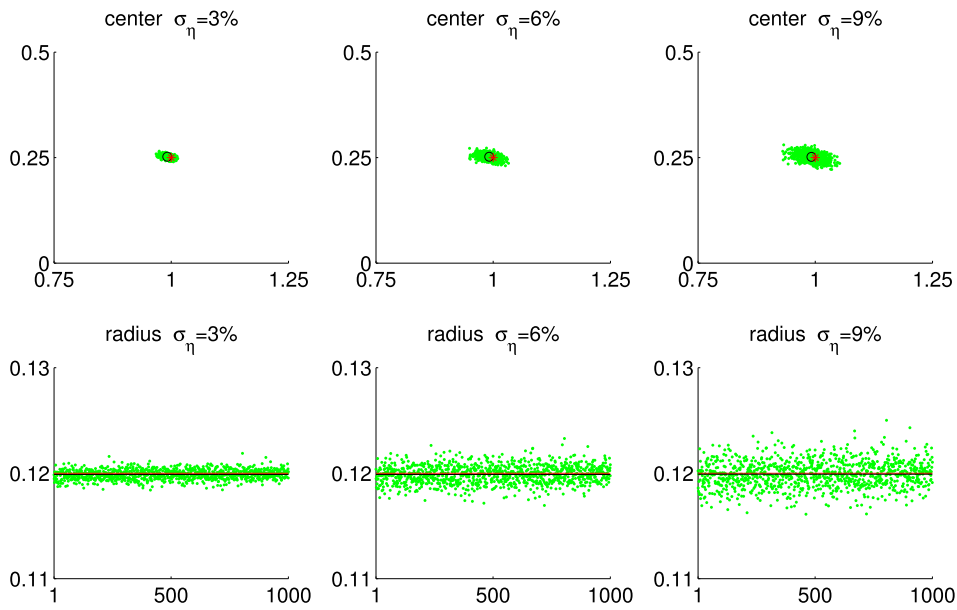
### 3.2. Line cracks

A crack or an open arc can be used to model the defects inside elastic and solid bodies such as bridge structures, aircraft engines and wings etc. Detection of such scatterers is important in safety and health assessment and is one of the fundamental topics in ultrasonic non-destructive testing. In this subsection, we want to recover a sound-soft crack of line-segment-type with the starting point at  $x = (x_1, x_2)^\top \in \mathbb{R}^2$  and the ending point  $y = (y_1, y_2)^\top \in \mathbb{R}^2$ . Hence, such line cracks can be characterized by  $N = 4$  parameters:

$$\mathbf{Z} := (z_1, z_2, z_3, z_4)^\top = (x_1, x_2, y_1, y_2)^\top. \quad (3.10)$$

Unlike the scattering from disks, we do not have an analytical expression of the far-field pattern corresponding to a line crack. Below we describe the integral equation method to solve the forward scattering problem, following the numerical scheme of [28] for general cracks. Denote by  $\Gamma \subset \mathbb{R}^2$  an open arc of class  $C^3$  in 2D, which can be parameterized as

$$\Gamma = \{z(s) : s \in [-1, 1]\}. \quad (3.11)$$



**Figure 11.** Reconstructions of the center (top) and radius (bottom) of a disk at different noise levels  $\sigma_\eta = 3\%$  (left),  $6\%$  (middle),  $9\%$  (right). The red star \* (resp. line) is the accurate center (resp. radius), the green dots · are the numerical reconstructions with each sample of observation noise and the black  $\circ$  (resp. line) is the mean of reconstructed centers (resp. radii) with 1000 samples of observation noise.

Using the integral equation method, the solution  $u^{\text{sc}}$  to the Helmholtz equation (1.2) in  $\mathbb{R}^2 \setminus \Gamma$  can be expressed as a single-layer potential ([8])

$$u^{\text{sc}}(x) = \int_{\Gamma} \Phi(x, y) \varphi(y) \, ds(y), \quad x \in \mathbb{R}^2 \setminus \Gamma, \quad (3.12)$$

where  $\Phi(x, y)$  is the fundamental solution to the Helmholtz equation in two dimensions given by

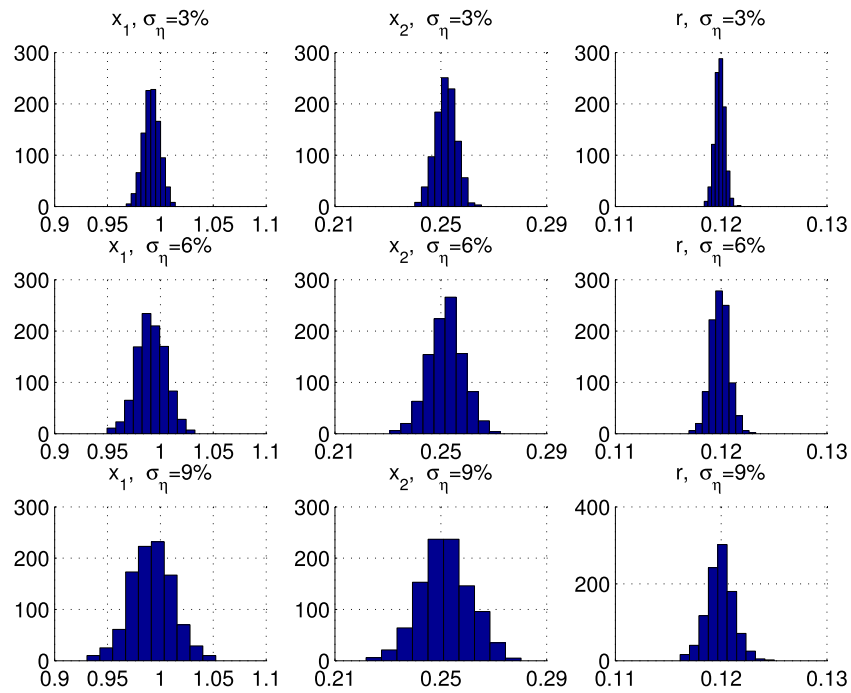
$$\Phi(x, y) := \frac{i}{4} H_0^{(1)}(k|x - y|), \quad x, y \in \mathbb{R}^2, x \neq y. \quad (3.13)$$

Due to the Dirichlet boundary condition (1.3) on  $\Gamma$ , the unknown density function  $\varphi$  is sought as a solution to the integral equation

$$\int_{\Gamma} \Phi(x, y) \varphi(y) \, ds(y) = f(x), \quad x \in \Gamma, \quad f = -u^{\text{in}}. \quad (3.14)$$

Once  $\varphi$  is calculated from (3.14), the far-field pattern could be expressed in the form

$$u^\infty(\hat{x}) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\Gamma} e^{-ik\hat{x} \cdot y} \varphi(y) \, ds(y), \quad \hat{x} \in \mathbb{S}. \quad (3.15)$$



**Figure 12.** Histogram of 1000 numerical reconstructions of  $x_1, x_2, r$  with each sample of observation noise at different noise levels  $\sigma_\eta = 3\%$  (top),  $\sigma_\eta = 6\%$  (middle),  $\sigma_\eta = 9\%$  (bottom).

To describe the numerical scheme of [28], we first introduce two functions defined on  $\mathbb{R} \times \mathbb{R}$  as follows

$$H_1(t, \tau) := \begin{cases} J_0(k|z(\cos t) - z(\cos \tau)|) - 1, & t \neq \tau, \\ 0, & t = \tau, \end{cases} \quad (3.16)$$

and

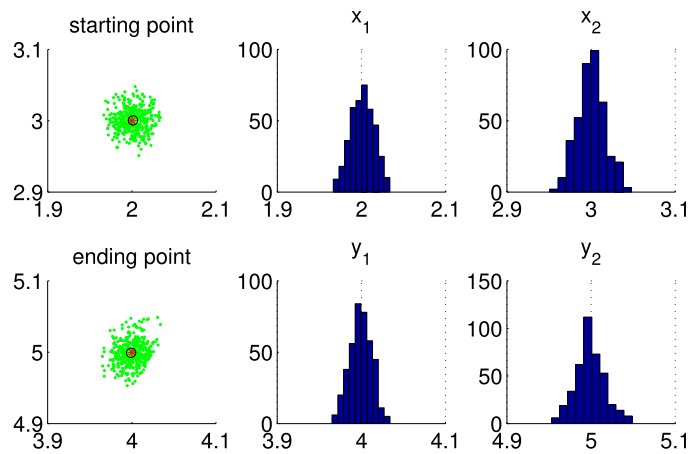
$$H_2(t, \tau) := \begin{cases} \frac{\pi}{i} H_0^{(1)}(k|z(\cos t) - z(\cos \tau)|) - \{1 + H_1(t, \tau)\} \ln \left( \frac{4}{e^2} [\cos t - \cos \tau]^2 \right), & t \neq \tau, \\ \frac{\pi}{i} + 2C + 2 \ln \left\{ \frac{ke}{4} |z'(\cos t)| \right\}, & t = \tau. \end{cases} \quad (3.17)$$

Here,  $C \approx 0.577216$  is the Euler's constant. Then the integral equation (3.14) can be rephrased as

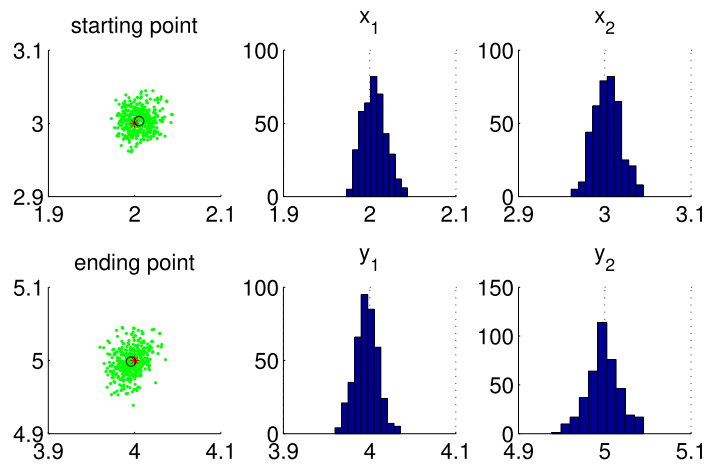
$$\frac{1}{2\pi} \int_0^{2\pi} K(t, \tau) \psi(\tau) \, ds(\tau) = g(t), \quad g(t) := -2f(z(\cos t)), \quad (3.18)$$

for  $t \in [0, 2\pi]$ . Here,

$$K(t, \tau) = \left\{ 1 + \sin^2 \frac{t-\tau}{2} K_1(t, \tau) \right\} \ln \left( \frac{4}{e} \sin^2 \frac{t-\tau}{2} \right) + \frac{1}{2} H_2(t, \tau), \quad (3.19)$$



**Figure 13.** Trace of the iterations in the MCMC method with one sample of the observation noise at  $\sigma_\eta = 3\%$ . The accurate line crack is  $(\hat{x}_1, \hat{x}_2, \hat{y}_1, \hat{y}_2)^\top = (2, 3, 4, 5)^\top$ .

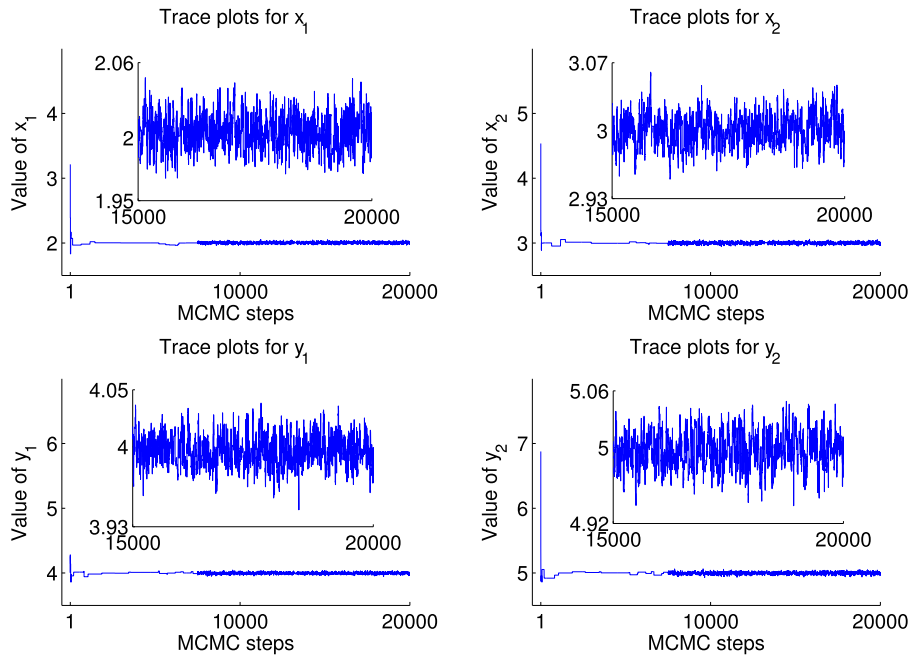


**Figure 14.** Histogram (center, right) and scatterer plot (left) of the reconstructions of starting point (top) and ending point (bottom) of a line crack with the idea setting. The red star \* denotes the accurate point, the green dots · are the numerical points, and the black o is the mean of numerical points.

$$K_1(t, \tau) = \begin{cases} \frac{H_1(t, \tau)}{\sin^2((t - \tau)/2)}, & t \neq \tau, \\ -k^2 \sin^2(t) |z'(\cos t)|^2, & t = \tau, \end{cases} \quad (3.20)$$

$$K_2(t, \tau) = \frac{1}{2} H_2(t, \tau). \quad (3.21)$$

The quadrature method [7] can be employed to discretize the integral equation (3.18), based on the trigonometric interpolation with  $2n$  equidistant nodal points  $t_j := \frac{j\pi}{n}, j = 0, 1, \dots, 2n - 1$ . Then the unknown solution  $\psi$  to the integral equation (3.18) can be approximated by



**Figure 15.** Histogram (center, right) and scatterer plot (left) of reconstructions of starting point (top) and ending point (bottom) of a line crack with one sample of observation noise at  $\sigma_\eta = 3\%$ . The red star \* denotes the accurate point, the green dots · are the numerical points, and the black  $\circ$  is the mean of numerical points.

**Table 4.** The mean, standard deviation and relative error of  $x_1, x_2, y_1, y_2$  vs noise coefficient  $\sigma_\eta$ .

$\sigma_\eta$	Mean	Standard deviation ( $10^{-3}$ )	Relative error (%)
3%	1.9994, 3.0055, 3.9992, 5.0029	3.5, 4.4, 3.4, 5.1	0.03, 0.18, 0.02, 0.06
6%	1.9989, 3.0056, 4.0013, 5.0133	6.2, 7.7, 6.1, 8.8	0.06, 0.19, 0.03, 0.27
9%	1.9936, 2.9994, 3.9968, 5.0081	8.6, 11.4, 8.9, 12.5	0.32, 0.02, 0.08, 0.16

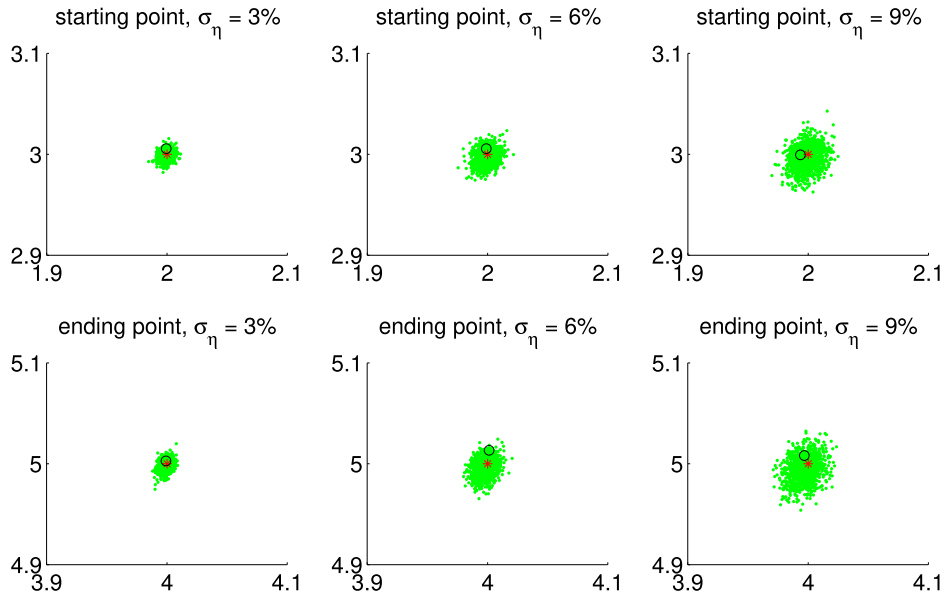
the  $2n$  discrete nodal values  $\{\psi_j = \psi(t_j)\}_{j=0}^{2n-1}$ . Since  $\psi_k = \psi_{2n-k}, k = 1, 2, \dots, n-1$  with  $\psi_j = |\sin(t_j)| |z'(\cos(t_j))| \varphi(z(\cos(t_j)))$ , it suffices to compute the  $n+1$  discrete nodal values  $\{\psi_j\}_{j=0}^n$  from the following  $(n+1) \times (n+1)$  algebraic system

$$\sum_{j=0}^{2n-1} \psi_j \left\{ R_{|k-j|} + F_{|k-j|} K_1(t_k, t_j) + \frac{1}{2n} K_2(t_k, t_j) \right\} = g(t_k), \quad k = 0, 1, \dots, n, \tag{3.22}$$

with

$$R_j := \frac{1}{2n} \left\{ c_0 + 2 \sum_{m=1}^{n-1} c_m \cos \frac{mj\pi}{n} + (-1)^j c_n \right\}, \quad c_m := -\frac{1}{\max\{1, |m|\}}, \tag{3.23}$$





**Figure 16.** Reconstructions of starting point (top) and ending point (bottom) of the line crack at different noise levels  $\sigma_\eta = 3\%$  (top),  $\sigma_\eta = 6\%$  (middle),  $\sigma_\eta = 9\%$  (bottom). The red star \* denotes the accurate point, the green dots · are the numerical reconstructions with each sample of observation noise, and the black  $\circ$  is the mean of numerical reconstructions with 1000 observation noises.

$$F_j := \frac{1}{2n} \left\{ \gamma_0 + 2 \sum_{m=1}^{n-1} \gamma_m \cos \frac{mj\pi}{n} + (-1)^j \gamma_n \right\}, \quad \gamma_m := \frac{1}{4} (2c_m - c_{m+1} - c_{m-1}). \quad (3.24)$$

Note that there are totally  $n + 1$  unknown discrete nodal values  $\{\psi_j\}_{j=0}^n$  in (3.22), because  $\psi_j = \psi_{n-|n-j|}$  for all  $j = 0, 1, \dots, 2n - 1$ . Now the far-field pattern can be approximated by

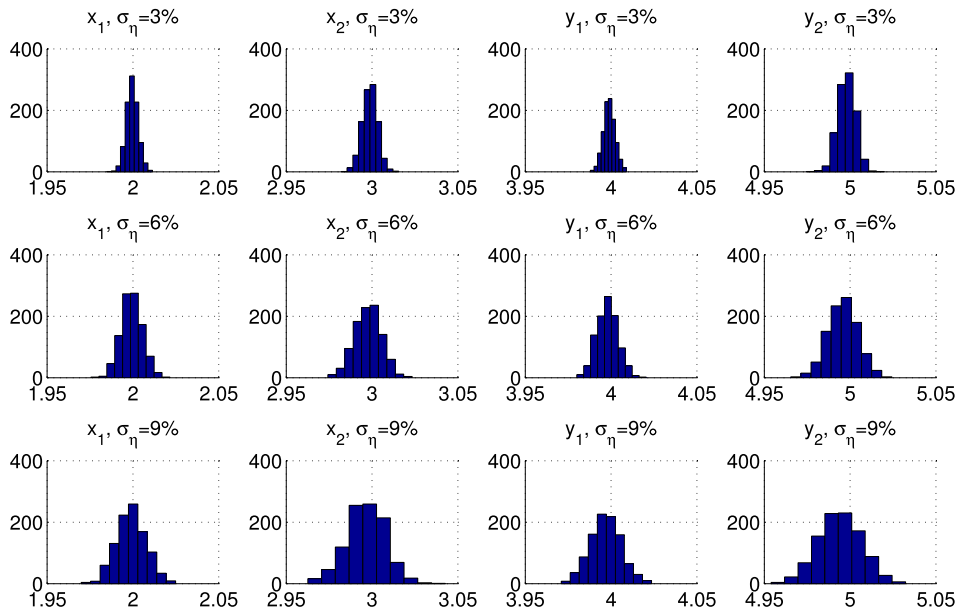
$$u^\infty(\hat{x}) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_0^{2\pi} e^{-ik\hat{x} \cdot z(\cos \tau)} \psi(\tau) d\tau \approx \frac{\pi e^{i\pi/4}}{n\sqrt{8\pi k}} \sum_{j=0}^{2n-1} e^{-ik\hat{x} \cdot z(\cos t_j)} \psi_j, \quad \hat{x} \in \mathbb{S}. \quad (3.25)$$

If the right hand side of (3.22) (or (3.14)) is given by the incident wave (1.7), we obtain the far-field pattern  $u^\infty(\hat{x}; \mathbf{Z}, \mathbf{d}_0, \mathbf{d}_\ell, k)$  where  $\mathbf{Z}$  denotes the crack parameter (3.10).

To set the parameters  $\beta$  and  $\Sigma_{\text{pcn}}$ , we let  $\Sigma_{\text{pcn}} = \mathbf{I}$  be the identity matrix, which is the same as the case of sound-soft disks. However, in this section the proposal variance coefficient  $\beta$  is not a fixed number, but is taken as a random variable. This suggests that a random proposal variance is adopted to reconstruct line cracks. Then the proposal takes the form

$$\mathbf{X} = \mathbf{m}_{\text{pr}} + (1 - \beta_j^2)^{1/2} (\mathbf{Z}_j - \mathbf{m}_{\text{pr}}) + \beta_j \omega, \quad \omega \sim N(\mathbf{0}, \mathbf{I}), \quad (3.26)$$

and the proposal variance coefficients  $\beta_j$  need to be updated by the formulas (2.16) and (2.17). It should be noted that, the MCMC method with a fixed proposal variance coefficient converges



**Figure 17.** Histogram of 1000 numerical reconstructions of  $x_1, x_2, y_1, y_2$  with each sample of the observation noise at the noise level  $\sigma_\eta = 3\%$  (top),  $\sigma_\eta = 6\%$  (middle),  $\sigma_\eta = 9\%$  (bottom).

slowly or even does not converge after a large number of iterations, which is in contrast to the efficient MCMC method for recovering disks. This could partly be due to the number of reconstructed parameters, which is four in the line crack case while three for a disk. The trace of the iterations of MCMC (shown in figure 13) verifies the efficiency of the random proposal variance. In the first 10 000 iteration steps, the trace converges fast but always drops into some fixed states, when the proposal variance coefficients  $\beta_j$  are not appropriately updated. Numerics show that the trace can converge to and oscillate around the accurate state only after a large number of iterations.

As in the previous subsection, we set some computational parameters as follows:

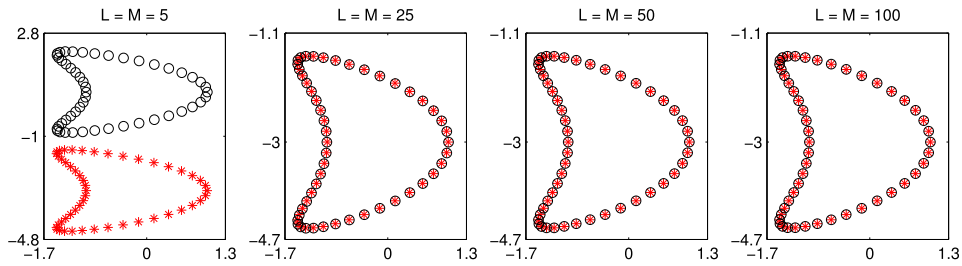
- The wave number  $k$ , the incident directions  $\mathbf{d}_\ell, \ell = 0, 1, \dots, L$ , the observation directions  $\hat{\mathbf{x}}_m, m = 1, 2, \dots, M$ ,  $\sigma_{\text{pr}}$  of the prior distribution and the observation pollution  $\eta^\ell, \ell = 1, \dots, L$  are given as same as those for recovering disks;
- We choose  $L = M = 40$ ;
- In the algorithm 1, we choose  $J_1 = 18\,000, J_2 = 5, J_3 = 401$ ;
- The accurate obstacle is  $\hat{\mathbf{Z}} = (\hat{x}_1, \hat{x}_2, \hat{y}_1, \hat{y}_2)^\top = (2, 3, 4, 5)^\top$ , that is, a line segment with the starting point  $(2, 3)^\top$  and the ending point  $(4, 5)^\top$ ;
- We assume the initial guess is a line segment with the starting point  $(0, 0)^\top$  and the ending point  $(1, 1)^\top$ . Then the mean of the prior distribution is  $\mathbf{m}_{\text{pr}} = (0, 0, 1, 1)^\top$ .

Unfortunately we do not have the uniqueness result analogous to theorem 3.1 for recovering cracks. A local uniqueness result for general cracks was proved in [27] using a single far-field pattern with information. In the idea setting ( $\omega_m^\ell = 0$  and  $\sigma_\eta = 3\%, \ell = 1, 2, \dots, L, m = 1, 2, \dots, M$ ), the numerical approximations of the crack parameters are exhibited in figure 14. The mean solutions of the starting and ending points are  $(2.0012, 3.0005), (3.9990, 4.9992)$ ,

Q2

**Table 5.** Reconstruction and Hausdorff distance (HD) vs  $L(M = L)$ .

$L$	Reconstruction	HD
5	-0.6441, 0.6287, 1.0013, 0.6460, 1.4934, -0.0051	3.6185
25	-0.6488, -3.0037, 1.0047, 0.6479, 1.5030, -0.0043	0.0021
50	-0.6499, -3.0001, 1.0000, 0.6503, 1.4998, -0.0009	$1.30 \times 10^{-4}$
100	-0.6497, -2.9995, 1.0000, 0.6508, 1.4999, 0.0006	$6.98 \times 10^{-5}$

**Figure 18.** Reconstructions of a kite-shaped obstacle at the idea setting with different  $L(M = L)$ . The red star \* denotes the exact boundary and the black circle  $\circ\sqrt{i}$  denotes the reconstructed boundary.

the standard deviations are (0.0139, 0.0164), (0.0129, 0.0170) and the relative errors are (0.06%, 0.02%), (0.02%, 0.02%).

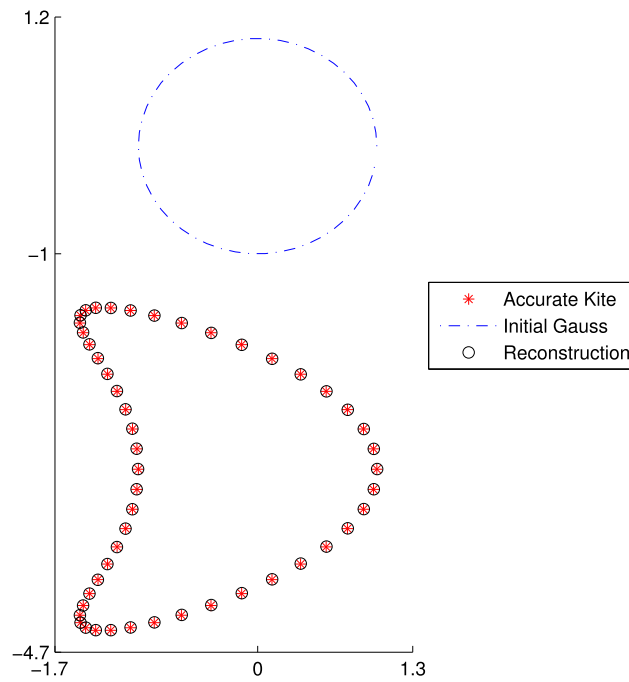
Setting the noise coefficient  $\sigma_\eta = 3\%$ , we generate one general sample of the observation noise by the formula (3.9), which takes the form of an  $M \times L$  matrix. The numerical approximations from the polluted observations is exhibited in figure 15. The mean solutions of the numerical starting and ending points are (2.0052, 3.0031), (3.9958, 4.9986), the standard deviations are (0.0135, 0.0155), (0.0130, 0.0184) and the relative errors are (0.26%, 0.10%), (0.10%, 0.03%). Besides, the trace of the iterations in MCMC are shown in figure 13.

As done for recovering disks, we also demonstrate the robustness of the numerical scheme with 1000 samples of the observation noise at different levels  $\sigma_\eta = 3\%, 6\%, 9\%$ . For each sample of the observation noise, one can gain a corresponding reconstruction of the parameters  $x_1, x_2, y_1, y_2$ . Hence, we can perform statistical analysis over totally 1000 reconstructions of  $x_1, x_2, y_1, y_2$ . The corresponding results are exhibited in table 4, figures 16 and 17. From these reconstructed parameters, we can draw almost the same conclusions as those for determining a sound-soft disk. The mean and relative error are robust against the noise pollution, but the standard deviation is very sensitive to the noisy level. It follows that the phaseless data with less noise give rise to a more reliable reconstruction result.

### 3.3. Kite-shaped obstacle

In this subsection, we consider the following kite-shaped sound-soft obstacle  $D_1$ ,

$$\partial D_1 = \{x(t) = (-0.65 + \cos t + 0.65 \cos(2t), -3 + 1.5 \sin t)^\top, \quad 0 \leq t \leq 2\pi\}, \quad (3.27)$$



**Figure 19.** Reconstruction of a kite-shaped obstacle with  $L = M = 50$  at the idea setting.

which is a benchmark acoustically impenetrable scatterer in inverse scattering problems. Obviously, the boundary of  $D_1$  can be parameterized by six parameters

$$\mathbf{Z} := (z_1, z_2, \dots, z_6)^\top. \quad (3.28)$$

Suppose that the exact parameters are given by  $\hat{\mathbf{Z}} = (\hat{z}_1, \hat{z}_2, \dots, \hat{z}_6)^\top = (-0.65, -3, 1, 0.65, 1.5, 0)^\top$ . To calculate numerical solutions of the forward problem, we adopt the MATLAB code given by [37, chapter 8]. Since there are six unknown parameters, the random proposal variance is adopted as same as in recovering line cracks. As in the previous subsections, we set the computational settings as follows.

- The wave number is  $k = 2$ ;
- The directions of plane incident waves are

$$\mathbf{d}_\ell = (\cos \theta_\ell, \sin \theta_\ell)^\top, \quad \theta_\ell = 2\pi\ell/(L+1), \quad \ell = 0, 1, \dots, L; \quad (3.29)$$

- The observation directions are

$$\hat{\mathbf{x}}_m = (\cos \theta_m, \sin \theta_m)^\top, \quad \theta_m = -\pi + 2\pi(m-1)/M, \quad m = 1, 2, \dots, M; \quad (3.30)$$

- Unless otherwise specified, we choose  $L = M = 50$ ;
- In the algorithm 1, we choose  $J_1 = 180\,000$ ,  $J_2 = 20$ ,  $J_3 = 1001$ ;

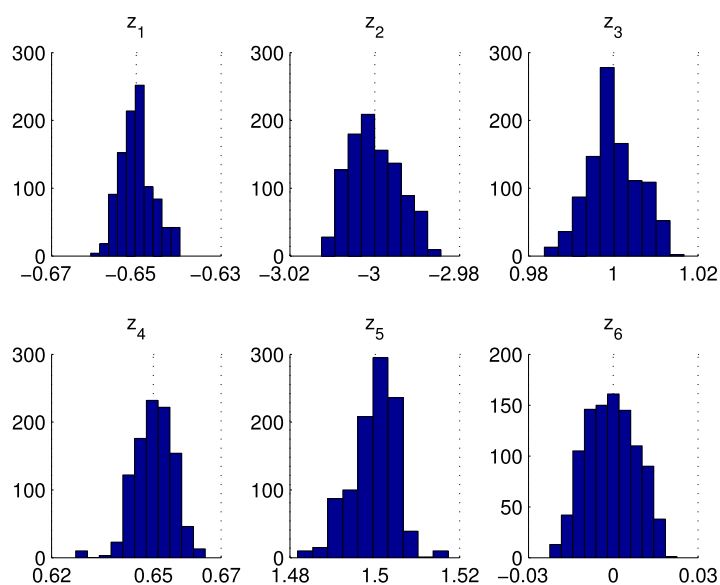


Figure 20. Histogram of the selected 1001 states with  $L = M = 50$  at the idea setting.

Table 6. Numerical solutions vs  $\sigma_\eta$  with  $L = M = 50$ .

$\sigma_\eta$	Mean of reconstructions	Mean of HD	SD of HD
3%	-0.6496, -2.9988, 0.9988, 0.6506, 1.4988, 0.0010	0.0039	0.0035
6%	-0.6493, -2.9978, 0.9975, 0.6514, 1.4975, 0.0020	0.0062	0.0054
9%	-0.6482, -2.9960, 0.9959, 0.6517, 1.4962, 0.0031	0.0088	0.0073

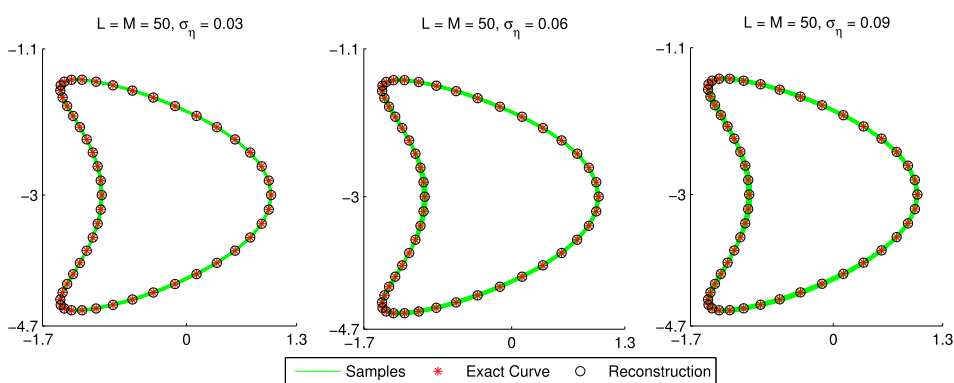
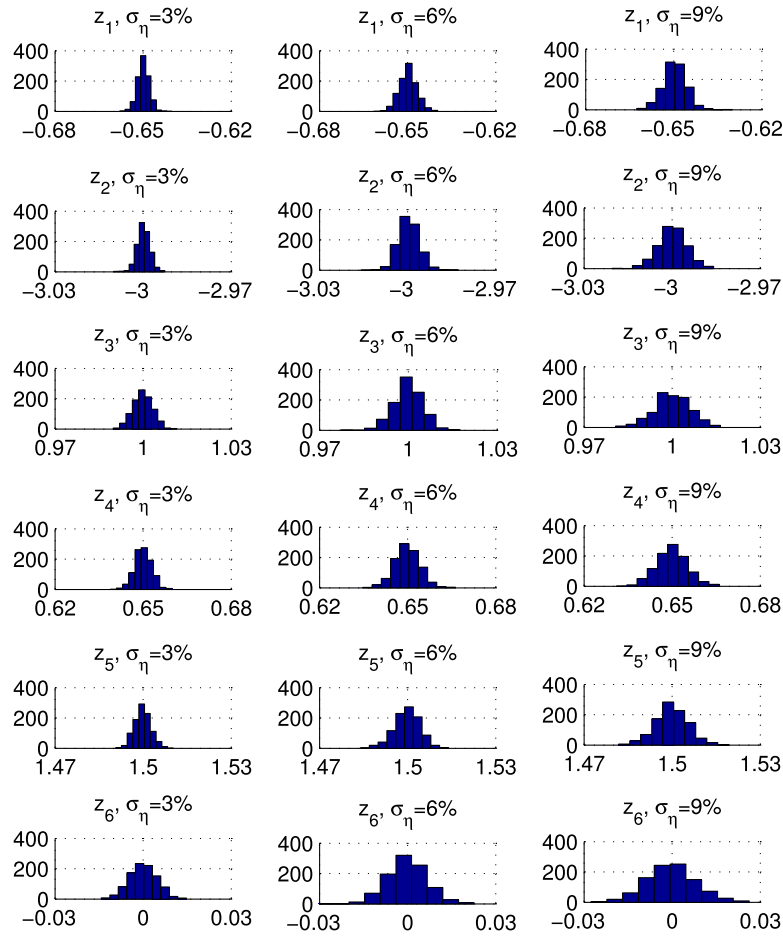


Figure 21. Reconstructions of a kite-shaped obstacle with  $L = M = 50$  at  $\sigma_\eta = 3\%$  (left),  $6\%$  (center),  $9\%$  (right).



**Figure 22.** Histogram of these 1000 reconstructions of the six obstacle parameters of the kite-shaped obstacle with  $L = M = 50$  at  $\sigma_\eta = 3\%$  (left),  $6\%$  (center),  $9\%$  (right).

- Since the initial guess is assumed to be a unit circle centered at the origin, the mean of the prior distribution  $\pi_{\text{pr}}$  is  $\mathbf{m}_{\text{pr}} = (0, 1, 0, 0, 1, 0)^\top$ . In this example, we also assume  $\sigma_{\text{pr}} = 1$ ;
- The observation pollution  $\eta^\ell$ ,  $\ell = 1, \dots, L$ , are given as same as those for recovering disks.

Since the sixth exact obstacle parameter  $\hat{z}_6$  is 0, we can not use the relative error to evaluate the accuracy of numerical results in this example. Instead, the Hausdorff distance (HD) is chosen to compute the distance between reconstructed and exact boundaries. Recall that the Hausdorff distance between two obstacles  $\partial D_2$  and  $\partial D_3$  is defined by

$$d_H(\partial D_2, \partial D_3) := \max \left\{ \sup_{x \in \partial D_2} \inf_{y \in \partial D_3} |x - y|, \sup_{y \in \partial D_3} \inf_{x \in \partial D_2} |y - x| \right\}. \quad (3.31)$$

In the first part, we consider the ideal setting ( $\omega_m^\ell = 0$  and  $\sigma_\eta = 3\%$ ,  $\ell = 1, 2, \dots, L$ ,  $m = 1, 2, \dots, M$ ) of observations to investigate the accuracy of our numerical method. As done for recovering disks, we discuss the accuracy of numerical solutions for different choice of  $L$  ( $M = L$ ). In table 5 and figure 18, we exhibit the numerical reconstructions and the Hausdorff

distances (HD) between the numerical reconstructions and the exact boundaries. We find that the reconstructed parameters are getting more accurate as the number of incident waves and observation directions becomes larger. The Hausdorff distance is less than 0.001 if we choose  $L = M = 50, 100$ . The numerical solution with  $L = M = 5$  is unreliable as illustrated in the figure 18.

To show that this method is not sensitive to initial guess, we exhibit the initial guess, the exact boundary and numerical reconstruction in figure 19, where we set  $L = M = 50$ . We can obtain an accurate numerical solution with the Hausdorff distance (HD) being  $1.30 \times 10^{-4}$ , even if the initial guess of the obstacle is separated from the exact one.

We draw the histogram of the selected 1001 sates with  $L = M = 50$  in the figure 20, which are used to construct the posterior density. From figure 20, we conclude that the posterior density is a Gaussian distribution and the mean of the posterior density approaches the exact obstacle parameters.

Then we consider the practical setting with noise-polluted phaseless far-field data. As done for recovering disks and line cracks, we also demonstrate robustness of the numerical scheme with 1000 samples of the observation noise at different noise levels  $\sigma_\eta = 3\%, 6\%, 9\%$  with  $L = M = 50$ . For every sample of the observation noise, one can gain a corresponding reconstruction of the parameters  $z_1, z_2, \dots, z_6$ , and then gain the corresponding Hausdorff distance (HD) between the numerical reconstruction and the exact boundary. Then we can use standard statistical tools to analyze these 1000 reconstructions to discuss robustness of our numerical scheme. In table 6, we exhibit the mean of reconstructions, the mean and standard deviation (SD) of Hausdorff distances (HD) at different noise levels. In figure 21, we describe the 1000 reconstructions at different noise levels  $\sigma_\eta$ . In figure 22, we show the histogram of these 1000 reconstructions, which correspond to the 1000 samples of the observation noise. We can find that the mean and standard deviation (SD) of Hausdorff distances (HD) become larger as the noise level  $\sigma_\eta$  is getting bigger. However, our inversion scheme is still robust against the noise pollution, since the mean and standard deviation (SD) of Hausdorff distances (HD) are very small as shown in table 6. Further, the phaseless data with less noise give rise to a more reliable reconstruction result.

#### 4. Conclusion

In this paper, we propose the Bayesian approach to inverse acoustic scattering from sound-soft disks, line cracks and kite-shaped obstacles with phaseless far-field data. Motivated by [41], the incoming waves are properly chosen in order to break the translational invariance of the far-field patten. Uniqueness of the inverse solution is proven for recovering a disk. When the Gaussian prior measure is given, we discuss well posedness of the posterior measure based on regularity properties of the deterministic direct scattering problem. Our numerics verify the efficiency of the preconditioned Crank–Nicolson algorithm with the random proposal variance. Further, increasing the number of incident and observation directions would lead to more accurate and reliable reconstructions. It is shown that the Bayesian method is robust for phaseless inverse scattering problems with respect to the observation noise. In this paper the obstacle boundary can be easily parameterized in a finite dimensional space. Our future efforts will be devoted to recovering the shape and physical properties of more general acoustic obstacles with a large number of unknown parameters from phaseless far-field patterns. Note that in this paper the number of unknown obstacle parameters is not larger than six, which has reduced the computational cost. For more complex scatterers, the increased computational cost of the Markov chain Monte Carlo method needs to be improved, for example, by combining the Gibbs sampling method and the stochastic surrogate model of the forward solver. Besides the idea



of using superposition of two plane wave, one can also make use of a single spherical incident wave within the Bayesian framework. Research outcomes along these directions will be reported in our forthcoming publications.

## ORCID iDs

Zhipeng Yang  <https://orcid.org/0000-0002-1886-8866>

Guanghui Hu  <https://orcid.org/0000-0002-8485-9896>

## References

- [1] Ammari H, Chow Y T and Zou J 2016 Phased and phaseless domain reconstructions in the inverse scattering problem via scattering coefficients *SIAM J. Appl. Math.* **76** 1000–30
- [2] Bao G, Li P and Lv J 2013 Numerical solution of an inverse diffraction grating problem from phaseless data *J. Opt. Soc. Am. A* **30** 293–9
- [3] Baussard A, Prémel D and Venard O 2001 A Bayesian approach for solving inverse scattering from microwave laboratory-controlled data *Inverse Problems* **17** 1659
- [4] Brooks S, Gelman A, Jones G and Meng X 2011 *Handbook of Markov Chain Monte Carlo* (Boca Raton, FL: CRC Press)
- [5] Bui-Thanh T and Ghattas O 2014 An analysis of infinite dimensional Bayesian inverse shape acoustic scattering and its numerical approximation *SIAM/ASA J. Uncertain. Quantification* **2** 203–22
- [6] Cakoni F and Colton D 2014 *A Qualitative Approach to Inverse Scattering Theory* (Berlin: Springer Science & Business Media) vol 188
- [7] Chapko R and Kress R 1993 *On a Quadrature Method for a Logarithmic Integral Equation of the First Kind* (Singapore: World Scientific) vol 2
- [8] Colton D and Kress R 2013 *Inverse Acoustic and Electromagnetic Scattering Theory* 3rd edn (Berlin: Springer) vol 93
- [9] Cotter S L, Roberts G O, Stuart A M and White D 2013 MCMC methods for functions: Modifying old algorithms to make them faster *Stat. Sci.* **28** 424–46
- [10] Fitzpatrick B G 1991 Bayesian analysis in inverse problems *Inverse Problems* **7** 675
- [11] Gamerman D and Lopes H F 2006 *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference* 2nd edn (London: Chapman and Hall/CRC)
- [12] Gao P, Dong H and Ma F 2018 Inverse scattering via nonlinear integral equations method for a sound-soft crack with phaseless data *Appl. Math.* **63** 149–65
- [13] Geyer C J 1992 Practical Markov Chain Monte Carlo *Stat. Sci.* **7** 473–83
- [14] Harris I and Rome S 2017 Near field imaging of small isotropic and extended anisotropic scatterers *Appl. Anal.* **96** 1713–36
- [15] Hastings W K 1970 Monte Carlo sampling methods using Markov chains and their applications *Biometrika* **57** 97–109
- [16] Iglesias M A, Lin K and Stuart A M 2014 Well-posed Bayesian geometric inverse problems arising in subsurface flow *Inverse Problems* **30** 114001
- [17] Iivanyshyn O 2007 Shape reconstruction of acoustic obstacles from the modulus of the far field pattern *Inverse Probl. Imag.* **1** 609–22
- [18] Iivanyshyn O and Kress R 2006 Nonlinear integral equations in inverse obstacle scattering *Mathematical Methods in Scattering Theory and Biomedical Engineering* ed D I Fotiatis and C V Massalas pp 39–50
- [19] Iivanyshyn O and Kress R 2008 Inverse scattering for planar cracks via nonlinear integral equations *Math. Methods Appl. Sci.* **31** 1221–32
- [20] Iivanyshyn O and Kress R 2010 Identification of sound-soft 3D obstacles from phaseless data *Inverse Probl. Imag.* **4** 131–49
- [21] Ji X, Liu X and Zhang B 2019 Inverse acoustic scattering with phaseless far field data: uniqueness, phase retrieval, and direct sampling methods *SIAM J. Imag. Sci.* **12** 1163–89
- [22] Kaipio J and Somersalo E 2006 *Statistical and Computational Inverse Problems* (Berlin: Springer Science & Business Media) vol 160

- [23] Kirsch A 2011 *An Introduction to the Mathematical Theory of Inverse Problems* (Berlin: Springer Science & Business Media) vol 120
- [24] Kirsch A and Grinberg N 2008 *The Factorization Method for Inverse Problems*. (New York: Oxford University Press)
- [25] Klibanov M V 2014 Phaseless inverse scattering problems in three dimensions *SIAM J. Appl. Math.* **74** 392–410
- [26] Klibanov M V 2017 A phaseless inverse scattering problem for the 3-D helmholtz equation *Inverse Probl. Imag.* **11** 263–76
- [27] Kress R 1995 Fréchet differentiability of the far field operator for scattering from a crack *J. Inverse Ill-Posed Probl.* **3** 305–13
- [28] Kress R 1995 Inverse scattering from an open arc *Math. Methods Appl. Sci.* **18** 267–93
- [29] Kress R and Rundell W 1997 Inverse obstacle scattering with modulus of the far field pattern as data *Inverse Problems in Medical Imaging and Nondestructive Testing* ed H W Engl et al pp 75–92
- [30] Kress R and Rundell W 2005 Nonlinear integral equations and the iterative solution for an inverse boundary value problem *Inverse Problems* **21** 1207–23
- [31] Law K, Stuart A M and Zygalakis K 2015 *Data Assimilation: A Mathematical Introduction* (Texts in Applied Mathematics vol 62) (Berlin: Springer International Publishing)
- [32] Li Z, Deng Z and Sun J 2019 Limited aperture inverse scattering problems using Bayesian approach and extended sampling method (arXiv:1905.12222)
- [33] Liu J, Liu Y and Sun J 2019 An inverse medium problem using Stekloff eigenvalues and a bayesian approach (arXiv:1903.05787)
- [34] Liu X and Zhang B 2010 Unique determination of a sound-soft ball by the modulus of a single far field datum *J. Math. Anal. Appl.* **365** 619–24
- [35] Maleki M H and Devaney A J 1993 Phase-retrieval and intensity-only reconstruction algorithms for optical diffraction tomography *J. Opt. Soc. Am. A* **10** 1086–92
- [36] Metropolis N, Rosenbluth A W, Rosenbluth M N, Teller A H and Teller E 1953 Equation of state calculations by fast computing machines *J. Chem. Phys.* **21** 1087–92
- [37] Nakamura G and Potthast R 2015 *Inverse Modeling: An Introduction to the Theory and Methods of Inverse Problems and Data Assimilation* (Bristol UK: IOP Publishing) pp 2053–563
- [38] Rick D 2010 *Probability: Theory and Examples* 4th edn (Cambridge: Cambridge University Press)
- [39] Stuart A M 2010 Inverse problems: a Bayesian perspective *Acta Numer.* **19** 451–559
- [40] Wang Y, Ma F and Zheng E 2015 Bayesian method for shape reconstruction in the inverse interior scattering problem *Math. Probl Eng.* 2015
- [41] Xu X, Zhang B and Zhang H 2018 Uniqueness in inverse scattering problems with phaseless far-field data at a fixed frequency *SIAM J. Appl. Math.* **78** 1737–53
- [42] Zhang B and Zhang H 2018 Fast imaging of scattering obstacles from phaseless far-field measurements at a fixed frequency *Inverse Problems* **34** 104005
- [43] Zhang D and Guo Y 2018 Uniqueness results on phaseless inverse acoustic scattering with a reference ball *Inverse Problems* **34** 085002