

An efficient Bayesian neural network surrogate algorithm for shape detection

M. Ganesh¹

S. C. Hawkins²

N. Kordzakhia³

S. Unicomb⁴

(Received 27 December 2020; revised 2 December 2021)

Abstract

We present an efficient Bayesian algorithm for identifying the shape of an object from noisy far field data. The data is obtained by illuminating the object with one or more incident waves. Bayes' theorem provides a framework to find a posterior distribution of the parameters that determine the shape of the scatterer. We compute the distribution using the Markov Chain Monte Carlo (MCMC) method with a Gibbs sampler. The principal novelty of this work is to replace the forward far-field-ansatz wave model (in an unbounded region) in the MCMC sampling with a neural-network-based surrogate that is hundreds of times faster to evaluate. We demonstrate the accuracy and efficiency of our algorithm by constructing the distributions, medians and confidence intervals of non-convex shapes using a Gaussian random circle prior.

DOI:10.21914/anziamj.v62.16110, © Austral. Mathematical Soc. 2022. Published 2022-01-12, as part of the Proceedings of the 19th Biennial Computational Techniques and Applications Conference. ISSN 1445-8810. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to the DOI for this article.

Contents

1	Introduction	C113
2	Wave scattering forward model	C116
3	BIE reformulation and discretisation	C117
4	Hierarchical Bayesian model	C118
5	Neural-network-based surrogate model	C120
6	Numerical results	C122
7	Conclusion	C125

1 Introduction

We present an efficient Bayesian algorithm for solving the inverse problem of identifying the shape of a two-dimensional object using far field data obtained by illuminating the object with one or more plane waves. Our algorithm reconstructs the shape of the object in the form of a high-dimensional data-based probability distribution and this probability distribution is interpreted by Markov Chain Monte Carlo sampling. The key to the efficiency of our algorithm is to perform the sampling using a feed-forward neural network surrogate for the wave scattering model, which is constructed offline and can subsequently be evaluated online much faster than the full wave scattering model itself.

The far-field-data-based inverse problem considered in this article is a classical one and has been investigated by several authors, especially since the development, in recent decades, of high-order algorithms to solve the associated forward wave propagation model. In particular, we refer to the celebrated book by Colton and Kress [2] for the extensive developments, over decades, of

both the theory and computational methods from a wave scattering modeling perspective, including existence and uniqueness results.

The computational methods for the far-field-data-based inverse problem are mainly based on formulating the problem as a deterministic nonlinear equation and applying iterative methods, or constructing and sampling particular indicator functions for the shape [2]. These computational approaches typically provide deterministic approximate solutions to inverse problems with noisy data. In this work we take the novel approach of using Bayesian inversion to determine probability distributions for the shape of the object by modeling uncertainties in the observed far field data.

Bayesian algorithms have been established for inverse problems of partial differential equations (PDEs) posed on bounded domains [12], but they have not been widely applied to the inverse problems in scattering theory modeled by PDEs posed on unbounded regions. Recent articles [4, 7, 10] (and related references) extended the Bayesian framework to a class of unbounded region wave propagation inverse models. The limited literature for this inverse problem is probably due to the enormous computational challenge; Bayesian algorithms typically require a very large number of evaluations of the forward model to sample the high-dimensional posterior distribution, and wave scattering models are notoriously expensive to evaluate. In this work we address that issue by replacing—in the sampling of the posterior—the wave scattering model by a feed-forward neural-network-based surrogate model. Neural network (NN) based approximations to physics informed (PI) processes are of substantial recent interest [11].

For PDE based PINN-type modeling [11], the unknown in the PDE is approximated by a NN and the approximation is computed by minimizing a loss-function (LF) that is designed by the constraint to match the known information in the PDE with the unknown parts of the PDE. The PINN-LF requires application of the governing PDE operator and given constraints (such as the boundary condition) on the NN, and the final approximation is obtained using optimization techniques to minimize the PINN-LF. A symbolic representation

of the function that measures accuracy is crucial for the optimization process in NN computations to ensure the validity of the NN approximations. The minimization of the PINN-LF process is the most expensive part of the PINN and for the PDE models the main approach is to apply *automatic differentiation* (AD) methods to evaluate the PDE operator at large sampled locations in the PDE domain. Avoiding the AD and symbolic setup of the PINN-LF remains a challenging problem to ensure the efficiency of computing the NN approximations, compared to standard numerical approximations such as using the finite element approximations based on the variational formulation of the PDE that involves domain integral operators.

The investigated PINN-type modeling, with AD-based LF setup with sampling in the PDE domain, is not practical for the wave inverse problems considered in this article because the far field data is modeled using the homogeneous Helmholtz PDE posed on an unbounded region in \mathbb{R}^2 . Using the fundamental solution of the Helmholtz PDE operator and an integral ansatz of the scattered field (that involves only unknowns on the boundary of the scatterer) the problem can be reduced to boundary integral equations (BIEs) posed only on the boundary of the scatterer [2]. The BIE operator (involving certain types of singular integrals) cannot be evaluated symbolically for non-trivial scatterer shapes and hence the current PINN-type approaches are not practical, even for the part of the forward Helmholtz model considered in this article. Accordingly, we avoid such PINN-type variants in this article and, instead, we use our NN only as a surrogate, which we construct offline using a supervised training regime in which the NN is passed training data obtained from the BIE. The amount of training data required is less dimension-sensitive than for other polynomial based surrogate models that we developed for low-dimensional parameter models [4, 7].

In Section 2 we present the wave scattering forward model and in Section 3 we briefly describe a BIE reformulation of the model and discretisation. In Section 4 we describe the hierarchical Bayesian inversion method, and in Section 5 we describe our NN-based surrogate model. Numerical results in Section 6 demonstrate the effectiveness and efficiency of our approach.

2 Wave scattering forward model

We consider the interaction of a time harmonic incident plane wave

$$\mathbf{u}^i(\mathbf{x}) = e^{i\mathbf{k}\mathbf{x}\cdot\hat{\mathbf{d}}} \quad (1)$$

with a two dimensional sound-soft scatterer $\mathbf{D} \subseteq \mathbb{R}^2$. Here $\mathbf{k} = 2\pi/\lambda$ is the wavenumber and the unit vector $\hat{\mathbf{d}}$ is the direction of propagation of the plane wave. We assume that the wavelength λ of the incident wave is of the same order as the diameter of the scatterer.

The interaction of the incident wave with the scatterer induces a scattered wave \mathbf{u} , which satisfies the Helmholtz equation exterior to the closure $\overline{\mathbf{D}}$ of \mathbf{D} ,

$$\Delta\mathbf{u} + k^2\mathbf{u} = 0, \quad \mathbf{x} \in \mathbb{R}^2 \setminus \overline{\mathbf{D}}, \quad (2)$$

and the Sommerfeld radiation condition (SRC) [2, equation (3.108)]. The sound-soft characteristic of the scatterer implies that the total field $\mathbf{u} + \mathbf{u}^i$ vanishes on the boundary $\partial\mathbf{D}$ of \mathbf{D} [2], leading to the Dirichlet boundary condition

$$\mathbf{u}(\mathbf{x}) + \mathbf{u}^i(\mathbf{x}) = 0, \quad \mathbf{x} \in \partial\mathbf{D}, \quad (3)$$

where \mathbf{u}^i is the incident plane wave (1).

Using the SRC, the radiating scattered field is decomposed as

$$\mathbf{u}(\mathbf{x}) = \frac{e^{ik|\mathbf{x}|}}{\sqrt{|\mathbf{x}|}} \left[\mathbf{u}^\infty(\hat{\mathbf{x}}) + \mathcal{O}\left(\frac{1}{|\mathbf{x}|}\right) \right], \quad \hat{\mathbf{x}} = \frac{\mathbf{x}}{|\mathbf{x}|} \in \partial\mathbf{B}, \quad (4)$$

where $\partial\mathbf{B}$ denotes the unit circle, and the distance-independent function $\mathbf{u}^\infty(\hat{\mathbf{x}})$ is known as the far field of \mathbf{D} . In practical experiments the typical quantity of interest, measured at relatively long distances from the scatterer, is the acoustic cross section of \mathbf{D} ,

$$\sigma(\hat{\mathbf{x}}) = |\mathbf{u}^\infty(\hat{\mathbf{x}})|^2, \quad \hat{\mathbf{x}} \in \partial\mathbf{B}. \quad (5)$$

For a given incident direction $\widehat{\mathbf{d}}$ and scatterer \mathbf{D} , the forward problem is to compute the scattered field \mathbf{u} that satisfies the scattering problem (2)–(3) with the SRC, and the associated acoustic cross section. In this work we are interested in solving the corresponding inverse problem of computing the scatterer shape \mathbf{D} from acoustic cross section data. To that end we assume that \mathbf{D} is star shaped (without loss of generality with star-centre at the origin) and we parametrise the boundary curve $\partial\mathbf{D}$ using the local polar coordinates representation

$$\mathbf{x}(\theta) = r(\theta) \widehat{\mathbf{x}}(\theta), \quad \theta \in [0, 2\pi), \quad (6)$$

where $\widehat{\mathbf{x}}(\theta) = (\cos \theta, \sin \theta)$. We assume that the radius $r(\theta) > 0$, and it is therefore convenient to express the radius in the form

$$r(\theta) = r_0 e^{s(\theta)}, \quad \theta \in [0, 2\pi), \quad (7)$$

where r_0 is a constant whose value we assign in Section 4. We parametrise the log-radius $s(\theta)$ as

$$s(\theta) = \sum_{j=0}^N a_j \cos j\theta + \sum_{j=1}^N b_j \sin j\theta, \quad \theta \in [0, 2\pi), \quad (8)$$

where $\mathbf{a}_0, \dots, \mathbf{a}_N$ and $\mathbf{b}_1, \dots, \mathbf{b}_N$ are unknown parameters to be determined from the acoustic cross section data.

3 BIE reformulation and discretisation

Given a scatterer \mathbf{D} with known parametrisation (6)–(8), we solve the forward problem (2)–(3) using a combined field ansatz [2]

$$\mathbf{u}(\mathbf{x}) = \int_{\partial\mathbf{D}} \left[\frac{\partial \mathbf{G}}{\partial \mathbf{n}(\mathbf{y})}(\mathbf{x}, \mathbf{y}) - ik\mathbf{G}(\mathbf{x}, \mathbf{y}) \right] \phi(\mathbf{y}) ds(\mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^2 \setminus \overline{\mathbf{D}}, \quad (9)$$

which satisfies the Helmholtz PDE (2) and the SRC. Here $\mathbf{G}(\mathbf{x}, \mathbf{y})$ is the free space Green's function for the two-dimensional Helmholtz equation [2,

Section 3.5] and $\mathbf{n}(\mathbf{y})$ is the unit outward normal to ∂D at \mathbf{y} . Taking the limit of (9) as \mathbf{x} approaches ∂D using the jump relations for the single- and double-layer potentials [2, Theorem 3.1], and forcing the ansatz (9) to satisfy the boundary condition (3) shows that the unknown boundary density ϕ in (9) should be the solution of the second kind BIE [2]

$$\phi(\mathbf{x}) + 2 \int_{\partial D} \left[\frac{\partial G}{\partial \mathbf{n}(\mathbf{y})}(\mathbf{x}, \mathbf{y}) - ikG(\mathbf{x}, \mathbf{y}) \right] \phi(\mathbf{y}) \, ds(\mathbf{y}) = -2u^i(\mathbf{x}), \quad \mathbf{x} \in \partial D. \quad (10)$$

This combined field integral equation (CFIE) has a unique solution for all wavenumbers $k \in \mathbb{R}$ [2]. Once the solution ϕ of the CFIE (10) is obtained, applying the ansatz (9), the corresponding far field is computed using

$$u^\infty(\hat{\mathbf{x}}) = \frac{e^{i\pi/4}}{\sqrt{8\pi k}} \int_{\partial D} \left[\frac{\partial e^{-ik\hat{\mathbf{x}} \cdot \mathbf{y}}}{\partial \mathbf{n}(\mathbf{y})} - ike^{-ik\hat{\mathbf{x}} \cdot \mathbf{y}} \right] \phi(\mathbf{y}) \, ds(\mathbf{y}), \quad \hat{\mathbf{x}} \in \partial B. \quad (11)$$

In this work we compute high-order approximate solutions of (10) using efficient implementations from our open-source software [5] for both smooth and non-smooth curves ∂D . In particular, for the Nyström method [2] used by Ganesh and Hawkins [5] for smooth curves, the discretised form of the BIE (10) is a linear system

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad (12)$$

where the matrix \mathbf{A} is the discretisation of the integral operators on the left hand side of (10), the vector \mathbf{b} is the discretisation of the right hand side, and \mathbf{x} is a vector whose entries approximate the values of ϕ at the discretisation points. Since the integral operator depends on the scatterer boundary ∂D , the matrix \mathbf{A} must be recomputed if the boundary ∂D is changed.

4 Hierarchical Bayesian model

In the remainder of the article we consider the acoustic cross section σ associated with the solution \mathbf{u} of (2)–(3) as a function of the parameters

associated with the boundary parametrisation (6)–(8). To precisely define the inverse problem we introduce the nonlinear operator $F: \mathbb{R}^{2N+1} \rightarrow C(\partial B)$,

$$F(\boldsymbol{\omega}) = \sigma,$$

where $\boldsymbol{\omega} = (\mathbf{a}_0, \dots, \mathbf{a}_N, \mathbf{b}_1, \dots, \mathbf{b}_N) \in \mathbb{R}^{2N+1}$ parametrises ∂D via (6)–(8) and σ is computed using (5) from the solution \mathbf{u} of (2)–(3). That is, the forward model operator F maps a parametrized scatterer to the acoustic cross section (induced by the incident wave impinging on the scatterer). In this article, our focus is on the inverse problem: Find $\boldsymbol{\omega}$ such that

$$F(\boldsymbol{\omega}) = \hat{f}, \quad (13)$$

for given noisy acoustic cross section data $\hat{f} \in C(\partial B)$. Later we focus on the particular practically important case where (13) is required to hold only for the acoustic cross section evaluated in a discrete subset of ∂B , comprising a *finite* number of measurement direction points.

Our approach is to model the parameters $\boldsymbol{\omega}$ as a multivariate random variable in a probability space Ω_0 with *prior* probability distribution π_0 . The probability space and associated probability distribution are chosen to encapsulate all available knowledge of the scatterer D , and we give a concrete example in Section 5. The expected value of a function $f(\boldsymbol{\omega})$ with respect to the probability distribution π_0 is

$$\mathbb{E}[f] = \int_{\Omega_0} f(\boldsymbol{\omega}) d\pi_0(\boldsymbol{\omega}). \quad (14)$$

The Monte Carlo approximation to $\mathbb{E}[f]$ is

$$\mathbb{E}_m[f] = \frac{1}{m} \sum_{k=1}^m f(\boldsymbol{\omega}_k), \quad (15)$$

where $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_m$ are m independent samples of $\boldsymbol{\omega}$.

We assume that the noise in the data is random. The data is

$$\hat{\mathbf{f}} = \mathbf{f} + \boldsymbol{\eta}, \quad (16)$$

where \mathbf{f} is in the range of F and the noise $\boldsymbol{\eta}$ has probability distribution π . Then Bayes' theorem provides a data-conditioned *posterior* probability distribution $\pi_{\hat{\mathbf{f}}}(\boldsymbol{\omega}, \sigma)$ for $(\boldsymbol{\omega}, \sigma)$ satisfying

$$\pi_{\hat{\mathbf{f}}}(\boldsymbol{\omega}, \sigma) \propto \pi(\hat{\mathbf{f}} - F(\boldsymbol{\omega})|\sigma) \pi_0(\boldsymbol{\omega}) \pi_1(\sigma). \quad (17)$$

The constant of proportionality in (17) normalises $\pi_{\hat{\mathbf{f}}}$ but is not required when the posterior distribution is sampled using a Markov Chain Monte Carlo (MCMC) method. In this work we sample the posterior distribution (17) by MCMC method using the Gibbs algorithm because it readily provides marginal posterior probability distributions for each of the parameters in $\boldsymbol{\omega}$.

In practice we discretise (13) by requiring it to hold for N_θ equally spaced points $\hat{\mathbf{x}}(\theta)$ corresponding to discrete angles $\theta = \theta_1, \dots, \theta_{N_\theta}$. Accordingly $\hat{\mathbf{f}}$ reduces to a vector and $\boldsymbol{\eta}$ reduces to a multivariate random variable. We assume that the components of $\boldsymbol{\eta}$ have identical Gaussian distributions with variance σ^2 , where σ is a random variable with probability distribution π_1 .

5 Neural-network-based surrogate model

For the parameter prior $\pi_0(\boldsymbol{\omega})$ in (17) we make use of a class of random particles that have been well studied in the literature [6, 8, 9, 13] and utilised as models of atmospheric particles such as ice and dust. In particular, we assign to the coefficients in (8) the distributions $\mathbf{a}_j, \mathbf{b}_j \sim \mathcal{N}(\mathbf{0}, \delta_k^2)$ with

$$\delta_j = \begin{cases} \mathbf{a}, & \text{for } j = 0, \\ 0, & \text{for } j = 1, \\ \alpha \mathbf{j}^{-\nu}, & \text{for } j \geq 2, \end{cases} \quad (18)$$

where $\mathbf{a} = \log(10)/10$. Here ν acts as a regularisation parameter that controls the smoothness of the particle boundary and α is a normalisation parameter

chosen so that

$$\sum_{j=2}^N \delta_j^2 = \delta^2, \quad \delta^2 = \log(1 + \sigma^2),$$

where σ^2 determines the variance of the radius of the particles. Thus the prior π_0 involves parametrising $\partial\mathbf{D}$ using $\mathbf{d} = 2N - 1$ stochastic dimensions.

We set the scaling parameter $r_0 = 1/\sqrt{1 + \sigma^2}$ in (7) and, following arguments by Ganesh and Hawkins [6], the radius of $\partial\mathbf{D}$ has mean $\mathbb{E}[r(\boldsymbol{\theta})] = e^{\alpha^2/2} = e^{\delta_0^2/2}$, and covariance

$$\text{Cov}[r(\boldsymbol{\theta}_1), r(\boldsymbol{\theta}_2)] = e^{\text{Cov}[s(\boldsymbol{\theta}_1), s(\boldsymbol{\theta}_2)]} - e^{\delta_0^2},$$

where

$$\text{Cov}[s(\boldsymbol{\theta}_1), s(\boldsymbol{\theta}_2)] = \delta_0^2 + \sum_{j=0}^N \delta_j^2 \cos j(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2).$$

Each sample of the posterior (17) using the Gibbs sampler requires hundreds of evaluations of the forward model $F(\boldsymbol{\omega})$ for each of the $\mathbf{d} + 1$ dimensions of $(\boldsymbol{\omega}, \sigma)$. Here the forward model requires assembly and solution of the linear system (12), and we show in Section 6 that the CPU time is unfeasibly large to sample directly using the full forward model.

In previous works for $\mathbf{d} \leq 4$ we accelerated the MCMC sampling by replacing the full forward model with a cheap-to-compute surrogate model based on an approximation of $F(\boldsymbol{\omega})$ using degree L polynomials [4, 7]. In this work, motivated by the challenges of using polynomial-based surrogates for problems with high stochastic dimension, we develop a neural-network-based surrogate

$$F(\boldsymbol{\omega}) \approx Q_M^{\text{ML}}(\boldsymbol{\omega}). \quad (19)$$

The surrogate model Q_M^{ML} is computed by a feed-forward neural network with a hidden layer of size N_θ . The activation function for the hidden layer is the hyperbolic-tangent sigmoid function, and the activation function for the output layer is linear. We train the neural network using training data

$$(\boldsymbol{\omega}_m, F(\boldsymbol{\omega}_m)), \quad m = 1, \dots, M, \quad (20)$$

where $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_M$ are M independent samples from the prior π_0 . The neural network is trained using the Levenberg–Marquardt algorithm in the Matlab (2017b) Neural Network Toolbox (the `train` function).

6 Numerical results

We demonstrate our Bayesian NN (BNN) surrogate algorithm with noisy test data obtained by simulating far fields from two smooth non-convex shapes and a non-smooth scatterer: a trefoil, a pinched ball and a square. The smooth case simulations use the BIE Nyström method discussed in Section 3, but we avoid the ‘inverse crime’ by using a different parametrisation of the shapes than can be obtained using (6)–(8). The simulation for the square case uses the T-matrix computed using TMATROM [5]. In all our experiments we use parametrisations with order $N = 6$ and stochastic dimension $\mathbf{d} = 11$ and, except where stated otherwise, we fix the wavenumber $k = \pi$ with corresponding wavelength $\lambda = 2$, so that the scatterers are approximately one wavelength in diameter.

To improve the reconstructions we generalise the forward model in Section 4 by combining data from four plane waves with incident directions $\hat{\mathbf{d}} = \pm(1, 0), \pm(0, 1)$. The implementation of the Bayesian algorithm and the surrogate models is identical. For each incident wave we measure the acoustic cross section at N_θ equally spaced angles, so that our total data dimension is $4N_\theta$. Except where stated otherwise we fix $N_\theta = 12$ and our total data dimension is 48. We generate corresponding noisy data by adding Gaussian noise with standard deviation two percent of the maximum data value.

In Table 1 we demonstrate the performance of our BNN surrogate model with respect to the amount of training data by tabulating the CPU time for generating the training data, the CPU time for training the model, and the relative error in the surrogate model. For fair comparison with the non-surrogate version of the model, in our numerical experiments we avoided parallelization of the training data generation. Matlab automatically parallelises the NN

Table 1: CPU time and relative error for the NN surrogate, with training data size M .

M	CPU fun. (s)	CPU train. (s)	rel. error (ρ_{1000})
200	10.5	68.5	2.3×10^{-1}
400	20.6	133.3	1.9×10^{-1}

training. The CPU times in this section are obtained using a a four-core 2.6 GHz Intel i7 CPU. In contrast to the typical NN framework, the domain for our surrogate model is the (unbounded) set of outcomes in the probability space associated with the prior, and it is appropriate and robust to measure the error in the norm induced by the probability measure at a larger number ($m > M$) of samples than the training data size M in (20). Accordingly, we estimate the relative error in the surrogate model $Q(\boldsymbol{\omega})$ using the Monte Carlo method,

$$\rho_m = \sqrt{\frac{\frac{1}{m} \sum_{j=1}^m (F(\boldsymbol{\omega}_j) - Q(\boldsymbol{\omega}_j))^2}{\frac{1}{m} \sum_{j=1}^m (F(\boldsymbol{\omega}_j))^2}},$$

where $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_m$ are m independent samples of the prior distribution.

In Figure 1 we visualise the computed posterior distributions for our three shapes by plotting the median of 10 000 samples and indicating the 95.45% confidence interval. The CPU time required for drawing 10 000 samples with the $M = 400$ neural network is about 23 minutes; the estimated CPU time required for the 10 000 samples using the full wave scattering BIE model would be more than 80 days.

The non-smooth square geometry demonstrates the capability of our algorithm to identify challenging shapes. This shape is challenging because it is not in the set of smooth particles described by the prior in Section 5, and sub-wavelength features like the corners are difficult to resolve because of the uncertainty principle [1]. In Figure 2 we demonstrate that using more data (larger N_θ) and using higher frequency incident waves can both improve the reconstruction of the square shape.

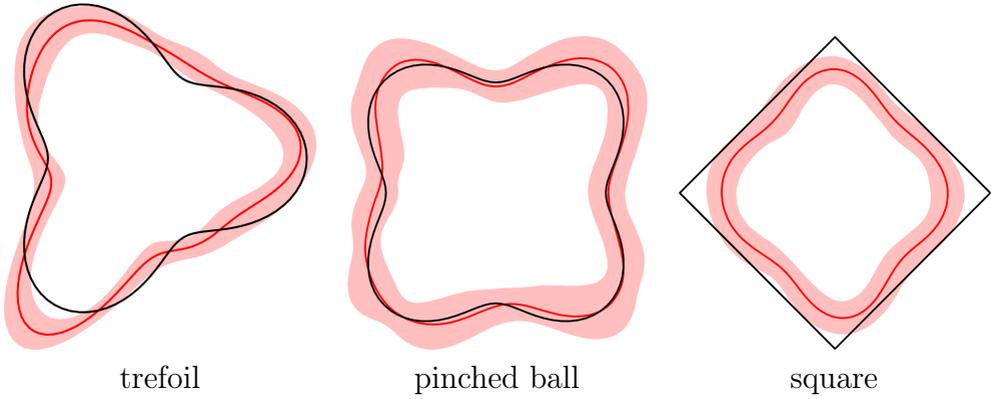


Figure 1: Reconstructions of various shapes using wavelength $\lambda = 2$ and $N_\theta = 12$. The true shapes (black lines) and visualisations of the corresponding computed posterior distributions (17) showing the median (red line) and the 95.45% confidence interval (shading). The posterior is computed using the neural-network-based surrogate with training data of length $M = 400$.

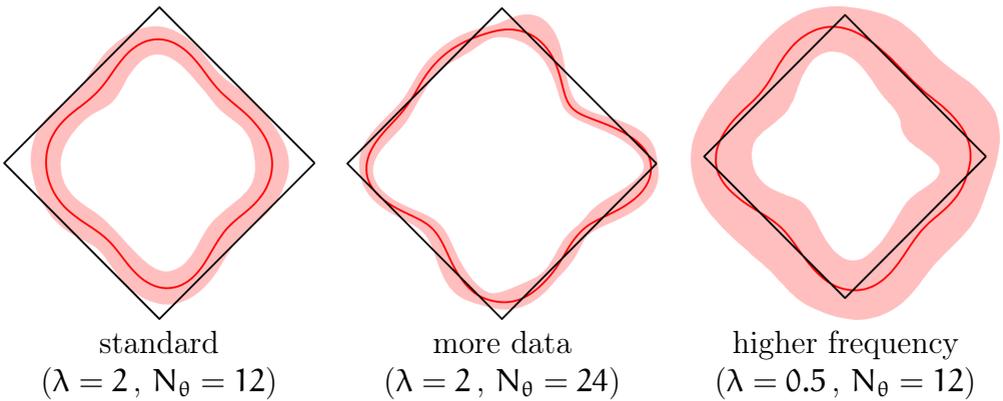


Figure 2: Reconstructions of the square using more data or higher frequency. The true shapes (black lines) and visualisations of the corresponding computed posterior distributions (17) showing the median (red line) and the 95.45% confidence interval (shading). The posterior is computed using the neural-network-based surrogate with training data of length $M = 400$.

7 Conclusion

We have computationally explored the advantages of using NN-based approximations, within the Bayesian setting, to efficiently solve a two dimensional (2D) inverse wave propagation model in an unbounded region exterior to a scatterer. The standard Bayesian framework for the inverse model requires simulation of a large number of forward wave problems in the unbounded region and hence, in general, is computationally prohibitive. We have demonstrated that our Bayesian NN surrogate algorithm facilitates massive reduction in the CPU time, thereby making such simulations practical. Our approach provides a framework for future extension to the three dimensional (3D) problem, and the scalability depends on a similar surrogate for the 3D forward model, using algorithms and software such as those developed by Ganesh and Hawkins [5, 7].

A theoretical understanding of NN approximations in the context of real-world model simulations is an open problem. A basic theoretical framework for the future development of numerical analysis for PDE-based models can be found in the recent survey by DeVore et al. [3] and references therein. Such analysis needs to be first developed for forward PDE models on bounded domains, based on substantial PINN-based computational results observed in recent years. Future NN-based computational developments for unbounded regions will be needed to extend the numerical analysis to inverse models such as those considered in this article.

References

- [1] Y. Chen. “Inverse scattering via Heisenberg’s uncertainty principle”. In: *Inv. Prob.* 13 (1997), pp. 253–282. DOI: [10.1088/0266-5611/13/2/005](https://doi.org/10.1088/0266-5611/13/2/005) (cit. on p. C123).
- [2] D. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*. 4th Edition. Vol. 93. Applied Mathematical Sciences.

- Springer, 2019. DOI: [10.1007/978-3-030-30351-8](https://doi.org/10.1007/978-3-030-30351-8) (cit. on pp. [C113](#), [C114](#), [C115](#), [C116](#), [C117](#), [C118](#)).
- [3] R. DeVore, B. Hanin, and G. Petrova. “Neural Network Approximation”. In: *Acta Num.* 30 (2021), pp. 327–444. DOI: [10.1017/S0962492921000052](https://doi.org/10.1017/S0962492921000052) (cit. on p. [C125](#)).
- [4] M. Ganesh and S. C. Hawkins. “A reduced-order-model Bayesian obstacle detection algorithm”. In: *2018 MATRIX Annals*. Ed. by J. de Gier et al. Springer, 2020, pp. 17–27. DOI: [10.1007/978-3-030-38230-8_2](https://doi.org/10.1007/978-3-030-38230-8_2) (cit. on pp. [C114](#), [C115](#), [C121](#)).
- [5] M. Ganesh and S. C. Hawkins. “Algorithm 975: TMATROM—A T-matrix reduced order model software”. In: *ACM Trans. Math. Softw.* 44.9 (2017), pp. 1–18. DOI: [10.1145/3054945](https://doi.org/10.1145/3054945) (cit. on pp. [C118](#), [C122](#), [C125](#)).
- [6] M. Ganesh and S. C. Hawkins. “Scattering by stochastic boundaries: hybrid low- and high-order quantification algorithms”. In: *ANZIAM J.* 56 (2016), pp. C312–C338. DOI: [10.21914/anziamj.v56i0.9313](https://doi.org/10.21914/anziamj.v56i0.9313) (cit. on pp. [C120](#), [C121](#)).
- [7] M. Ganesh, S. C. Hawkins, and D. Volkov. “An efficient algorithm for a class of stochastic forward and inverse Maxwell models in \mathbb{R}^3 ”. In: *J. Comput. Phys.* 398 (2019), p. 108881. DOI: [10.1016/j.jcp.2019.108881](https://doi.org/10.1016/j.jcp.2019.108881) (cit. on pp. [C114](#), [C115](#), [C121](#), [C125](#)).
- [8] L. Lamberg, K. Muinonen, J. Ylönen, and K. Lumme. “Spectral estimation of Gaussian random circles and spheres”. In: *J. Comput. Appl. Math.* 136 (2001), pp. 109–121. DOI: [10.1016/S0377-0427\(00\)00578-1](https://doi.org/10.1016/S0377-0427(00)00578-1) (cit. on p. [C120](#)).
- [9] T. Nousiainen and G. M. McFarquhar. “Light scattering by quasi-spherical ice crystals”. In: *J. Atmos. Sci.* 61 (2004), pp. 2229–2248. DOI: [10.1175/1520-0469\(2004\)061<2229:LSBQIC>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<2229:LSBQIC>2.0.CO;2) (cit. on p. [C120](#)).

- [10] A. Palafox, M. A. Capistrán, and J. A. Christen. “Point cloud-based scatterer approximation and affine invariant sampling in the inverse scattering problem”. In: *Math. Meth. Appl. Sci.* 40 (2017), pp. 3393–3403. DOI: [10.1002/mma.4056](https://doi.org/10.1002/mma.4056) (cit. on p. [C114](#)).
- [11] M. Raissi, P. Perdikaris, and G. E. Karniadakis. “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations”. In: *J. Comput. Phys.* 378 (2019), pp. 686–707. DOI: [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045) (cit. on p. [C114](#)).
- [12] A. C. Stuart. “Inverse problems: A Bayesian perspective”. In: *Acta Numer.* 19 (2010), pp. 451–559. DOI: [10.1017/S0962492910000061](https://doi.org/10.1017/S0962492910000061) (cit. on p. [C114](#)).
- [13] B. Veihelmann, T. Nousiainen, M. Kahnert, and W. J. van der Zande. “Light scattering by small feldspar particles simulated using the Gaussian random sphere geometry”. In: *J. Quant. Spectro. Rad. Trans.* 100 (2006), pp. 393–405. DOI: [10.1016/j.jqsrt.2005.11.053](https://doi.org/10.1016/j.jqsrt.2005.11.053) (cit. on p. [C120](#)).

Author addresses

1. **M. Ganesh**, Department of Applied Mathematics and Statistics, Colorado School of Mines, USA.
2. **S. C. Hawkins**, School of Mathematical and Physical Sciences, Macquarie University, NSW 2109, AUSTRALIA.
<mailto:stuart.hawkins@mq.edu.au>
3. **N. Kordzakhia**, School of Mathematical and Physical Sciences, Macquarie University, NSW 2109, AUSTRALIA.
4. **S. Unicom**, School of Mathematical and Physical Sciences, Macquarie University, NSW 2109, AUSTRALIA.