FLATIRON INSTITUTE
Center for Computational Mathematics

A Neural Network Warm-Start Approach for Inverse Scattering Problems

Jiequn Han
Center of Computational Mathematics
Flatiron Institute, Simons Foundation

Deep Learning and Numerical PDEs
June 23, 2023

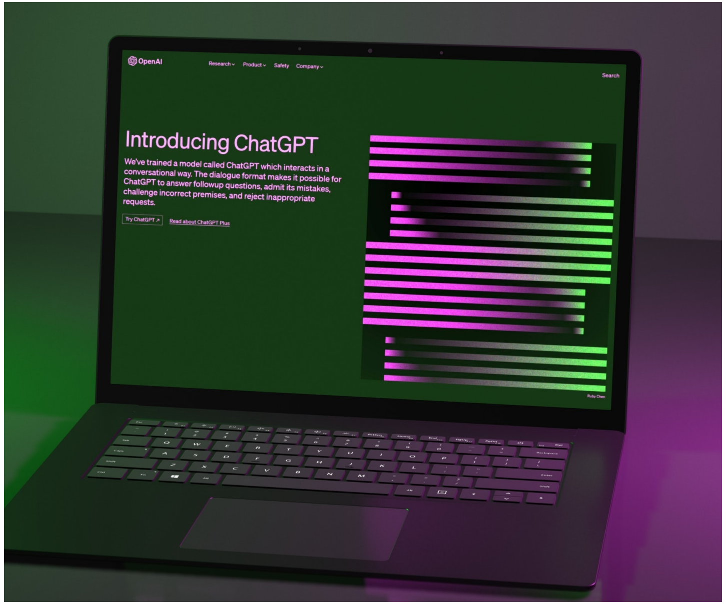
Mo Zhou

Manas Rachh

Carlos Borges

Leslie Greengard

ChatGPT - For Math?



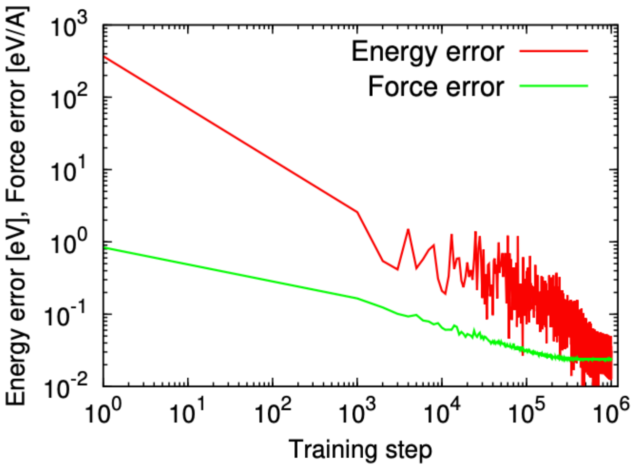
Terence Tao
@tao@mathstodon.xyz

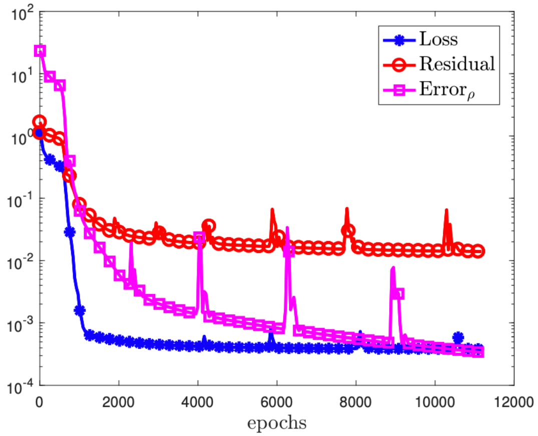
Here is another interesting example (using a SMS interface created by my son). The AI generated "proof" of the infinitide of primes os not quite correct, though this is only apparent on a closer reading, and in fact the argument can be fixed to give a proof that I for one had not seen before: the AI argument *does* imply that the infinitude of squarefree numbers implies the infinitude of primes, and the former statement can be proven by a standard sieve argument.

In general, I am finding that while these #AI tools do not directly assist me in core tasks such as trying to attack an unsolved mathematical problem, they are quite useful for a wide variety of peripheral (but still work-related) tasks (though often with some manual tweaking afterwards).

Deep learning solutions work to some extent, but not entirely!

An "Accuracy Wall" of Deep Learning





Accuracy wall: trained neural networks cannot achieve accuracy close to machine precision due to limitations of

1. architecture (approximation error)
2. data (noise and distribution)
3. optimization (non-convexity and the nature of stochastic optimization)

Is the "Accuracy Wall" a Concern?

No, if

- The accuracy of previous methods is rather low, and the accuracy improvement brought by deep learning is substantial. Example: AlphaFold

- The required accuracy is relatively low, achievable for deep learning with much lower computation cost after training. Example: DeePMD ("chemical accuracy")

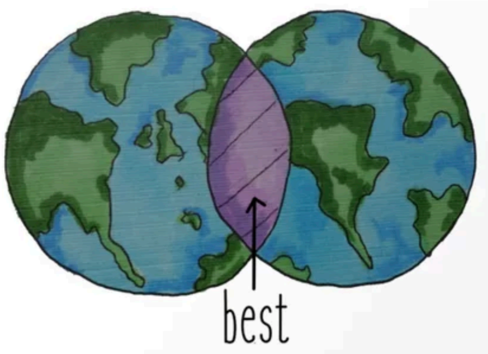
⚠️ Yes, if the problem requires high precision, which is common in many physical simulations and related inference problems.

FLATIRON INSTITUTE
Center for Computational Mathematics

Best of Both Worlds: Warm-Start Approach

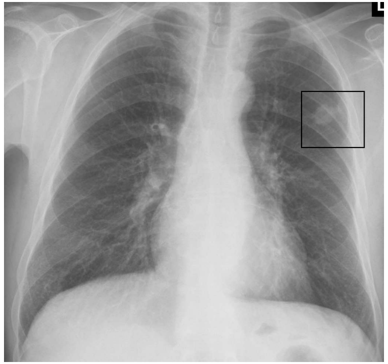
Classical numerical methods have been developed to achieve high accuracy but often face other limitations. Can deep learning help? How?

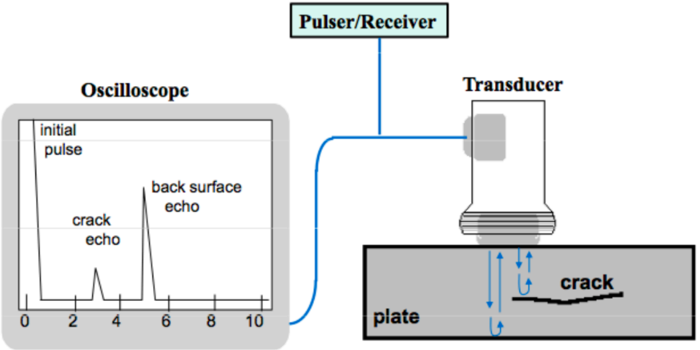
Neural network warm-start: use deep learning to provide proper initial guess that allows classical solvers to find a good solution more easily and quickly.



Inverse scattering problem: providing initial guess to help identify the basin of attraction of the true solution

Inverse Scattering Problem



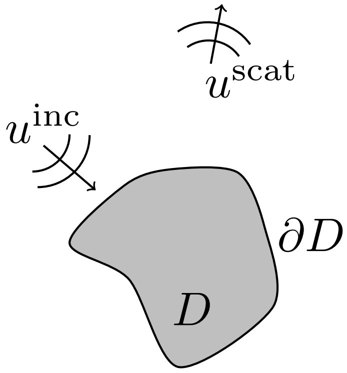


Inverse scattering problems arise in various fields including medical imaging, non-destructive testing, sensing, oil/gas prospecting, radar and sonar.

In these problems, a set of measured data is collected from experiments with the goal of reconstructing the object or its properties.

Forward Scattering Problem

$$\begin{cases} \Delta u^{\text{scat}} + k^2 u^{\text{scat}} = 0, \quad \text{in } \mathbb{R}^2 \setminus \overline{D} & \text{(Helmholtz equation)} \\ u^{\text{scat}} + u^{\text{inc}} = 0 \quad \text{on } \partial D & \text{(physical property of the scatter)} \\ \lim_{r\to\infty} r^{1/2} \left(\frac{\partial u^{\text{scat}}}{\partial r} - iku^{\text{scat}} \right) = 0 & \text{(Sommerfeld radiation condition)} \end{cases}$$



Input:

D: scatter object

$u^{\text{inc}} = e^{ikx\cdot d}$: incident plane wave $(\mathbb{R}^2 \to \mathbb{C})$

k : frequency/wavenumber; d: incident direction

Output:

u^{scat} : scattered field $(\mathbb{R}^2 \to \mathbb{C})$

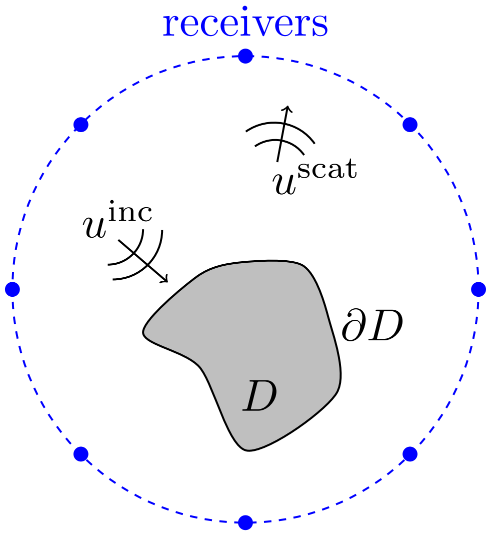
Inverse Scattering Problem

$$\begin{cases} \Delta u^{\text{scat}} + k^2 u^{\text{scat}} = 0, \quad \text{in } \mathbb{R}^2 \setminus \overline{D} \\ u^{\text{scat}} + u^{\text{inc}} = 0 \quad \text{on } \partial D \\ \lim_{r\to\infty} r^{1/2}\left(\frac{\partial u^{\text{scat}}}{\partial r} - iku^{\text{scat}}\right) = 0 \end{cases}$$

(Helmholtz equation)

(physical property of the scatter)

(Sommerfeld radiation condition)



Input:

multiple incident waves u^{inc} with different incident directions d, at frequency k corresponding measurement $u^{\text{meas}}_{k,d} \in \mathbb{C}^{N_t}$

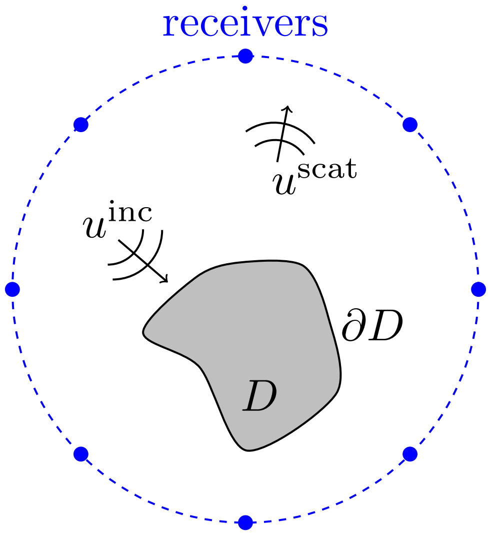
Output:

∂D: boundary/shape of the scatter

Optimization Formulation

Given a single frequency k, N_t receivers, and N_d incident directions ($d_\ell, l = 1,\ldots, N_d$), let $\mathcal{F}_k$ denote the forward scattering operator

$$\mathcal{F}_k(\partial D) = u^{\text{meas}} := [u^{\text{meas}}_{k,d_1}; \ldots, u^{\text{meas}}_{k,d_{N_d}}] \in \mathbb{C}^{N_t N_d}$$



We then search for the shape that minimizes

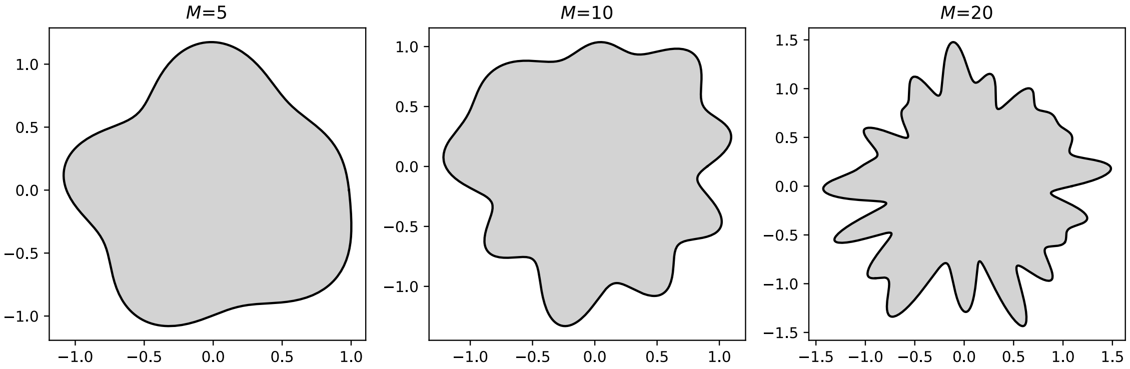
$$\widetilde{\partial D} = \arg\min_{\partial D} \|u^{\text{meas}} - \mathcal{F}_k(\partial D)\|^2 .$$

To make it computable, we need to further constrain the optimization space.

Star-Shaped Family

Restricting the boundary to the star-shaped family (in polar coordinates (r,θ)):

$$r(\theta;\boldsymbol{c}) = c_0 + \sum_{m=1}^{M} \left(c_m \cos(m\theta) + c_{m+M} \sin(m\theta)\right), \quad \boldsymbol{c} \in \mathbb{R}^{2M+1}$$



we have the finite-dimensional optimization problem

$$\tilde{\boldsymbol{c}} = \underset{\substack{\boldsymbol{c} \\ r(t;\boldsymbol{c})>0\ \forall t\in[0,2\pi)}}{\arg\min} \left\| u^{\text{meas}} - \mathcal{F}_k(\boldsymbol{c}) \right\|^2.$$

Gauss-Newton Method

We use the Gauss-Newton method to solve

$$\tilde{\boldsymbol{c}} = \underset{\substack{\boldsymbol{c} \\ r(t;\boldsymbol{c})>0 \ \forall t\in[0,2\pi)}}{\arg\min} \left\| u^{\text{meas}} - \mathcal{F}_k(\boldsymbol{c}) \right\|^2 .$$

We iteratively compute

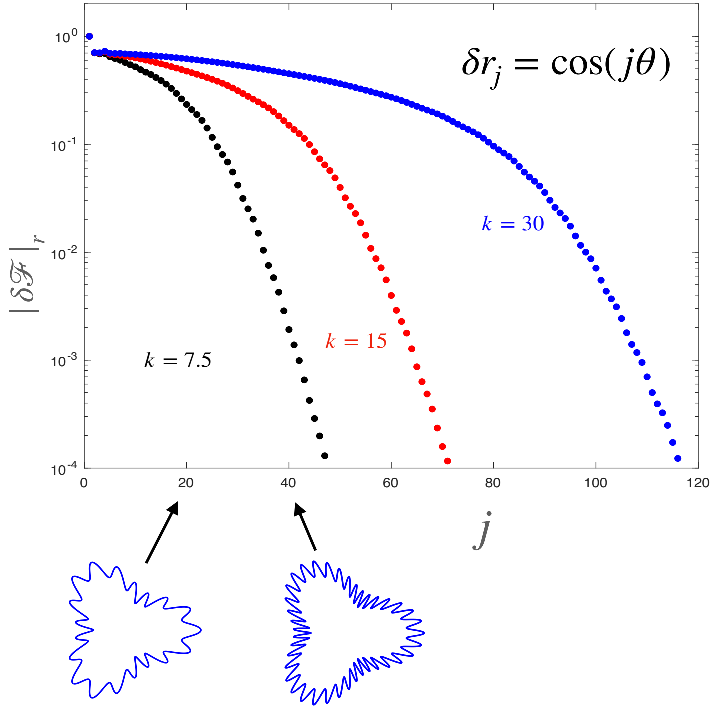
$$\boldsymbol{c}^{(j+1)} = \boldsymbol{c}^{(j)} + \delta\boldsymbol{c}^{(j)}$$

$$[\,\mathrm{Re}(J)\,;\mathrm{Im}(J)\,]\delta\boldsymbol{c}^{(j)} = [\,\mathrm{Re}(u^{\text{meas}} - \mathcal{F}_k(\boldsymbol{c}^{(j)}))\,;\,\mathrm{Im}(u^{\text{meas}} - \mathcal{F}_k(\boldsymbol{c}^{(j)}))\,]$$

$J \in \mathbb{C}^{(N_t N_d)\times(2M+1)}$ is the Fréchet derivative matrix $\partial\mathcal{F}_k$ (Jacobian matrix).

The product $\partial\mathcal{F}_k(\boldsymbol{c})\delta\boldsymbol{c}$ can be solved from the Helmholtz equation with a different boundary condition (adjoint method).

Ill-Posedness in Optimization



Magnitude of Fréchet derivative $|\delta \mathscr{F}|_r$ for $\delta r_j := \cos(j\theta)$

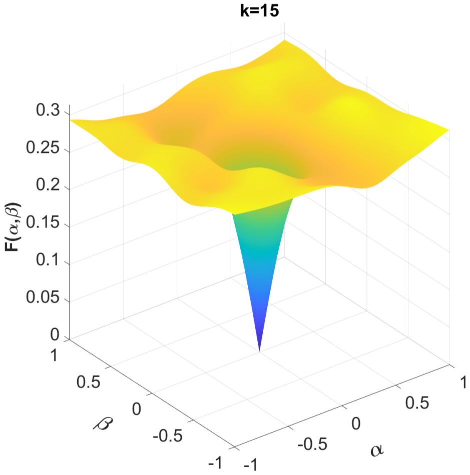
High-frequency waves are needed to recover small-scale features (Heisenberg's uncertainty principle)

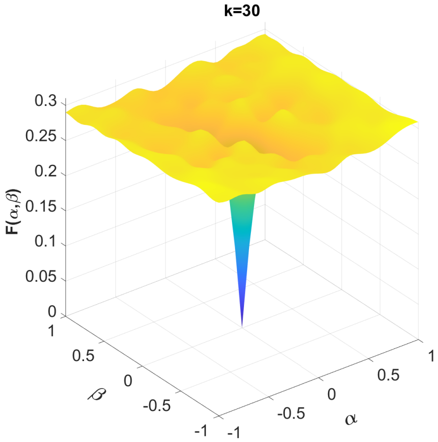
For a fixed frequency k

- The Fréchet derivative is $O(1)$ if the oscillation in the perturbation is less than $O(k)$.
- The Fréchet derivative decreases exponentially fast for oscillation with higher frequency.

Non-convexity in Optimization

However, as frequency increases, the loss landscape becomes more and more non-convex, with more and more bad local minima.





The Gauss-Newton method's success highly relies on the quality of the initial guess.

Non-convexity in Optimization (continued)

Recursive linearization*: using multifrequency measurement data to address non-convexity in the inverse problems, a continuation procedure in frequency.

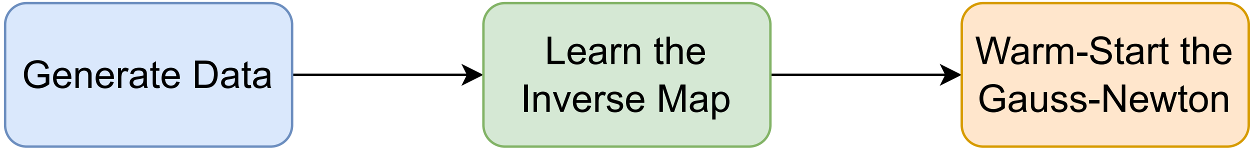
However, in experiments, it is quite infeasible to generate incident waves and collect measurements at low frequencies.

All the computation before the highest frequency essentially only contributes to the initial guess at the highest frequency.

Deep neural networks can provide an initial guess of high-quality!

* *Inverse scattering via Heisenberg's uncertainty principle*, Chen (1997)

Neural Network Warm-Start Approach



1. Generate data by solving forward scattering problems with different star-shaped obstacles parameterized by Fourier coefficients $\boldsymbol{c}$.

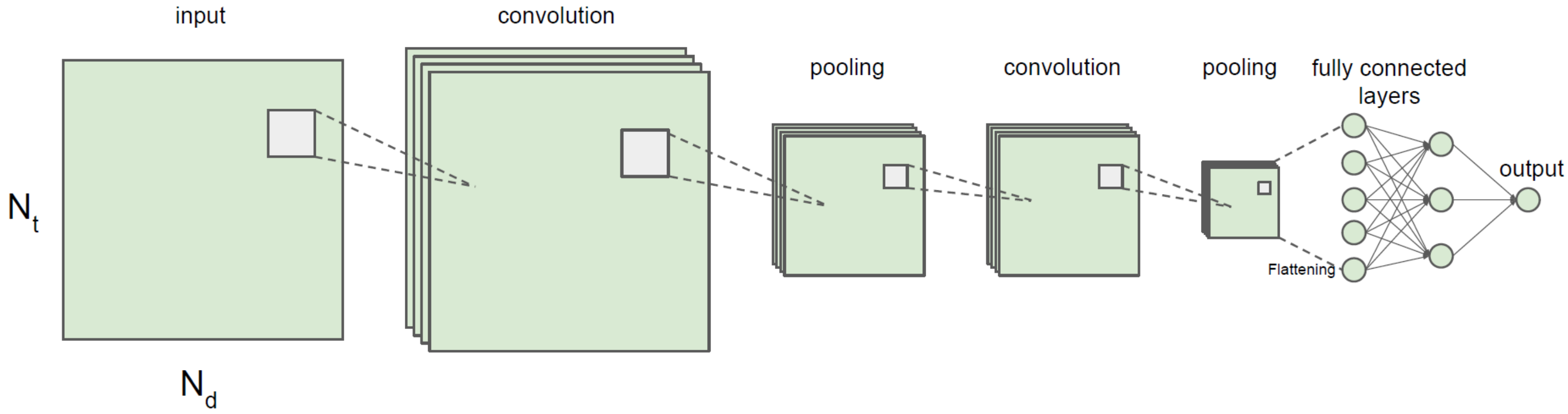
2. Train a neural network to approximate the inverse map that takes the measurement as input and predicts $\boldsymbol{c}$.

3. Given a new measurement, predict $\boldsymbol{c}$ as the initial guess to warm-start the Gauss-Newton method.

A neural network warm-start approach for the inverse acoustic obstacle scattering problem, Zhou, Han, Rachh, and Borges, arXiv (2022)

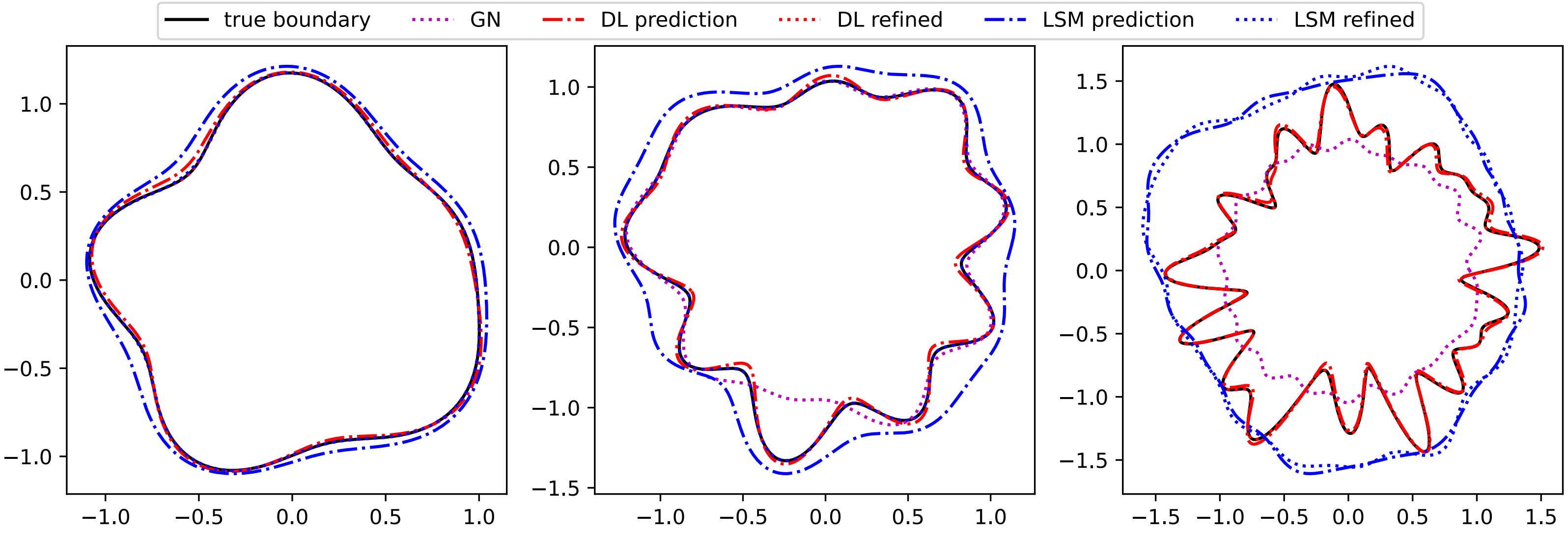
Learning the Inverse Map through CNN

Input: $\mathbb{R}^{N_t \times N_d}$, output: $\mathbb{R}^{2M+1}$. The structure of the convolutional neural network: 2 layers of convolution-ReLu-pooling, followed by 2 fully connected hidden layers.



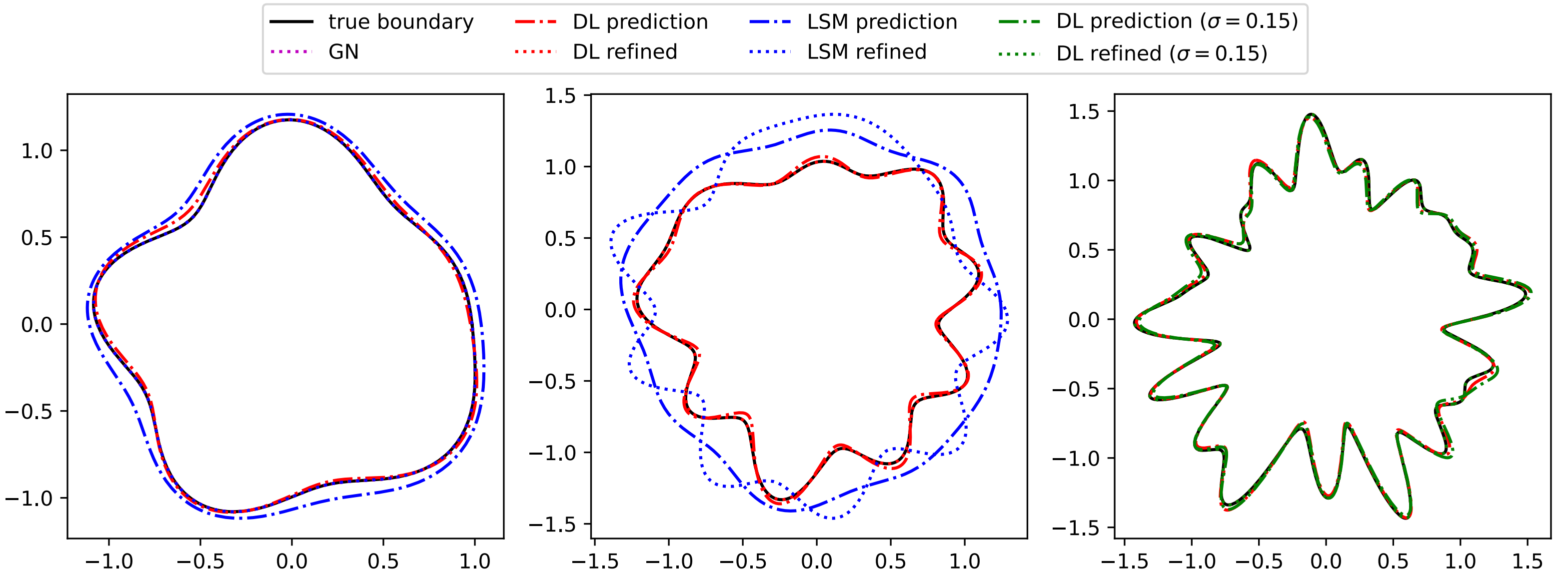
We consider single frequency inverse scattering problems with varying frequencies k and set $M \sim O(k)$, leading to different difficulties.

Results of Full-Aperture Noiseless Data



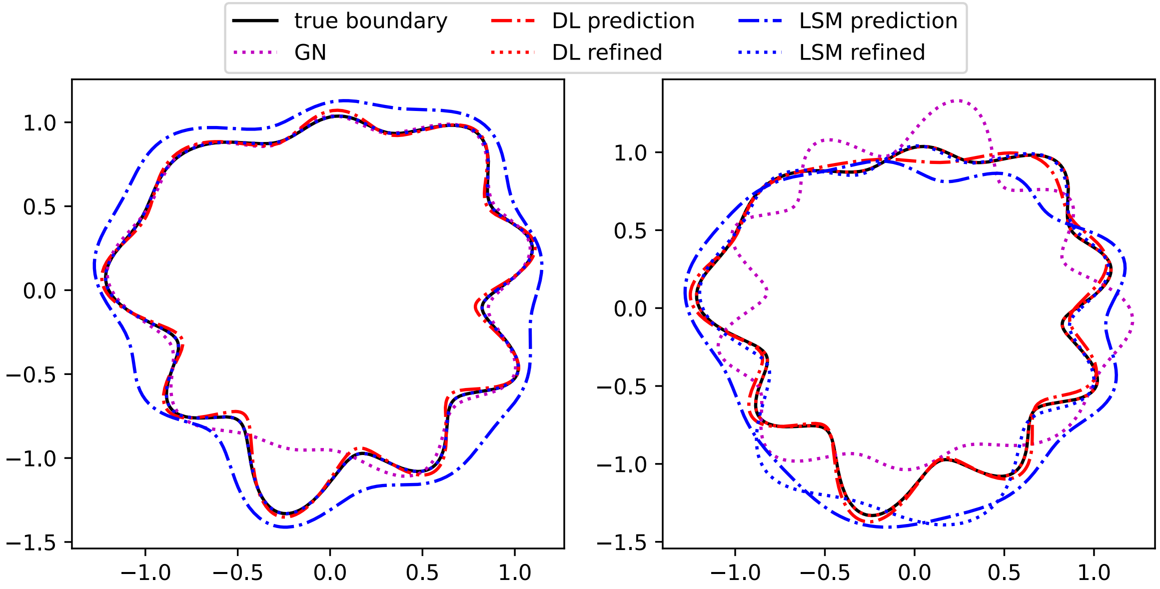
	GN	LSM prediction	LSM refined	DL prediction	DL refined
Left: $M = 5,\ k = 5$	10.78%	2.17%	0.40%	5.40%	**0**
Middle: $M = 10,\ k = 10$	21.29%	13.04%	7.72%	3.57%	**0**
Right: $M = 20,\ k = 30$	24.53%	32.21%	32.31%	4.30%	**1.32%**

Results of Measurement Noise



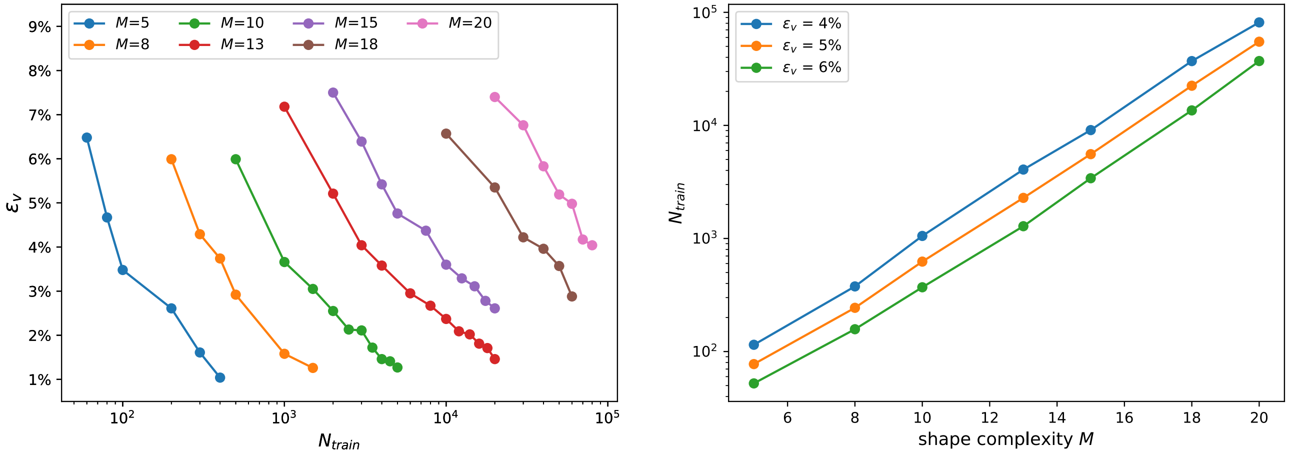
	GN	LSM prediction	LSM refined	DL prediction	DL refined
Left: $M = 5$, $k = 5$, $\sigma = 0.05$	10.78%	8.72%	4.33%	5.40%	**0.083%**
Middle: $M = 10$, $k = 10$, $\sigma = 0.05$	21.29%	22.10%	21.30%	3.57%	**0.054%**
Right: $M = 20$, $k = 30$, $\sigma = 0.05$	24.54%	34.41%	34.57%	4.30%	**1.35%**
Right: $M = 20$, $k = 30$, $\sigma = 0.15$	24.55%	49.54%	49.42%	4.30%	**1.39%**

Results of Partial Aperture Data



	GN	LSM prediction	LSM refined	DL prediction	DL refined
Left: $M = 10$, $k = 10$, full data	21.29%	13.04%	7.72%	3.57%	**0**
Right: $M = 10$, $k = 10$, partial data	19.46%	21.37%	18.38%	6.02%	**0**

How Much Can We Scale?



We need exponentially many samples of training data in terms of shape complexity/ frequency - A purely data-driven method is doomed to limited success.

neural network approximation | accurate PDE solving

frequency

Toward a General Inverse Solver

Goal: a robust inverse solver for a broad range of scatter objects using diverse measurement formats.

1. Apply to inverse medium scattering problems to infer the variations in sound speed inside the scatter.
2. Design inverse maps to handle diverse measurement formats (multifrequency, backscatter, phaseless, etc.).
3. Explore other network architectures with the inverse map as an inductive bias.

Journal of Machine Learning (JML)

Journal of Machine Learning

Editor-in-Chief
Weinan E

Editorial Board

Electronic:2790-2048 ISSN Print:2790-203X

Aims and Scope

Journal of Machine Learning (JML) publishes high quality research papers in all areas of machine learning, including innovative algorithms of machine learning, theories of machine learning, important applications of machine learning in AI, natural sciences, social sciences, and engineering etc. The journal emphasizes a balanced coverage of both theory and practice. The journal is published in a timely fashion in electronic form. All articles in JML are open-access and there is no charge for the authors.