

SOLVING ILL-POSED HELMHOLTZ PROBLEMS  
WITH PHYSICS-INFORMED NEURAL NETWORKS\*

MIHAI NECHITA<sup>†‡</sup>

**Abstract.** We consider the unique continuation (data assimilation) problem for the Helmholtz equation and study its numerical approximation based on physics-informed neural networks (PINNs). Exploiting the conditional stability of the problem, we first give a bound on the generalization error of PINNs. We then present numerical experiments in 2d for different frequencies and for geometric configurations with different stability bounds for the continuation problem. The results show that vanilla PINNs provide good approximations even for noisy data in configurations with robust stability (both low and moderate frequencies), but may struggle otherwise. This indicates that more sophisticated techniques are needed to obtain PINNs that are frequency-robust for inverse problems subject to the Helmholtz equation.

**MSC.** 65N12, 65N20, 68T07.

**Keywords.** ill-posed problems, inverse problems, unique continuation, data assimilation, Helmholtz equation, physics-informed neural networks.

1. INTRODUCTION

Physics-informed neural networks [16] have recently emerged as an effective way of approximating ill-posed/inverse problems for partial differential equations (PDEs), which are challenging to solve numerically due their inherent instability [11]. One of their main advantages is the versatility with which they can solve both forward and inverse problems. As reported in [9] for well-posed boundary value problems, PINNs are not able, however, to outperform finite elements methods for low dimensional forward problems (2d and 3d).

In this paper we explore PINNs for an important class of inverse problems, namely the unique continuation (data assimilation) problem in which measurements are available in a subset of the domain and no boundary conditions are given. The goal is to find a solution that extends the data subject to the PDE

---

\*This work was supported by the project "The Development of Advanced and Applicative Research Competencies in the Logic of STEAM + Health" /POCU/993/6/13/153310, project co-financed by the European Social Fund through The Romanian Operational Programme Human Capital 2014-2020.

<sup>†</sup>Tiberiu Popoviciu Institute of Numerical Analysis, Romanian Academy, Cluj-Napoca, Romania, e-mail: [mihai.nechita@ictp.acad.ro](mailto:mihai.nechita@ictp.acad.ro), [orcid.org/0000-0002-6448-912X](https://orcid.org/0000-0002-6448-912X).

<sup>‡</sup>Department of Mathematics, Babeş-Bolyai University, Cluj-Napoca, Romania.

considered. Numerical methods for such problems typically consider regularization at the continuous level (Tikhonov or quasi-reversibility). More recently, other kinds of methods have been proposed that make use of regularization at the discrete level in the framework of stabilized finite element methods (FEM). In this latter approach, conditional stability estimates can be used to prove error bounds and convergence in terms of the degree of ill-posedness and the approximation order, see *e.g.* [3, 5, 15]. Conditional stability estimates have also been employed in [14] to obtain bounds for the generalization error of PINNs solving unique continuation problems for Poisson, Stokes, heat and wave equations.

We are interested in the ill-posed unique continuation problem for the Helmholtz equation, whose stability properties (in terms of the frequency) depend on the geometry of the data set relative to the geometry of the target set where the solution is reconstructed [6]. We consider PINNs for approximating this problem and we first give a bound in Section 3 on the generalization error that takes into account the frequency dependence. We then explore numerically in Section 4 how different frequencies and geometric configuration impact the effectiveness of PINNs. We see that PINNs give good approximations when the frequency is low and the target set is inside the convex hull of the data set (robust stability in terms of frequency as discussed in Section 2.1). This includes results where data perturbations are present. When the target set is outside of the convex hull of the data set (stability highly sensitive to frequency) or when the frequency is large we observe poor results with vanilla PINNs. This indicates that more sophisticated techniques are needed to obtain PINNs that are robust in frequency for inverse problems subject to the Helmholtz equation.

## 2. UNIQUE CONTINUATION FOR THE HELMHOLTZ EQUATION

Let  $\Omega \subset \mathbb{R}^2$  be a bounded domain (open and connected), and let  $\omega \subset \Omega$  be an open and connected subset. We consider the ill-posed unique continuation problem for the Helmholtz equation: find  $u \in H^1(\Omega)$  such that

$$(1) \quad \begin{cases} -\Delta u - k^2 u = f & \text{in } \Omega, \\ u = g & \text{in } \omega, \end{cases}$$

with wave number  $k > 0$ , source term  $f \in L^2(\Omega)$  and datum  $g \in H^1(\omega)$ . In other words, partial measurements are given in a subset  $\omega \subset \Omega$  of the domain for the solution of the Helmholtz equation, while no boundary conditions are prescribed. If problem (1) has a solution, then its uniqueness is guaranteed by the unique continuation principle for elliptic operators. We will assume that the function  $g$  is the restriction on  $\omega$  of a solution to the Helmholtz equation with source term  $f$ .

Problem (1) is ill-posed in the sense of Hadamard: there is no uniform stability with respect to the data  $f$  and  $g$ , see *e.g.* [1] or [11]. A small

perturbation in data can lead to a large change in the solution. Nonetheless, assuming an additional a priori bound, the solution can be bounded by the data in the following way.

**2.1. Conditional stability.** We denote by  $B \subset \Omega$  a target set containing  $\omega$  such that  $B \setminus \bar{\omega} \subset \Omega$ , *i.e.*  $B \setminus \omega$  does not touch the boundary of  $\Omega$ . A standard result for quantitative unique continuation of elliptic operators, see *e.g.* [1], states that there exist constants  $C_{st}(k) > 0$  and  $\alpha \in (0, 1)$  such that the following Hölder stability estimate holds

$$(2) \quad \|u\|_{H^1(B)} \leq C_{st}(k) \left( \|f\|_{L^2(\Omega)} + \|g\|_{H^1(\omega)} \right)^\alpha \|u\|_{H^1(\Omega)}^{1-\alpha},$$

for any  $u \in H^1(\Omega)$  satisfying (1), where the stability constant  $C_{st}(k)$  depends on the frequency  $k$ . The exponent  $\alpha \in (0, 1)$  encodes the degree of ill-posedness for the continuation problem: as  $\alpha < 1$  decreases the Hölder stability deteriorates, while  $\alpha = 1$  would give a well-posed problem with Lipschitz stability. Both  $C_{st}$  and  $\alpha$  depend on the geometric configuration in a nontrivial way, and we give details below regarding  $C_{st}$ .

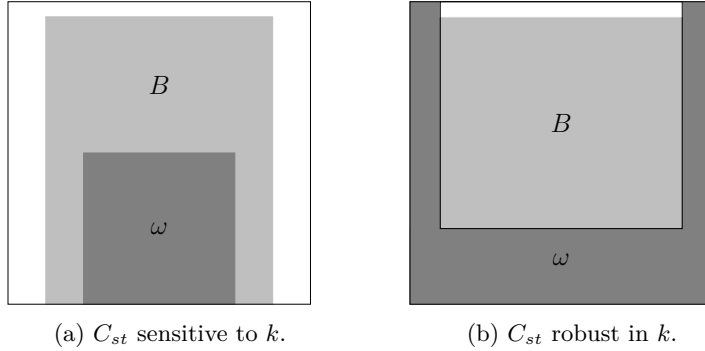


Fig. 1. Frequency dependence for the stability constant. Data set  $\omega$  (dark grey) and target set  $B$  (light grey).  $\Omega$  is the whole square.

An important aspect for the stability of this ill-posed Helmholtz problem is the dependence of the stability constant  $C_{st}(k)$  on the frequency  $k$ . For example, when there is a straight line that intersects  $B$  but not  $\bar{\omega}$ , as in Fig. 1a, it was proven in [6, Example 4] that for any  $N \in \mathbb{N}$ ,  $C_{st}(k) \leq k^N$  cannot hold uniformly in  $k$ ; this means that the stability constant in (2) grows superpolynomially in the frequency. Also, for three-ball inequalities (where  $\omega, B, \Omega$  are concentric balls) it was recently shown in [2] that in the maximum norm  $C_{st}(k)$  grows exponentially in  $k$  and this dependence is optimal.

Bounds that have a different behavior with respect to the frequency can be obtained under a convexity condition of the target domain  $B$  relative to the data set  $\omega$ , essentially that  $B$  is included in the convex hull of  $\omega$ , as for example in Fig. 1b. Such a condition was first considered in [10], where it was

shown that the stability of the solution in the  $L^2$ -norm can actually improve in a certain sense as the frequency increases. In this vein, similar results that make use of a convexity condition were proven in [6] for a particular geometric setting prototypical for continuation inside the convex hull of  $\omega$ . It was shown in [6, Corollary 2] that there exist constants  $C > 0$  and  $\alpha \in (0, 1)$  such that

$$(3) \quad \|u\|_{H^1(B)} \leq C(\|f\|_{L^2(\Omega)} + \|g\|_{H^1(\omega)})^\alpha \|u\|_{H^1(\Omega)}^{1-\alpha},$$

for any  $u \in H^1(\Omega)$  satisfying (1). Moreover, the norms in which data is measured can be weakened [6, Corollary 3 and Lemma 2]: there exist constants  $C > 0$  and  $\alpha \in (0, 1]$  such that

$$(4) \quad \begin{aligned} \|\nabla u\|_{L^2(B)} + k \|u\|_{L^2(B)} &\leq Ck \left( \|f\|_{H^{-1}(\Omega)} + \|g\|_{L^2(\omega)} \right)^\alpha \left( \|f\|_{H^{-1}(\Omega)} + \|u\|_{L^2(\Omega)} \right)^{1-\alpha} \\ &\leq Ck \left( \|f\|_{L^2(\Omega)} + \|g\|_{L^2(\omega)} \right)^\alpha \left( \|f\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} \right)^{1-\alpha}, \end{aligned}$$

for any  $u \in H^1(\Omega)$  satisfying (1). Note that the bound is robust in the  $L^2$ -norm, while the frequency dependence is linear for the  $H^1$ -seminorm. Apart from the good dependence on the frequency, this kind of estimate is particularly suitable for numerical analysis since it can be directly applied to the error equation.

REMARK 1. If the target set is the whole domain  $\Omega$ , then a global stability estimate holds with the modulus of continuity being logarithmic  $|\log(\cdot)|^{-\alpha}$  instead of Hölder-type  $|\cdot|^\alpha$ .  $\square$

### 3. PHYSICS-INFORMED NEURAL NETWORKS (PINNS)

Let us first recall the setup of a feed-forward neural network, which is based on successively applying linear and nonlinear transformations to the inputs. Let  $\mathcal{N}^L(\mathbf{x}) : \mathbb{R}^{d_{\text{in}}} \rightarrow \mathbb{R}^{d_{\text{out}}}$  be an  $L$ -layer neural network with  $(L - 1)$  hidden layers and  $N_\ell$  neurons in the  $\ell$ -th layer ( $N_0 = d_{\text{in}}$ ,  $N_L = d_{\text{out}}$ ). Let us denote the weight matrix and bias vector in the  $\ell$ -th layer by  $\mathbf{W}^\ell \in \mathbb{R}^{N_\ell \times N_{\ell-1}}$  and  $\mathbf{b}^\ell \in \mathbb{R}^{N_\ell}$ , respectively. For a nonlinear activation function  $\sigma$  applied element wise, the feed-forward neural network is given by:

$$\begin{aligned} \text{input layer: } \mathcal{N}^0(\mathbf{x}) &:= \mathbf{x} \in \mathbb{R}^{d_{\text{in}}}, \\ \text{hidden layers: } \mathcal{N}^\ell(\mathbf{x}) &:= \sigma(\mathbf{W}^\ell \mathcal{N}^{\ell-1}(\mathbf{x}) + \mathbf{b}^\ell) \in \mathbb{R}^{N_\ell}, \quad \text{for } 1 \leq \ell \leq L - 1, \\ \text{output layer: } \mathcal{N}^L(\mathbf{x}) &:= \mathbf{W}^L \mathcal{N}^{L-1}(\mathbf{x}) + \mathbf{b}^L \in \mathbb{R}^{d_{\text{out}}}. \end{aligned}$$

We denote the parameters of the network (weights and biases) by

$$\theta := \{(\mathbf{W}^1, \mathbf{b}^1), \dots, (\mathbf{W}^L, \mathbf{b}^L)\},$$

and the parameter set by  $\Theta$ , hence

$$\theta \in \Theta \subset \mathbb{R}^M, \quad M = \sum_{l=1}^L N_l(N_{l-1} + 1).$$

The output of a feed-forward neural network,

$$u_\theta(\mathbf{x}) := \mathcal{N}^L(\mathbf{x}),$$

depends on the tuning parameter  $\theta$ . Training the network means using an optimization algorithm to find the parameters  $\theta \in \Theta$  that minimize a certain loss function  $\mathcal{L}_\theta$ .

Let us now briefly describe physics-informed neural networks (PINNs) as introduced in [16], based on feed-forward neural networks. The idea is to consider a loss function with two components: the residual of the differential equation and the data residual. For typical boundary value problems, the data residual contains boundary conditions, initial conditions, and any other available measurements of the solution. In our case, since no boundary conditions are given, the data residual will only involve the interior partial measurements.

To solve the ill-posed unique continuation problem (1) with PINNs we consider the residual of the Helmholtz equation

$$\mathring{\mathcal{R}}_\theta := -\Delta u_\theta - k^2 u_\theta - f \text{ in } \Omega$$

and the data residual

$$\mathring{\mathcal{R}}_{\theta,d} := u_\theta - g \text{ in } \omega,$$

and aim to minimize the two residuals over the admissible set of tuning parameters  $\Theta$  by considering the loss function

$$\mathring{\mathcal{L}}_\theta := \|\mathring{\mathcal{R}}_\theta\|_{L^2(\Omega)} + \|\mathring{\mathcal{R}}_{\theta,d}\|_{L^2(\omega)}.$$

To approximate the integrals in the setting of a neural network, we introduce the following collocation points:  $\mathcal{T}_\Omega := \{\mathbf{x}_i^\Omega\}_{i=1}^{|\mathcal{T}_\Omega|}$  for the whole domain  $\Omega$  and  $\mathcal{T}_\omega := \{\mathbf{x}_i^\omega\}_{i=1}^{|\mathcal{T}_\omega|}$  for the data region  $\omega \subset \Omega$  in which observations of the solution are given as  $\{g_i(\mathbf{x}^\omega)\}_{i=1}^{|\mathcal{T}_\omega|}$ . The loss function for the PINN is hence defined as

$$(5) \quad \mathcal{L}_\theta := \mathcal{R}_\theta + \mathcal{R}_{\theta,d},$$

with the PDE residual

$$(6) \quad \mathcal{R}_\theta := \frac{1}{N_\Omega} \sum_{\mathbf{x} \in \mathcal{T}_\Omega} \left| (\Delta u_\theta + k^2 u_\theta + f)(\mathbf{x}) \right|^2,$$

and the data residual

$$(7) \quad \mathcal{R}_{\theta,d} := \frac{1}{N_\omega} \sum_{\mathbf{x} \in \mathcal{T}_\omega} |(u_\theta - g)(\mathbf{x})|^2.$$

By training the neural network we aim to

$$(8) \quad \text{find } \theta^* \in \Theta \text{ such that } \theta^* = \arg \min_{\theta \in \Theta} \mathcal{L}_\theta.$$

We will denote by

$$u^* = u_{\theta^*}$$

the PINN solution obtained by this algorithm, see Fig. 2 for a sketch.

As the loss function is highly non-linear and non-convex, we will use the standard approach of minimizing by gradient-based methods such as ADAM [12] or L-BFGS [7], more details are given in Section 4.

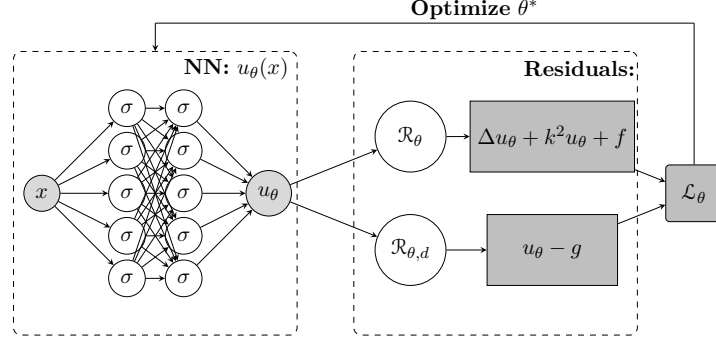


Fig. 2. Representation of a PINN for solving (1), adapted from [8].

**3.1. Estimating the generalization error.** Such PINNs for unique continuation problems have been considered in [14] where the authors prove for the first time estimates for the generalization error by employing conditional stability estimates; details of the analysis and numerical examples were given for the Poisson, Stokes, heat and wave equations. In brief, one uses conditional stability estimates to bound the error in a target domain (generalization error) by the residuals, which are then bounded in terms of the training error and the number of training samples. We now adapt the analysis in [14, Theorem 2.4] to the case of the Helmholtz equation by using the stability results discussed in Section 2.1.

**THEOREM 2.** *Let  $f \in C^{n-2}(\Omega)$  and  $g \in C^n(\omega)$ , with continuous extensions of the functions and derivatives up to the boundaries, with  $n \geq 2$ . Let  $u \in H^1(\Omega)$  be the solution of the unique continuation problem (1). Let  $u^* = u_{\theta^*} \in C^n(\Omega)$  be a PINN solution generated by training (8). Consider a target set  $B \subset \Omega$  containing  $\omega$  satisfying the convexity condition in [6, Corollary 2]. Then there exist constants  $C_{st}(k) > 0$  and  $\alpha \in (0, 1)$  such that the generalization error*

$$(9) \quad \mathcal{E}_G(B) := \|\nabla(u - u^*)\|_{H^1(B)} + k\|u - u^*\|_{L^2(B)}$$

is bounded by

$$(10) \quad \mathcal{E}_G(B) \leq C_{st}(k) \left( \|u\|_{L^2(\Omega)}^{1-\alpha} + \|u^*\|_{L^2(\Omega)}^{1-\alpha} + \mathcal{E}_{\Omega,T}^{1-\alpha} + C_q^{\frac{1-\alpha}{2}} N_\Omega^{-\frac{\tau(1-\alpha)}{2}} \right) \\ \times \left( \mathcal{E}_{\Omega,T}^\alpha + \mathcal{E}_{\omega,T}^\alpha + C_q^{\frac{\alpha}{2}} N_\Omega^{-\frac{\tau\alpha}{2}} + C_{qd}^{\frac{\alpha}{2}} N_\omega^{-\frac{\tau_d\alpha}{2}} \right),$$

where  $C_{st}(k)$  grows polynomially in  $k$ , the training errors are

$$(11) \quad \mathcal{E}_{\Omega,T} = \left( \sum_{j=1}^{N_{\Omega}} w_j |\mathring{\mathcal{R}}_{\theta^*}(\mathbf{x}_j^{\Omega})|^2 \right)^{\frac{1}{2}}, \quad \mathcal{E}_{\omega,T} = \left( \sum_{i=1}^{N_{\omega}} w_i |\mathring{\mathcal{R}}_{\theta^*,d}(\mathbf{x}_i^{\omega})|^2 \right)^{\frac{1}{2}},$$

with constants  $C_q$  and  $C_{qd}$  given by the quadrature bounds, and some  $\tau, \tau_d > 0$ .

*Proof.* Consider the error  $\hat{u} = u^* - u \in H^1(\Omega)$  which satisfies

$$\begin{aligned} -\Delta \hat{u} - k^2 \hat{u} &= \mathring{\mathcal{R}}_{\theta^*}, & \text{in } \Omega, \\ \hat{u} &= \mathring{\mathcal{R}}_{\theta^*,d}, & \text{in } \omega, \end{aligned}$$

in a weak sense. Applying the conditional stability estimate (4) we obtain

$$\begin{aligned} \mathcal{E}_G(B) &\leq \\ &\leq C_{st}(k) \left( \|\mathring{\mathcal{R}}_{\theta^*}\|_{L^2(\Omega)} + \|\mathring{\mathcal{R}}_{\theta^*,d}\|_{L^2(\omega)} \right)^{\alpha} \left( \|\mathring{\mathcal{R}}_{\theta^*}\|_{L^2(\Omega)} + \|\hat{u}\|_{L^2(\Omega)} \right)^{1-\alpha} \\ &\leq C_{st}(k) \left( \|\mathring{\mathcal{R}}_{\theta^*}\|_{L^2(\Omega)} + \|\mathring{\mathcal{R}}_{\theta^*,d}\|_{L^2(\omega)} \right)^{\alpha} \left( \|\mathring{\mathcal{R}}_{\theta^*}\|_{L^2(\Omega)} + \|u\|_{L^2(\Omega)} + \|u^*\|_{L^2(\Omega)} \right)^{1-\alpha}. \end{aligned}$$

The conclusion follows by assuming quadrature rules with approximation bounds as in [14, Eqs (2.9) and (2.11)], where for some  $\tau, \tau_d > 0$  one has that

$$(12) \quad \begin{aligned} \|\mathring{\mathcal{R}}_{\theta^*}\|_{L^2(\Omega)} &\leq \mathcal{E}_{\Omega,T} + C_q^{\frac{1}{2}} N_{\Omega}^{-\frac{\tau}{2}}, \\ \|\mathring{\mathcal{R}}_{\theta^*,d}\|_{L^2(\omega)} &\leq \mathcal{E}_{\omega,T} + C_{qd}^{\frac{1}{2}} N_{\omega}^{-\frac{\tau_d}{2}}, \end{aligned}$$

with constants  $C_q = C_q \left( \|\mathring{\mathcal{R}}_{\theta^*}\|_{C^{n-2}(\Omega)} \right)$  and  $C_{qd} = C_{qd} \left( \|\mathring{\mathcal{R}}_{\theta^*,d}\|_{C^n(\omega)} \right)$ .  $\square$

REMARK 3. As discussed in Section 2.1, if the target domain  $B$  is outside the convex hull of the measurement domain  $\omega$ , the constant  $C_{st}(k)$  in the conditional stability estimate might depend exponentially on the wave number  $k$ . In that case, the same behaviour will appear in the bound of the generalization error in Theorem 2, which indicates that obtaining good approximations for this ill-posed problem with PINNs can be very challenging for high frequencies.  $\square$

REMARK 4. If the target set is the whole domain  $\Omega$ , then one can obtain a global bound on the generalization error by using a global logarithmic stability estimate with an implicit dependence on the wave number.  $\square$

#### 4. NUMERICAL EXPERIMENTS

We present numerical experiments for the Helmholtz unique continuation problem (1) solved with physics-informed neural networks (PINNs) described in Section 3 and implemented using the open-source library DeepXDE [13]. The test case we will focus on has been considered as a benchmark test for this problem solved with primal-dual stabilized finite element methods: with a conforming discretization in [6] and with a high-order hybridized discontinuous

Galerkin method in [4]. It represents the Helmholtz version of the classical Hadamard example for ill-posed elliptic equations.

Let  $n \in \mathbb{N}$  and consider the Cauchy problem

$$(13) \quad \begin{cases} \Delta u + k^2 u = 0 & \text{in } \Omega := (0, 1) \times (0, 1), \\ u(x, 0) = 0 & \text{for } x \in [0, 1], \\ u_y(x, 0) = \sin(nx) & \text{for } x \in [0, 1], \end{cases}$$

whose solution for  $n > k$  is given by

$$(14) \quad u(x, y) = \frac{1}{\sqrt{n^2 - k^2}} \sin(nx) \sinh(\sqrt{n^2 - k^2}y),$$

for  $n = k$  by  $u(x, y) = \sin(kx)y$ , and for  $n < k$  by

$$u(x, y) = \frac{1}{\sqrt{k^2 - n^2}} \sin(nx) \sin(\sqrt{k^2 - n^2}y).$$

For such Hadamard-type solutions, we consider the interior datum  $g = u|_\omega$  and study two geometric configurations of  $\omega$  and  $B$ : one with frequency-robust stability bounds and one exponentially sensitive to the frequency, as discussed in Section 2.1, namely

$$(15) \quad \omega = \Omega \setminus [0, 0.865] \times [0.125, 0.875], \quad B = \Omega \setminus [0, 0.125] \times [0.125, 0.875],$$

similar to Fig. 1b, and

$$(16) \quad \omega = (0.25, 0.75) \times (0, 0.5), \quad B = (0.125, 0.875) \times (0, 0.875),$$

sketched in Fig. 1a.

To assess the effect of increasing the frequency, we will take exact solutions (14) with  $n = 5, k = 1$  and  $n = 7, k = 5$ , both having  $\sqrt{n^2 - k^2} = \sqrt{24}$ .

**Hyper-parameters.** The PINNs described in Section 3 need to be configured with the following hyper-parameters: number of hidden layers  $L - 1$  (depth), number of neurons in each hidden layer  $\ell$  (width), learning rate  $\lambda$ , activation function  $\sigma$ . In order to find good configurations for these, we draw on previous numerical experiments for forward and inverse problems presented in the DeepXDE tutorials [13] and the numerical experiments for unique continuation subject to the Poisson equation in [14]. We also validate the choices by comparing them with the results given by the hyper-parameter optimization in [8], which uses Gaussian processes-based Bayesian optimization. For this we consider the search space  $L - 1 \in [4, 50], N_\ell \in [20, 150], \lambda \in [10^{-4}, 10^{-2}], \sigma \in \{\sin, \tanh\}$ . In these ways we obtain some optimal or near-optimal configurations to which the results presented below correspond. From different numerical experiments, it turns out that a good choice for the activation function is  $\sigma = \sin$  and for the learning rate  $\lambda = 10^{-3}$ , which will be fixed from this point onwards.

**Training.** We sample  $N = N_\Omega = N_\omega$  points on Cartesian grids from the PDE domain  $\Omega$  and the measurement domain  $\omega$ . We train the model for 50000 iterations with the ADAM optimizer [12] and then we train again with L-BFGS



[7], following [13]. Since PINNs approximations are obtained by non-convex minimization which in general does not have a unique solution, the method might converge to different solutions depending on the network's initial values. We use the standard strategy of training the PINNs with different random starting values for the optimizer and average the errors over 30 retrainings. We note that the smallest training error is considerably better.

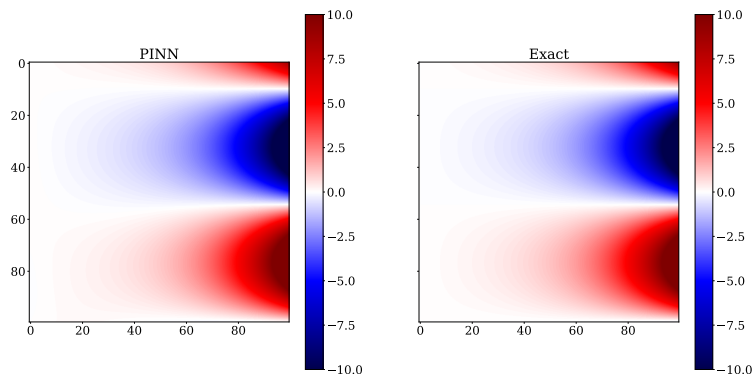


Fig. 3. PINN solution and exact solution (14) with  $n = 7, k = 5$ .

**Results.** We first show in Fig. 3 the exact solution  $u$  for the Hadamard-type function (14) with  $n = 7, k = 5$ , and the PINN solution  $u^*$  when considering a network generated with  $N = 400$  training points,  $L - 1 = 4$  hidden layers and  $N_\ell = 24$  neurons in each layer. As discussed above, the activation function is  $\sigma = \sin$  and the learning rate  $\lambda = 10^{-3}$ .

$N$	$L - 1$	$N_\ell$	$\ u - u^*\ _{L^2(B)}$	$\ u - u^*\ _{H^1(B)}$
400	4	24	3.35%	3.77%
1600	4	24	1.87%	2.70
6400	4	24	1.45%	2.31%

Table 1. Geometry with good stability (15). Relative percentage generalization errors in  $B$ ,  $n = 5$  and  $k = 1$ .

We study the efficiency of PINNs by considering the  $L^2$  and  $H^1$  relative percentage errors for the number of training points  $N = 20^2, 40^2, 80^2$  in the geometric configuration (15) with robust bounds in Theorem 2. Table 1 shows the errors in the target domain  $B$ , while Table 2 shows the global errors. In both tables we observe similar behaviours for the error. Moreover, we notice that even for very few training points the generalization errors are small (around 3%). Note that the reported errors correspond to averages over 30

retrainings. The generalization errors corresponding to the network with the smallest training error are typically smaller by a factor of 3 to 5. The training error in these examples is in between  $10^{-3}$  and  $10^{-2}$ . The training of such a network is very fast, on a single Quadro GPU it takes about 1-2 minutes.

$N$	$L - 1$	$N_\ell$	$\ u - u^*\ _{L^2(\Omega)}$	$\ u - u^*\ _{H^1(\Omega)}$
400	4	24	4.58%	5.15%
1600	4	24	2.89%	3.61%
6400	4	24	2.22%	3.01%

Table 2. Geometry with good stability (15). Relative percentage generalization errors in  $\Omega$ ,  $n = 5$  and  $k = 1$ .

$N$	$L - 1$	$N_\ell$	$\ u - u^*\ _{L^2(\Omega)}$	$\ u - u^*\ _{H^1(\Omega)}$
400	4	24	4.73%	5.17%
1600	4	24	3.34%	4.06%
6400	4	24	2.38%	3.19%

Table 3. Geometry with good stability (15). Relative percentage generalization errors in  $\Omega$ ,  $n = 5$  and  $k = 1$ . Perturbed data.

We also test the performance of the method when the measurements are perturbed with random values sampled from the normal distribution with zero mean and standard deviation 0.01 (representing approximately 2.5% of the average of the exact solution (14) with  $n = 5, k = 1$  considered in that example). The results given in Table 3 show that the numerical approximation is robust to such perturbations in data.

In Table 4 we keep the same (convex) geometric configuration but increase the frequency to  $k = 5$ , with  $n = 7$  such that  $\sqrt{n^2 - k^2} = \sqrt{24}$  as before. We observe that the approximation improves as the frequency increases: the errors decrease compared to Table 2. This is probably an artefact of the particular conditions of the experiment, but we note that such a surprising phenomenon for inverse Helmholtz problems has been previously noticed and theoretically discussed in [10]. However, when increasing the wave number  $k > 10$  we report that the PINNs no longer provide a good approximation to the solution.

$N$	$L - 1$	$N_\ell$	$\ u - u^*\ _{L^2(\Omega)}$	$\ u - u^*\ _{H^1(\Omega)}$
400	4	24	2.03%	3.31%
1600	4	24	1.48%	2.04%
6400	4	24	1.27%	1.77%

Table 4. Geometry with good stability (15). Relative percentage generalization errors in  $\Omega$ ,  $n = 7$  and  $k = 5$ .

We now consider continuation outside the convex hull of the data set in (16). We observe in Table 5 that even though the wave number is small  $k = 1$ , the PINNs fail to find a good approximation of the solution. This could be caused by a bad search space for hyper-parameter optimization or it could be an inherent limitation of vanilla PINNs.

$N$	$L - 1$	$N_\ell$	$\ u - u^*\ _{L^2(\Omega)}$	$\ u - u^*\ _{H^1(\Omega)}$
400	4	24	46.1%	71.4%
1600	4	24	41.8%	66.6%
25600	4	24	39.4%	62.5%

Table 5. Geometry with bad stability (16). Relative percentage generalization errors in  $\Omega$ ,  $n = 5$  and  $k = 1$ .

## REFERENCES

- [1] G. ALESSANDRINI, L. RONDI, E. ROSSET, AND S. VESSELLA, *The stability for the Cauchy problem for elliptic equations*, Inverse Problems, **25** (2009), p. 123004, <https://doi.org/10.1088/0266-5611/25/12/123004>.
- [2] S. M. BERGE AND E. MALINNIKOVA, *On the three ball theorem for solutions of the Helmholtz equation*, Complex Anal. Synerg., **7** (2021), pp. 1–10, <https://doi.org/10.1007/s40627-021-00070-3>.
- [3] E. BURMAN, *Stabilized finite element methods for nonsymmetric, noncoercive, and ill-posed problems. Part I: elliptic equations*, SIAM J. Sci. Comput., **35** (2013), pp. A2752–A2780, <https://doi.org/10.1137/130916862>.
- [4] E. BURMAN, G. DELAY, AND A. ERN, *A hybridized high-order method for unique continuation subject to the Helmholtz equation*, SIAM J. Numer. Anal., **59** (2021), pp. 2368–2392, <https://doi.org/10.1137/20M1375619>.
- [5] E. BURMAN, P. HANSBO, AND M. G. LARSON, *Solving ill-posed control problems by stabilized finite element methods: an alternative to Tikhonov regularization*, Inverse Problems, **34** (2018), p. 035004, <https://doi.org/10.1088/1361-6420/aaa32b>.
- [6] E. BURMAN, M. NECHITA, AND L. OKSANEN, *Unique continuation for the Helmholtz equation using stabilized finite element methods*, J. Math. Pures Appl., **129** (2019), pp. 1–22, <https://doi.org/10.1016/j.matpur.2018.10.003>.
- [7] R. H. BYRD, P. LU, J. NOCEDAL, AND C. ZHU, *A limited memory algorithm for bound constrained optimization*, SIAM J. Sci. Comput., **16** (1995), pp. 1190–1208, <https://doi.org/10.1137/0916069>.
- [8] P. ESCAPIL-INCHAUSPÉ AND G. A. RUZ, *Hyper-parameter tuning of physics-informed neural networks: Application to Helmholtz problems*, arXiv preprint arXiv:2205.06704, (2022).
- [9] T. G. GROSSMANN, U. J. KOMOROWSKA, J. LATZ, AND C.-B. SCHÖNLIEB, *Can physics-informed neural networks beat the finite element method?*, arXiv preprint arXiv:2302.04107, (2023).
- [10] T. HRYCAK AND V. ISAKOV, *Increased stability in the continuation of solutions to the Helmholtz equation*, Inverse Problems, **20** (2004), pp. 697–712, <https://doi.org/10.1088/0266-5611/23/4/019>.
- [11] V. ISAKOV, *Inverse problems for partial differential equations*, vol. 127 of Applied Mathematical Sciences, Springer, 3rd ed., 2017, <https://doi.org/10.1007/978-3-319-51658-5>.

- 
- [12] D. P. KINGMA AND J. BA, *Adam: A method for stochastic optimization*, arXiv preprint arXiv:1412.6980, (2014).
  - [13] L. LU, X. MENG, Z. MAO, AND G. E. KARNIADAKIS, *DeepXDE: A deep learning library for solving differential equations*, SIAM Rev., **63** (2021), pp. 208–228, <https://doi.org/10.1137/19M1274067>.
  - [14] S. MISHRA AND R. MOLINARO, *Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs*, IMA J. Numer. Anal., **42** (2022), pp. 981–1022, <https://doi.org/10.1093/imanum/drab032>.
  - [15] M. NECHITA, *Unique continuation problems and stabilised finite element methods*, PhD thesis, University College London, 2020, <https://discovery.ucl.ac.uk/id/eprint/10113065/>.
  - [16] 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*, J. Comp. Phys., **378** (2019), pp. 686–707, <https://doi.org/10.1016/j.jcp.2018.10.045>.

Received by the editors: February 21, 2023; accepted: June 29, 2023; published online: July 5, 2023.