

# PHYSICS-INFORMED NEURAL NETWORKS FOR LEARNING THE HOMOGENIZED COEFFICIENTS OF MULTISCALE ELLIPTIC EQUATIONS

JUN SUR R PARK\* AND XUEYU ZHU†

**Abstract.** Multiscale elliptic equations with scale separation are often approximated by the corresponding homogenized equations with slowly varying homogenized coefficients (the G-limit). The traditional homogenization techniques typically rely on the periodicity of the multiscale coefficients, thus finding the G-limits often requires sophisticated techniques in more general settings even when multiscale coefficient is known, if possible. Alternatively, we propose a simple approach to estimate the G-limits from (noisy-free or noisy) multiscale solution data, either from the existing forward multiscale solvers or sensor measurements. By casting this problem into an inverse problem, our approach adopts physics-informed neural networks (PINNs) algorithm to estimate the G-limits from the multiscale solution data by leveraging a priori knowledge of the underlying homogenized equations. Unlike the existing approaches, our approach does not rely on the periodicity assumption or the known multiscale coefficient during the learning stage, allowing us to estimate homogenized coefficients in more general settings beyond the periodic setting. We demonstrate that the proposed approach can deliver reasonable and accurate approximations to the G-limits as well as homogenized solutions through several benchmark problems.

**Key words.** physics-informed neural network, multiscale, homogenization, G-limit, G-convergence

**1. Introduction.** A wide range of scientific and engineering problems involve multiple scales due to the heterogeneity of the media properties. Direct numerical simulation for multiscale problems, such as multiscale elliptic equations, is typically computationally demanding due to the finescale fluctuation of the media properties. A major effort has been made in past decades to approximate a multiscale equation by the corresponding homogenized equation, whose coefficient, known as the homogenized coefficient or G-limit [26, 27], does not depend on the fine scale. The resulting solution is referred to as the homogenized solution. However, deriving the homogenized equations requires the computation of the G-limits, which is a difficult task for general problems. For standard periodic or locally periodic problems, there are several homogenization methods to find the G-limits, such as two-scale and multiscale convergence [4, 6], but they can be computationally demanding as they often involve a large number of local problem computations. Additionally, if the periodicity assumption does not hold, the standard homogenization methods are not directly applicable, and non-trivial extensions are usually needed if possible. As a result, deriving the homogenized models from the first principles remains challenging for general homogenization problems.

Alternatively, there has been a surge of interest in data-driven learning the effective macroscale model from available measurements or simulated data. In [29], coarse-grained nonlocal models are learned from synthetic high-fidelity (multiscale) data by recovering the sign-changing kernels. In [10], physics-informed neural networks (PINNs) were employed to retrieve the effective permittivity parameters from scattering data in inverse scattering electromagnetic problems. In [7], a neural network algorithm coupled with an equation-free method has been developed to approximate homogenized solution of a time-dependent multiscale problem using simulated mul-

\*Department of Mathematics, University of Iowa, Iowa City, IA 52246. USA. Email: junsur-park@uiowa.edu.

†Department of Mathematics, University of Iowa, Iowa City, IA 52246. USA. Email: xueyu-zhu@uiowa.edu.

tiscale solution data. Regarding the homogenization on multiscale elliptic equations, there have been several inversion approaches related to the homogenization problems in the past several years. The authors in [13, 1, 2] recovered multiscale coefficients from (noisy) multiscale solution data using corresponding homogenized models based on numerical homogenization techniques - the finite element heterogeneous method (FE-HMM) to reduce the computational cost of their forward problems. A Bayesian estimation has been developed to reconstruct the slowly varying parts of the multiscale coefficients from the noisy measurement of multiscale solution data in [21]. Nonetheless, the majority of the existing methods assume that the multiscale coefficients are periodic. More general multiscale coefficients such as non-periodic coefficients are considered in [15]. The authors separated the oscillations of the multiscale coefficients from the weak  $L^2$  limits of them and recovered the part of G-limits from the contributions of the oscillations. However, they required the multiscale coefficients to be known during the inversion stage. In addition, the existing inversion methods often require specialist knowledge, such as numerical multiscale methods or homogenization theory, which can be difficult for application practitioners. These limitations motivate the development of simple and flexible algorithms for the homogenization of multiscale elliptic equations with scale separations in more general settings.

The goal in this paper is to develop a simple and flexible framework to learn the G-limits and corresponding homogenized solutions simultaneously for multiscale elliptic equations, given multiscale solution data. Unlike other approaches, our approach does not require the periodicity of the multiscale coefficient or a known multiscale coefficients during the learning stage. Instead, we assume that the (simulated or measured) solution data of the multiscale equations are available and the structure of corresponding homogenized equations are known. We mainly consider the following two possible scenarios:

- *Noise-free data*: In this scenario, we assume that the traditional homogenization methods may not be applicable, e.g., in non-periodic cases, but the multiscale solution data can be generated by the existing forward solver of the multiscale problem with a known multiscale coefficient. Our goal is to estimate the corresponding G-limit and the homogenized solutions.
- *Noisy data*: In this case, we consider that noisy multiscale solution data (from a specific medium with a fixed finescale size  $\epsilon$ ) can be collected by sensors. We aim to learn the G-limit of the unknown multiscale coefficient and corresponding homogenized solution as they can serve as good approximations to the effective behavior of the multiscale coefficient when  $\epsilon$  is sufficiently small.

Specifically, we adopt one emerging scientific machine learning framework - the physics-informed neural network (PINNs) for our problem. They have been successfully used for approximating solutions to both forward and inverse problems regarding PDEs [19, 10, 24]. One key component of PINNs is to provide neural network approximations to the solutions of forward or inverse problems by incorporating prior physics knowledge into the loss functional. This feature turns out to be beneficial for our current setting. Since the multiscale solution data often contains rapid oscillations or noise, estimating the G-limits and homogenized solutions from the multiscale or random fluctuations is a fundamental challenge. To address these issues, we trained the neural networks to approximate the G-limit and the corresponding homogenized solution for the elliptic homogenized equations based on the multiscale solution data. By incorporating the corresponding homogenized equation into the loss function, PINNs

can encourage the neural network to capture the slowly varying parts of the multiscale solution data.

It is worth noting that collecting a large number of the multiscale solution data containing sufficient finescale information is in general difficult, especially when the finescale parameter  $\epsilon$  is very small. In addition, the measurements by sensors are often corrupted by noises that dominate the finescale fluctuations. Nonetheless, we found that our approach does not require dense sampling of the multiscale solution data in space in order to retain the detailed finescale information as we are only interested in the macroscopic (homogenized) behavior of the multiscale solution data. With the prior knowledge of the structure of the homogenized equation, PINNs can provide an effective regularization that can cope with the noise and the multiscale features in the data. We demonstrate the applicability and performance of our approach via several benchmark examples with both noise-free and noisy data.

The paper is organized as follows. In Section 2, we introduce the concepts of G-convergence and G-limit, and the formulation of the inverse problem. In Section 3, we briefly review the physics-informed neural networks (PINNs) and adopt them in our context. Finally, we demonstrate the performance of the proposed methods with several numerical examples, including locally periodic, non-periodic, non-standard, and random homogenization cases.

**2. Background and Problem Setup.** In this section, we first introduce the definition of G-convergence and G-limit in the homogenized equations, given the multiscale elliptic equations. Then we discuss the convergence of homogenization in a special case where the periodic multiscale coefficients are given. Finally, we formulate the inverse problem to learn the G-limits and the corresponding homogenized solutions.

**2.1. G-convergence and G-limit.** We first briefly review the general theory of homogenization and introduce the notion of the G-convergence and G-limit (homogenized coefficient). Let us consider a sequence of the following second order multiscale elliptic equations:

$$\begin{aligned} -\operatorname{div}\left(A^\epsilon(x)\nabla u^\epsilon(x)\right) &= f(x) \quad \text{in } \Omega, \\ u^\epsilon(x) &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.1}$$

where  $\Omega \in \mathbb{R}^N$  is the domain and  $A^\epsilon : \Omega \rightarrow \mathbb{R}^{N \times N}$  is a symmetric multiscale coefficient with finescale size  $\epsilon$ . We consider the sequence of coefficients  $A^\epsilon(x)$  and the corresponding solution  $u^\epsilon(x)$  of (2.1). The G-convergence of the sequence  $A^\epsilon(x)$  is defined as follows [26, 27]:

**DEFINITION 2.1.** *A sequence of coefficient  $A^\epsilon(x)$  in (2.1) is said to G-converge to a limit  $A^*(x)$  as  $\epsilon$  tends to 0, if the sequence of solution  $u^\epsilon(x)$  converges weakly in  $H_0^1(\Omega)$  to  $u_0(x)$ , the unique solution of the following homogenized equation,*

$$\begin{aligned} -\operatorname{div}\left(A^*(x)\nabla u_0(x)\right) &= f(x) \quad \text{in } \Omega, \\ u_0(x) &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.2}$$

for any source term  $f(x)$ . The limit matrix  $A^*(x)$  is called the G-limit of  $A^\epsilon(x)$ .

We now define the following class of matrices.

DEFINITION 2.2. A matrix function  $A(x)$  is said to belong to  $E(\alpha, \beta, \Omega)$  if the followings are satisfied for some  $\alpha, \beta > 0$ .

$$\begin{aligned} A(x) &\in L^\infty(\Omega)^{N \times N}, \\ A(x)k \cdot k &\geq \alpha|k|^2, \text{ for all } k \in \mathbb{R}^N, \text{ a.e. } x \in \Omega \\ |A(x)k| &\leq \beta|k|, \text{ for all } k \in \mathbb{R}^N, \text{ a.e. } x \in \Omega. \end{aligned} \quad (2.3)$$

We have the following theorem that justifies the definition of G-convergence [11].

THEOREM 2.3. Let  $A^\epsilon(x)$  be a sequence of functions that belong to  $E(\alpha, \beta, \Omega)$ . Then there exist a function  $A^*(x) \in E(\alpha, \beta, \Omega)$  such that  $A^\epsilon(x)$  G-converges to  $A^*(x)$  up to subsequence.

The following theorem guarantees the uniqueness of the G-limit.

THEOREM 2.4. The G-limit of a G-converging sequence is unique.

*Proof.* See [11, Section 7]  $\square$

The following remark provides important properties of the G-limit and one motivation for the recovery of the G-limit.

REMARK 2.1. The G-limit  $A^*(x)$  does not depend on the source term  $f(x)$  by definition. It is also known that it also does not depend on the boundary conditions [5, Chapter 1]. Thus, the G-limit recovered with specific source term  $f(x)$  and the boundary condition  $g(x)$  in (2.1) can be reused with the different source terms in the same medium.

From Theorem 2.3, we know that a well-posed homogenized limit (2.2) exists, but in general, there is no systemic way to find the explicit formula for the G-limit  $A^*(x)$ . In addition, the G-convergence is only guaranteed up to a subsequence in the theorem. For (locally) periodic media, the G-convergence is well studied and the G-limit can be computed by the periodic homogenization methods [22, 17]. We remark that even though it might not be clear that how to construct the explicit form of the the G-limit  $A^*(x)$  in general, (2.2) does provide the structure of the homogenized equation served as a generic priori knowledge for PINNs.

**2.2. Homogenization for periodic media.** In this section, we present the outline and the convergence results of the standard periodic homogenization. We let  $\Omega \in \mathbb{R}^N$  be a bounded domain and  $Y$  be a unit cube in  $\mathbb{R}^N$ . We consider the homogenization of the following multiscale elliptic equation:

$$\begin{aligned} -\operatorname{div} \left( A^\epsilon(x) \nabla u^\epsilon(x) \right) &= f(x) \text{ in } \Omega, \\ u^\epsilon(x) &= 0 \text{ on } \partial\Omega. \end{aligned} \quad (2.4)$$

Here,  $\epsilon$  represents the fine scale of the system. The coefficient has the scale separation and is defined by  $A^\epsilon(x) = A(x, \frac{x}{\epsilon})$ , where  $A(x, y)$  is  $Y$ -periodic with respect to the fast variable  $y$ . Thus, we consider the coefficient  $A^\epsilon(x)$  with smooth finescale oscillations. We further assume that  $A^\epsilon(x)$  is in  $C^\infty(\Omega)$  and uniformly positive, i.e.,  $A^\epsilon(x) > c > 0$  for some constant  $c$ .

We consider the following two-scale asymptotic expansion of the solution  $u^\epsilon(x)$ .

$$u^\epsilon(x) = u_0(x) + \epsilon u_1(x, \frac{x}{\epsilon}) + \epsilon^2 u_2(x, \frac{x}{\epsilon}) + \dots, \quad (2.5)$$

where  $u_i(x, \frac{x}{\epsilon})$ , ( $i = 1, 2, \dots$ ) are  $Y$ -periodic with respect to  $y = \frac{x}{\epsilon}$ . We can derive the

following homogenized equation with the G-limit  $A^*(x)$  using the above expansion:

$$\begin{aligned} -\operatorname{div}\left(A^*(x)\nabla u_0(x)\right) &= f(x) \quad \text{in } \Omega, \\ u_0(x) &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{2.6}$$

where the G-limit  $A^*(x)$  is defined as follows:

$$A_{ij}^*(x) = \int_Y A(x, y) \left( \delta_{ij} + \frac{\partial \chi^j(x, y)}{\partial y_i} \right) dy, \tag{2.7}$$

where  $\chi^i(x, y)$  is the solution of the following *cell problem*:

$$\operatorname{div}_y \left( A(x, y) \nabla_y \chi^i(x, y) \right) = -\operatorname{div}_y (A(x, y) e^i), \tag{2.8}$$

on  $Y$  with periodic boundary condition. Here,  $e^i$  is the standard basis vector in  $\mathbb{R}^n$ .

This homogenized equation does not depend on the fine scale  $\epsilon$  and the solution  $u_0(x)$  represents the macroscopic behavior of the solution  $u^\epsilon(x)$  to the multiscale equation (2.4) when  $\epsilon$  is sufficiently small. This can be rigorously explained by the following theorem on the convergence of the multiscale solution  $u^\epsilon(x)$  to the homogenized solution  $u_0$  [22].

**THEOREM 2.5.** *Assume  $A^\epsilon(x) \in L^\infty(\Omega)$ ,  $f(x) \in L^2(\Omega)$ . Let  $u^\epsilon(x)$  and  $u_0(x)$  be the solutions to (2.4) and (2.6) respectively. Then as  $\epsilon \rightarrow 0$ , the sequence  $u^\epsilon(x)$  converges weakly in  $H^1(\Omega)$  to  $u_0(x)$ .*

Above result is obtained under minimal regularity assumptions on the multiscale coefficient and the source term. However, in practice, it is often the case that we can achieve the strong convergence of the multiscale solution to the homogenized solution. For example, we have the following convergence estimates [22, Chapter 6].

**REMARK 2.2.** *Let  $u^\epsilon(x)$  and  $u_0(x)$  be the solutions to (2.4) and (2.6) respectively. Assuming  $A^\epsilon(x)$  and  $f(x)$  are smooth, we have the following convergence rate.*

$$\|u^\epsilon(x) - u_0(x)\|_{L^\infty(\Omega)} \leq C\epsilon, \tag{2.9}$$

where  $C > 0$  is independent of  $\epsilon$ .

In (locally) periodic media, we solve the cell problems (2.8) to compute the G-limit when the explicit form of the multiscale coefficient  $A^\epsilon(x)$  is known. Without the periodicity assumption or a known multiscale coefficient  $A^\epsilon(x)$ , traditional homogenization methods are typically not applicable. Despite the fact that the convergence result (2.9) holds only for periodic cases, we can still expect the multiscale solution data to be close to the homogenized solution for sufficiently small  $\epsilon$  even if the periodic assumption is violated. This motivates us to utilize multiscale solution data as a surrogate for the corresponding homogenized solution data for more general scenarios beyond the periodicity assumptions.

**2.3. Inverse Problem formulation.** Equipped with the background knowledge introduced above, we now consider the multiscale elliptic equations (2.1) and assume a well-posed homogenized limit (2.2) exists. We also assume the multiscale coefficient is smooth, but no geometric assumptions, such as periodicity, are required.

In this work, we consider the following inverse problem setting: given a set of observations/data points, our goal is to learn the G-limit  $A^*(x)$  and the homogenized

solution  $u_0(x)$  of the homogenized limit (2.2). As we mentioned before, the homogenized solution data are often not available. Instead, we utilize the multiscale solution data of the equation (2.1) as a surrogate for the homogenized solution data.

We remark that even though the multiscale solution data are close to the homogenized solution in most of the regions in our domain for sufficiently small  $\epsilon$ , our solution data contains multiscale or noise fluctuations that do not present in the homogenized solution. This introduces additional difficulties because one needs to approximate the slowly varying functions from multiscale solution. It is preferable for a method to be less sensitive to these finescale oscillations and the noise in our multiscale solutions data. Motivated by recent developments of the physics-informed neural networks (PINNs) [19, 24], we propose to develop PINNs for estimating the G-limits, which not only simultaneously match the measurements/data while respecting the underlying physics in the problem, but also provide an effective regularization to mitigate the adversarial effects due to the multiscale fluctuations or noise in the data.

**3. Method.** Next, we will briefly review physics-informed neural networks (PINNs) [19, 10] and adopt it to tackle the inverse problems to learn the G-limits in the homogenized equation (2.2), given the corresponding multiscale solution data.

**3.1. Feed-forward neural network.** We shall use feed-forward neural networks to approximate the solution  $u_0(x)$  and the effective coefficients  $A^*(x)$  in (2.2). The feed-forward neural network with  $L$  layers and  $N_l$  neurons in the  $l$ th layer is a function  $\mathcal{N}_\theta(x) : \mathbb{R}^{N_0} \rightarrow \mathbb{R}^{N_L}$  defined by

$$\begin{aligned}\mathcal{N}_\theta(x) &= W^L \mathcal{N}^{L-1}(x) + b^L, \\ \mathcal{N}^l(x) &= \sigma(W^l \mathcal{N}^{l-1}(x) + b^l), \\ \mathcal{N}^1(x) &= W^1 x + b^1,\end{aligned}\tag{3.1}$$

for  $1 < l < L$ . The matrix  $W^l \in \mathbb{R}^{N_{l-1} \times N_l}$  and the vector  $b^l \in \mathbb{R}^{N_l}$  represent the weight and bias in  $l$ -th layer and  $\sigma$  is a nonlinear activation function, such as ReLU function, the hyperbolic tangent function, and the sine function [14]. We further define the set of tunable weights and biases of the neural network,  $\theta = \{W^l, b^l\}$  for  $1 \leq l \leq L$ .

**3.2. PINNs for inverse problems.** For ease of presentation, we consider the following partial differential equation for the solution  $u(x)$  with an unknown coefficient  $A(x)$ :

$$\mathcal{F}[u(x); A(x)] = 0, \text{ in } \Omega,\tag{3.2}$$

given a Dirichlet boundary condition,

$$u(x) = g(x), \text{ on } \partial\Omega.\tag{3.3}$$

Given the observation data on the solution  $u(x)$  is available, we are interested in recovering the unknown coefficient  $A(x)$  and the entire solution field  $u(x)$ .

PINNs employ the feed-forward networks  $\mathcal{N}_{\theta_u}(x)$  and  $\mathcal{N}_{\theta_A}(x)$  to approximate the solution and the unknown coefficients respectively, where  $\theta_u$  and  $\theta_A$  represent the trainable network parameters for each network. Then we train the networks to get the approximations  $\hat{u}(x)$  for solutions and  $\hat{A}(x)$  for unknown coefficients by minimizing the following loss functional, including data misfit, the PDE residual loss (3.2) and the boundary condition loss (3.3) over the training set  $\mathcal{T}$ :

$$\mathcal{L}(\theta_u, \theta_A; \mathcal{T}) = \lambda_r \mathcal{L}_r(\theta_u, \theta_A; \mathcal{T}_r) + \lambda_d \mathcal{L}_d(\theta_u, \theta_A; \mathcal{T}_d) + \lambda_b \mathcal{L}_b(\theta_u, \theta_A; \mathcal{T}_b),\tag{3.4}$$

where

$$\begin{aligned}\mathcal{L}_r(\theta_u, \theta_A; \mathcal{T}_r) &= \frac{1}{|\mathcal{T}_r|} \sum_{x_i^r \in \mathcal{T}_r} \left| \mathcal{F} \left( \hat{u}(x_i^r); \hat{A}(x_i^r) \right) \right|^2, \\ \mathcal{L}_d(\theta_u, \theta_A; \mathcal{T}_d) &= \frac{1}{|\mathcal{T}_d|} \sum_{x_i^d \in \mathcal{T}_d} |\hat{u}(x_i^d) - u(x_i^d)|^2, \\ \mathcal{L}_b(\theta_u, \theta_A; \mathcal{T}_b) &= \frac{1}{|\mathcal{T}_b|} \sum_{x_i^b \in \mathcal{T}_b} |\hat{u}(x_i^b) - g(x_i^b)|^2,\end{aligned}\tag{3.5}$$

where  $\lambda_r$ ,  $\lambda_d$  and  $\lambda_b$  denote the weights for each loss term. The training points  $\mathcal{T} = \mathcal{T}_r \cup \mathcal{T}_d \cup \mathcal{T}_b$ .  $\mathcal{T}_d$ ,  $\mathcal{T}_r$ , and  $\mathcal{T}_b$  denote data/measurement points, PDE residual points, and boundary data points. Both  $\mathcal{T}_r \in \Omega$ , and  $\mathcal{T}_b \in \partial\Omega$  are predefined and can be chosen from mesh grid points or randomly. The parameters  $\theta_u$  and  $\theta_A$  can be found by minimizing the loss function (3.4), and the resulting networks  $\hat{u}(x)$  and  $\hat{A}(x)$  are the approximations to the solution  $u(x)$  and the coefficient  $A(x)$  of the equation (3.2).

**3.3. Learning the G-limits via PINNs.** Following that, we adopt the PINNs framework to tackle the inverse problem of estimating the G-limit  $A^*(x)$  for the multiscale elliptic equation (2.1). One issue is that the measurements of the homogenized solution are often not available. Motivated by the convergence results for periodic media in (2.9), we employ the multiscale solution data  $u^\epsilon$  of the multiscale equation (2.1) as the training data, which is expected to be a good surrogate for the homogenized solution data when  $\epsilon$  is sufficiently small.

We construct two feed-forward neural networks  $\hat{A}^*(x) = \mathcal{N}_{\theta_{A^*}}(x)$  and  $\hat{u}_0(x) = \mathcal{N}_{\theta_{u_0}}(x)$  to approximate the G-limit and the solution of the homogenized equation (2.2). Since we consider Dirichlet boundary condition (3.3) in this work, the boundary condition can be embedded into the neural network exactly. Specifically, we follow the approach suggested in [20] by modifying the solution network output  $\mathcal{N}_{\theta_{u_0}}$ :

$$\hat{u}_0(x) = g(x) + l(x)\mathcal{N}_{\theta_{u_0}},\tag{3.6}$$

where  $u_0(x) = g(x)$  is a Dirichlet boundary condition, and  $l(x)$  is a function that satisfies the following conditions.

$$l(x) = 0 \text{ on } \partial\Omega, \quad l(x) > 0 \text{ in } \Omega - \partial\Omega.\tag{3.7}$$

With a simple domain, we can analytically choose  $l(x)$  [18]. For example, for the domain  $[a, b]^2$ , we can choose  $l(x) = (x_1 - a)(b - x_1)(x_2 - a)(b - x_2)$ , where  $x = (x_1, x_2)$ .

We then seek a set of network parameters  $\theta_{u_0}$  and  $\theta_{A^*}$  that minimize the loss function defined as follows:

$$\mathcal{L}(\theta_{u_0}, \theta_{A^*}; \mathcal{T}) = \lambda_r \mathcal{L}_r(\theta_{u_0}, \theta_{A^*}; \mathcal{T}_r) + \lambda_d \mathcal{L}_d(\theta_{u_0}, \theta_{A^*}; \mathcal{T}_d),\tag{3.8}$$

where

$$\begin{aligned}\mathcal{L}_r(\theta_{u_0}, \theta_{A^*}; \mathcal{T}_r) &= \frac{1}{|\mathcal{T}_r|} \sum_{x_i^r \in \mathcal{T}_r} \left| \operatorname{div} \left( \hat{A}^*(x_i^r) \nabla \hat{u}_0(x_i^r) \right) + f(x_i^r) \right|^2, \\ \mathcal{L}_d(\theta_{u_0}, \theta_{A^*}; \mathcal{T}_d) &= \frac{1}{|\mathcal{T}_d|} \sum_{x_i^d \in \mathcal{T}_d} |\hat{u}_0(x_i^d) - u^\epsilon(x_i^d)|^2.\end{aligned}\tag{3.9}$$

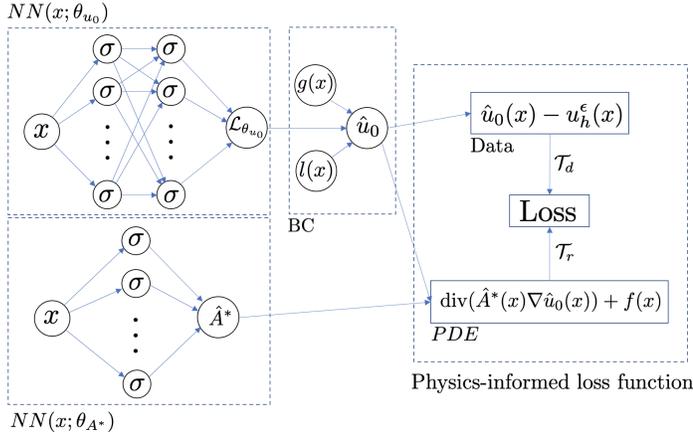


Fig. 3.1: The schematic architecture of PINNs for learning the G-limit  $A^*(x)$  in the homogenized equation (2.2) by the neural network  $\hat{A}^*(x)$ . The boundary condition  $u_0(x) = g(x)$  is strictly imposed using (3.6).

Here,  $u^\epsilon(x_i^d)$  denotes the (noise-free/noisy) multiscale solution data at  $x_i^d$ . Note that since the neural network  $\hat{u}_0(x)$  satisfies the boundary condition exactly, there are only two terms in the loss function. The first term (3.8) encourages the neural network to respect the homogenized equation (2.2). The second term makes sure that the approximated homogenized solution is not far from the multiscale solution data  $u^\epsilon(x)$ . The choice of the regularization parameters  $\lambda_r$  and  $\lambda_d$  could affect the training performance considerably. We adopted the adaptive weight techniques [28] in this work. In summary, Figure 3.1 presents the schematic design of the PINNs for our problem.

**4. Numerical Examples.** In this section, we present several numerical examples to illustrate the effectiveness and applicability of our method, including the elliptic equations with locally periodic, non-periodic, and ergodic random multiscale coefficients. The noise-free measurements are generated from the multiscale solution data of the multiscale elliptic equation by the underlying forward FEM simulation. For noisy scenario, we corrupt the noise-free data with independent, and identically distributed normal noise with different noise levels.

To estimate the accuracy of the recovered G-limit  $\hat{A}^*(x)$  and the homogenized solutions  $\hat{u}_0(x)$ , we use the following relative  $L^2$ -errors computed over a predefined mesh grid in spatial domain:

$$e_{\hat{A}^*} = \frac{\|\hat{A}^*(x) - A^*(x)\|_{L^2(\Omega)}}{\|A^*(x)\|_{L^2(\Omega)}}, \quad e_{\hat{u}_0} = \frac{\|\hat{u}_0(x) - u_{0,h}(x)\|_{L^2(\Omega)}}{\|u_{0,h}(x)\|_{L^2(\Omega)}}. \quad (4.1)$$

Here,  $A^*(x)$  is the reference G-limit that is either exact or pre-computed by FEM via traditional homogenization methods. The reference homogenized solutions,  $u_{0,h}(x)$ , are computed by FEM using the reference G-limits.

During the training stage, we alternatively use ADAM and L-BFGS as suggested in [19, 25]. A hyperbolic tangent function is used as the activation in all examples. In addition, the architectural parameters of neural network were tuned to achieve reasonable results. The architecture parameters and other hyperparameters used for each

example are listed in *Table A.1* in the appendix. Advanced hyperparameter selection techniques can further improve the results, which, however, is not the focus of this work. In addition, the multiple restarts approach is adopted to prevent the results from being affected by how the weights are (randomly) initialized. More specifically, we train the nets with a number of random initialization using Glorot normal initializer, and report the best possible results for each example. All examples are carried out on Google’s Colab [9] using the library SciANN [16].

**4.1. Homogenization of a slowly varying periodic coefficient.** To test the basic capability of our proposed method, we first consider the following multiscale elliptic equation with a slowly varying periodic coefficient:

$$-\frac{d}{dx} \left( \frac{1+x^2}{2+\sin(2\pi\frac{x}{\epsilon})} \frac{d}{dx} u^\epsilon(x) \right) = \cos(\pi x) \quad \text{in } \Omega = [0, 1], \quad (4.2)$$

$$u^\epsilon(0) = u^\epsilon(1) = 0.$$

In this example, the permeability coefficient depends on both  $x$  and  $\frac{x}{\epsilon}$ , and is periodic with respect to  $\frac{x}{\epsilon}$ . The analytical G-limit is known as  $A^*(x) = \frac{x^2+1}{2}$ . We compute the reference homogenized solution  $u_{0,h}(x)$  using FEM via the exact G-limit.

To generate the noise-free data, we compute the multiscale solution to the equation (4.2) for each finescale parameter value  $\epsilon$  by FEM with mesh size  $h = 1/10^5$  and obtain equally spaced data sampled from the multiscale solution as the training data. For noisy data, we corrupt the measurements with different noise levels. The architecture parameters and other hyperparameters of PINNs are listed in *Table A.1* in appendix. The relative  $L^2$  errors for both G-limit and homogenized solution are computed on a mesh with size  $h = 1/10^5$ .

We first plot the relative  $L^2$  errors with respect to the size of training data with  $\epsilon = 2^{-7}$  in Figure 4.1. For the noise-free case, the proposed method can achieve the errors at the level of  $\mathcal{O}(10^{-3})$  for the G-limit and  $\mathcal{O}(10^{-4})$  for the homogenized solution. As the data set was enriched, the error level saturated. With noisy data, the error increases with the noise level and can be reduced as additional data are available, particularly for a high noise level. With 5% noise, the relative errors for the approximated G-limit and solution are roughly 4% and 1% respectively, given enough data. To further demonstrate the performance of the method, Figure 4.2 plots the G-limits and the homogenized solution recovered by PINNs for  $\epsilon = 2^{-7}$ , where both G-limit and homogenized solution are well approximated under the different noise levels. As shown in Figure 4.2e and 4.2f, even when the multiscale data contain non-negligible random fluctuations, PINNs can still capture the macroscopic variation of the data thanks to the regularization provided by the homogenized equation.

To investigate the impacts of the finescale parameter  $\epsilon$  of the multiscale data on the performance of our algorithm, we plot the relative errors for the G-limit and the homogenized solution with different values of  $\epsilon$  in Figure 4.3. For each  $\epsilon$  value,  $|\mathcal{T}_d| = 160$  multiscale solution data collected at fixed spatial locations are used. In noise-free cases, the errors for homogenized solutions tend to decrease when  $\epsilon$  becomes smaller. This is expected as multiscale solution data are closer to the homogenized solution for smaller  $\epsilon$ .

Figure 4.4 further shows the learned G-limit and homogenized solution with different finescale parameters  $\epsilon$ . We can observe that the multiscale data converge to the reference homogenized solution as  $\epsilon$  becomes smaller. For example, when  $\epsilon = 2^{-7}$ , our data almost overlap with the reference homogenized solution (Figure 4.4f) and

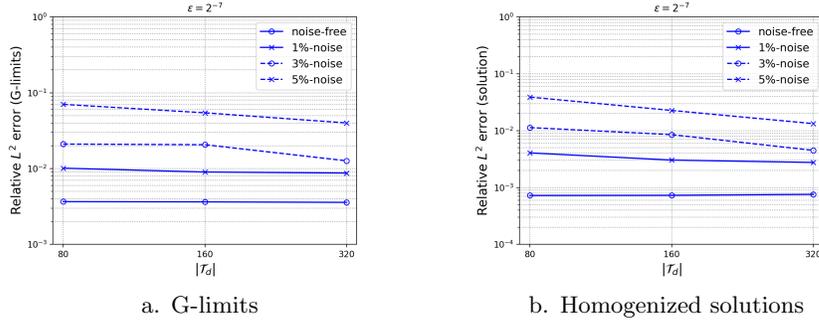


Fig. 4.1: Problem (4.2): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different number of multiscale data points corrupted by different noise levels for  $\epsilon = 2^{-7}$  and the number of PDE residual points is  $|\mathcal{T}_r| = |\mathcal{T}_d| + 30$ .

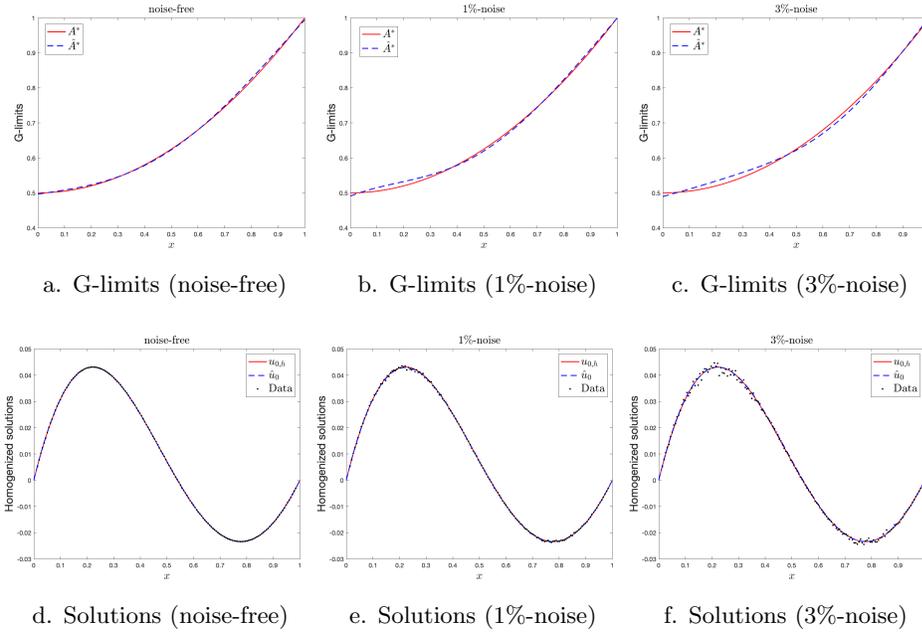


Fig. 4.2: Error results for problem (4.2): comparison of the reference solutions ( $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts ( $\hat{A}^*(x)$  and  $\hat{u}_0(x)$ ) learned by PINNs with different noise levels in the data. (a. b. c.): the G-limit; (d. e. f.): the homogenized solution, where  $\epsilon = 2^{-7}$  and the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$ ,  $|\mathcal{T}_r| = 190$ .

both G-limit and the homogenized solution learned by PINNs agree very well with their references. Furthermore, despite the presence of noticeable multiscale oscillations in the data, PINNs can still provide reasonably good results for larger epsilons ( $\epsilon = 2^{-3}$ ,  $2^{-5}$ ). This is because the proposed PINN tends to promote the smooth macroscale behavior of the data rather than their microscale fluctuations shown in Figure 4.4d and 4.4e.

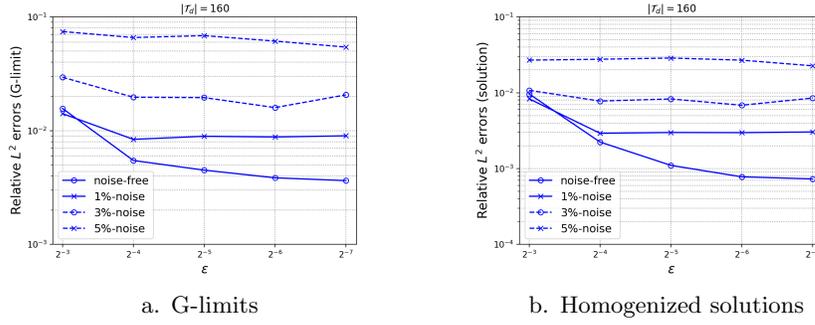


Fig. 4.3: Error results for problem (4.2): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different finescale parameter  $\epsilon$  and noise levels, when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$  and  $|\mathcal{T}_r| = 190$ .

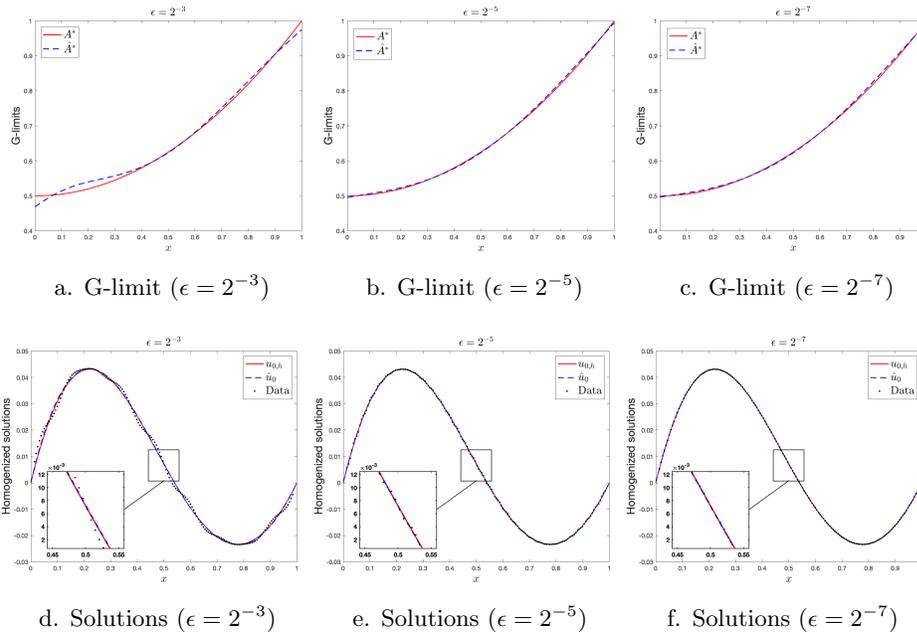


Fig. 4.4: Problem (4.2) with noise-free data: comparison of the reference solutions ( $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts ( $\hat{A}^*(x)$  and  $\hat{u}_0(x)$ ) learned by PINNs with different values of  $\epsilon$ . (a. b. c.): the G-limit; (d. e. f.): the homogenized solution, when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$ ,  $|\mathcal{T}_r| = 190$ .

For noisy scenarios, the approximation quality deteriorates with the noise level as seen in Figure 4.3. We also observe that the impact of the finescale size  $\epsilon$  of the medium becomes negligible once the noise level is large enough, suggesting that the magnitude of the noises is dominant over the multiscale oscillations in our data. Nonetheless, our approach can still provide reasonably good approximations under a mild noise level.

**4.2. Homogenization of a heavily oscillatory coefficient.** Next, we consider the following elliptic equation with a heavily oscillatory permeability coefficient introduced in [12]:

$$-\frac{d}{dx} \left( A^\epsilon(x) \frac{d}{dx} u^\epsilon(x) \right) = 3 + \sin(x) \quad \text{in } \Omega \in [0, 1], \quad (4.3)$$

$$u^\epsilon(0) = 0, \quad u^\epsilon(1) = 0,$$

where  $A^\epsilon(x) = \int_Y \left( 1 + \frac{1}{2} \sin \left( \left( y + \frac{1}{2\epsilon} \sin \left( \pi \sqrt{\frac{2}{\epsilon}} x \right) \right)^2 \right) \right) e^{y(1+\sin x)} dy$ . The coefficient  $A^\epsilon(x)$  is quite oscillatory. Figure 4.5 illustrates the multiscale coefficients  $A^\epsilon(x)$  and the effective coefficients  $A^*(x)$  for  $\epsilon = 2^{-3}, 2^{-5}$ . Due to strong oscillations in the coefficients, direct numerical simulation of this problem is very expensive when the formula for  $A^\epsilon(x)$  is known. This homogenization problem is in general challenging: (1) The explicit integral of the multiscale coefficient is not available. (2) This problem cannot be handled by the traditional homogenization method, such as the two-scale convergence method, because the oscillations in  $A^\epsilon(x)$  cannot be captured by any test functions admissible for the two-scale convergence [4]. For this example, it can be shown that the analytical G-limit coincides with the weak  $L^2$  limit of  $A^\epsilon(x)$  given by  $A^*(x) = \frac{e^{(1+\sin x)} - 1}{1 + \sin x}$  [12], but this is not the case in general [5, Chapter1].

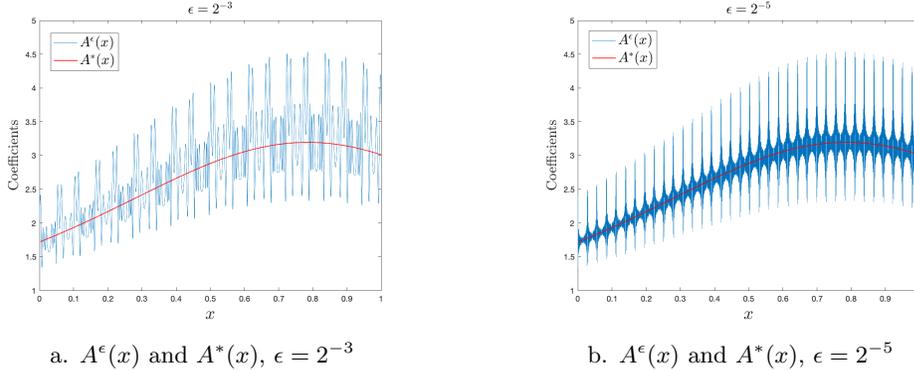


Fig. 4.5: The G-limits  $A^*(x)$  and multiscale coefficients  $A^\epsilon(x)$  with  $\epsilon = 2^{-3}$  and  $\epsilon = 2^{-5}$  for problem (4.3)

The synthetic training data set are equally spaced sampled from the multiscale solution for each finescale parameter value  $\epsilon$  computed by FEM with a mesh size of  $h = 1/10^5$ . The reference homogenized solution is computed by FEM with the same mesh based on the analytic G-limit. The architecture parameters and other hyperparameters of PINNs are listed in *Table A.1* in appendix. The relative  $L^2$  errors for both G-limit and homogenized solution are computed based on the same mesh aforementioned.

We first consider the case with a relatively small finescale size  $\epsilon = 2^{-7}$ . Figure 4.6 presents the relative  $L^2$  errors for the G-limit and homogenized solution with respect to the number of multiscale solution data. With noise-free data, we can achieve an error level of  $\mathcal{O}(10^{-3})$  for both homogenized coefficient and the homogenized solution.

It appears that 80 multiscale data are enough to obtain good approximations. With a high noise level, PINNs can still achieve satisfactory approximations when the data set is large enough. This can be further supported by the corresponding G-limit and homogenized solution obtained by PINNs under different levels of noise corruptions in Figure 4.7. It is clear that the proposed method can still capture the G-limit and the smooth homogenized solution accurately under mild noise corruptions. This can be further evidenced by Figure 4.7e and 4.7f where the learned homogenized solutions tend to fit the macroscopic behavior of the noisy data that is close to the reference solution.

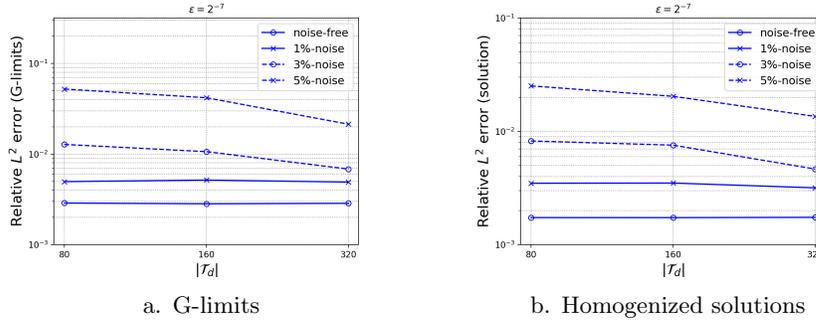


Fig. 4.6: Error results for problem (4.3): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different number of multiscale data corrupted by different noise levels, when  $\epsilon = 2^{-7}$  and the number of PDE residual points is  $|\mathcal{T}_r| = |\mathcal{T}_d| + 30$ .

Figure 4.8 presents errors of the estimated G-limit and the homogenized solution with respect to the finescale parameter  $\epsilon$  and 160 multiscale solution data collected at fixed spatial locations for all  $\epsilon$ , i.e.,  $|\mathcal{T}_d| = 160$ . As expected, better approximation of the G-limit and the homogenized solution can be delivered as  $\epsilon$  becomes smaller in noise-free case. In addition, Figure 4.9 again shows that the learned homogenized solutions tend to fit the multiscale solution data. We note that even if the finescale oscillations in our data are not visible in the figures, a relatively large  $\epsilon$  ( $=2^{-3}, 2^{-5}$ ) could result in the non-negligible difference between the reference homogenized solution and our data. As a result, approximations of G-limits and homogenized solutions are less accurate but satisfactory for larger  $\epsilon$ . In contrast to noise-free scenarios, the finescale size  $\epsilon$  has much less impact on both learned G-limit and the homogenized solution as the noise level increases as shown in Figure 4.8.

**4.3. Homogenization of a 2D non-periodic coefficient.** We next consider the following 2D multiscale elliptic equation with a non-periodic coefficient introduced in [23]:

$$\begin{aligned} -\operatorname{div}\left(A^\epsilon(x) \cdot \nabla u^\epsilon(x)\right) &= 1 \quad \text{in } \Omega = [1, 2]^2 \\ u^\epsilon(x) &= 0 \quad \text{on } \partial\Omega, \end{aligned} \quad (4.4)$$

where  $A^\epsilon(x) = \left(1 + 0.9 \sin(2\pi \frac{x_1}{\epsilon}) \sin(2\pi \frac{x_2}{\epsilon})\right)$ . Figure 4.10 illustrates the multiscale coefficient  $A^\epsilon(x)$  when  $\epsilon = 2^{-3}$ . The G-limit  $A^* = \begin{pmatrix} A_{11}(x) & A_{12}(x) \\ A_{21}(x) & A_{22}(x) \end{pmatrix}$  that is a  $2 \times 2$

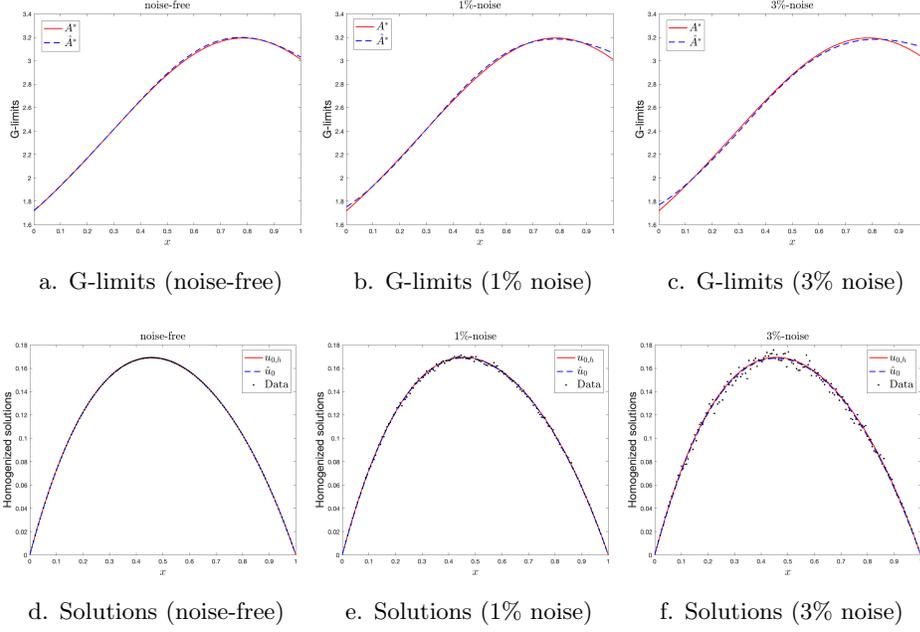


Fig. 4.7: Problem (4.3): comparison of the reference solutions ( $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts ( $\hat{A}^*(x)$  and  $\hat{u}_0(x)$ ) learned by PINNs with different noise levels in the data. (a. b. c.): the G-limit; (d. e. f.): the homogenized solution, when  $\epsilon = 2^{-7}$  and the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$ ,  $|\mathcal{T}_r| = 190$ .

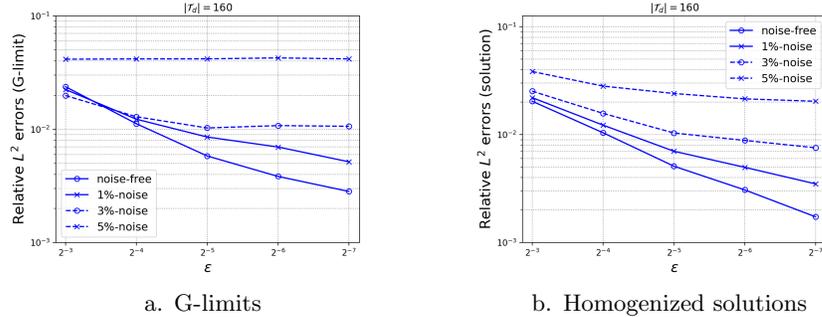


Fig. 4.8: Error results for problem (4.3): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different finescale parameter  $\epsilon$  and noise levels of data, when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$  and  $|\mathcal{T}_r| = 190$ .

matrix function, can be found via the  $\lambda$ -scale convergence technique [23]. Since  $A^\epsilon(x)$  is periodic with respect to  $x_1$ , we know that the G-limit only depends on  $x_2$ , i.e.,  $A^*(x) = A^*(x_2)$ . We assume a priori that the non-diagonal entries of the G-limit are zeros i.e.,  $A_{12}(x_2) = A_{21}(x_2) = 0$ . Therefore, we shall only approximate the diagonal entries of the G-limit. The reference G-limit is computed by the  $\lambda$ -scale convergence method. Specifically, we solved local cell problems at 129 equidistant points of  $x_2$  and

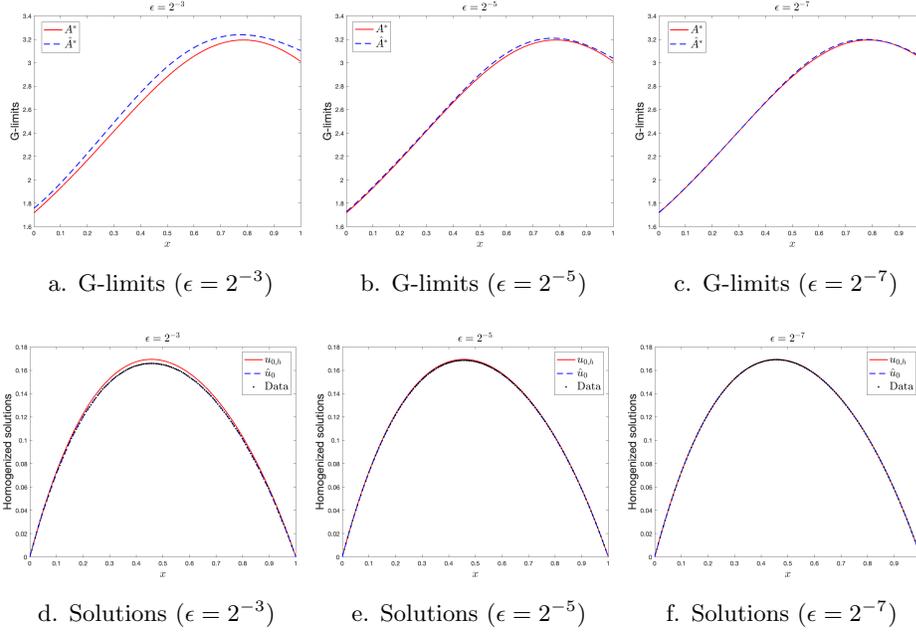


Fig. 4.9: Problem (4.3) with noise-free data: comparison of the reference solutions ( $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts ( $\hat{A}^*(x)$  and  $\hat{u}_0(x)$ ) learned by PINNs with different values of  $\epsilon$ . (a. b. c.): the G-limit; (d. e. f.): the homogenized solution, when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$ ,  $|\mathcal{T}_r| = 190$ .

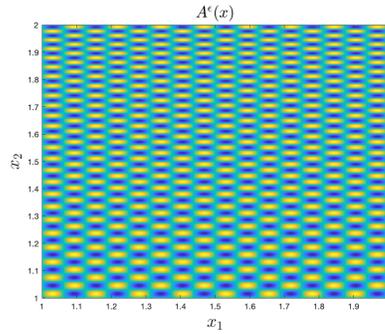


Fig. 4.10: The multiscale permeability coefficient  $A^\epsilon(x)$  in (4.4) when  $\epsilon = 2^{-3}$ .

each problem is solved by FEM with a mesh size of  $h = 1/1000$ . With the reference G-limit, we computed the reference homogenized solution by FEM with a mesh size  $h = 1/128$ .

The training data are equally spaced sampled from the multiscale solution for each finescale parameter value  $\epsilon$  obtained by the forward FEM simulation of the problem (4.4) with a fine mesh size  $1/8000$ . The architecture parameters and other hyperparameters of PINNs are listed in *Table A.1* in appendix. We compute the errors on a mesh with size  $h = 1/128$  in the spatial domain.

Figure 4.11 presents the error convergence of the G-limit and the homogenized solutions for  $\epsilon = 2^{-7}$  with different numbers of training data. With noise-free data, 400 data points appear to be sufficient to obtain good approximations with errors of less than  $10^{-3}$  for both G-limit and homogenized solution. Increasing the amount of training data helps improve the accuracy for noisy data cases. Even with 5%-noise in the data, our proposed method can still achieve an error less than  $\mathcal{O}(10^{-2})$  for both coefficient and the solution, when the number of available data is large enough. We also plot the G-limits and the homogenized solutions at  $x_2 = 1.25$  shown in Figure 4.12 when  $\epsilon = 2^{-7}$ . Both diagonal entries of the G-limit and homogenized solution agree well with their references. Again, we observe that the learned solutions tend to fit the macroscale behaviors of the data, even when non-negligible noises present in 4.12e and 4.12f.

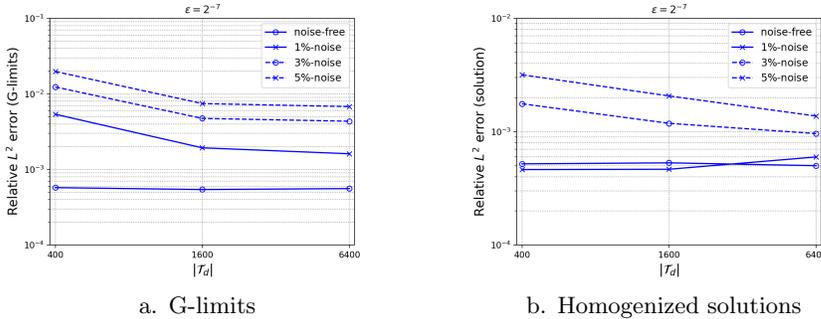


Fig. 4.11: Error results for problem (4.4): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different number of multiscale data corrupted by different noise levels, when  $\epsilon = 2^{-7}$  and the number of PDE residual points is  $|\mathcal{T}_r| = |\mathcal{T}_d|$ .

Figure 4.13 shows the relative  $L^2$  errors for the learned G-limits and the homogenized solutions for different finescale parameter  $\epsilon$  and 1600 multiscale solution data collected at fixed spatial locations for all  $\epsilon$ , i.e.,  $|\mathcal{T}_d| = 1600$ . For noise-free scenarios, the error decays as the finescale size  $\epsilon$  decreases. To further examine this effect, we plot the corresponding the learned G-limit and homogenized solutions in Figure 4.14. In particular, the approximation quality of G-limits appears to be more sensitive to the size of  $\epsilon$  for the noise-free case. When the noise dominates over the multiscale oscillations in the data, the results are no longer sensitive to the size of finescale. Nonetheless, we can still achieve the errors of less than 1% for both G-limit and the homogenized solutions under a 5% noise level.

To further highlight the performance of the proposed method, we also show the learned homogenized solution with the reference solutions in Figure 4.15. As we can see, our solutions agree very well with the reference solutions. Overall, our results show that when mild noise and multiscale fluctuations are presented in the data, the PINNs can provide good estimations of the G-limit for the 2D non-periodic example.

**4.4. Homogenization of an ergodic random coefficient.** Finally, we consider the following two-scale elliptic equation with an ergodic coefficient, inspired by

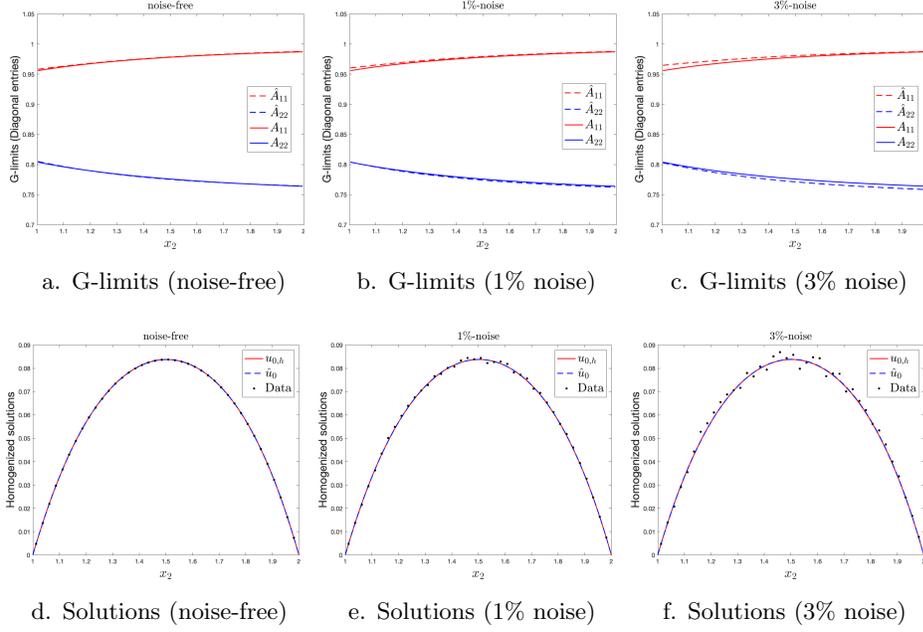


Fig. 4.12: Problem (4.4): comparison of the reference solutions (the diagonal entries of  $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts learned by PINNs with different noise levels. (a. b. c.): the G-limit; (d. e. f.): the homogenized solution at  $x_1 = 1.25$ , where  $\epsilon = 2^{-7}$ , the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 1600$ ,  $|\mathcal{T}_r| = 1600$ .

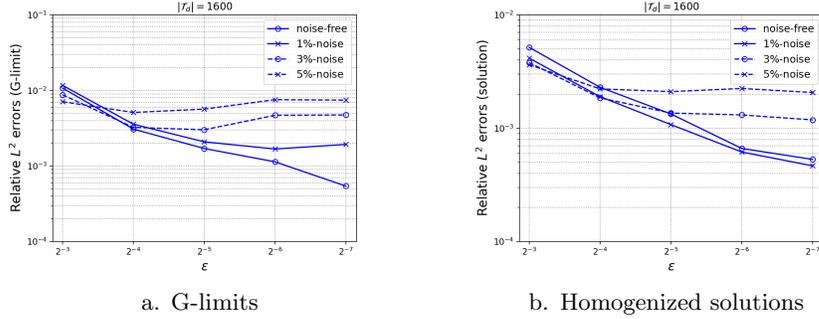


Fig. 4.13: Error results for problem (4.4): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different finescale parameter  $\epsilon$  and noise levels in the data when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 1600$  and  $|\mathcal{T}_r| = 1600$ .

the example in [8, Section 4.2]:

$$\begin{aligned}
 -\frac{d}{dx} \left( A^\epsilon(x, \omega) \cdot \frac{d}{dx} u^\epsilon(x) \right) &= 1 \quad \text{in } \Omega = [0, 1], \\
 u^\epsilon(0) &= u^\epsilon(1) = 0,
 \end{aligned} \tag{4.5}$$

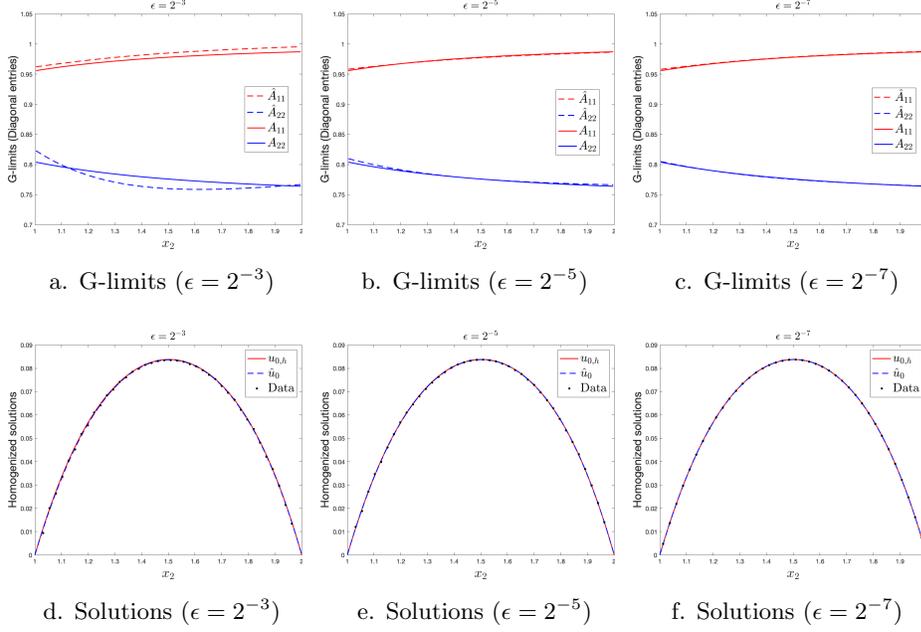


Fig. 4.14: Problem (4.4) with noise-free data: comparison of the reference coefficient and homogenized solutions (the diagonal entries of  $A^*(x)$  and  $u_{0,h}(x)$ ) and counterparts learned by PINNs with different values of  $\epsilon$ . (a. b. c.): the G-limit; (d. e. f.): the homogenized solution at  $x_1 = 1.25$ , where the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 1600$ ,  $|\mathcal{T}_r| = 1600$ .

where  $A^\epsilon(x, \omega) = A(x, T_\frac{x}{\epsilon}(\omega)) = 3.1 + (x + 1) \sin(2\pi(\omega_1 + \frac{x}{\epsilon})) + \sin(2\pi(\omega_2 + \sqrt{2}\frac{x}{\epsilon}))$  for  $\omega = (\omega_1, \omega_2)$  drawn from a uniform distribution over  $[0, 1]^2$ . Here the ergodic dynamical system  $T : \mathbb{R} \times \mathcal{Z} \rightarrow \mathcal{Z}$  is given by

$$T(x)\omega = \omega + (1, \sqrt{2})x. \quad (4.6)$$

Notably, the G-limit  $A^*(x)$  of this ergodic homogenization problem is deterministic and independent of the realization of  $\omega$  [17]. In this example, it is known that the exact G-limit is given by  $1/\mathbb{E}[1/A(x, \omega)]$ , where  $\mathbb{E}$  denotes the expectation with respect to the realizations of  $\omega$  [3]. Traditional approaches for ergodic homogenization usually first compute the local cell problems for many different realizations of the coefficient  $A^\epsilon(x, \omega)$  to obtain the realization dependent approximations to the G-limits. Then the G-limit can be approximated by taking its expectation. This procedure requires solving a lot of cell problems at many different points  $x$  with thousands of realizations of  $\omega$ . For PINNs, on the other hand, we just need to collect the multiscale solution data based on a single realization of the coefficient for PINNs.

Specifically, the reference G-limit  $A^*(x)$  is computed as the expectation by roughly 200,000 Monte Carlo samples over 2000 equidistant points in the spatial domain. Based on this G-limit, we compute the reference homogenized solution by FEM with mesh size  $h = 1/2000$ . For PINNs, we learned the G-limit based on only one re-

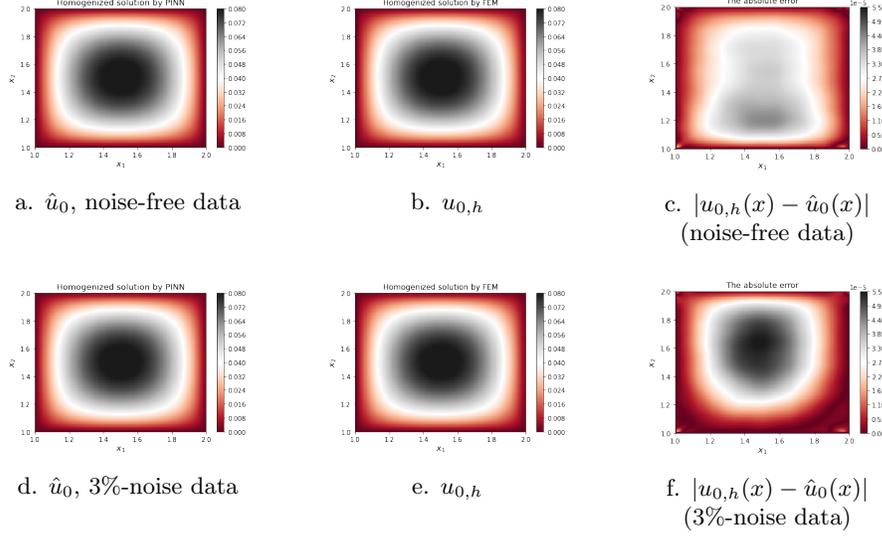


Fig. 4.15: The 2D Homogenized solutions of problem (4.4) obtained with noise-free and 3%-noise data. (a. d.): Homogenized solution obtained by PINNs; (b. e.): the reference homogenized solution; (c. f.): the absolute error between the two solutions, when  $\epsilon = 2^{-7}$ , the number of multiscale solution data and PDE residual points used are  $|\mathcal{T}_d| = 1600$ ,  $|\mathcal{T}_r| = 1600$ .

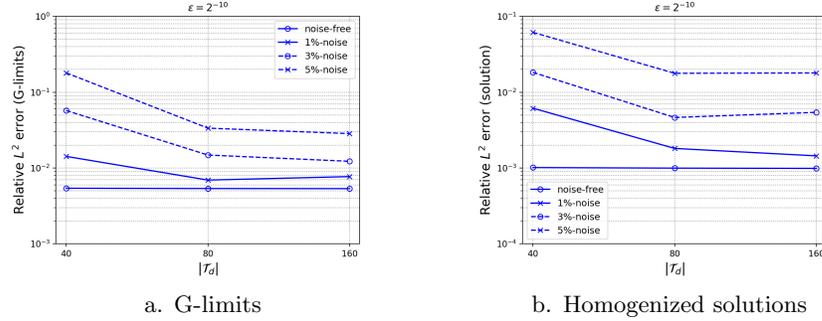


Fig. 4.16: Error results for problem (4.5): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different number of multiscale data corrupted by different noise levels for  $\epsilon = 2^{-10}$  and the number of PDE residual points is  $|\mathcal{T}_r| = |\mathcal{T}_d| + 20$ .

alization of  $\omega = (0.5, 0.5)$ . The training data are equally spaced sampled from the multiscale solution for each finescale parameter value  $\epsilon$  computed by FEM with a mesh size  $h = 1/10^5$ . The architecture parameters and other hyperparameters of PINNs are listed in *Table A.1* in the appendix. The relative  $L^2$  errors are computed using a mesh of size  $h = 1/2000$ .

Figure 4.16 shows the error convergence of the learned G-limits and homogenized solution for  $\epsilon = 2^{-10}$ . For the noise-free case, an error level  $\mathcal{O}(10^{-3})$  for both G-limit

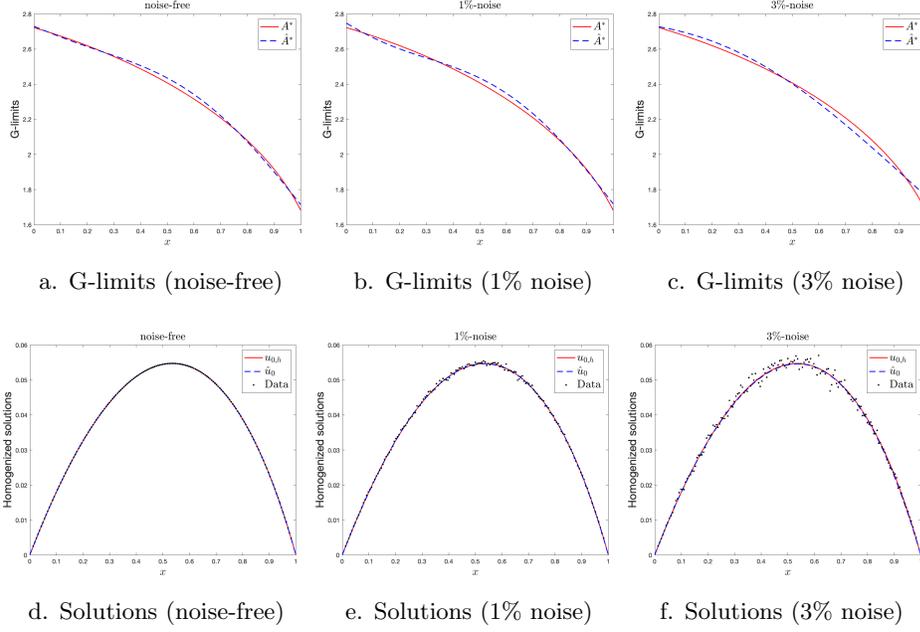


Fig. 4.17: Problem (4.5): comparison of the reference coefficient and solutions ( $A^*(x)$  and  $u_{0,h}(x)$ ) and the counterparts learned by PINNs with different noise levels in the data. (a. b. c.): the G-limit; (d. e. f.): the homogenized solution, where  $\epsilon = 2^{-10}$  and the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$ ,  $|\mathcal{T}_r| = 180$ .

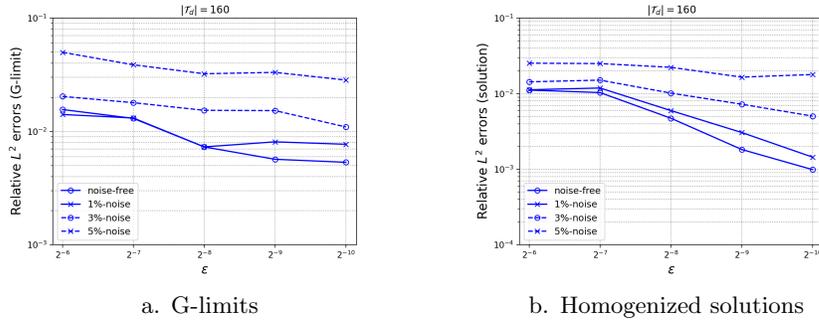


Fig. 4.18: Error results for problem (4.5): the relative  $L^2$  errors for the G-limits and the homogenized solutions with different finescale parameter  $\epsilon$  and noise levels in the data when the number of multiscale data and PDE residual points are  $|\mathcal{T}_d| = 160$  and  $|\mathcal{T}_r| = 180$ .

and the homogenized solution can be achieved. For noisy data, while the errors for the G-limit and homogenized solution tend to stagnate after more than 80 multiscale data points are used, we can still achieve errors of  $\mathcal{O}(10^{-2})$  for both G-limit and the homogenized solution with 5%-noise corruption in the data.

We also compare the learned G-limits and the homogenized solutions with their references for  $\epsilon = 2^{-10}$  and  $|\mathcal{T}_d| = 160$  data with different noise levels in Figure 4.17. While the learned G-limit is close to the reference coefficient, the approximated

homogenized solution almost overlaps with the data. Again, we observed PINNs tend to learn the macroscopic behavior of the noisy data that is close to the reference homogenized solution.

The error results with finescale parameter  $\epsilon$  are presented in Figure 4.18. For both noise-free and noisy scenarios, the errors tend to decay as the finescale size  $\epsilon$  decreases, particularly for homogenized solutions. The effect of  $\epsilon$  is less pronounced when the noise level is high because the noise dominates over the fine scale size of  $\epsilon$ . Notably, with 5%-noise corruption, we can still achieve the relative errors less than 5% for both G-limit and homogenized solution by incorporating the corresponding homogenized equation.

**5. Conclusion.** In this paper, we proposed a simple and flexible approach to estimate the G-limit and approximate the homogenized solution for multiscale elliptic equations from data, by adopting physics-informed neural networks (PINNs). Due to the lack of the homogenized solution data or measurements, we employ the multiscale solution data as the surrogate of the homogenized solution. Despite the rapid multiscale and noisy fluctuations presented in the data, we demonstrated that PINNs are capable to effectively extract the macroscopic (homogenized) behavior from data and provide good approximations to the G-limits and the homogenized solution. The applicability and performance of the method have been demonstrated through a number of different benchmark problems. Finally, we remark that except for the assumption of the existence and structure of the homogenized equation, our approach does not rely on the periodicity or the explicit formula of the underlying multiscale coefficient during the learning stage, which can be applicable to more general settings beyond periodic cases.

**Acknowledgments.** XZ was supported by Simons Foundation.

**Appendix A. Hyperparameters used in each numerical example.**

Table A.1: Hyperparameters used for each numerical example: For each example, the learning rate is decayed when the training loss plateaus.

1D locally periodic coefficient (Section 4.1)						
NN depth		NN width		Initial learning rate	# of epochs	Batch size
$\hat{A}^*(x)$	$\hat{u}_0(x)$	$\hat{A}^*(x)$	$\hat{u}_0(x)$			
3	3	30	30	0.001	40000	64
1D heavily oscillatory coefficient (Section 4.2)						
NN depth		NN width		Initial learning rate	# of epochs	Batch size
$\hat{A}^*(x)$	$\hat{u}_0(x)$	$\hat{A}^*(x)$	$\hat{u}_0(x)$			
3	3	50	50	0.0001	80000	64
2D non-periodic coefficient (Section 4.3)						
NN depth		NN width		Initial learning rate	# of epochs	Batch size
$\hat{A}_{ii}^*(x)$	$\hat{u}_0(x)$	$\hat{A}_{ii}^*(x)$	$\hat{u}_0(x)$			
2	4	40	45	0.001	100000	200
1D ergodic random coefficient (Section 4.4)						
NN depth		NN width		Initial learning rate	# of epochs	Batch size
$\hat{A}^*(x)$	$\hat{u}_0(x)$	$\hat{A}^*(x)$	$\hat{u}_0(x)$			
2	3	10	30	0.001	60000	64

## REFERENCES

- [1] Assyr Abdulle and Andrea Di Blasio. A bayesian numerical homogenization method for elliptic multiscale inverse problems. *SIAM/ASA Journal on Uncertainty Quantification*, 8(1):414–450, 2020.
- [2] Assyr Abdulle, Giacomo Garegnani, and Andrea Zanoni. Ensemble kalman filter for multiscale inverse problems. *Multiscale Modeling & Simulation*, 18(4):1565–1594, 2020.
- [3] Alen Alexanderian. A primer on homogenization of elliptic pdes with stationary and ergodic random coefficient functions. *arXiv preprint arXiv:1408.5827*, 2014.
- [4] Grégoire Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.
- [5] Grégoire Allaire. *Shape optimization by the homogenization method*, volume 146. Springer Science & Business Media, 2012.
- [6] Grégoire Allaire and Marc Briane. Multiscale convergence and reiterated homogenisation. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 126(2):297–342, 1996.
- [7] Hassan Arbabi, Judith E Bunder, Giovanni Samaey, Anthony J Roberts, and Ioannis G Kevrekidis. Linking machine learning with multiscale numerics: data-driven discovery of homogenized equations. *Jom*, 72(12):4444–4457, 2020.
- [8] Donald L Brown and Viet Ha Hoang. A hierarchical finite element monte carlo method for stochastic two-scale elliptic equations. *Journal of Computational and Applied Mathematics*, 323:16–35, 2017.
- [9] Tiago Carneiro, Raul Victor Medeiros Da Nóbrega, Thiago Nepomuceno, Gui-Bin Bian, Victor Hugo C De Albuquerque, and Pedro Pedrosa Reboucas Filho. Performance analysis of google colaboyatory as a tool for accelerating deep learning applications. *IEEE Access*, 6:61677–61685, 2018.
- [10] Yuyao Chen, Lu Lu, George Em Karniadakis, and Luca Dal Negro. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Optics Express*, 28(8):11618–11633, 2020.
- [11] Anneliese Defranceschi. An introduction to homogenization and g-convergence. *School on Homogenization, Lecture notes of the courses held at ICTP, Trieste*, pages 63–122, 1993.
- [12] Liselott Flodén. *G-convergence and homogenization of some sequences of monotone differential operators*. PhD thesis, Mittuniversitetet, 2009.
- [13] Christina Frederick and Bjorn Engquist. Numerical methods for multiscale inverse problems. *arXiv preprint arXiv:1401.2431*, 2014.
- [14] Ian Goodfellow, Yoshua Bengio, Aaron Courville, and Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.
- [15] Marten Gulliksson, Anders Holmbom, Jens Persson, and Ye Zhang. A separating oscillation method of recovering the g-limit in standard and non-standard homogenization problems. *Inverse Problems*, 32(2):025005, 2016.
- [16] Ehsan Haghghat and Ruben Juanes. Sciann: A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks. *Computer Methods in Applied Mechanics and Engineering*, 373:113552, 2021.
- [17] V. Jikov, S. Kozlov, and O. Oleinik. *Homogenization of differential operators and integral functionals*. Springer Science & Business Media, 2012.
- [18] Isaac E Lagaris, Aristidis Likas, and Dimitrios I Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5):987–1000, 1998.
- [19] Lu Lu, Xuhui Meng, Zhiping Mao, and George E Karniadakis. Deepxde: A deep learning library for solving differential equations. *arXiv preprint arXiv:1907.04502*, 2019.
- [20] Lu Lu, Raphael Pestourie, Wenjie Yao, Zhicheng Wang, Francesc Verdugo, and Steven G Johnson. Physics-informed neural networks with hard constraints for inverse design. *arXiv preprint arXiv:2102.04626*, 2021.
- [21] James Nolen and George Papanicolaou. Fine scale uncertainty in parameter estimation for elliptic equations. *Inverse Problems*, 25(11):115021, 2009.
- [22] G Papanicolau, A Bensoussan, and J-L Lions. *Asymptotic analysis for periodic structures*. Elsevier, 1978.
- [23] Jens Persson. *Selected topics in homogenization*. PhD thesis, Mittuniversitetet, 2012.
- [24] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
- [25] Yeonjong Shin, Jerome Darbon, and George Em Karniadakis. On the convergence and gener-

- alization of physics informed neural networks. *arXiv e-prints*, pages arXiv-2004, 2020.
- [26] Sergio Spagnolo. Sul limite delle soluzioni di problemi di cauchy relativi all'equazione del calore. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 21(4):657–699, 1967.
- [27] Sergio Spagnolo. Convergence in energy for elliptic operators. In *Numerical Solution of Partial Differential Equations-III*, pages 469–499. Elsevier, 1976.
- [28] Sifan Wang, Xinling Yu, and Paris Perdikaris. When and why pinns fail to train: A neural tangent kernel perspective. *arXiv preprint arXiv:2007.14527*, 2020.
- [29] Huaiqian You, Yue Yu, Nathaniel Trask, Mamikon Gulian, and Marta D'Elia. Data-driven learning of nonlocal physics from high-fidelity synthetic data. *Computer Methods in Applied Mechanics and Engineering*, 374:113553, 2021.