

Int-Deep: A Deep Learning Initialized Iterative Method for Nonlinear Problems

Jianguo Huang* Hao-qin Wang† Haizhao Yang‡

Abstract

This paper focuses on proposing a deep learning initialized iterative method (Int-Deep) for low-dimensional nonlinear partial differential equations (PDEs). The corresponding framework consists of two phases. In the first phase, an expectation minimization problem formulated from a given nonlinear PDE is approximately resolved with mesh-free deep neural networks to parametrize the solution space. In the second phase, a solution ansatz of the finite element method to solve the given PDE is obtained from the approximate solution in the first phase, and the ansatz can serve as a good initial guess such that Newton’s method for solving the nonlinear PDE is able to converge to the ground truth solution with high-accuracy quickly. Systematic theoretical analysis is provided to justify the Int-Deep framework for several classes of problems. Numerical results show that the Int-Deep outperforms existing purely deep learning-based methods or traditional iterative methods (e.g., Newton’s method and the Picard iteration method).

1 Introduction

This paper is concerned with the efficient numerical solution of nonlinear partial differential equations (PDEs) including a class of eigenvalue problems as special cases, which is a ubiquitous and important topic in science and engineering [1, 2, 3, 4, 5, 6, 7, 8]. As far as we know, there have developed many traditional and typical numerical methods in this area, e.g., the finite difference method, the spectral method, and the finite element method [9]. The first two methods are generally used for solving problems over regular domains while the latter one is particularly suitable for solving problems over irregular domains [10, 11]. To achieve the numerical solution with the desired accuracy, one is often required to numerically solve the discrete problem formulated as a large-scale nonlinear system of nonlinear equations, which is time-consuming. In this case, one of the most critical issues is to choose the feasible initial guess so that the numerical solver (e.g. Newton’s method) is convergent. On the other hand, for reducing the computational cost, two grid methods are thereby devised in [12, 13], which only require to solve a small-sized nonlinear system arising from the finite element discretization based on a coarse triangulation. However, the difficulty is

* School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China.

E-mail: jghuang@sjtu.edu.cn

† School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, China.

E-mail: wanghaoqin@sjtu.edu.cn

‡ Current address: Department of Mathematics, Purdue University, West Lafayette, IN 47907 US

Part of the work was done in Department of Mathematics, National University of Singapore, Singapore, 119076

E-mail: matyh@nus.edu.sg

that we even can not ensure if the nonlinear system has a solution in theory if the mesh size of the coarse triangulation is large enough.

Recently, science and engineering have undergone a revolution driven by the success of deep learning techniques that originated in computer science. This revolution also includes broad applications of deep learning in computational and applied mathematics. Many new branches in scientific computing have emerged based on deep learning in the past five years including new methods for solving nonlinear PDEs. There are mainly two kinds of deep learning approaches for solving nonlinear PDEs: mesh-based [14, 15, 16, 17, 18, 19, 20] and mesh-free [21, 22, 23, 24, 25, 26, 27]. In the mesh-based methods, deep neural networks (DNNs) are constructed to approximate the solution operator of a PDE, e.g., seeking a DNN that approximates the map mapping the coefficients (or initial/boundary conditions) of a PDE to the corresponding solution. After construction, the DNN can be applied to solve a specific class of PDEs efficiently. In the mesh-free methods, DNNs are applied as the parametrization of the solution space of a PDE; then the solution of the PDE is identified via seeking a DNN that fits the constraints of the PDE in the least squares sense or minimizes a variational problem formulated from PDEs. The key to the success of these approaches is the universal approximation capacity of DNNs [28, 28, 29, 30, 31, 32] even without the curse of dimensionality for a large class of functions [29, 33, 34, 35].

Though deep learning has made it possible to solve high-dimensional problems, which is a significant breakthrough in scientific computing, to the best of our knowledge, the advantage of deep learning approaches over traditional methods in the low-dimensional region is still not clear yet. The main concern is the computational efficiency of these frameworks: the number of iterations in deep learning methods is usually large or the accuracy is very limited (e.g., typically 10^{-2} to 10^{-4} relative error). In order to overcome this difficulty, one is tempted to set up a more efficient neural network architecture (e.g., incorporating physical information in the structure designing [36, 37, 38, 18]), designing a more advanced learning algorithm for deep learning training [39, 40], or using a solution ansatz according to prior knowledge [41, 42, 43]. However, the overall performance of these frameworks for nonlinear problems without any prior knowledge may still not be very convincingly efficient.

This paper proposes the Int-Deep framework from a new point of view for designing highly efficient solvers of low-dimensional nonlinear PDEs with a finite element accuracy leveraging both the advantages of traditional algorithms and deep learning approaches. The Int-Deep framework consists of two phases. In the first phase, an approximate solution to the given nonlinear PDE is obtained via deep learning approaches using DNNs of size $O(1)$ and $O(100)$ iterations, where $O(\cdot)$ means that the prefactor is independent of the final target accuracy in the Int-Deep framework, i.e., the accuracy of finite element methods. In particular, based on variational principles, we propose new methods to formulate the problem of solving nonlinear PDEs into an unconstrained minimization problem of an expectation over a function space parametrized via DNNs, which can be solved efficiently via batch stochastic gradient descent (SGD) methods due to the special form of expectation. Unlike previous methods in which the form of expectation is only derived for nonlinear PDEs related to variational equations, our proposed method can also handle those related to variational inequalities and eigenvalue problems, providing a unified variational framework for a wider range of nonlinear problems.

In the second phase, the approximate solution provided by deep learning can serve as a good initial guess such that traditional iterative methods (e.g., Newton’s method for solving nonlinear

PDEs and the power method for eigenvalue problems) converge quickly to the ground truth solution with high-accuracy. The hybrid algorithm substantially reduces the learning cost of deep learning approach while keeping the quality of initial guesses for traditional iterative methods; good initial guesses enable traditional iterative methods to converge in $O(\log(\frac{1}{\epsilon}))$ iterations to the ϵ precision of finite element methods. In each iteration of the traditional iterative method, the nonlinear problem has been linearized and hence traditional fast solvers for linear systems can be applied depending on the underlying structure of the linear system. Therefore, as we shall see in the numerical section, the Int-Deep framework outperforms existing purely deep learning-based methods or traditional iterative methods, e.g., Newton’s method and Picard iteration. Furthermore, systematic theoretical analysis is provided to characterize the conditions under which the Int-Deep framework converges, serving as a trial to change current deep learning research from trial-and-error to a suite of methods informed by a principled design methodology.

This paper will be organized as follows. In Section 2, we briefly review the definitions of DNNs. In Section 3, we introduce the expectation minimization framework for deep learning-based PDE solvers in the first phase of the Int-Deep framework. In Section 4, as the second phase of Int-Deep, traditional iterative methods armed with good initial guesses provided by deep learning approaches will be introduced together with its theoretical convergence analysis, for clarity of presentation the proofs of which are left in Appendix. In Section 5, a set of numerical examples will be provided to demonstrate the efficiency of the proposed framework and to justify our theoretical analysis. Finally in Section 6, we summarize our paper with a short discussion.

2 Deep Neural Networks (DNNs)

Mathematically, DNNs are a form of highly non-linear function parametrization via function compositions using simple non-linear functions [44]. The validity of such an approximation method can be ensured by the universal approximation theorems of DNNs in [28, 28, 29, 30, 31, 32]. In this paper, we will focus on the following two neural networks for computation.

The first one is the so-called fully connected feed-forward neural network (FNN), which is a function in the form of a composition of L simple nonlinear functions as follows:

$$\phi(\mathbf{x}; \boldsymbol{\theta}) := \mathbf{a}^T \mathbf{h}_L \circ \mathbf{h}_{L-1} \circ \cdots \circ \mathbf{h}_1(\mathbf{x}),$$

where $\mathbf{h}_\ell(\mathbf{x}) = \sigma(\mathbf{W}_\ell \mathbf{x} + \mathbf{b}_\ell)$ with $\mathbf{W}_\ell \in \mathbb{R}^{N_\ell \times N_{\ell-1}}$, $\mathbf{b}_\ell \in \mathbb{R}^{N_\ell}$ for $\ell = 1, \dots, L$, $\mathbf{a} \in \mathbb{R}^{N_L}$, σ is a non-linear activation function, e.g., a rectified linear unit (ReLU) $\sigma(x) = \max\{x, 0\}$ or hyperbolic tangent function $\tanh(x)$. Each \mathbf{h}_ℓ is referred as a hidden layer, N_ℓ is the width of the ℓ -th layer, and L is called the depth of the FNN. Throughout this paper, we consider $N_\ell = N$ in the numerical implementation of FNNs for the purpose of simplicity. In the above formulation, $\boldsymbol{\theta} := \{\mathbf{a}, \mathbf{W}_\ell, \mathbf{b}_\ell : 1 \leq \ell \leq L\}$ denotes the set of all parameters in ϕ , which uniquely determines the underlying neural network.

The second one is the so-called residual neural network (ResNet) defined recursively as follows:

$$\mathbf{h}_0 = \mathbf{V}\mathbf{x}, \quad \mathbf{g}_\ell = \sigma(\mathbf{W}_\ell \mathbf{h}_{\ell-1} + \mathbf{b}_\ell), \quad \mathbf{h}_\ell = \bar{\mathbf{U}}_\ell \mathbf{h}_{\ell-1} + \mathbf{U}_\ell \mathbf{g}_\ell, \quad \ell = 1, \dots, L, \quad \phi(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{a}^T \mathbf{h}_L,$$

where $\mathbf{V} \in \mathbb{R}^{N_0 \times d}$, $\mathbf{W}_\ell \in \mathbb{R}^{N_\ell \times N_0}$, $\mathbf{U}_\ell \in \mathbb{R}^{N_0 \times N_\ell}$, $\mathbf{b}_\ell \in \mathbb{R}^{N_\ell}$ for $\ell = 1, \dots, L$, $\mathbf{a} \in \mathbb{R}^{N_0}$. Throughout this paper, we consider $N_\ell = N$ and \mathbf{U}_ℓ is set as the identity matrix in the numerical implementation of ResNets for the purpose of simplicity. Furthermore, we set $\bar{\mathbf{U}}_\ell$ as the identity matrix when ℓ is

even and set $\bar{U}_\ell = 0$ when ℓ is odd, i.e., we use skip connections every two layers for the purpose of fair comparison with the Deep Ritz method in [45] later in our numerical section.

3 Phase I of Int-Deep: Variational Formulas for Deep Learning

In this section, we will present existing and our new analysis for reformulating nonlinear PDEs including eigenvalue problems to the minimization of expectation that can be solved by SGD. The analysis works for PDEs related to variational equations and variational inequalities, the latter of which is of special interest since there might be no available literature discussing this case to the best of our knowledge. Hence, our analysis could serve as a good reference for a wide range of problems.

Throughout this paper, we will use standard symbols and notations for Sobolev spaces and their norm/semi-norms; we refer the reader to the reference [46] for details. Moreover, the standard $L^2(D)$ -inner product for a bounded domain D is denoted by $(\cdot, \cdot)_D$. If the domain D is the solution domain Ω , we will drop out the dependence of Ω in all Sobolev norms/seminorms when there is no confusion caused.

3.1 PDE Solvers Based on DNNs

With the definition of DNNs introduced in the previous section, we are now ready to introduce deep learning-based PDE solvers. The general idea of these solvers is to treat DNNs as an efficient parametrization of the solution space of a PDE and the solution of the PDE is identified via seeking a DNN that fits the constraints of the PDE in the least squares sense or minimizes the variational minimization problem related to the PDE.

For simplicity, let us use a PDE defined on a domain Ω in a compact form with equality constraints to illustrate the main idea, e.g.,

$$\begin{aligned}\mathcal{D}(u) &= f \quad \text{in } \Omega, \\ \mathcal{B}(u) &= g \quad \text{on } \partial\Omega,\end{aligned}\tag{3.1}$$

where \mathcal{D} is a differential operator and \mathcal{B} is the operator for specifying an appropriate boundary condition.

In the least squares type methods (LSM), a DNN $\phi(\mathbf{x}; \boldsymbol{\theta}^*)$ is constructed to approximate the solution $u(\mathbf{x})$ for $\mathbf{x} \in \Omega$ via minimizing the square loss

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) := \mathbb{E}_{\mathbf{x} \in \Omega} \left[|\mathcal{D}\phi(\mathbf{x}; \boldsymbol{\theta}) - f(\mathbf{x})|^2 \right] + \lambda \mathbb{E}_{\mathbf{x} \in \partial\Omega} \left[|\mathcal{B}\phi(\mathbf{x}; \boldsymbol{\theta}) - g(\mathbf{x})|^2 \right], \tag{3.2}$$

with a positive parameter λ . These methods probably date back to 1990's (e.g., see [47, 41]) and were revisited recently [21, 24, 25, 26, 27] due to the significant development of GPU computing that accelerates DNN computation. But so far there are no sufficiently efficient solvers in the sense that solving the optimization problems in (3.2) is expensive or the accuracy of DNN solutions is limited.

In the variational type methods (VM), the PDE in (3.1) is reformulated as a minimization problem in a variational form

$$u^* = \arg \min_{u \in H} J(u), \tag{3.3}$$

where the Hilbert space H is an admissible space, and $J(u)$ is a nonlinear functional over H . Then, the solution space H is parametrized via DNNs, i.e., $H \approx \{\phi(\mathbf{x}; \boldsymbol{\theta})\}_{\boldsymbol{\theta}}$, where ϕ is a DNN with a fixed depth L and width N . After parametrization, the variational minimization problem in (3.3) is approximated by the following problem:

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} J(\phi(\mathbf{x}; \boldsymbol{\theta})). \quad (3.4)$$

In general, $J(\phi(\mathbf{x}; \boldsymbol{\theta}))$ can be formulated as the sum of several integrals over several sets $\{\Omega_i\}_{i=1}^p$, each of which corresponds to one equation in the PDE:

$$J(\phi(\mathbf{x}; \boldsymbol{\theta})) = \sum_{i=1}^p \int_{\Omega_i} F_i(\mathbf{x}; \boldsymbol{\theta}) \, dx = \sum_{i=1}^p |\Omega_i| \mathbb{E}_{\boldsymbol{\xi}_i}[F_i(\boldsymbol{\xi}_i; \boldsymbol{\theta})], \quad (3.5)$$

where $F_i(\cdot; \boldsymbol{\theta})$ is a function coming from the variational form of a constrain on $\phi(\mathbf{x}; \boldsymbol{\theta})$, $\boldsymbol{\xi}_i$ is a random vector produced by the uniform distribution over Ω_i , $|\Omega_i|$ denotes the measure of Ω_i , and $\mathbb{E}_{\boldsymbol{\xi}_i}[\cdot]$ stands for the mathematical expectation with respect to the random vector $\boldsymbol{\xi}_i$. In addition, all the random vectors $\boldsymbol{\xi}_i$ ($1 \leq i \leq p$) are mutually independent. Based on the formulation (3.5), the problem (3.4) can be expressed as

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \sum_{i=1}^p |\Omega_i| \mathbb{E}_{\boldsymbol{\xi}_i}[F_i(\boldsymbol{\xi}_i; \boldsymbol{\theta})]. \quad (3.6)$$

Both the LSM in (3.2) and VM in (3.4) can be reformulated to the expectation minimization problem in (3.6), which can be solved by the stochastic gradient descent (SGD) method or its variants (e.g., Adam [48]). In this paper, we refer to (3.6) as the expectation minimization framework for PDE solvers based on deep learning.

Although the convergence of SGD for minimizing the expectation in (3.6) is still an active research topic, empirical success shows that SGD can provide a good approximate local minimizer of (3.2) and (3.6). This completes the algorithm of using deep learning to solve nonlinear PDEs with equality constraints.

We would like to emphasize that when the PDE in (3.1) is nonlinear, its solutions might be the saddle points of (3.5) making it very challenging to identify its solutions via minimizing (3.5). Hence, for PDEs associated with variational equations, we will use the LSM in (3.2) for numerical computation. It deserves to point out that (3.2) is also regarded as a variational formulation, referred to as the least-squares variational principle in [49] in contrast with the usual Ritz variational principle.

3.2 Minimization Problems for Variational Inequalities

Variational inequalities are a class of important nonlinear problems, frequently encountered in various industrial and engineering applications. We refer the reader to [50, 51] for details about the mathematical theory and numerical methods for such problems. Unlike previous methods in which the form of expectation is only derived for nonlinear PDEs with variational equations, our expectation minimization framework is also suitable for variational inequalities

The abstract framework of an elliptic variational inequality of the second kind can be described as follows (cf. [51]). Find $u \in H$ such that

$$a(u, v - u) + j(v) - j(u) \geq \langle f, v - u \rangle, \quad v \in H, \quad (3.7)$$

where H is a Hilbert space equipped with the norm $\|\cdot\|_H$, and $\langle \cdot, \cdot \rangle$ stands for the duality pairing between H' and H , with H' being the dual space of H ; $a(\cdot, \cdot)$ is a continuous, coercive and symmetric bilinear form over H ; $j(\cdot) : H \rightarrow \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is a proper, convex and lower semi-continuous functional.

As shown in [51], the above problem has a unique solution under the stated conditions on the problem data. Moreover, it can be reformulated as the following minimization problem:

$$u = \arg \min_{v \in H} J(v) = \frac{1}{2}a(v, v) - \langle f, v \rangle + j(v), \quad (3.8)$$

which naturally falls into our expectation minimization framework in (3.6).

To illustrate the above framework more clearly, let us discuss in detail the simplified friction problem, which is a typical elliptic variational inequality of the second kind. In this case, the nonlinear PDE is given by

$$\begin{aligned} -\Delta u + u &= f \quad \text{in } \Omega, \\ |\partial_{\mathbf{n}} u| &\leq g, \quad u \partial_{\mathbf{n}} u + g|u| = 0 \quad \text{on } \Gamma_C, \\ u &= 0 \quad \text{on } \Gamma_D, \end{aligned} \quad (3.9)$$

where $\Omega \subset \mathbb{R}^2$ is a bounded domain with the Lipschitz boundary $\partial\Omega$, and \mathbf{n} is the unit outward normal to $\partial\Omega$; $\Gamma_C \subset \partial\Omega$ denotes the friction boundary on which some friction condition is imposed, and $\Gamma_D = \partial\Omega \setminus \Gamma_C$; $f \in L^2(\Omega)$ and $g \in L^2(\Gamma_C)$ are two given functions.

After some direct manipulation, the problem (3.9) can be expressed as an elliptic variational inequality in the form (3.7) or (3.8) by choosing

$$H = V_D = \{\phi \in H^1(\Omega) : \phi = 0 \text{ on } \Gamma_D\}$$

and

$$a(\phi, \chi) = (\nabla \phi, \nabla \chi)_\Omega + (\phi, \chi)_\Omega, \quad j(\phi) = (g, |\phi|)_{\Gamma_C}, \quad \phi, \chi \in V_D.$$

Hence, the minimization problem of the nonlinear PDE (3.9) is given by

$$u = \arg \min_{\phi \in V_D} J_1(\phi), \quad (3.10)$$

where

$$\begin{aligned} J_1(\phi) &= \frac{1}{2}\|\phi\|_1^2 - (f, \phi)_\Omega + j(\phi) \\ &= |\Omega| \mathbb{E}_{\boldsymbol{\xi}_1} \left[\frac{1}{2} (|\nabla_{\boldsymbol{\xi}_1} \phi(\boldsymbol{\xi}_1; \boldsymbol{\theta})|^2 + \phi^2(\boldsymbol{\xi}_1; \boldsymbol{\theta})) - f(\boldsymbol{\xi}_1) \phi(\boldsymbol{\xi}_1; \boldsymbol{\theta}) \right] + |\Gamma_C| \mathbb{E}_{\boldsymbol{\xi}_2} [g(\boldsymbol{\xi}_2) |\phi(\boldsymbol{\xi}_2; \boldsymbol{\theta})|]. \end{aligned}$$

We remark that if the friction condition is dropped out, the problem (3.9) reduces to the standard elliptic equation with homogeneous boundary conditions. In this case, the minimization problem (3.10) is nothing but the Ritz variational principle related to the admissible space $V = H_0^1(\Omega)$.

We can also use the penalty method to remove the constrained condition in the admissible space V_D , giving rise to an easier unconstrained minimization problem for deep learning. Then, the problem (3.10) is modified as

$$u = \arg \min_{\phi \in V_1} J_2(\phi), \quad (3.11)$$

where

$$V_1 = H^1(\Omega) \quad J_2(\phi) = J_1(\phi) + \gamma \|\phi\|_{0,\Gamma_D}^2 = |\Omega| \mathbb{E}_{\xi_1} \left[\frac{1}{2} (|\nabla_{\xi_1} \phi(\xi_1; \theta)|^2 + \phi^2(\xi_1; \theta)) - f(\xi_1) \phi(\xi_1; \theta) \right] \\ + |\Gamma_C| \mathbb{E}_{\xi_2} [g(\xi_2) |\phi(\xi_2; \theta)|] + |\Gamma_D| \mathbb{E}_{\xi_3} [\gamma \phi^2(\xi_3; \theta)],$$

with γ denoting a penalty parameter to be determined feasibly. The minimization problem in (3.11) is of the form of expectation minimization in (3.6).

The other way to avoid the constraint of an admissible function will be shown in the next subsection.

3.3 Minimization Problems for Eigenvalue Problems

At last, let us discuss how to evaluate the smallest eigenvalue and its eigenfunction for a positive self-adjoint differential operator using the expectation minimization framework. The problem under discussion reads

$$\begin{aligned} -\nabla(p(\mathbf{x})\nabla u) + q(\mathbf{x})u &= \lambda u \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{3.12}$$

where $p(\mathbf{x}) \in C^1(\bar{\Omega})$ and there exist two positive constants $p_1 \geq p_0$ such that $p_0 \leq p(\mathbf{x}) \leq p_1$ for all $\mathbf{x} \in \Omega$, and $q(\mathbf{x}) \in C(\bar{\Omega})$ is nonnegative over Ω .

We first show the solution (λ, u) is governed by the following variational problem

$$u = \arg \min_{\phi \in V} J_3(\phi), \tag{3.13}$$

where $V = H_0^1(\Omega)$ as before, and

$$J_3(\phi) = \frac{a(\phi, \phi)}{\|\phi\|_0^2} + \gamma(\phi(\mathbf{x}_0) - 1)^2; \quad a(\phi, \chi) = \int_{\Omega} [p(\mathbf{x})\nabla\phi \cdot \nabla\chi + q(\mathbf{x})\phi\chi] dx, \quad \phi, \chi \in V;$$

γ is any given positive number, and \mathbf{x}_0 is any given point in the interior of Ω .

As a matter of fact, if $u_* \in V$ is a solution, we have by the variational principle for eigenvalue problems [52] that

$$u_* = \arg \min_{\phi \in V} \frac{a(\phi, \phi)}{\|\phi\|_0^2}. \tag{3.14}$$

Then we choose a constant α such that $(\alpha u_*)(\mathbf{x}_0) = 1$. Hence, if write $u = \alpha u_*$, we find from (3.14) that, for all $\phi \in V$,

$$J_3(u) \leq \frac{a(\phi, \phi)}{\|\phi\|_0^2} \leq J_3(\phi),$$

as required. The converse can be proved similarly.

Furthermore, it is easy to check that

$$J_3(\phi) = \frac{a(\phi(\mathbf{x}, \boldsymbol{\theta}), \phi(\mathbf{x}, \boldsymbol{\theta}))}{\|\phi(\mathbf{x}, \boldsymbol{\theta})\|_0^2} + \gamma(\phi(\mathbf{x}_0, \boldsymbol{\theta}) - 1)^2 = \mathbb{E}_{\xi, \boldsymbol{\eta}} \left[\frac{a(\phi(\xi, \boldsymbol{\theta}), \phi(\xi, \boldsymbol{\theta}))}{\phi^2(\boldsymbol{\eta}, \boldsymbol{\theta})} + \gamma(\phi(\mathbf{x}_0, \boldsymbol{\theta}) - 1)^2 \right],$$

where ξ and $\boldsymbol{\eta}$ are i.i.d. random vectors, and

$$a(\phi(\xi, \boldsymbol{\theta}), \phi(\xi, \boldsymbol{\theta})) = a(\phi(\mathbf{x}, \boldsymbol{\theta}), \phi(\mathbf{x}, \boldsymbol{\theta}))|_{\mathbf{x}=\xi}.$$

In fact, (3.13) is a constrained minimization problem, that is

$$\begin{aligned} u &= \arg \min_{\phi \in V_1} J_3(\phi), \\ \text{s.t.} \quad u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where $V_1 = H^1(\Omega)$ as before. One can invoke the penalty method to eliminate the constraint condition as used in the last subsection. Here, we exploit another idea from Jens Berg and Kaj Nyström [21] to reformulate the above problem as an unconstrained minimization one. To this end, we construct the neural network function as follows:

$$\phi(\mathbf{x}; \boldsymbol{\theta}) = B(\mathbf{x})\psi(\mathbf{x}; \boldsymbol{\theta}).$$

where $B(\mathbf{x})$ is a known smooth function such that the boundary $\partial\Omega$ can be parametrized by $B(\mathbf{x}) = 0$, and $\psi(\mathbf{x}, \boldsymbol{\theta})$ is any function in V_1 . So (3.13) is equivalent to

$$u = \arg \min_{\psi \in V_1} J_3(B(\mathbf{x})\psi(\mathbf{x}; \boldsymbol{\theta})), \quad (3.15)$$

which enjoys the form of expectation minimization in (3.6).

Once we have obtained the eigenfunction $u(x)$, then the corresponding eigenvalue is $\lambda = a(u, u)/\|u\|_0^2$. Note that the variational formulation developed here is different from the one devised by E and Yu [45] and the expectation form in our formulation makes it easier to implement SGD.

4 Phase II of Int-Deep: Traditional Iterative Methods

We propose to use deep learning solutions as initial guesses so as to achieve a high-accurate solution by traditional iterative methods in a few iterations (e.g., Newton's method for solving nonlinear systems [53] or the two grid methods in the context of finite elements [12, 13, 54]). In fact, these ideas have led to high-performance methods for solving semilinear elliptic problems and eigenvalue problems, with the effectiveness shown by mathematical theory and numerical simulation. The key analysis of this idea is to characterize the conditions under which deep learning solutions can help traditional iterative methods converge quickly. We will provide several classes of examples and the corresponding analysis to support this idea as follows.

4.1 Semilinear Elliptic Equations with Equality Constrains

Consider the following semilinear elliptic equation

$$\begin{cases} -\Delta u + f(\mathbf{x}, u) = 0 & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \quad (4.1)$$

where Ω is a bounded convex polygon \mathbb{R}^d ($d = 1, 2$), and $f(\mathbf{x}, u)$ is a sufficiently smooth function. For simplicity, we use $f(u)$ for $f(\mathbf{x}, u)$ and f' for f_u in what follows.

Let $V = H_0^1(\Omega)$. Then the variational formulation of problem (4.1) is to find $u \in V$ such that

$$a(u, \chi) := (\nabla u, \nabla \chi) + (f(u), \chi) = 0, \quad \chi \in V. \quad (4.2)$$

For any $v \in L^\infty(\Omega)$, define

$$a_v(\phi, \chi) := (\nabla \phi, \nabla \chi) + (f'(v)\phi, \chi), \quad \phi, \chi \in V. \quad (4.3)$$

As in [12, 13], we assume that problem (4.1) (equivalently, problem (4.2)) satisfies the following conditions:

A1 Any solution u of (4.1) has the regularity $u \in W^{2,\infty}(\Omega)$.

A2 Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.1). Then there exists a positive constant C^u such that

$$C^u \|\phi\|_1 \leq \sup_{\chi \in V} \frac{a_u(\phi, \chi)}{\|\chi\|_1}, \quad \phi \in V.$$

We next consider the finite element method for solving problem (4.2). Let \mathcal{T}_h be a quasi-uniform and shape regular triangulation of Ω into K . Write $h_K = \text{diam}(K)$, and $h = \max_{K \in \mathcal{T}_h} h_K$. Introduce the Courant element space by

$$V_h = \{\phi \in C(\bar{\Omega}) : \phi|_K \in \mathbb{P}_1(K)\} \cap V.$$

where $\mathbb{P}_1(K)$ denotes the function space consisting of all linear polynomials over K . Then the finite element method is to find $u^h \in V_h$ such that

$$a(u^h, \chi) = 0, \quad \chi \in V_h. \quad (4.4)$$

Now, let us introduce Int-Deep for solving the problem in (4.4). Concretely speaking, we choose $I_h u^{DL}$ as the initial guess, where u^{DL} denotes the numerical solution obtained by the deep learning algorithm, and I_h is the usual nodal interpolation operator (cf. [55, 56]).

Algorithm 1 A hybrid Newton's method for semilinear problems

Input: the target accuracy ϵ , the maximum number of iterations N_{\max} , the approximate solution in a form of a DNN u^{DL} in Phase I of Int-Deep.

Output: $u^h = u_{k+1}^h$.

Initialization: Let $u_0^h = I_h u^{DL}$, $k = 0$, and $e_k = 1$;

while $e_k < \epsilon$ and $k < N_{\max}$ **do**

Find $v_k^h \in V_h$ such that

$$(\nabla v_k^h, \nabla \chi) + (f'(u_k^h) v_k^h, \chi) = -(\nabla u_k^h, \nabla \chi) - (f(u_k^h), \chi), \quad \chi \in V_h.$$

Let $u_{k+1}^h = u_k^h + v_k^h$.

$$e_{k+1} = \|u_{k+1}^h - u_k^h\|_0 / \|u_k^h\|_0, \quad k = k + 1.$$

end while

Next, we turn to discuss the convergence of the Algorithm 1. In the theorem below, we introduce a discrete maximum norm $\|\cdot\|_{0,\infty,h}$ to quantify errors. Let Ω_h consist of all the vertices of the triangulation \mathcal{T}_h of Ω . Then for any $v \in C(\bar{\Omega})$, $\|v\|_{0,\infty,h} = \max_{\mathbf{x} \in \Omega_h} |v(\mathbf{x})|$.

Theorem 4.1. *Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.1) and u_k^h be the function sequence formed by Algorithm 1. Assume $u_k^h \in B(u)$ for $k = 0, 1, \dots$. Write $\delta = \|u - u^{DL}\|_{0,\infty,h}$. Then there exist two positive constant $\bar{h}_0 < h_3$ and δ_0 such that if $h < \bar{h}_0$ and $\delta < \delta_0$, there holds*

$$\begin{aligned} \|u - u_k^h\|_{0,\infty} &\lesssim h^2 + \beta_1^{2^k} && \text{for } d = 1, \\ \|u - u_k^h\|_{0,\infty} &\lesssim h^2 |\log(h)| + h^{-2/p} \beta_2^{2^k} && \text{for } d = 2, \end{aligned}$$

where h_3 is a positive constant given in Lemma A.4,

$$\beta_1 = c_0 c_1 (h^2 + \delta) < 1, \quad \beta_2 = c_2 c_3 (h^2 |\log(h)| + \delta) < 1,$$

with c_i ($0 \leq i \leq 3$) as positive constants.

The proof of Theorem 4.1 can be found in Appendix A. According to the exponential convergence in β_1 and β_2 in the above theorem, one can use only a few Newton's iterations to achieve a numerical solution with the same accuracy as in the finite element solution u^h . The forthcoming numerical results will demonstrate this theoretical estimate. Moreover, the analysis developed here can be applied to high order finite element methods.

4.2 Eigenvalue Problems

In this subsection, we propose and analyze Int-Deep for solving eigenvalue problems in the similar spirit of the two grid method due to Xu and Zhou [54]. For simplicity, let us consider the following problem

$$\begin{cases} -\Delta u = \lambda u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \quad (4.5)$$

The variational problem for (4.5) is given as follows. Find the smallest number λ and a nonzero function $u \in V$ such that

$$a(u, \chi) = \lambda(u, \chi), \quad \chi \in V, \quad (4.6)$$

where $V = H_0^1(\Omega)$ and $a(u, \chi) = (\nabla u, \nabla \chi)$. Without loss of generality, assume $\|u\|_0 = 1$.

In the discretization, we use the same finite element space V_h as given in the last subsection. Hence, the finite element method for (4.6) is to find the smallest number λ^h and a nonzero function $u^h \in V_h$ such that

$$a(u^h, \chi) = \lambda^h(u^h, \chi), \quad \chi \in V_h. \quad (4.7)$$

Now, let us introduce Int-Deep for solving the problem in (4.7). Concretely speaking, we choose $I_h u^{DL}$ as the initial guess, where u^{DL} denotes the numerical solution obtained by the deep learning algorithm, and I_h is the usual nodal interpolation operator. The iterative scheme is motivated by the two grid method for eigenvalue problems due to Xu and Zhou [54]. The method is essentially the Power method for eigenvalue problems and the eigenvalue computation is accelerated by the Rayleigh quotient.

Algorithm 2 Int-Deep for eigenvalue problems

Input: the target accuracy ϵ , the maximum number of iterations N_{\max} , the approximate solution in a form of a DNN u^{DL} in Phase I of Int-Deep.

Output: $u^h = u_{k+1}^h$.

Initialization: Let $u_0^h = I_h u^{DL} / \|I_h u^{DL}\|_0$, $\lambda_0^h = |u_0^h|_1^2 / \|u_0^h\|_0^2$, $k = 0$, $e_k = 1$;

while $e_k < \epsilon$ and $k < N_{\max}$ **do**

Find $u_{k+1}^h \in V_h$ such that

$$\begin{aligned} a(u_{k+1}^h, \chi) &= \lambda_k^h(u_k^h, \chi), \quad \chi \in V_h, \\ \lambda_{k+1}^h &= \frac{|u_{k+1}^h|_1^2}{\|u_{k+1}^h\|_0^2}; \end{aligned}$$

$e_{k+1} = \|u_{k+1}^h - u_k^h\|_0 / \|u_k^h\|_0$, $k = k + 1$.

end while

To analyze the convergence of Algorithm 2, we need to introduce the elliptic projection operator as follows. For $u \in V$, find $P_h u \in V_h$ such that

$$a(P_h u, \chi) = a(u, \chi), \quad \chi \in V_h.$$

The estimates of this operator are well-known (cf. [55, 56]). We also need the following important result (cf. [57, 58])

Lemma 4.1. *Let (λ, u) be an eigenpair of (4.6). For any $\phi \in H_0^1(\Omega) \setminus \{0\}$, there holds*

$$\frac{a(\phi, \phi)}{(\phi, \phi)} - \lambda = \frac{a(\phi - u, \phi - u)}{(\phi, \phi)} - \lambda \frac{(\phi - u, \phi - u)}{(\phi, \phi)}.$$

With the help of the above results, we can derive the following theorem.

Theorem 4.2. *Let λ be the smallest eigenvalue of (4.5) and $u \in H^2(\Omega)$ be the corresponding eigenvector with $\|u\|_0 = 1$, respectively. Denote $\bar{\delta} = \max\{|\lambda - \lambda_0^h|, \|u - u^{DL}\|_{0,\infty,h}\}$. Assume that the sequence $\|u_k^h\|_0$ is bounded below and above by two positive constants ε_0 and ε_1 , respectively. Then if $\beta_3 = c_4 c_6 (\bar{\delta} + h^2) < 1$ where c_4 and c_6 are two positive constants, there holds*

$$|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \lesssim \beta_3^{2^k} + 2^{k+1} h^2. \quad (4.8)$$

The proof of Theorem 4.2 can be found in Appendix B. Theorem 4.2 shows that if the deep learning solution is sufficiently accurate (almost independent of the mesh size h) and the mesh size h is sufficiently small, the proposed Int-Deep framework can converge quickly to a precision balancing two error terms in (4.8). To make a good balance, we are tempted to use only a few iterations during numerical simulation, which is sufficient as we will show in the numerical section. Deriving more effective error estimates for Algorithm 2 would be interesting future work. Moreover, the above error analysis can also be applied to high order finite element methods.

5 Numerical Experiments

This section consists of two parts. In the first part, we provide various numerical examples to illustrate the performance of the proposed deep learning-based methods. As we shall see, deep learning approaches are capable of providing approximate solutions to nonlinear problems in $O(100)$ iterations. However, continuing the iteration cannot further improve accuracy. In the second part, we will investigate the numerical performance of the Int-Deep framework in terms of network hyperparameters and the convergence analysis in the previous section. As we shall see, DNNs of size $O(1)$ and trained with $O(100)$ iterations can provide good initial guesses enabling traditional iterative methods to converge in $O(\log(\frac{1}{\epsilon}))$ iterations to the ϵ precision of finite element methods.

Without especial explanation, we always use the ResNet of width 50 and depth 6 with an activation function $\sigma(x) = \max\{x^3, 0\}$. Neural networks are trained by Adam optimizer [48] with a learning rate $\eta = 1e - 03$. The batch size is 512 for all 1D examples and 1024 for all 2D examples. Deep learning algorithms are implemented by Python 3.7 using PyTorch 1.0. and a single NVIDIA Quadro P6000 GPU card. All finite element methods are implemented in MATLAB 2018b in an Intel Core i7, 2.6GHz CPU on a personal laptop with a 32GB RAM.

Let us recall and define notations in this section. Suppose a problem is defined on a domain Ω and Ω_h is the point set consisting of all vertices of the uniform triangulation \mathcal{T}_h with a mesh

size h of Ω . For any $v \in C(\bar{\Omega})$, where $\bar{\Omega}$ is the closure of Ω , we define the discrete maximum norm $\|\cdot\|_{0,\infty,h}$ via

$$\|v\|_{0,\infty,h} = \max_{\mathbf{x} \in \Omega_h} |v(\mathbf{x})|.$$

Suppose u is the exact solution to a problem and u_k^h is the approximation evaluated by the Int-Deep framework in the k -th iteration in the second phase, we apply the absolute error

$$e_k^h := \|u - u_k^h\|_{0,\infty,h}$$

to measure the accuracy of our framework in the k -th iteration in the second phase. e^h is the error after the maximum number of iterations or convergence in Newton's method. Similarly, u_k^{DL} denotes the approximate solution by the deep learning method and e_k^{DL} denotes its error. After completing the iterations, u^{DL} and u^h are used as the approximate solution by deep learning and traditional iterative methods. In eigenvalue problems, let λ denote the target eigenvalue, λ^{DL} (and λ_k^{DL} in the k -th iteration) denote the approximate one by deep learning, λ^h (and λ_k^h in the k -th iteration) denote the approximate one by traditional methods. #Epoch means the number of epochs in the Adam for deep learning and #K stands for the number of iterations in the traditional iterative methods in Int-Deep. We always apply the Courant element method to solve the weak form in Algorithm 1 and Algorithm 2.

5.1 Phase I of Int-Deep: Deep Learning Methods

Let us first provide numerical examples to illustrate the performance of the proposed variational formulations for deep learning in Section 3.

5.1.1 Linear PDEs

Example 5.1. Consider the second-order linear elliptic equation with the Dirichlet boundary condition in one dimension:

$$\begin{cases} -\frac{d}{dx} \left(p(x) \frac{du}{dx} \right) + x^2 u(x) = f(x), & x \in (-1, 1), \\ u = 0, & x = -1 \text{ or } 1, \end{cases}$$

with $p(x) = 1 + x^2$ and $f(x) = \pi^2(1 + x^2) \sin(\pi x) - 2\pi x \cos(\pi x) + x^2 \sin(\pi x)$. The exact solution of this problem is

$$u(x) = \sin(\pi x).$$

In this experiment, the variational formulation in (3.6) is applied in the deep learning method. To evaluate the test error, we adopt a mesh size $h = \frac{2}{1024}$ to generate the uniform triangulation \mathcal{T}_h as the test locations. The test errors during training are summarized in Table 1 and Figure 1.

Table 1: Example 5.1

#Epoch	e_k^{DL}
300	1.0e-3
500	2.4e-3
1000	1.1e-3
5000	2.1e-3
10000	5.7e-3

Table 2: Example 5.2.

#Epoch	e_k^{DL}
300	6.5e-2
500	5.7e-2
1000	6.8e-3
5000	4.3e-3
10000	2.2e-3

Table 3: Example 5.3.

#Epoch	$ \lambda - \lambda^{DL} $	$\ u - u^{DL}\ _{0,\infty,h}$
300	1.3e-1	4.7e-2
500	4.4e-2	2.3e-2
1000	3.5e-2	2.5e-2
5000	2.0e-3	1.0e-2
10000	4.1e-3	1.3e-2

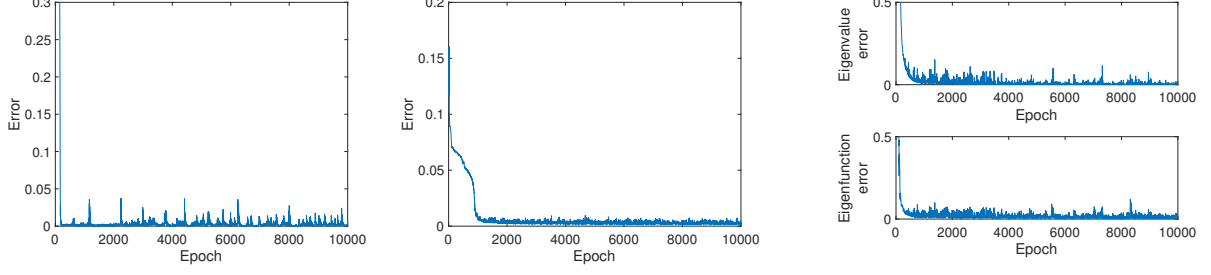


Figure 1: e_k^{DL} of Example 5.1. **Figure 2:** e_k^{DL} of Example 5.2. **Figure 3:** Test errors of Example 5.3.

5.1.2 Variational Inequalities

Example 5.2. Consider the simplified friction problem

$$\begin{cases} -\Delta u + u = f & \text{in } \Omega, \\ |\partial_{\mathbf{n}} u| \leq g, \quad u \partial_{\mathbf{n}} u + g|u| = 0 & \text{on } \Gamma_C, \\ u = 0 & \text{on } \Gamma_D, \end{cases} \quad (5.1)$$

where \mathbf{n} is the outer normal vector, $f \in L^2(\Omega)$ and $g \in L^2(\Gamma_C)$ are given functions. Γ_C is a subset of $\partial\Omega$ called friction boundary and $\Gamma_D = \partial\Omega \setminus \Gamma_C$. Let $\Omega = (0, 1)^2$ and $\Gamma_C = \{1\} \times [0, 1]$. We set $g = 1$ and choose f such that the problem has an exact solution $u(x, y) = (\sin x - x \sin 1) \sin(2\pi y)$.

In this experiment, the variational formula in (3.11) with a penalty constant $\gamma = 500$ is applied in the deep learning method. To evaluate the test error, we adopt a mesh size $h = \frac{1}{128}$ to generate the uniform triangulation \mathcal{T}_h as the test locations. The test errors during training are summarized in Table 2 and Figure 2.

5.1.3 Eigenvalue Problem

Example 5.3. Consider the following problem

$$\begin{cases} -u'' = \lambda u, & x \in (0, 1), \\ u = 0, & x = 0 \text{ or } 1. \end{cases}$$

The smallest eigenvalue is π^2 and the corresponding eigenvector is $u(x) = \sin(\pi(x - 1))$.

In this experiment, the variational formula in (3.15) with $x_0 = 0.5$ and $\gamma = 100$ is applied in the deep learning method. To evaluate the test error, we adopt a mesh size $h = \frac{1}{256}$ to generate the uniform triangulation \mathcal{T}_h as the test locations. The test errors during training are summarized in Table 3 and Figure 3.

5.2 Phase II of Int-Deep: Traditional Iterative Methods

Now we use the approximate solution by deep learning as the initial guess in traditional iterative methods. We'll show that DNNs of size $O(1)$ and trained with $O(100)$ iterations are good enough for traditional iterative methods to converge in at most $O(\log(\frac{1}{\epsilon}))$ iterations to the ϵ precision of finite element methods.

5.2.1 Semilinear PDEs

We consider semilinear elliptic equations with homogeneous Dirichlet boundary conditions to demonstrate the efficiency of the Int-Deep framework in Algorithm 1 (deep learning combined with Newton’s method) and to verify Theorem 4.1.

Example 5.4. Consider the following semilinear elliptic equations

$$\begin{cases} -\Delta u - (u-1)^3 + (u+2)^2 = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where $\Omega = (0,1)^d$ with $d = 1$ or 2 . We choose f such that the problem has an exact solution $u(x) = 3\sin(2\pi x)$ for $d = 1$ and $u(x) = 3\sin(2\pi x)\sin(2\pi y)$ for $d = 2$.

Case $d = 1$.

First of all, we show that the Newton’s method cannot converge to a solution without a good initial guess u_0 and deep learning can provide a good initial guess. We apply the Newton’s iteration in Algorithm 1 with several types of initial guesses u_0 listed in Table 4 with parameters $h = \frac{1}{1024}$, $N_{\max} = 15$, and $\epsilon = 0.01 \times h^2$. u^{DL} is obtained via the deep learning approach based on the variational formula (3.6) with $\lambda = 500$ and 200 epoches in the Adam. Without good knowledge of the ground truth solution or u^{DL} , Newton’s method fails to converge to a good numerical solution.

Table 4: The performance of Newton’s method with different initial guesses u_0 . u is the ground truth solution and ω stands for a Gaussian random noise with mean zero and unit variance.

u_0	#K	e_0^h	e^h	u_0	#K	e_0^h	e^h
1	5	4.0e+0	2.0e+0	ω	6	5.1e+0	3.0e+0
2	5	5.0e+0	2.0e+0	1 + ω	5	7.0e+0	2.0e+0
5	15	8.0e+0	1.3e+5	-1 + ω	15	6.4e+0	5.3e+4
-1	15	4.0e+0	9.5e+4	u + ω	6	3.7e+0	1.8e-5
-2	12	5.0e+00	4.7e+0	$u + 2.5 \times \omega$	15	9.6e+0	5.7e+3
-5	15	8.0e+0	6.5e+4	u^{DL}	5	3.0e-01	1.8e-5

Besides, Picard’s iteration is a global convergence algorithm for few problems. However, it can not work for this problem. We apply Picard’s iteration with several types of initial guesses u_0 listed in Table 5 with parameters $h = \frac{1}{1024}$, $N_{\max} = 15$, and $\epsilon = 0.01 \times h^2$.

Table 5: The performance of Picard’s method with different initial guesses u_0 . u is the ground truth solution and ω stands for a Gaussian random noise with mean zero and unit variance.

u_0	#K	e_0^h	e^h	u_0	#K	e_0^h	e^h
1	15	4.0e+0	2.0e+0	ω	15	6.0e+0	2.0e+0
0	15	3.0e+0	2.0e+0	$u + \omega$	10	1.9e+0	nan

Next, we show that the initial guess by deep learning enables Newton’s method to converge to a solution with the precision of finite element methods, i.e., the numerical convergence order in terms of h defined by

$$\text{order} := \left| \frac{\log_2 e_k^h}{\log_2 e_k^{h/2}} \right| \quad (5.2)$$

is 2 as proved by Theorem 4.1. For the purpose of convenience, we choose $h = 2^{-\ell}$ for different integers ℓ 's and let $e_{h^2} = h^2$ as the theoretical precision of finite element methods. We repeatedly apply the same deep learning method as in Table 4 to generate different initial guesses u^{DL} for different h 's. Table 6 summarizes the performance of Algorithm 1 and numerical results verify the accuracy and the convergence order of the Int-Deep framework, even though the number of epoches in deep learning is $O(100)$ and the initial error is very large.

Table 6: The performance of Int-Deep in Algorithm 1 for Example 5.4 in 1D with different mesh sizes h .

h	e_{h^2}	# epoch	e^{DL}	#K	e^h	order
2^{-7}	6.1e-5	200	3.0e-1	5	1.2e-3	-
2^{-8}	1.5e-5	200	3.0e-1	5	2.9e-4	2.0
2^{-9}	3.8e-6	200	3.0e-1	5	7.3e-5	2.0
2^{-10}	9.5e-7	200	3.0e-1	5	1.8e-5	2.0
2^{-11}	2.4e-7	200	3.0e-1	5	4.6e-6	2.0
2^{-12}	6.0e-8	200	3.0e-1	5	1.1e-6	2.0
2^{-13}	1.5e-8	200	3.0e-1	5	2.8e-7	2.0
2^{-14}	3.7e-9	200	3.0e-1	5	7.1e-8	2.0

Finally, we show that the performance of the Int-Deep in Algorithm 1 is independent of the size of DNNs, which is supported by the numerical results in Table 7.

Table 7: The performance of Int-Deep in Algorithm 1 for Example 5.4 in 1D with different DNN width N and depth L when $h = \frac{1}{1024}$. T^{DL} denotes the running time of GPU for deep learning training in Pytorch.

L \ N	10				30				50			
	e_0^h	e^h	#K	T^{DL}	e_0^h	e^h	#K	T^{DL}	e_0^h	e^h	#K	T^{DL}
2	3.7e+0	1.8e-5	8	42.0s	2.3e-1	1.8e-5	5	42.0s	5.2e-2	1.8e-5	4	46.9s
4	1.1e-1	1.8e-5	4	48.1s	1.6e-2	1.8e-5	3	49.0s	1.5e-2	1.8e-5	3	48.7s
6	9.8e-2	1.8e-5	4	45.1s	8.5e-3	1.8e-5	3	50.4s	5.3e-3	1.8e-5	3	51.8s

Case $d = 2$.

Again, we show that the initial guess by deep learning enables Newton's method to converge to a solution with the precision of finite element methods, i.e., the numerical convergence order in terms of h is almost 2 as proved by Theorem 4.1. We repeatedly apply the variational formula in (3.6) with $\lambda = 500$ and deep learning to generate different initial guesses u^{DL} for different h 's. Let $N_{\max} = 15$ and $\epsilon = 0.01 \times h^2$ in Algorithm 1. Table 8 summarizes the performance of Algorithm 1 and numerical results verify the accuracy and the convergence order of the Int-Deep framework, even though the number of epochs in deep learning is $O(100)$ and the initial error is very large.

Table 8: The performance of Int-Deep in Algorithm 1 for Example 5.4 in 2D with different mesh sizes h .

h	e_{h^2}	# epoch	e^{DL}	#K	e^h	order
2^{-4}	1.1e-2	200	2.2e+0	5	2.0e-1	-
2^{-5}	3.4e-3	200	2.2e+0	5	5.9e-2	1.8
2^{-6}	1.0e-3	200	2.2e+0	5	1.5e-2	1.9
2^{-7}	3.0e-4	200	2.2e+0	6	3.9e-3	2.0
2^{-8}	8.5e-5	200	2.2e+0	6	9.8e-4	2.0

Finally, we show that the performance of the Int-Deep in Algorithm 1 is independent of the size of DNNs, which is supported by the numerical results in Table 9.

Table 9: The performance of Int-Deep in Algorithm 1 for Example 5.4 in 1D with different DNN width N and depth L when $h = \frac{1}{128}$. T^{DL} denotes the running time of GPU for deep learning training in Pytorch.

L \ N	10				30				50			
	e_0^h	e^h	#K	T^{DL}	e_0^h	e^h	#K	T^{DL}	e_0^h	e^h	#K	T^{DL}
2	3.2e+0	3.9e-3	6	48.2s	2.9e+0	3.9e-3	6	48.7s	3.3e+0	3.9e-3	6	55.5s
4	4.1e+0	3.9e-3	6	53.9s	2.5e+0	3.9e-3	5	52.3s	7.6e-1	3.9e-3	4	61.1s
6	3.5e+0	3.9e-3	6	66.1s	1.2e+0	3.9e-3	4	58.9s	1.1e+0	3.9e-3	4	59.9s

5.2.2 Eigenvalue problem

Here, we demonstrate the efficiency of the Int-Deep framework in Algorithm 2 (deep learning combined with the power method) and to verify Theorem 4.2. λ^h is the approximation to λ .

Example 5.5. Consider the following problem

$$\begin{aligned} -\Delta u &= \lambda u \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where $\Omega = (0, 1)^d$ for $d = 1$ and 2 . The smallest eigenvalue is $d\pi^2$. The corresponding eigenfunction is $u(x) = \sin(\pi(x - 1))$ for $d = 1$, and $u(x, y) = \sin(\pi(x - 1)) \sin(\pi(y - 1))$ for $d = 2$.

Case $d = 1$.

We show that the initial guess by deep learning enables the power method to quickly converge to a solution with the precision of finite element methods, i.e., the numerical convergence order in terms of h is almost 2 as proved by Theorem 4.2 if the number of iterations is $O(1)$. We repeatedly apply the variational formula in (3.15) (with $x_0 = 0.5$ and $\gamma = 100$) and deep learning to generate different initial guesses u^{DL} for different h 's. Let $N_{\max} = 10$ and $\epsilon = h^2$ in Algorithm 2. Table 10 summarizes the performance of Algorithm 2 and numerical results verify the accuracy and the convergence order of the Int-Deep framework, even though the number of epochs in deep learning is $O(100)$ and the initial error is very large.

Table 10: The performance of Int-Deep in Algorithm 2 for Example 5.5 in 1D with different mesh sizes h .

h	#epoch	e^{DL}	#K	$ \lambda - \lambda^h $	eigenvalue order	e^h	eigenfunction order
2^{-5}	300	1.1e-2	3	7.9e-3	-	8.1e-4	-
2^{-6}	300	1.1e-2	4	2.0e-3	2.0	2.0e-4	2.0
2^{-7}	300	1.1e-2	5	5.0e-4	2.0	5.1e-5	2.0
2^{-8}	300	1.1e-2	6	1.2e-4	2.0	1.3e-5	2.0
2^{-9}	300	1.1e-2	7	3.0e-5	2.0	3.2e-6	2.0

Next, we show that the performance of Int-Deep in Algorithm 2 is independent of the size of DNNs, which is supported by the numerical results in Table 11.

Table 11: The performance of Int-Deep in Algorithm 2 for Example 5.5 in 1D with different DNN width N and depth L when $h = \frac{1}{512}$. T^{DL} denotes the running time of GPU for deep learning traning in Pytorch.

L \ N	10				30				50			
	$ \lambda - \lambda^h $	e^h	#K	T^{DL}	$ \lambda - \lambda^h $	e^h	#K	T^{DL}	$ \lambda - \lambda^h $	e^h	#K	T^{DL}
2	1.8e-5	4.4e-6	8	50.0s	3.0e-5	3.e-6	7	40.4s	3.2e-5	3.4e-6	7	40.9s
4	2.5e-5	3.4e-6	8	42.7s	2.0e-5	4.2e-6	8	41.9s	3.0e-5	3.2e-6	7	42.3s
6	3.2e-5	3.2e-6	7	42.9s	3.0e-5	3.2e-6	7	42.4s	3.1e-5	3.2e-6	7	43.2s

Case $d = 2$.

Again, we show that the initial guess by deep learning enables the power method to quickly converge to a solution with the precision of finite element methods, i.e., the numerical convergence order in terms of h is almost 2 as proved by Theorem 4.2 if the number of iterations is $O(1)$. We repeatedly apply the variational formula in (3.15) (with $\mathbf{x}_0 = (0.5, 0.5)$ and $\gamma = 100$) and deep learning to generate different initial guesses u^{DL} for different h 's. Let $N_{\max} = 10$ and $\epsilon = h^2$ in Algorithm 2. Table 12 summarizes the performance of Algorithm 2 and numerical results verify the accuracy and the convergence order of the Int-Deep framework, even though the number of epochs in deep learning is $O(100)$ and the initial error is very large.

Table 12: The performance of Int-Deep in Algorithm 2 for Example 5.5 in 2D with different mesh sizes h .

h	#epoch	e^{DL}	#K	$ \lambda - \lambda^h $	eigenvalue order	e^h	eigenfunction order
2^{-4}	300	5.6e-2	4	1.9e-1	-	6.6e-3	-
2^{-5}	300	5.6e-2	5	4.8e-2	2.0	1.7e-3	2.0
2^{-6}	300	5.7e-2	7	1.2e-2	2.0	4.2e-4	2.0
2^{-7}	300	5.7e-2	8	3.0e-3	2.0	1.1e-4	2.0
2^{-8}	300	5.7e-2	10	7.4e-4	2.0	2.6e-5	2.0

Next, we show that the performance of Int-Deep in Algorithm 2 is independent of the size of DNNs, which is supported by the numerical results in Table 11.

Table 13: The performance of Int-Deep in Algorithm 2 for Example 5.5 in 2D with different DNN width N and depth L when $h = \frac{1}{128}$. T^{DL} denotes the running time of GPU for deep learning traning in Pytorch.

L \ N	10				30				50			
	$ \lambda - \lambda^h $	e^h	#K	T^{DL}	$ \lambda - \lambda^h $	e^h	#K	T^{DL}	$ \lambda - \lambda^h $	e^h	#K	T^{DL}
2	3.0e-3	1.1e-4	10	47.2s	3.0e-3	1.1e-4	9	43.4s	3.0e-3	1.1e-4	6	43.8s
4	3.0e-3	1.1e-4	8	44.1s	3.0e-3	1.1e-4	8	45.2s	3.0e-3	1.1e-4	7	49.8s
6	3.0e-3	1.1e-4	9	51.2s	3.0e-3	1.1e-4	8	46.3s	3.0e-3	1.1e-4	7	51.1s

6 Consclusion

This paper proposed the Int-Deep framework from a new point of view for designing highly efficient solvers of low-dimensional nonlinear PDEs with a finite element accuracy leveraging both the advantages of traditional algorithms and deep learning approaches. The Int-Deep framework consists of two phases. In the first phase, an approximate solution to the given nonlinear PDE is obtained via deep learning approaches using DNNs of size $O(1)$ and $O(100)$ iterations. In the second phase, the approximate solution provided by deep learning can serve as a good initial guess such that traditional iterative methods converge in $O(\log(\frac{1}{\epsilon}))$ iterations to the ϵ precision of finite element methods. The Int-Deep framework outperforms existing purely deep learning-based methods or traditional iterative methods.

In particular, based on variational principles, we propose new methods to formulate the problem of solving nonlinear PDEs into an unconstrained minimization problem of an expectation over a function space parametrized via DNNs, which can be solved efficiently via batch stochastic gradient descent (SGD) methods due to the special form of expectation. Unlike previous methods in which the form of expectation is only derived for nonlinear PDEs related to variational equations, our proposed method can also handle variational inequalities and eigenvalue problems, providing a unified variational framework for a wider range of nonlinear problems. With the good initialization given by deep learning, we hope to reduce the difficulty for designing an efficient traditional iterative algorithm for variational inequalities. We will leave this as a future work.

Acknowledgments. J. H. was partially supported by NSFC (Grant No.11571237). H. Y. was partially supported by Grant R-146-000-251-133 in the Department of Mathematics at the National University of Singapore, by the Ministry of Education in Singapore under the grant MOE2018-T2-2-147, and the start-up grant of the Department of Mathematics at Purdue University.

References

- [1] A. L. Fetter, Vortices in an imperfect bose gas. i. the condensate, Phys. Rev. 138 (1965) A429–A437. doi:10.1103/PhysRev.138.A429.
URL <https://link.aps.org/doi/10.1103/PhysRev.138.A429>
- [2] J. R. Anglin, W. Ketterle, Boseinstein condensation of atomic gases, Nature 416 (2002) 211–218. doi:10.1038/416211a.
- [3] P. Hohenberg, W. Kohn, Inhomogeneous electron gas, Phys. Rev. 136 (1964) B864–B871.

- doi:10.1103/PhysRev.136.B864.
 URL <https://link.aps.org/doi/10.1103/PhysRev.136.B864>
- [4] W. Kohn, L. J. Sham, Self-consistent equations including exchange and correlation effects, *Phys. Rev.* 140 (1965) A1133–A1138. doi:10.1103/PhysRev.140.A1133.
 URL <https://link.aps.org/doi/10.1103/PhysRev.140.A1133>
 - [5] C. Basdevant, M. Deville, P. Haldenwang, J. Lacroix, J. Ouazzani, R. Peyret, P. Orlandi, A. Patera, Spectral and finite difference solutions of the burgers equation, *Computers & Fluids* 14 (1) (1986) 23 – 41. doi:[https://doi.org/10.1016/0045-7930\(86\)90036-8](https://doi.org/10.1016/0045-7930(86)90036-8).
 URL <http://www.sciencedirect.com/science/article/pii/0045793086900368>
 - [6] R. Miura, The kortewegdevries equation: A survey of results, *SIAM Review* 18 (3) (1976) 412–459. arXiv:<https://doi.org/10.1137/1018076>, doi:10.1137/1018076.
 URL <https://doi.org/10.1137/1018076>
 - [7] J. M. Hyman, B. Nicolaenko, The kuramoto-sivashinsky equation: A bridge between pde’s and dynamical systems, *Physica D: Nonlinear Phenomena* 18 (1) (1986) 113 – 126. doi:[https://doi.org/10.1016/0167-2789\(86\)90166-1](https://doi.org/10.1016/0167-2789(86)90166-1).
 URL <http://www.sciencedirect.com/science/article/pii/0167278986901661>
 - [8] A. J. Chorin, Numerical solution of the navier-stokes equations, *Mathematics of Computation* 22 (104) (1968) 745–762.
 URL <http://www.jstor.org/stable/2004575>
 - [9] A. Quarteroni, A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer, New York, 1994.
 - [10] S. Bartels, *Numerical Methods for Nonlinear Partial Differential Equations*, Springer International Publishing, Cham, 2015.
 - [11] T. R. L. Zienkiewicz, O. C., J. Zhu, *Numerical Approximation of Partial Differential Equations*, Elsevier, Amsterdam, 2005.
 - [12] J. Xu, A novel two-grid method for semilinear elliptic equations, *SIAM Journal on Scientific Computing* 15 (1) (1994) 231–237.
 - [13] J. Xu, Two-grid discretization techniques for linear and nonlinear pdes, *SIAM Journal on Numerical Analysis* 33 (5) (1996) 1759–1777.
 - [14] Mingui Sun, Xiaopu Yan, R. J. Sclabassi, Solving partial differential equations in real-time using artificial neural network signal processing as an alternative to finite-element analysis, in: *International Conference on Neural Networks and Signal Processing*, 2003. Proceedings of the 2003, Vol. 1, 2003, pp. 381–384 Vol.1. doi:10.1109/ICNNSP.2003.1279289.
 - [15] W. Tang, T. Shan, X. Dang, M. Li, F. Yang, S. Xu, J. Wu, Study on a poisson’s equation solver based on deep learning technique, in: *2017 IEEE Electrical Design of Advanced Packaging and Systems Symposium (EDAPS)*, 2017, pp. 1–3.

- [16] Y. Khoo, J. Lu, L. Ying, Solving Parametric PDE Problems with Artificial Neural Networks, arXiv e-prints (2017) arXiv:1707.03351.
- [17] Y. Khoo, L. Ying, SwitchNet: a neural network model for forward and inverse scattering problems, arXiv e-prints (2018) arXiv:1810.09675arXiv:arXiv:1810.09675.
- [18] Y. Fan, L. Lin, L. Ying, L. Zepeda-Nunez, A multiscale neural network based on hierarchical matrices, arXiv:1807.01883.
- [19] Y. Fan, J. Feliu-Fabà, L. Lin, L. Ying, L. Zepeda-Núñez, A multiscale neural network based on hierarchical nested bases, Research in the Mathematical Sciences 6 (2) (2019) 21. doi:10.1007/s40687-019-0183-3.
URL <https://doi.org/10.1007/s40687-019-0183-3>
- [20] Y. Fan, C. O. Bohorquez, L. Ying, Bcr-net: A neural network based on the nonstandard wavelet form, J. Comput. Physics 384 (2019) 1–15.
- [21] J. Berg, K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, Neurocomputing 317 (2018) 28 – 41. doi:<https://doi.org/10.1016/j.neucom.2018.06.056>.
URL <http://www.sciencedirect.com/science/article/pii/S092523121830794X>
- [22] G. Carleo, M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science 355 (2017) 602–606.
- [23] J. Han, A. Jentzen, W. E, Solving high-dimensional partial differential equations using deep learning, Proceedings of the National Academy of Sciences 115 (34) (2018) 8505–8510. arXiv:<https://www.pnas.org/content/115/34/8505.full.pdf>, doi:10.1073/pnas.1718942115.
URL <https://www.pnas.org/content/115/34/8505>
- [24] Y. Khoo, J. Lu, L. Ying, Solving for high-dimensional committor functions using artificial neural networks, Research in the Mathematical Sciences 6 (1) (2018) 1. doi:10.1007/s40687-018-0160-2.
URL <https://doi.org/10.1007/s40687-018-0160-2>
- [25] K. Rudd, S. Ferrari, A constrained integration (cint) approach to solving partial differential equations using artificial neural networks, Neurocomputing 155 (2015) 277 – 285.
- [26] J. Sirignano, K. Spiliopoulos, DGM: A deep learning algorithm for solving partial differential equations, Journal of Computational Physics 375 (2018) 1339 – 1364. doi:<https://doi.org/10.1016/j.jcp.2018.08.029>.
URL <http://www.sciencedirect.com/science/article/pii/S0021999118305527>
- [27] M. Raissi, P. Perdikaris, G. 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 (2019) 686 – 707.
- [28] V. Kůrková, Kolmogorov’s theorem and multilayer neural networks, Neural Netw. 5 (1992) 501–506.

- [29] A. Barron, Universal approximation bounds for superpositions of a sigmoidal function, *IEEE Trans. Inf. Theory* 39 (1993) 930–945.
- [30] D. Yarotsky, Error bounds for approximations with deep ReLU networks, *Neural Netw.* 94 (2017) 103–114.
- [31] D. Yarotsky, Optimal approximation of continuous functions by very deep ReLU networks, in: S. Bubeck, V. Perchet, P. Rigollet (Eds.), *31st Annual Conference on Learning Theory, Proc. Mach. Learn. Res.* 75, 2018, pp. 1–11.
- [32] Z. Shen, H. Yang, S. Zhang, Deep Network Approximation Characterized by Number of Neurons, *arXiv e-prints* (2019) arXiv:1906.05497.
- [33] H. Montanelli, Q. Du, New error bounds for deep ReLU networks using sparse grids, *SIAM J. Math. Data Sci.* 1 (2019) 78–92.
- [34] H. Montanelli, H. Yang, Deep relu networks overcome the curse of dimensionality for bandlimited functions, arXiv:1903.00735.
- [35] H. Montanelli, H. Yang, Error bounds for deep relu networks using the kolmogorov–arnold superposition theorem, arXiv:1906.11945.
- [36] K. Shao, J. Chen, Z. Zhao, D. H. Zhang, Communication: Fitting potential energy surfaces with fundamental invariant neural network, *The Journal of Chemical Physics* 145 (7) (2016) 071101. arXiv:<https://doi.org/10.1063/1.4961454>, doi:10.1063/1.4961454.
URL <https://doi.org/10.1063/1.4961454>
- [37] C. Xie, X. Zhu, D. R. Yarkony, H. Guo, Permutation invariant polynomial neural network approach to fitting potential energy surfaces. iv. coupled diabatic potential energy matrices, *The Journal of Chemical Physics* 149 (14) (2018) 144107. arXiv:<https://doi.org/10.1063/1.5054310>, doi:10.1063/1.5054310.
URL <https://doi.org/10.1063/1.5054310>
- [38] L. Zhang, J. Han, H. Wang, W. A. Saidi, R. Car, E. Weinan, End-to-end symmetry preserving inter-atomic potential energy model for finite and extended systems, in: *Proceedings of the 32Nd International Conference on Neural Information Processing Systems, NIPS’18*, Curran Associates Inc., USA, 2018, pp. 4441–4451.
URL <http://dl.acm.org/citation.cfm?id=3327345.3327356>
- [39] A. Meade, A. Fernandez, Solution of nonlinear ordinary differential equations by feedforward neural networks, *Mathematical and Computer Modelling* 20 (9) (1994) 19 – 44. doi:[https://doi.org/10.1016/0895-7177\(94\)00160-X](https://doi.org/10.1016/0895-7177(94)00160-X).
URL <http://www.sciencedirect.com/science/article/pii/089571779400160X>
- [40] Y. Zang, G. Bao, X. Ye, H. Zhou, Weak adversarial networks for high-dimensional partial differential equations, arXiv:1907.08272 [math.NA].
- [41] I. E. Lagaris, A. Likas, D. I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Transactions on Neural Networks* 9 (5) (1998) 987–1000. doi:10.1109/72.712178.

- [42] R. S. Beidokhti, A. Malek, Solving initial-boundary value problems for systems of partial differential equations using neural networks and optimization techniques, *Journal of the Franklin Institute* 346 (9) (2009) 898 – 913. doi:<https://doi.org/10.1016/j.jfranklin.2009.05.003>.
URL <http://www.sciencedirect.com/science/article/pii/S0016003209000878>
- [43] K. S. McFall, J. R. Mahan, Artificial neural network method for solution of boundary value problems with exact satisfaction of arbitrary boundary conditions, *IEEE Transactions on Neural Networks* 20 (8) (2009) 1221–1233. doi:10.1109/TNN.2009.2020735.
- [44] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, Adaptive Computation and Machine Learning, MIT Press, Cambridge, MA, 2016.
- [45] W. E, B. Yu, The deep ritz method: A deep learning-based numerical algorithm for solving variational problems, *Communications in Mathematics and Statistics* 6 (1) (2018) 1–12.
- [46] R. A. Adams, *Sobolev Spaces*, Academic Press, New York, 1975.
- [47] M. W. M. G. Dissanayake, N. Phan-Thien, Neural-network-based approximations for solving partial differential equations, *Communications in Numerical Methods in Engineering* 10 (3) (1994) 195–201. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/cnm.1640100303>, doi:10.1002/cnm.1640100303.
URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.1640100303>
- [48] D. P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization, arXiv e-prints (2014) arXiv:1412.6980.
- [49] P. B. Bochev, M. D. Gunzburger, *Least-Squares Finite Element Methods*, Springer, New York, 2009.
- [50] G. Duvaut, J. L. Lions, *Inequalities in Mechanics and Physics*, Springer, Berlin, 1976.
- [51] R. Glowinski, *Numerical methods for nonlinear variational problems*, Springer-Verlag, New York, 1984.
- [52] A. Ambrosetti, A. Malchiodi, *Nonlinear analysis and semilinear elliptic problems*, Cambridge University Press, Cambridge, 2007.
- [53] J. Stoer, R. Bulirsch, *Introduction to Numerical Analysis*, 3rd Edition, Springer-Verlag, New York, 2002.
- [54] J. Xu, A. Zhou, A two-grid discretization scheme for eigenvalue problems, *Mathematics of Computation* 70 (233) (2001) 17–25.
- [55] S. C. Brenner, L. R. Scott, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.
- [56] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978.

- [57] I. Babuška, J. E. Osborn, Finite element-galerkin approximation of the eigenvalues and eigenvectors of selfadjoint problems, *Math. Comp.* (1989) 275–297.
- [58] I. Babuška, J. Osborn, Eigenvalue problems, in: *Handbook of Numerical Analysis*, Vol. II, North-Holland, Amsterdam, 1991.
- [59] Q. D. Zhu, Q. Lin, *The Hyperconvergence Theory of Finite Elements*, Hunan Science and Technology Publishing House, Changsha, 1989.

Appendix

A Proof of Theorem 4.1

Making use of the assumption A2 in Section 4.1, we immediately have the following result (see [12]).

Lemma A.1. *Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.1). There exist two positive constant C_1^u and C_2^u such that if a function v satisfies that $\|v - u\|_{0,\infty} \leq C_1^u$, then*

$$C_2^u \|\phi\|_1 \leq \sup_{\chi \in V} \frac{a_v(\phi, \chi)}{\|\chi\|_1}, \quad \phi \in V.$$

The mathematical analysis for the finite element method (4.4) is very technical and has been established in [13]. In the following two lemmas, we collect some results which will be used frequently later on.

Lemma A.2. *Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.2). Then there exists a constant $h_0 > 0$ and a positive constant C_3^u independent of h such that if $h < h_0$ and a function v satisfies $\|v - u\|_{0,\infty} \leq C_1^u$, then*

$$C_3^u \|\phi\|_1 \leq \sup_{\chi \in V_h} \frac{a_v(\phi, \chi)}{\|\chi\|_1}, \quad \phi \in V_h.$$

This lemma can be derived by using Lemma A.1 and the arguments for proving Lemma 2.2 in [13].

Lemma A.3. *Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.2). Then there exists a positive constant $h_1 < h_0$ such that if $h < h_1$, the finite element method (4.4) has exactly one solution u^h satisfying*

$$\|u - u^h\|_1 \leq C_4^u$$

for a positive constant C_4^u independent of h . Moreover,

$$\lim_{h \rightarrow 0^+} \|u - u^h\|_1 = 0.$$

In what follows, to simplify the presentation, for any two quantities a and b , we write “ $a \lesssim b$ ” for “ $a \leq Cb$ ”, where C is a generic positive constant independent of h , which may take different values at different occurrences. Moreover, any symbols C or c (with or without superscript or subscript) denote positive generic constants independent of the finite element mesh size h .

Lemma A.4. *Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.2) and u^h be the finite element method of (4.4). Then, there exists a positive constant $h_3 < h_2$ such that if $h < h_3$,*

$$\begin{aligned} \|u - u^h\|_{0,\infty} &\lesssim h^2 && \text{when } d = 1, \\ \|u - u^h\|_{0,\infty} &\lesssim h^2 |\log(h)| && \text{when } d = 2. \end{aligned}$$

Proof. The estimate for $d = 2$ is given in [13]. We will derive the estimate for $d = 1$ following some ideas in [13]. To this end, we first introduce an elliptic projection operator P_h such that if $u \in V$, then $P_h u \in V_h$ satisfies

$$(\nabla(P_h u), \nabla \chi) = (\nabla u, \nabla \chi), \quad \chi \in V_h. \tag{A.1}$$

It is shown in [59] that

$$\|u - P_h u\|_{0,\infty} \lesssim \|u - P_h u\|_1 \lesssim h^2 |u|_{2,\infty}. \quad (\text{A.2})$$

On the other hand, the finite element solution u^h satisfies

$$(\nabla u^h, \nabla \chi) + (f(u^h), \chi) = 0, \quad \chi \in V_h.$$

Recalling the relation (A.1) and using (4.2) gives

$$(\nabla(P_h u), \nabla \chi) = (\nabla u, \nabla \chi) = -(f(u), \chi), \quad \chi \in V_h.$$

Hence, subtracting the last two equations and using Taylor's expansion yield

$$a_{P_h u}((u^h - P_h u), \chi) = -\frac{1}{2}(f''(\zeta)(u^h - P_h u)^2, \chi) - (f(P_h u) - f(u), \chi), \quad \chi \in V_h,$$

where $\zeta = \nu(x)u^h(x) + (1 - \nu(x))(P_h u)(x)$ for some $\nu(x) \in (0, 1)$. We have by the estimate (A.2) that there exists a positive constant $\bar{h} < h_2$ such that if $h < \bar{h}$, $\|P_h u - u\|_{0,\infty} \leq C_1^u$. In this case, it follows from Lemma A.2 and the last equation that

$$\begin{aligned} \|u^h - P_h u\|_1 &\lesssim \sup_{\chi \in V_h} \frac{a_{P_h u}((u^h - P_h u), \chi)}{\|\chi\|_1} \\ &\lesssim \|u^h - P_h u\|_{0,\infty}^2 + \|f(u) - f(P_h u)\|_{0,\infty} \\ &\lesssim \|u^h - P_h u\|_1^2 + \|u - P_h u\|_{0,\infty} \\ &\leq C\|u^h - P_h u\|_1^2 + C\|u - P_h u\|_{0,\infty}, \end{aligned} \quad (\text{A.3})$$

where $C > 0$ is a generic constant. On the other hand, it follows from Lemma A.3 and the estimate (A.2) that

$$\|u^h - P_h u\|_1 \leq \|u - u^h\|_1 + \|u - P_h u\|_1 \rightarrow 0 \quad \text{as } h \rightarrow 0^+.$$

Hence, there exists a positive constant $h_3 < \bar{h}$ such that if $h < h_3$, then $C\|u^h - P_h u\|_1 < 1/2$, which combined with (A.3) readily implies

$$\|u^h - P_h u\|_1 \lesssim \|u - P_h u\|_{0,\infty}.$$

Therefore, by the Sobolev embedding theorem and the estimate (A.2),

$$\begin{aligned} \|u - u^h\|_{0,\infty} &\leq \|u - P_h u\|_{0,\infty} + \|P_h u - u^h\|_{0,\infty} \\ &\lesssim \|u - P_h u\|_{0,\infty} + \|P_h u - u^h\|_1 \lesssim \|u - P_h u\|_{0,\infty} \lesssim h^2 |u|_{2,\infty}, \end{aligned}$$

as required. \square

Remark 1. As shown in [55], we require to make certain strong regularity assumption on the solution to problem (4.2) for $d = 3$, so as to derive the maximum norm estimate for the related finite element method. To avoid too technical treatment, we skip the further discussion in this case.

Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.2). Define

$$B(u) = \{v \in V : \|v - u\|_{0,\infty} \leq C_1^a\}.$$

The following result plays an important role in the convergence analysis of Int-Deep for solving the finite element method (4.4), which will be introduced later on.

Lemma A.5. Let $u \in W^{2,\infty}(\Omega)$ be a solution of (4.2). Let u^h be an approximation of u obtained by the finite element method (4.4). Assume that $h < h_3$ with h_3 the same as given in Lemma A.4. For any $v \in V_h \cap B(u)$, define a mapping $Tv = v + w$, where $w \in V_h$ is uniquely determined by

$$(\nabla w, \nabla \chi) + (f'(v)w, \chi) = -(\nabla v, \nabla \chi) - (f(v), \chi), \quad \chi \in V_h. \quad (\text{A.4})$$

Then there holds

$$\|u^h - Tv\|_1 \lesssim \|u^h - v\|_{0,p}^2 \lesssim \|u^h - v\|_1^2.$$

Proof. Recalling the definition (4.4), we know

$$(\nabla u^h, \nabla \chi) + (f(u^h), \chi) = 0, \quad \chi \in V_h,$$

Subtracting the above equation from (A.4) and reorganizing terms, we find

$$(\nabla E_h, \nabla \chi) + (f(u^h) - f(v) - f'(v)w, \chi) = 0, \quad \chi \in V_h,$$

where $E_h = u^h - Tv$. Furthermore, use Taylor's expansion to get

$$a_v(E_h, \chi) = -\frac{1}{2}(f''(\zeta_h)(u^h - v)^2, \chi), \quad \chi \in V_h,$$

where $\zeta = (1 - \nu(x))v(x) + \nu(x)u^h(x)$ with $\nu(x) \in (0, 1)$. Since $h < h_3$ and $v \in B(u)$, from Lemma A.3 it follows that $\|\zeta_h\|_{0,\infty}$ is uniformly bounded with respect to h . Hence, by the Hölder inequality and the Sobolev embedding theorem, for $p > 2$ and any $\chi \in V_h$,

$$\begin{aligned} a_v(E_h, \chi) &\lesssim \int_{\Omega} (u^h - v)^2 |\chi| \, dx \\ &\lesssim \|(u^h - v)^2\|_{0,p/2} \|\chi\|_{0,p/p-2} \lesssim \|u^h - v\|_{0,p}^2 \|\chi\|_1, \end{aligned}$$

where the generic constant is independent of h but depends on p . The combination of the last estimate with (A.2) immediately implies

$$\|E_h\|_1 \lesssim \|u^h - v\|_{0,p}^2 \lesssim \|u^h - v\|_1^2,$$

as required. □

Now we are ready to prove Theorem 4.1.

Proof. Denote $E_k^h = u^h - u_k^h$. Applying Lemma (A.5) gives rise to

$$\|E_{k+1}^h\|_1 \lesssim \|E_k^h\|_{0,p}^2 \lesssim \|E_k^h\|_1^2. \quad (\text{A.5})$$

When $d = 1$, By the Sobolev embedding theorem and (A.5),

$$\|E_{k+1}^h\|_{0,\infty} \lesssim \|E_{k+1}^h\|_1 \lesssim \|E_k^h\|_{0,p}^2 \lesssim \|E_k^h\|_{0,\infty}^2.$$

This means there exists a positive constant c_1 such that $\|E_k^h\|_{0,\infty} \leq c_1 \|E_{k-1}^h\|_{0,\infty}^2$. Hence,

$$c_1 \|E_k^h\|_{0,\infty} \leq (c_1 \|E_{k-1}^h\|_{0,\infty})^2 \leq (c_1 \|E_0^h\|_{0,\infty})^{2^k}.$$

On the other hand, it follows from Lemma A.4 and error estimates for the interpolation operator I_h (cf. [55, 56]) that

$$\begin{aligned}\|E_0^h\|_{0,\infty} &= \|u^h - u_0^h\|_{0,\infty} \\ &\leq \|u^h - u\|_{0,\infty} + \|u - I_h u\|_{0,\infty} + \|I_h u - I_h u^{DL}\|_{0,\infty} \\ &\leq c_0(h^2 + \delta),\end{aligned}$$

where $c_0 > 0$ is a generic constant. Let $\beta_1 = c_1 c_0(h^2 + \delta)$. Then

$$\|E_k^h\|_{0,\infty} \leq \beta_1^{2^k} / c_1. \quad (\text{A.6})$$

When $d = 2$, for any given number $p > 2$, we have by the Sobolev embedding theorem and (A.5) that

$$\|E_{k+1}^h\|_{0,p} \lesssim \|E_{k+1}^h\|_1^2 \lesssim \|E_k^h\|_{0,p}^2,$$

which implies $c_2 \|E_k^h\|_{0,p} \leq (c_2 \|E_0^h\|_{0,p})^{2^k}$ for a generic positive constant c_2 . In addition, by Lemma A.4 and error estimates for I_h ,

$$\begin{aligned}\|E_0^h\|_{0,p} &= \|u^h - u_0^h\|_{0,p} \\ &\leq \|u^h - u\|_{0,p} + \|u - I_h u\|_{0,p} + \|I_h u - u_0^h\|_{0,p} \\ &\lesssim \|u^h - u\|_{0,\infty} + \|u - I_h u\|_{0,\infty} + \|I_h u - u_0^h\|_{0,\infty} \\ &\leq c_3(h^2 |\log h| + \delta),\end{aligned}$$

where $c_3 > 0$ is a generic constant dependent on p . Let $\beta_2 = c_2 c_3(h^2 |\log h| + \delta)$. Then

$$\|E_k^h\|_{0,p} \leq \beta_2^{2^k} / c_2,$$

and further by the inverse inequality for finite elements,

$$\|E_k^h\|_{0,\infty} \lesssim h^{-2/p} \|E_k^h\|_{0,p} \lesssim h^{-2/p} \beta_2^{2^k}. \quad (\text{A.7})$$

Now, we have by (A.6), (A.7) and Lemma A.4 that

$$\|u - u_k^h\|_{0,\infty} \leq \|u - u^h\|_{0,\infty} + \|u^h - u_k^h\|_{0,\infty} \lesssim h^2 + \beta_1^{2^k} \quad \text{for } d = 1,$$

and

$$\|u - u_k^h\|_{0,\infty} \leq \|u - u^h\|_{0,\infty} + \|u^h - u_k^h\|_{0,\infty} \lesssim h^2 |\log h| + h^{-2/p} \beta_2^{2^k} \quad \text{for } d = 2.$$

The proof is complete. □

B Proof of Theorem 4.2

Proof. For any $\phi \in V_h$, we know

$$\begin{aligned}a(P_h u - u_{k+1}^h, \chi) &= a(u, \chi) - a(u_{k+1}^h, \chi) = \lambda(u, \chi) - \lambda_k^h(u_k^h, \chi) \\ &= (\lambda - \lambda_k^h)(u, \chi) + \lambda_k^h(u - u_k^h, \chi).\end{aligned}$$

Choosing $\chi = P_h u - u_{k+1}^h$, we have by the coerciveness of $a(\cdot, \cdot)$ that

$$\begin{aligned} \alpha_0 \|P_h u - u_{k+1}^h\|_1^2 &\leq a(P_h u - u_{k+1}^h, P_h u - u_{k+1}^h) \\ &= |\lambda - \lambda_k^h|(u, P_h u - u_{k+1}^h) + |\lambda_k^h|(u - u_k^h, P_h u - u_{k+1}^h) \\ &\lesssim \left(|\lambda - \lambda_k^h| \|u\|_{-1} + |\lambda_k^h - \lambda + \lambda| \|u - u_k^h\|_{-1} \right) \|P_h u - u_{k+1}^h\|_1, \end{aligned}$$

which combined with the assumption $\|u_k^h\| \leq \varepsilon_1$ implies

$$\begin{aligned} \|P_h u - u_{k+1}^h\|_1 &\lesssim |\lambda - \lambda_k^h| \|u\|_0 + \left(|\lambda - \lambda_k^h| + |\lambda| \right) \|u - u_k^h\|_0 \\ &\lesssim |\lambda - \lambda_k^h| + |\lambda - \lambda_k^h| \left(\|u\|_0 + \|u_k^h\|_0 \right) + \|u - u_k^h\|_0 \\ &\lesssim |\lambda - \lambda_k^h| + \|u - u_k^h\|_0. \end{aligned}$$

Hence, by the error estimate for P_h and the triangle inequality,

$$\|u - u_{k+1}^h\|_1 \leq \|u - P_h u\|_1 + \|P_h u - u_{k+1}^h\|_1 \lesssim h + |\lambda - \lambda_k^h| + \|u - u_k^h\|_0. \quad (\text{B.1})$$

Using the duality argument to the equation determining u_{k+1}^h , we deduce from (B.1) and the Cauchy inequality that

$$\begin{aligned} \|u - u_{k+1}^h\|_0 &\lesssim h \|u - u_{k+1}^h\|_1 \lesssim h^2 + h \left(|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \right) \\ &\lesssim h^2 + \left(|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \right)^2. \end{aligned} \quad (\text{B.2})$$

By Lemma 4.1 and assumption $\|u_{k+1}^h\|_0 \geq \varepsilon_0 > 0$,

$$|\lambda - \lambda_{k+1}^h| = \left| |u_{k+1}^h - u|_1^2 - \lambda \|u_{k+1}^h - u\|_0^2 \right| / \|u_{k+1}^h\|_0^2 \lesssim \|u_{k+1}^h - u\|_1^2.$$

Inserting this into (B.1) gives

$$|\lambda - \lambda_{k+1}^h| \lesssim \|u - u_{k+1}^h\|_1^2 \lesssim \left(h + |\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \right)^2 \lesssim h^2 + \left(|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \right)^2,$$

which combined with (B.2) implies

$$|\lambda - \lambda_{k+1}^h| + \|u - u_{k+1}^h\|_0 \lesssim h^2 + \left(|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \right)^2. \quad (\text{B.3})$$

Let $e_k = |\lambda - \lambda_k^h| + \|u - u_k^h\|_0$. Then the estimate (B.3) can be expressed as

$$e_{k+1} \leq \bar{c}_4 h^2 + c_4 e_k^2,$$

where c_4 and \bar{c}_4 are two positive generic constants. The above inequality can be expressed as

$$c_4 e_{k+1} \leq c_5 h^2 + (c_4 e_k)^2, \quad (\text{B.4})$$

where $c_5 = c_4 \bar{c}_4$.

On the other hand, by assumption and using error estimates of I_h ,

$$e_0 = |\lambda - \lambda_0^h| + \|u - u_0^h\|_0 \leq |\lambda - \lambda_0^h| + \|u - I_h u\|_0 + \|I_h u - u_0^h\|_0 \leq c_6(\bar{\delta} + h^2),$$

where $c_6 > 0$ is a generic constant. In addition, using the estimate (B.4) and a direct manipulation we know

$$c_4 e_k \leq (c_4 e_0)^{2^k} + c_5 h^2 (2^{k+1} - k),$$

which readily implies

$$|\lambda - \lambda_k^h| + \|u - u_k^h\|_0 \lesssim \beta_3^{2^k} + 2^{k+1} h^2.$$

The proof is complete. \square