



---

*Research article*

## Robust error estimates of PINN in one-dimensional boundary value problems for linear elliptic equations

Jihahm Yoo<sup>1</sup> and Haesung Lee<sup>2,\*</sup>

<sup>1</sup> Korea Science Academy of KAIST, Busan 47162, Republic of Korea

<sup>2</sup> Department of Mathematics and Big Data Science, Kumoh National Institute of Technology, Gumi, Gyeongsangbuk-do 39177, Republic of Korea

\* **Correspondence:** Email: fthslt@kumoh.ac.kr, fthslt14@gmail.com.

**Abstract:** In this paper, we study physics-informed neural networks (PINN) to approximate solutions to one-dimensional boundary value problems for linear elliptic equations and establish robust error estimates of PINN regardless of the quantities of the coefficients. In particular, we rigorously demonstrate the existence and uniqueness of solutions using the Sobolev space theory based on a variational approach. Deriving  $L^2$ -contraction estimates, we show that the error, defined as the mean square of the differences between the true solution and our trial function at the sample points, is dominated by the training loss. Furthermore, we show that as the quantities of the coefficients for the differential equation increase, the error-to-loss ratio rapidly decreases. Our theoretical and experimental results confirm the robustness of the error regardless of the quantities of the coefficients.

**Keywords:** Sobolev spaces; boundary value problems; existence and uniqueness; physics-informed neural networks (PINN);  $L^2$ -contraction estimates; error estimates

**Mathematics Subject Classification:** Primary: 34B05, 35A15; Secondary: 68T07, 65L10

---

## 1. Introduction

### 1.1. Literature review

Neural network-based methods have recently gained significant attention for solving problems in mathematical sciences, particularly in numerical analysis. There has been active research for neural network-based methods to approximate solutions to various differential equations (see [25] and references therein). For instance, if a true solution to a differential equation is not explicitly known, one can train neural networks by minimizing a training loss associated with satisfying the differential equation, without relying on any actual data points that the solution fulfills. These neural networks

trained to satisfy differential equations are called physics-informed neural networks (PINN).

The basic idea for PINN may originate from [7, 13]. Particularly, in [7] one-layer neural network is used to approximate a solution to a special form of the two-dimensional Poisson equation with homogeneous boundary conditions. In [12], instead of using a neural network directly, the authors use a trial function which is a slight variation in neural networks forcing to satisfy the initial or boundary conditions. Indeed, it was discussed in [17, 23] that using a trial function forcing to satisfy the boundary condition is superior to using a standard neural network in terms of performance such as acceleration of convergence. Recently in [3, 18], PINN-based methods have been potent tools for various differential equations solvers with the advent of efficient utilities such as TensorFlow and the help of advanced GPU capability. In particular, [18] demonstrates that the solution of the two-dimensional Navier-Stokes equation and the Burgers equation can be approximated very effectively using PINN.

In addition to the success of empirical analysis in PINN, there has been various theoretical research to show the convergence of neural networks to a solution for each case of differential equations. One of the most important ingredients of using neural networks to approximate solutions to differential equations is the universal approximation theorem (see [4, 9, 21]). Although the universal approximation theorem mathematically guarantees that any continuous function on a compact interval can be uniformly approximated by neural networks, the theorem is crucially based on a functional analytic result, the Stone-Weierstrass theorem (see [20, Theorem 7.32]), making it difficult to find an approximating neural network explicitly. To overcome this limitation, one can use a posteriori analysis based on the stability estimates of PDEs. Indeed, one can show that the error related to the difference between the true solution and the neural network is controlled by the training loss which we desire to minimize. For this type of research, we refer to [1] which presents theoretical aspects of error estimates for the Navier-Stokes equation through stability estimates based on a functional analytic approach. We also refer to [22, 24] which importantly use the Schauder estimates and  $H^2$ -estimates for their error analysis, respectively. We mention [15, 16], which provide error estimates in PINN in the context of inverse and forward problems, respectively, within an abstract framework. Additionally, we address [5, 6], which rigorously derive error estimates in PINN, with explicit convergence rates for the Kolmogorov equations and the Navier-Stokes equations, respectively. More recently, [26] shows error estimates in PINN for initial value problems for ordinary differential equations, and [11] develops a new error analysis in PINN based on classical finite element approximation. In [27], a new abstract framework unifying forward and inverse problems in PINN is developed, and  $H^{1/2}$ -error estimates are obtained by utilizing C ea's lemma and the coercivity of bilinear forms. In [1, 15–17, 22, 24, 27], while the error estimates in PINN show that the error decreases as the loss decreases, the lack of explicit calculation of the constants in these estimates makes it challenging to determine precise upper bounds for the error-to-loss ratio.

## 1.2. Idea of the error estimates in PINN

Before explaining the main results in this paper, we first consider the following Dirichlet problem related to a general linear differential operator  $\mathcal{L}$  on a bounded open subset  $U$  of  $\mathbb{R}^d$  with sufficiently regular boundary  $\partial U$ :

$$\begin{cases} \mathcal{L}u = f & \text{in } U, \\ u = g & \text{on } \partial U, \end{cases} \quad (1.1)$$

where  $u$ ,  $f$  and  $g$  are included in some sufficiently regular function spaces. Assume that for each  $f \in L^2(U)$  and  $g \in L^2(\partial U)$  there exists a unique solution  $u$  to (1.1) such that the following  $L^2$ -error estimate holds:

$$\|u\|_{L^2(U)} \leq C_{\mathcal{L}} (\|f\|_{L^2(U)} + \|g\|_{L^2(\partial U)}),$$

where  $C_{\mathcal{L}} > 0$  is a constant independent of  $u$ ,  $f$  and  $g$ . Now let  $\Phi$  be a smooth trial function (for instance, one can choose  $\Phi$  as a neural network). Then, replacing the above  $u$ ,  $f$  and  $g$  by  $\Phi - u$ ,  $\mathcal{L}[\Phi] - f$  and  $\Phi - g$ , respectively, we get

$$\|\Phi - u\|_{L^2(U)} \leq C_{\mathcal{L}} (\|\mathcal{L}[\Phi] - f\|_{L^2(U)} + \|\Phi - g\|_{L^2(\partial U)}). \quad (1.2)$$

Then, as a consequence of Monte Carlo integration (see Proposition 2.3), we expect to obtain that for sufficiently large  $n \in \mathbb{N}$

$$\underbrace{\frac{1}{n} \sum_{i=1}^n |u(X_i) - \Phi(X_i)|^2}_{=: \text{Error}} \leq 2C_{\mathcal{L}}^2 \underbrace{\left( \frac{1}{n} \sum_{i=1}^n |\mathcal{L}[\Phi](X_i) - f(X_i)|^2 + \frac{1}{n} \sum_{i=1}^n |\Phi(Y_i) - g(Y_i)|^2 \right)}_{=: \text{Loss}}, \quad \text{very likely,}$$

where  $(X_i)_{i \geq 1}$  and  $(Y_i)_{i \geq 1}$  are sequences of independent and identically distributed random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  that have continuous uniform distributions on  $U$  and  $\partial U$ , respectively. Ultimately, we can expect a reduction of Error by attempting to minimize Loss.

Here one needs to observe the behavior of constant value  $C_{\mathcal{L}}$ . As the quantities of the coefficients of  $\mathcal{L}$  increase, one may expect that  $C_{\mathcal{L}}$  or Loss also increases. In that case, discrepancies in boundary values between the neural networks and the true solutions are expected to remain. Indeed, we observe in our numerical experiment that the Boundary Loss is larger than Error, regardless of the coefficient (see Example 3.7). To overcome boundary mismatch, we can use a trial function  $\Psi$  ( $\Psi = \Phi_A + \ell$  in our case, see Section 3.2) that exactly satisfies the boundary value and can obtain improved error analysis (cf. [17]). It has been difficult to find the existing literature that explicitly calculates the constant  $C_{\mathcal{L}}$ . In many cases, the  $L^2$ -error estimates are based on functional analytic results, making it difficult to explicitly compute the constant  $C_{\mathcal{L}}$  though its existence is guaranteed (cf. [27]). However, our main results, which will be explained in the next section, provide a concrete value for the constant  $C_{\mathcal{L}}$ .

### 1.3. Main results

Now let us deal with our main problem

$$\begin{cases} -\varepsilon \bar{y}'' + b\bar{y}' + c\bar{y} = f & \text{on } I = (x_1, x_2), \\ \bar{y}(x_1) = p, \quad \bar{y}(x_2) = q, \end{cases} \quad (1.3)$$

where  $\varepsilon > 0$ ,  $p, q \in \mathbb{R}$  and  $f, b, c \in C(\bar{I})$  with  $c \geq \lambda$  on  $I$  for some constant  $\lambda > 0$ . We first show rigorously the existence and uniqueness of a weak, strong, and classical solution  $\bar{y}$  to (1.3) by using the Sobolev space theory with a variational approach. Under the assumption above, we can obtain some error estimates (4.2) similarly to (1.2) (see Theorem 4.4), but the constant value  $C_{\mathcal{L}}$  highly depends

on  $\varepsilon > 0$  and the coefficients  $b$ , in fact,  $C_{\mathcal{L}} = \frac{1}{\lambda} \exp\left(\int_I \frac{1}{2\varepsilon} |b| dx\right)$ . On the other hand, if we additionally assume that for some constant  $\gamma > 0$

$$-\frac{1}{2}b' + c \geq \gamma \quad \text{in } I$$

and  $\Psi$  is a smooth function with  $\Psi(x_1) = p$ ,  $\Psi(x_2) = q$ , our main result is the following error estimate (see (5.3) in Section 5.2): for sufficiently large  $n \in \mathbb{N}$

$$\underbrace{\frac{1}{n} \sum_{i=1}^n |\bar{y}(X_i) - \Psi(X_i)|^2}_{=: \text{Error}} \leq \frac{1}{\gamma^2} \cdot \underbrace{\frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Psi](X_i))^2}_{=: \text{Loss}}, \quad \text{very likely.} \quad (1.4)$$

Indeed, (1.4) is an expression in the Monte Carlo integration for the below (see Proposition 2.3 and Theorem 5.5):

$$\left( \int_I |\bar{y} - \Psi|^2 dx \right)^{1/2} \leq C_{\mathcal{L}} \left( \int_I |\tilde{f} - \mathcal{L}[\Psi]|^2 dx \right)^{1/2},$$

where  $C_{\mathcal{L}} = \frac{1}{\gamma}$ . From the above estimates, we see that the trial function  $\Psi$  is not directly related to a neural network but needs to be a smooth function satisfying the boundary condition,  $\Psi(x_1) = p$  and  $\Psi(x_2) = q$ . However, for our a posteriori analysis, to make Loss practically small we will consider the form

$$\Psi(x) := -(x - x_1)(x - x_2)N(x; \theta) + \frac{q - p}{x_2 - x_1}(x - x_1) + p, \quad x \in I,$$

where  $N(x; \theta)$  is a neural network. The main observation in this paper is that  $\frac{\text{Error}}{\text{Loss}}$  dramatically decreases with the upper bound  $\frac{1}{\gamma^2}$  as  $\gamma$  increases (see Examples 4.5, 5.6, 5.7). For instance, as  $\gamma$  increases, Loss becomes large, so it is natural to imagine that Error may also increase. Actually, we can observe that as  $\gamma$  increases, Loss tends to increase, but Error remains robust since  $\frac{\text{Error}}{\text{Loss}}$ , with the upper bound  $\frac{1}{\gamma^2}$ , rapidly decreases. These observations are mathematically justified by our results where the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  is shown to be  $\frac{1}{\gamma^2}$ .

Another interesting feature in the error estimate (1.4) is that  $\frac{1}{\gamma^2}$  does not depend on  $\varepsilon > 0$ . If  $\varepsilon > 0$  is very small, then (1.3) is called a singularly perturbed convection-diffusion problem, which is known to be quite difficult to solve numerically (see [8, 10, 19] and references therein). However, in our estimates (1.4),  $\frac{\text{Error}}{\text{Loss}} (\leq \frac{1}{\gamma^2})$  is robust independent of  $\varepsilon$ , and Loss does not increase as  $\varepsilon$  decreases to 0. Thus, if  $\gamma$  is fixed, we observe that Error also would remain robust independent of  $\varepsilon$ . These expectations are specifically confirmed in our numerical experiments (see Examples 5.6, 5.7).

#### 1.4. Structure of this paper

This paper is structured as follows. Section 2 presents notations and conventions mainly used in this paper, especially basic results related to one-dimensional Sobolev spaces and probability theory. Section 3 derives energy estimates through a variational approach and rigorously shows the existence and uniqueness of solutions. In addition, we will conduct experiments using a standard neural network (PINN I) and a trial function that accurately satisfies the boundary conditions (PINN II) and will confirm that (PINN II) demonstrates superior performance. We will also observe in Example 3.8 that the energy estimate is not optimal in explaining the rapid decrease of the error-to-loss ratio as the quantity of the

zero-order term increases. In Section 4,  $L^2(I, \mu)$ -contraction estimates and related error estimates are derived, and in the subsequent numerical experiment (Example 4.2), we will analyze the error-to-loss ratio for three  $\varepsilon$ . In Section 5, we present our main error estimates (5.3) which improve the error estimates (4.2) by using  $L^2(I, dx)$ -contraction estimates. The two subsequent numerical experiments (Examples 5.6, 5.7) will effectively support our mathematical analysis. Section 6 briefly demonstrates the conclusions and discussions of this paper.

## 2. Notations and basic results

Here we explain some notations and basic results in this paper. Write  $a \vee b := \max(a, b)$ ,  $a \wedge b := \min(a, b)$ . Let  $I$  be a (possibly unbounded) open interval on  $\mathbb{R}$ . Let  $C(I)$  and  $C(\bar{I})$  denote the sets of all continuous functions on  $I$  and  $\bar{I}$ , respectively, and we write  $C^0(I) := C(I)$  and  $C^0(\bar{I}) := C(\bar{I})$ . For each  $k \in \mathbb{N} \cup \{\infty\}$ , denote by  $C^k(I)$  the set of all  $k$ -times continuously differentiable functions on  $I$ . Now let  $k \in \mathbb{N} \cup \{0, \infty\}$  and define  $C_0^k(I) := \{f \in C^k(I) : f \text{ has a compact support in } I\}$ , and write  $C_0(I) := C_0^0(I)$ . By the zero-extension on  $\mathbb{R}$ , it directly follows that  $C_0^k(I)$  is naturally extended to  $C_0^k(\mathbb{R})$ . Define for each  $k \in \mathbb{N} \cup \{\infty\}$

$$C^k(\bar{I}) := \{f \in C(\bar{I}) : \text{there exists a function } \tilde{f} \in C^k(\mathbb{R}) \text{ such that } \tilde{f}|_I = f \text{ on } I\}.$$

Let  $\nu$  be a locally finite measure on  $\mathbb{R}$ . Denote by  $dx$  the Lebesgue measure on  $\mathbb{R}$  and  $dx(I)$  by  $|I|$ . For each  $r \in [1, \infty)$ , denote by  $L^r(I, \nu)$  the space of all measurable functions  $f$  on  $I$  with  $\int_I |f|^r d\nu < \infty$  equipped with the norm  $\|f\|_{L^r(I, \nu)} := \left(\int_I |f|^r d\nu\right)^{1/r}$ . Denote by  $L^\infty(I, \nu)$  the space of all  $\nu$ -a.e. bounded measurable functions  $f$  on  $I$  equipped with the norm

$$\|f\|_{L^\infty(I, \nu)} := \inf\{c > 0 : \nu(\{|f| > c\}) = 0\}.$$

Write  $L^p(I) := L^p(I, dx)$  for each  $p \in [1, \infty]$ . For a locally integrable function  $u$  on  $I$ , if there exists a locally integrable function  $g$  such that  $\int_I u\varphi dx = -\int_I g\varphi dx$  for all  $\varphi \in C_0^1(I)$ , then we write  $u' = g$  and we call  $u'$  a weak derivative of  $u$  on  $I$ . Let  $p \in [1, \infty]$ . The Sobolev space  $H^{1,p}(I)$  is defined to be

$$H^{1,p}(I) = \{u \in L^p(I) : \text{there exists } g \in L^p(I) \text{ such that } u' = g \text{ on } I\},$$

equipped with the norm  $\|u\|_{H^{1,p}(I)} := \left(\|u\|_{L^p(I)}^p + \|u'\|_{L^p(I)}^p\right)^{1/p}$  if  $p \in [1, \infty)$  and  $\|u\|_{H^{1,\infty}(I)} := \|u\|_{L^\infty(I)} + \|u'\|_{L^\infty(I)}$ . Of course, one can also use the notation  $W^{1,p}(I)$  for  $H^{1,p}(I)$ , but here we stick to use the notation  $H^{1,p}(I)$ . We also define

$$H^{2,p}(I) := \left\{u \in H^{1,p}(I) : \text{there exists } h \in L^p(I) \text{ such that } (u')' = h \text{ on } I\right\}.$$

equipped with the norm  $\|u\|_{H^{2,p}(I)} := \left(\|u\|_{L^p(I)}^p + \|u'\|_{L^p(I)}^p + \|u''\|_{L^p(I)}^p\right)^{1/p}$  if  $p \in [1, \infty)$  and  $\|u\|_{H^{2,\infty}(I)} := \|u\|_{L^\infty(I)} + \|u'\|_{L^\infty(I)} + \|u''\|_{L^\infty(I)}$ . An important fact is that  $H^{1,p}(I)$  and  $H^{2,p}(I)$  are Banach spaces (see Proposition 8.1). For  $r \in [1, \infty)$ , denote by  $H_0^{1,r}(I)$  the closure of  $C_0^1(I)$  in  $H^{1,r}(I)$ . Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and  $X : \Omega \rightarrow \mathbb{R}$  be a random variable. Define

$$\mathbb{E}[X] := \int_{\Omega} X d\mathbb{P}, \quad \text{Var}[X] := \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

The following well-known results presented in [2, Theorem 8.2] are important in our analysis which describes a connection between weak and classical derivatives. Here we leave the statement for readers.

**Proposition 2.1.** Let  $u \in H^{1,p}(I)$ , where  $p \in [1, \infty]$  and  $I$  is a (possibly unbounded) open interval. Then, the following hold:

(i) There exists  $\tilde{u} \in C(\bar{I})$  such that  $\tilde{u} = u$  a.e. on  $I$  and that

$$\tilde{u}(y) - \tilde{u}(x) = \int_x^y u'(t)dt, \quad \text{for all } x, y \in \bar{I}.$$

(ii) If there exists  $v \in C(\bar{I})$  such that  $u' = v$  a.e. on  $I$ , then  $\tilde{u}$  as in (i) satisfies  $\tilde{u} \in C^1(\bar{I})$  and that

$$\lim_{h \rightarrow 0} \frac{\tilde{u}(x+h) - \tilde{u}(x)}{h} = v(x), \quad \text{for all } x \in \bar{I}.$$

(If  $x \in \partial\bar{I}$ , then the limit above is defined as the one-side limit). In other words, the classical derivative of  $\tilde{u}$  is the same as the weak derivative of  $u$  on  $I$ .

By the proposition above, every function  $u$  in the Sobolev space on an open interval  $I$  has a continuous version  $\tilde{u}$  on  $\bar{I}$ . Therefore, to get a meaning of  $u(x)$  for every point  $x \in \bar{I}$ , we will consider  $u$  as its (unique) continuous version  $\tilde{u}$  on  $\bar{I}$ . The following is also a well-known result ([2, Theorem 8.12]) that characterizes the space  $H_0^{1,r}(I)$

**Proposition 2.2.** Let  $\tilde{u} \in H^{1,r}(I) \cap C(\bar{I})$  with  $r \in [1, \infty)$ . Then,  $\tilde{u} \in H_0^{1,r}(I)$  if and only if  $\tilde{u} = 0$  on  $\partial\bar{I}$ .

Below are the basic probabilistic facts about well-known Monte Carlo integration, which will be used in our numerical analysis. We explain it by using the weak law of large numbers. For readers, we present the statement and its proofs.

**Proposition 2.3.** Let  $(X_i)_{i \geq 1}$  be a sequence of independent and identically distributed random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  that has a continuous uniform distribution on  $I$ . Let  $\Psi \in C(\bar{I})$  with  $\Psi \geq 0$  on  $\bar{I}$ . Then, for each  $n, k \in \mathbb{N}$

$$\mathbb{P} \left( \left| \frac{1}{|I|} \int_I \Psi dx - \frac{1}{n} \sum_{i=1}^n \Psi(X_i) \right| \leq \frac{k\beta}{\sqrt{n}} \right) \geq 1 - \frac{1}{k^2},$$

where  $\beta := \left( \frac{1}{|I|} \int_I \Psi^2 dx - \left( \frac{1}{|I|} \int_I \Psi dx \right)^2 \right)^{1/2}$ .

*Proof.* For each  $i \geq 1$ , we get  $\mathbb{E}[\Psi(X_i)] = \frac{1}{|I|} \int_I \Psi dx =: \alpha$  and

$$\text{Var}[\Psi(X_i)] = \mathbb{E}[\Psi(X_i)^2] - \mathbb{E}[\Psi(X_i)]^2 = \frac{1}{|I|} \int_I \Psi^2 dx - \left( \frac{1}{|I|} \int_I \Psi dx \right)^2 = \beta^2.$$

Define

$$\bar{Y}_n := \frac{1}{n} \sum_{i=1}^n \Psi(X_i).$$

Note that  $\mathbb{E}[\bar{Y}_n] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\Psi(X_i)] = \alpha$ . Since  $(\Psi(X_i))_{i \geq 1}$  is also a sequence of independent and identically distributed random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , we get

$$\text{Var}[\bar{Y}_n] = \frac{1}{n^2} \sum_{i=1}^n \text{Var}[\Psi(X_i)] = \frac{\beta^2}{n}.$$

Using Chebyshev's inequality, for each  $c > 0$  we get

$$\mathbb{P}(|\bar{Y}_n - \alpha| \geq c) \leq \frac{\beta^2}{c^2 n}.$$

By choosing  $c := \frac{k\beta}{\sqrt{n}}$  and considering the complement event, the assertion follows.  $\square$

Based on Proposition 2.3, for sufficiently large but fixed  $k \in \mathbb{N}$  if we choose  $n$  much larger than  $k$ , then we can write that

$$\mathbb{P}\left(\frac{1}{|I|} \int_I \Psi dx \approx \frac{1}{n} \sum_{i=1}^n \Psi(X_i)\right) \approx 1. \quad (2.1)$$

If (2.1) holds, then we will use the following notation that for sufficiently large  $n \in \mathbb{N}$

$$\frac{1}{|I|} \int_I \Psi dx = \frac{1}{n} \sum_{i=1}^n \Psi(X_i), \quad \text{very likely.} \quad (2.2)$$

### 3. Energy estimates

#### 3.1. Existence and uniqueness with energy estimates

**Lemma 3.1.** *Let  $u \in H_0^{1,1}(I)$  where  $I = (x_1, x_2)$  is a bounded open interval. Then, there exists  $\tilde{u} \in C_0(\mathbb{R}) \cap H_0^{1,1}(\mathbb{R})$  with  $\tilde{u} = 0$  on  $\mathbb{R} \setminus I$  such that  $\tilde{u} = u$  a.e. on  $I$  and*

$$\|u\|_{L^\infty(I)} = \|\tilde{u}\|_{L^\infty(\mathbb{R})} \leq \|u'\|_{L^1(I)}.$$

In particular, if  $u \in H^{1,2}(I)$ , then

$$|I|^{-1/2} \|u\|_{L^2(I)} \leq \|u\|_{L^\infty(I)} \leq |I|^{1/2} \|u'\|_{L^2(I)}.$$

*Proof.* Since  $u \in H_0^{1,1}(I)$ , there exists a sequence of functions  $(u_n)_{n \geq 1} \subset C_0^\infty(I)$  such that

$$\lim_{n \rightarrow \infty} u_n = u, \quad \text{in } H^{1,1}(I).$$

For each  $n \geq 1$ , let  $\bar{u}_n \in C_0^\infty(\mathbb{R})$  be the standard zero-extension of  $u_n$  on  $\mathbb{R}$ . Thus, there exists  $\bar{u} \in H_0^{1,1}(\mathbb{R})$  such that  $\bar{u}(x) = 0$  for a.e.  $x \in \mathbb{R} \setminus I$  and  $\lim_{n \rightarrow \infty} \bar{u}_n = \bar{u}$  in  $H^{1,1}(\mathbb{R})$ , and hence  $\bar{u} = u$  a.e. on  $I$ . Meanwhile, by Proposition 2.1(i) there exists  $\tilde{u} \in C(\mathbb{R})$  such that  $\tilde{u} = \bar{u}$  a.e. on  $\mathbb{R}$ . Thus,  $\tilde{u} \in C(\mathbb{R}) \cap H^{1,1}(\mathbb{R})$  satisfying that  $\tilde{u}(x) = 0$  for all  $x \in \mathbb{R} \setminus I$  and  $\tilde{u} = u$  a.e. on  $I$ . By the fundamental theorem of calculus (Proposition 2.1(i)), for each  $x \in I$

$$|\tilde{u}(x)| = \left| \tilde{u}(x_1) + \int_{x_1}^x \tilde{u}'(t) dt \right| = \left| \int_{x_1}^x \tilde{u}'(t) dt \right| \leq \int_I |\tilde{u}'(t)| dt = \|u'\|_{L^1(I)},$$

as desired. The rest follows from the Hölder inequality.  $\square$

We are now interested in solving the following Dirichlet boundary value problem for a one-dimensional second-order linear elliptic operator on  $I = (x_1, x_2)$ :

$$\begin{cases} -\varepsilon \tilde{y}'' + b\tilde{y}' + c\tilde{y} = \tilde{f} & \text{on } I, \\ \tilde{y}(x_1) = p, \quad \tilde{y}(x_2) = q, \end{cases} \quad (3.1)$$

where  $p, q \in \mathbb{R}$  and  $\varepsilon > 0$  are constants,  $b \in L^1(I)$  and  $c \in L^1(I)$  with  $c \geq 0$ . By using the linearity and considering an affine function defined by  $\ell(x) = \frac{p-q}{x_2-x_1}(x-x_1) + p$  and replacing  $\tilde{y}$  by  $y = \tilde{y} - \ell$  in (3.1), it is enough to study the existence and uniqueness of solutions to the following homogeneous boundary value problem:

$$\begin{cases} -\varepsilon y'' + by' + cy = \tilde{f} - b\ell' - c\ell & \text{on } I, \\ y(x_1) = 0, \quad y(x_2) = 0. \end{cases} \quad (3.2)$$

The main idea for solving (3.2) is to convert the non-divergence form of (3.2) to the symmetric divergence form and to apply the Lax-Milgram theorem on the Sobolev space which is  $H_0^{1,2}(I)$ .

**Theorem 3.2.** Let  $I = (x_1, x_2)$  be a bounded open interval in  $\mathbb{R}$  and let  $\varepsilon > 0$  be a constant,  $b \in L^1(I)$ ,  $c \in L^1(I)$  with  $c \geq 0$  and  $f \in L^1(I)$ . Then, the following hold:

(i) Let

$$\rho(x) := \exp\left(\int_{x_1}^x \frac{-b(s)}{\varepsilon} ds\right), \quad x \in \mathbb{R}, \quad (3.3)$$

where  $b$  is considered as the zero-extension of  $b \in L^1(I)$  on  $\mathbb{R}$ . Then, there exists a unique weak solution  $y \in H_0^{1,2}(I) \cap C(\bar{I})$  to

$$-(\varepsilon \rho y')' + \rho c y' = \rho f \quad \text{on } I, \quad (3.4)$$

i.e.,

$$\int_I \varepsilon y' \psi' \rho dx + \int_I c y \psi \rho dx = \int_I f \psi \rho dx, \quad \text{for all } \psi \in H_0^{1,2}(I). \quad (3.5)$$

(ii) Indeed,  $y \in H_0^{1,2}(I) \cap C(\bar{I})$  in (i) satisfies  $y \in H^{2,1}(I) \cap C^1(\bar{I})$  and  $y$  is a unique strong solution to

$$\begin{cases} -\varepsilon y'' + by' + cy = f & \text{on } I, \\ y(x_1) = 0, \quad y(x_2) = 0, \end{cases} \quad (3.6)$$

i.e.,

$$(-\varepsilon y'' + by' + cy)(x) = f(x), \quad \text{for a.e. } x \in I \quad \text{and} \quad y(x_1) = y(x_2) = 0. \quad (3.7)$$

(iii) If  $b, c, f \in C(\bar{I})$ , then  $y$  in (ii) satisfies  $y \in C^2(\bar{I})$  and  $y$  is a unique classical solution to (3.6), i.e.,

$$(-\varepsilon y'' + by' + cy)(x) = f(x), \quad \text{for all } x \in I \quad \text{and} \quad y(x_1) = y(x_2) = 0.$$



*Proof.* (i) Observe that by [2, Corollary 8.11],  $\rho \in H^{1,1}(I) \cap C(\bar{I})$  and that  $\rho(x) > 0$  for all  $x \in \bar{I}$ . Define a bilinear form  $\mathcal{E}$  on  $H_0^{1,2}(I) \times H_0^{1,2}(I)$  by

$$\mathcal{E}(u, v) := \int_I \varepsilon u' v' d\mu + \int_I cuv d\mu, \quad u, v \in H_0^{1,2}(I),$$

where  $\mu = \rho dx$ . Then, we obtain from Lemma 3.1 that for each  $u, v \in H_0^{1,2}(I)$

$$\begin{aligned} |\mathcal{E}(u, v)| &\leq \varepsilon \left( \int_I |u'|^2 d\mu \right)^{1/2} \left( \int_I |v'|^2 d\mu \right)^{1/2} + \|\rho c\|_{L^1(I)} \|u\|_{L^\infty(I)} \|v\|_{L^\infty(I)} \\ &\leq \varepsilon \left( \int_I |u'|^2 d\mu \right)^{1/2} \left( \int_I |v'|^2 d\mu \right)^{1/2} + |I| \cdot \|\rho c\|_{L^1(I)} \left( \int_I |u'|^2 dx \right)^{1/2} \left( \int_I |v'|^2 dx \right)^{1/2} \\ &\leq \underbrace{(\varepsilon + |I| \cdot \|c\|_{L^1(I)} \left( \max_{\bar{I}} \rho \right))}_{=:K} \left( \int_I |u'|^2 dx \right)^{1/2} \left( \int_I |v'|^2 dx \right)^{1/2} \\ &\leq K \|u\|_{H^{1,2}(I)} \|v\|_{H^{1,2}(I)}. \end{aligned}$$

Moreover,

$$\begin{aligned} \mathcal{E}(u, u) &\geq \varepsilon \int_I |u'|^2 d\mu \geq \varepsilon \left( \min_{\bar{I}} \rho \right) \frac{1}{2} \int_I |u'|^2 dx + \varepsilon \left( \min_{\bar{I}} \rho \right) \frac{1}{2} |I|^{-1} \|u\|_{L^\infty(I)}^2 \\ &\geq \varepsilon \left( \min_{\bar{I}} \rho \right) \frac{1}{2} \int_I |u'|^2 dx + \varepsilon \left( \min_{\bar{I}} \rho \right) \frac{1}{2} |I|^{-2} \int_I |u|^2 dx \\ &\geq \delta \|u\|_{H_0^{1,2}(I)}^2, \end{aligned}$$

where  $\delta = \frac{\varepsilon}{2} \min_{\bar{I}} \rho \cdot (1 \wedge |I|^{-2}) > 0$ . By the Lax-Milgram theorem (see [2, Corollary 5.8]), there exists a unique  $y \in H_0^{1,2}(I)$  such that

$$\mathcal{E}(y, \psi) = \int_I f \psi d\mu, \quad \text{for all } \psi \in H_0^{1,2}(I),$$

and hence (3.5) follows.

(ii) Observe that  $y' \in H^{1,1}(I) \cap C(\bar{I})$ . Indeed, (3.5) yields that  $\rho y' \in H^{1,1}(I) \cap C(\bar{I})$ . Since  $\frac{1}{\rho} \in H^{1,1}(I) \cap C(\bar{I})$ , we have  $y' \in H^{1,1}(I) \cap C(\bar{I})$  by the product rule, so that  $y \in C^1(\bar{I})$  by Proposition 2.1(ii). Moreover, (3.5) is equivalent to

$$\int_I (-\varepsilon y'' + by' + cy) \psi \rho dx = \int_I \left( -\varepsilon y'' - \frac{\varepsilon \rho'}{\rho} y' + cy \right) \psi \rho dx = \int_I f \psi \rho dx, \quad \text{for all } \psi \in H_0^{1,2}(I),$$

and hence (3.7) holds. Conversely, as the above equivalence if  $\bar{y}$  is a strong solution to (3.6), then  $\bar{y}$  is a weak solution to (3.4). Thus, the uniqueness of weak solutions as in (i) implies that  $\bar{y} = y$  on  $I$ , and hence the uniqueness of strong solutions to (3.6) is shown.

(iii) Assume that  $b, c, f \in C(\bar{I})$ . Then,  $\varepsilon y'' = by' + cy - f$  on  $I$ . Since  $by' + cy - f \in C(\bar{I})$ , we have  $y'' \in C(\bar{I})$  by Proposition 2.1(ii). Therefore,  $y \in C^2(\bar{I})$  is a unique classical solution to (3.6).  $\square$

**Theorem 3.3** (Energy estimates). *Let  $\varepsilon > 0$  be a constant,  $b \in L^1(I)$ ,  $c \in L^1(I)$  with  $c \geq 0$  on  $I$  and  $f \in L^1(I)$ , where  $I$  is a bounded open interval. Let  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  be a unique strong solution to (3.7) as in Theorem 3.2(ii). Then, it holds that*

$$\|y'\|_{L^2(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^{1/2} \|f\|_{L^1(I)}, \quad (3.8)$$

where  $\rho$  is the function defined in (3.3). In particular, if  $f \in L^2(I)$ , then

$$\begin{aligned} \|y\|_{L^2(I)} &\leq |I|^{1/2} \|y\|_{L^\infty(I)} \leq |I| \|y'\|_{L^2(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^{3/2} \|f\|_{L^1(I)} \\ &\leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^2 \|f\|_{L^2(I)}. \end{aligned} \quad (3.9)$$

*Proof.* Substituting  $y$  for  $\psi$  in (3.5) and using Lemma 3.1 and the Hölder inequality, we get

$$\begin{aligned} \varepsilon \left( \min_{\bar{I}} \rho \right) \int_I |y'|^2 dx &\leq \varepsilon \int_I |y'|^2 \rho dx \leq \varepsilon \int_I |y'|^2 \rho dx + \int_I c y^2 \rho dx = \int_I f y \rho dx \\ &\leq \left( \max_{\bar{I}} \rho \right) \|y\|_{L^\infty(I)} \|f\|_{L^1(I)} \leq \left( \max_{\bar{I}} \rho \right) |I|^{1/2} \left( \int_I |y'|^2 dx \right)^{1/2} \|f\|_{L^1(I)}, \end{aligned}$$

and hence (3.8) follows. The rest follows by Lemma 3.1 and the Hölder inequality.  $\square$

**Theorem 3.4.** *Let  $\varepsilon > 0$  be a constant,  $b \in L^1(I)$ ,  $c \in L^1(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^1(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $p, q \in \mathbb{R}$  be given. Then, there exists a unique strong solution  $\tilde{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  to the following boundary value problem*

$$\begin{cases} -\varepsilon \tilde{y}'' + b \tilde{y}' + c \tilde{y} = \tilde{f} & \text{on } I, \\ \tilde{y}(x_1) = p, \quad \tilde{y}(x_2) = q, \end{cases} \quad (3.10)$$

i.e.,

$$(-\varepsilon \tilde{y}'' + b \tilde{y}' + c \tilde{y})(x) = \tilde{f}(x), \quad \text{for a.e. } x \in I \quad \text{and} \quad \tilde{y}(x_1) = p, \quad \tilde{y}(x_2) = q.$$

In particular, if  $\tilde{f} \in L^2(I)$ , then  $\tilde{y}$  fulfills the following estimate:

$$\|\tilde{y}\|_{L^2(I)} \leq K_1 \|\tilde{f}\|_{L^2(I)} + K_2 (|p| \vee |q|),$$

where  $K_1 = \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^2$  and  $K_2 = \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) (|I|^{1/2} \|b\|_{L^1(I)} + |I|^{3/2} \|c\|_{L^1(I)}) + |I|^{1/2}$  and  $\rho$  is the function defined in (3.3). Moreover, if  $b, c \in C(\bar{I})$ , then  $\tilde{y} \in C^2(\bar{I})$  and it is a unique classical solution to (3.10).

*Proof.* Let  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1)+p$ ,  $x \in \mathbb{R}$ . Then, it is easy to check that  $\|\ell\|_{L^\infty(I)} \leq |p| \vee |q|$  and  $\|\ell'\|_{L^\infty(I)} = |I|^{-1} |p-q|$ . Then,  $\tilde{y} \in H^{2,1}(I) \cap C(\bar{I})$  is a strong solution to (3.10) if and only if  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C(\bar{I})$  is a strong solution to

$$\begin{cases} -\varepsilon y'' + b y' + c y = \tilde{f} - b \ell' - c \ell & \text{on } I, \\ y(x_1) = 0, \quad y(x_2) = 0, \end{cases} \quad (3.11)$$

where  $y = \bar{y} - \ell$ . Indeed, by Theorem 3.2 there exists a unique strong solution  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  to (3.11). Define  $\bar{y} := y + \ell$ . Then,  $\bar{y}$  is a unique strong solution  $\bar{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  to (3.10). Moreover, (3.9) in Theorem 3.3 implies that

$$|I|^{1/2} \|\bar{y} - \ell\|_{L^\infty(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^{3/2} \|\tilde{f} - b\ell' - c\ell\|_{L^1(I)}.$$

In particular, if  $\tilde{f} \in L^2(I)$ , then

$$\begin{aligned} \|\bar{y}\|_{L^2(I)} &\leq |I|^{1/2} \|\bar{y}\|_{L^\infty(I)} \leq |I|^{1/2} \|\bar{y} - \ell\|_{L^\infty(I)} + |I|^{1/2} \|\ell\|_{L^\infty(I)} \\ &\leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) \left( |I|^2 \|\tilde{f}\|_{L^2(I)} + |I|^{1/2} \|b\|_{L^1(I)} |q - p| + |I|^{3/2} \|c\|_{L^1(I)} (|p| \vee |q|) \right) + |I|^{1/2} (|p| \vee |q|) \\ &\leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^2 \cdot \|\tilde{f}\|_{L^2(I)} + \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) \left( |I|^{1/2} \|b\|_{L^1(I)} + |I|^{3/2} \|c\|_{L^1(I)} + |I|^{1/2} \right) (|p| \vee |q|), \end{aligned}$$

as desired.  $\square$

**Theorem 3.5.** Let  $\varepsilon > 0$  be a constant,  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $p, q \in \mathbb{R}$  be given and  $\bar{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi \in H^{2,2}(I) \cap C^1(\bar{I})$ . Then,

$$\|\bar{y} - \Phi\|_{L^2(I)} \leq K_1 \|\tilde{f} - \mathcal{L}[\Phi]\|_{L^2(I)} + K_2 \left( |p - \Phi(x_1)| \vee |q - \Phi(x_2)| \right),$$

where  $K_1, K_2 > 0$  are constants in Theorem 3.4 and  $\mathcal{L}$  is a differential operator defined by

$$\mathcal{L}[w] := -\varepsilon w'' + bw' + cw, \quad w \in H^{2,1}(I). \quad (3.12)$$

*Proof.* Let  $u := \bar{y} - \Phi$ . Then,  $u \in H^{2,2}(I) \cap C^1(\bar{I})$ . Since  $\tilde{f} - \mathcal{L}[\Phi] \in L^2(I)$ , it follows from Theorem 3.2 that  $u$  is a (unique) strong solution to the following problem:

$$\begin{cases} -\varepsilon u'' + bu' + cu = \tilde{f} - \mathcal{L}[\Phi] & \text{on } I, \\ u(x_1) = p - \Phi(x_1), \quad u(x_2) = q - \Phi(x_2). \end{cases}$$

By Theorem 3.4,

$$\|u\|_{L^2(I)} \leq K_1 \|\tilde{f} - \mathcal{L}[\Phi]\|_{L^2(I)} + K_2 \left( |p - \Phi(x_1)| \vee |q - \Phi(x_2)| \right),$$

as desired.  $\square$

**Theorem 3.6.** Let  $\varepsilon > 0$  be a constant,  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $p, q \in \mathbb{R}$  be given. Let  $\bar{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi_A \in H_0^{1,2}(I) \cap H^{2,2}(I) \cap C^1(\bar{I})$  and  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then,

$$\|\bar{y} - (\Phi_A + \ell)\|_{L^2(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) |I|^2 \|\tilde{f} - \mathcal{L}[\Phi_A + \ell]\|_{L^2(I)},$$

where  $\mathcal{L}$  is a differential operator defined by (3.12).

*Proof.* Since  $(\Phi_A + \ell)(x_1) = p$  and  $(\Phi_A + \ell)(x_2) = q$ , the assertion follows from Theorem 3.5 where  $\Phi$  there is replaced by  $\Phi_A + \ell$  here.  $\square$

### 3.2. Numerical experiments for PINN with energy estimates

Let  $\varepsilon > 0$  be a constant,  $b \in C(\bar{I})$ ,  $c \in C(\bar{I})$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in C(\bar{I})$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $p, q \in \mathbb{R}$  be given and  $\rho$  be the function defined in (3.3). Let  $\tilde{y} \in C^2(\bar{I})$  be the unique classical solution to (3.10) as in Theorem 3.4. Let  $(X_i)_{i \geq 1}$  be a sequence of independent and identically distributed random variables on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  that has a continuous uniform distribution on  $I$ . Let  $\Phi \in C^2(\bar{I})$ . Then, by using Theorem 3.5 with the Monte Carlo integration as in (2.2), we obtain that for sufficiently large  $n \in \mathbb{N}$

(PINN I):

$$\begin{aligned} \underbrace{\frac{1}{n} \sum_{i=1}^n |\tilde{y}(X_i) - \Phi(X_i)|^2}_{=: \text{Error}} &\leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right)^2 |I|^4 \cdot \left( \frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi](X_i))^2 \right) \\ &\quad + \left( \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) (\|b\|_{L^1(I)} + |I| \|c\|_{L^1(I)} + 1) \right)^2 \underbrace{\left( |p - \Phi(x_1)|^2 + |q - \Phi(x_2)|^2 \right)}_{=: \text{Boundary Loss}} \\ &\leq \left( \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right)^2 |I|^4 + \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right) (\|b\|_{L^1(I)} + |I| \|c\|_{L^1(I)} + 1) \right)^2 \times \\ &\quad \underbrace{\left( \frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi](X_i))^2 + |p - \Phi(x_1)|^2 + |q - \Phi(x_2)|^2 \right)}_{=: \text{Loss}}, \quad \text{very likely.} \end{aligned} \quad (3.13)$$

On the other hand, let  $\Phi_A \in H_0^{1,2}(I) \cap C^2(\bar{I})$  and  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then, by using Theorem 3.6 with the Monte Carlo integration as in (2.2), we find that for sufficiently large  $n \in \mathbb{N}$

(PINN II):

$$\underbrace{\frac{1}{n} \sum_{i=1}^n |\tilde{y}(X_i) - (\Phi_A + \ell)(X_i)|^2}_{=: \text{Error}} \leq \left( \frac{\max_{\bar{I}} \rho}{\varepsilon \min_{\bar{I}} \rho} \right)^2 |I|^4 \cdot \underbrace{\frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi_A + \ell](X_i))^2}_{=: \text{Loss}}, \quad \text{very likely.} \quad (3.14)$$

Let  $N(x; \theta)$  be a real-valued neural network defined on  $\mathbb{R}$  and let  $\chi(x) = -(x-x_1)(x-x_2)$ . In the following example, we will apply (PINN I) and (PINN II) where  $\Phi(x)$  and  $\Phi_A(x) + \ell(x)$  are replaced by  $N(x; \theta)$  and  $\chi(x)N(x; \theta) + \ell(x)$ , respectively. Both methods can reduce the error by reducing Loss, but (PINN I) has a limitation in that the discrepancy of the boundary value remains. In addition, the following example confirms that (PINN II) is superior to (PINN I) in terms of error reduction performance.

**Example 3.7.** Let  $I = (0, 1)$ . Let  $\varepsilon := 1$ ,  $c(x) = 10$ ,  $b(x) = x$  and  $\tilde{y}(x) = x(1-x)e^x$ ,  $x \in \mathbb{R}$ . Note that

$$\tilde{y}(x) = (1-x-x^2)e^x, \quad \tilde{y}'(x) = (-3x-x^2)e^x, \quad x \in \mathbb{R}.$$

Let  $\tilde{f}(x) = -\tilde{y}''(x) + b(x)\tilde{y}'(x) + c(x)\tilde{y}(x) = ((4+c)x - cx^2 - x^3)e^x$ ,  $x \in \mathbb{R}$ . Then,  $\tilde{y}$  is a unique strong solution to (3.10) with  $p = q = 0$ .

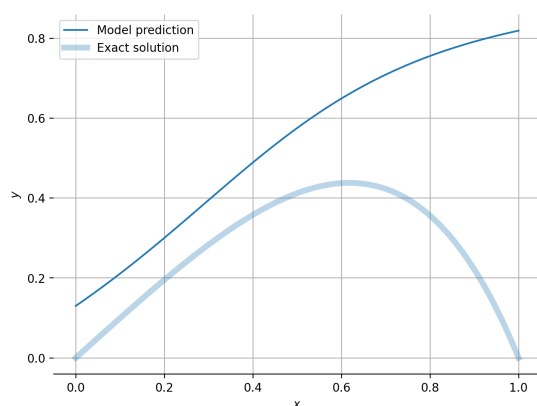
For the following experiment, a three-layer neural network with an architecture configured as (1, 16, 32, 1) was utilized. The hyperbolic tangent function ( $\tanh$ ) was used as the activation function throughout the network. The Adam optimizer was used with an initial learning rate of 0.1 and batch size of 10000. To optimize training, an exponential learning rate scheduler was attached, which methodically decreased the learning rate at each epoch to ensure convergence.

Figure 1 (a) describes the graph of the model prediction (neural network  $N(x; \theta)$ ) and the exact solution ( $\bar{y}(x)$ ) on  $I$  where the model prediction does not accurately capture the exact solution. Meanwhile, Figure 2 (a) describes the graph of the model prediction ( $x(1-x)N(x, \theta)$ ) and the exact solution ( $\bar{y}(x)$ ) for which the model prediction accurately captures the exact solution. Figure 1 (b) and Figure 2 (b) describe the graphs of Loss and Error along epoch increases, respectively, and we can observe that (PINN II) is more efficient than (PINN I) in reducing Error and Loss. Furthermore, Figure 3 describes that (PINN II) gives a more precise model prediction compared to (PINN I) in the perspective of Relative error. The Relative errors are defined as

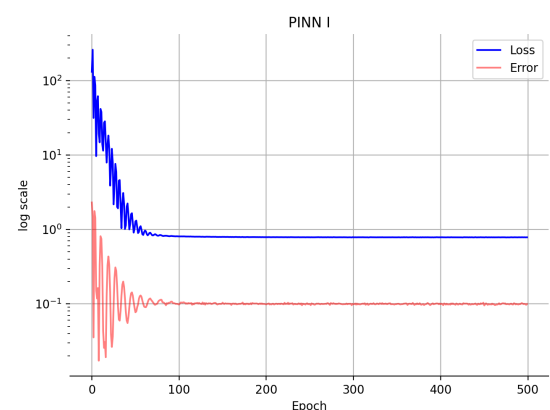
$$\text{Relative error (PINN I)} = \frac{\sum_{i=1}^n |\bar{y}(X_i) - \Phi(X_i)|^2}{\sum_{i=1}^n |\bar{y}(X_i)|^2},$$

$$\text{Relative error (PINN II)} = \frac{\sum_{i=1}^n |\bar{y}(X_i) - (\Phi_A + \ell)(X_i)|^2}{\sum_{i=1}^n |\bar{y}(X_i)|^2}.$$

In Example 3.7, we can observe that (PINN I) exhibits significant boundary discrepancies, where Boundary Loss is larger than Error. In contrast, (PINN II) accurately captures the exact solution. Therefore, from now on we exclusively use (PINN II) where the trial function  $\Phi_A + \ell$  exactly satisfies the boundary conditions of (3.10).

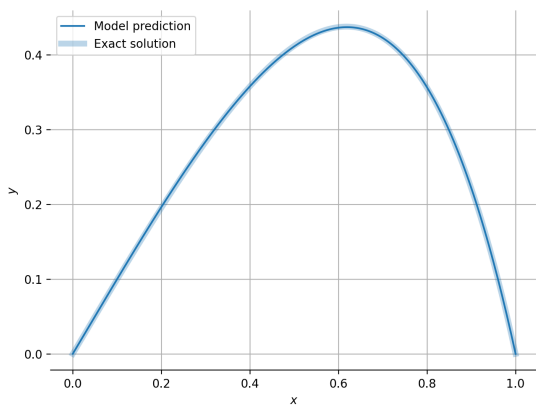


(a) Model prediction comparison with exact solution after 500 epochs (PINN I)

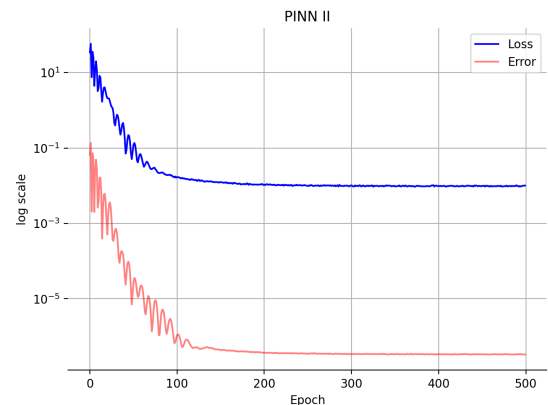


(b) The behaviors of Loss and Error as epoch increases (PINN I)

**Figure 1.** Not effective (PINN I).

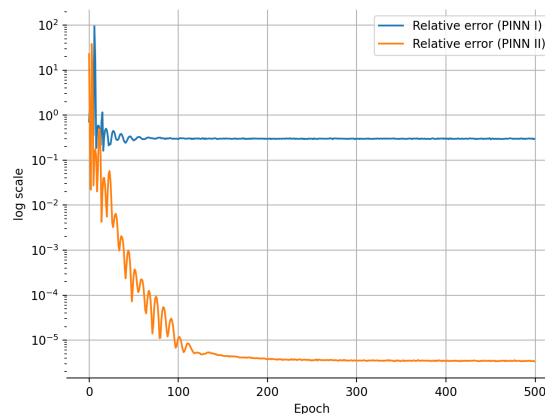


(a) Model prediction comparison with exact solution after 500 epochs (PINN II)



(b) The behaviors of Loss and Error as epoch increases (PINN II)

**Figure 2.** Successful (PINN II).



**Figure 3.** Comparison of Relative error between two methods as epoch increases.

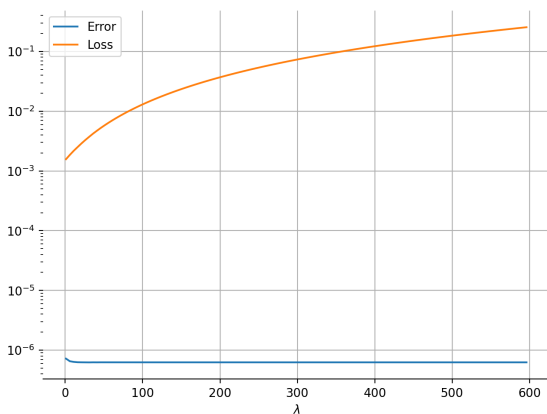
**Example 3.8.** Let  $I = (0, 1)$ . Let  $\varepsilon := 1$ ,  $c(x) = \lambda$ ,  $b(x) = \tan \frac{\pi x}{6}$ , and  $\tilde{y}(x) = \sin \frac{\pi x}{6}$ ,  $x \in \mathbb{R}$ . Note that

$$\tilde{y}'(x) = \frac{\pi}{6} \cos \frac{\pi x}{6}, \quad \tilde{y}''(x) = -\frac{\pi^2}{36} \sin \frac{\pi x}{6}, \quad x \in \mathbb{R}.$$

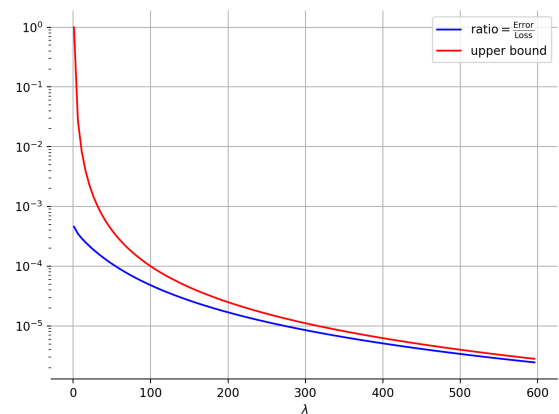
Let  $\tilde{f}(x) = -\tilde{y}''(x) + b(x)\tilde{y}'(x) + c(x)\tilde{y}(x) = \left(\frac{\pi^2}{36} + \frac{\pi}{6} + \lambda\right) \sin \frac{\pi x}{6}$ ,  $x \in \mathbb{R}$ . Then,  $\tilde{y}$  is a unique strong solution to (3.10) with  $p = 0$  and  $q = \frac{1}{2}$ .

For the following experiment, a three-layer neural network with an architecture configured as (1, 50, 50, 1) was utilized to describe a solution for a more complex differential equation. All experimental details are identical to those described in Example 3.7.

In Figure 4(a), increasing the zero-order coefficient  $\lambda$  leads to a more robust Error, although the Loss increases as well. We investigated this phenomenon and found that the ratio  $\frac{\text{Error}}{\text{Loss}}$  decreases, and this ratio is sharply bounded by the function  $\frac{1}{\lambda^2}$  in Figure 4(b). However, this observation is not explained by our energy estimates 3.14, as it does not account for zero-order terms. This discrepancy suggests the need for further investigation into this topic.



(a) Loss and Error as  $\lambda$  increases after 500 epochs as  $\lambda$  increases



(b)  $\frac{\text{Error}}{\text{Loss}}$  bounded by the upper bound  $\frac{1}{\lambda^2}$  as  $\lambda$  increases

**Figure 4.** As  $\lambda$  increases, Loss increases but Error remains robust due to the decrease of  $\frac{\text{Error}}{\text{Loss}}$ .

#### 4. Error analysis via $L^2(I, \mu)$ -contraction estimates

##### 4.1. $L^2(I, \mu)$ -contraction estimates

**Theorem 4.1** ( $L^2(I, \mu)$ -contraction estimates). Let  $\varepsilon > 0$  be a constant,  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq \lambda$  on  $I$  for some constant  $\lambda > 0$  and  $f \in L^2(I)$ , where  $I$  is a bounded open interval. Let  $\rho$  be a function defined in (3.3), where  $b$  above is considered as the zero extension of  $b$  in  $\mathbb{R}$ . Let  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  be a unique strong solution to (3.6) as in Theorem 3.2. Then,

$$\|y\|_{L^2(I, \mu)} \leq \frac{1}{\lambda} \|f\|_{L^2(I, \mu)}, \quad (4.1)$$

where  $\mu = \rho dx$ . In particular,

$$\|y\|_{L^2(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\min_{\bar{I}} \rho} \right)^{1/2} \frac{1}{\lambda} \|f\|_{L^2(I, \mu)} \leq \exp \left( \int_I \frac{1}{2\varepsilon} |b| dx \right) \frac{1}{\lambda} \|f\|_{L^2(I)}.$$

*Proof.* By Theorem 3.2, the unique strong solution  $y$  to (3.6) satisfies that

$$\int_I \varepsilon y' \psi' \rho dx + \int_I c y \psi \rho dx = \int_I f \psi \rho dx, \quad \text{for all } \psi \in H_0^{1,2}(I).$$

Substituting  $y$  for  $\psi$ , we get

$$\lambda \int_I y^2 \rho dx \leq \int_I \varepsilon |y'|^2 \rho dx + \int_I c y^2 \rho dx = \int_I f y \rho dx \leq \left( \int_I f^2 \rho dx \right)^{1/2} \left( \int_I y^2 \rho dx \right)^{1/2},$$

and hence the result follows.  $\square$

**Theorem 4.2.** Let  $p, q \in \mathbb{R}$ ,  $\varepsilon > 0$  be constants,  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq \lambda$  on  $I = (x_1, x_2)$  for some constant  $\lambda > 0$  and  $f \in L^2(I)$ , where  $I$  is a bounded open interval. Let  $\rho$  be a function defined in (3.3). Then, the unique strong solution  $\bar{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  to (3.10) as in Theorem 3.4 satisfies

$$\|\bar{y}\|_{L^2(I,\mu)} \leq \frac{1}{\lambda} \|\tilde{f}\|_{L^2(I,\mu)} + K_3(|p| \vee |q|),$$

where  $\mu = \rho dx$  and  $K_3 = \frac{1}{\lambda} \|b\|_{L^2(I,\mu)} |I|^{-1} + \frac{1}{\lambda} \|c\|_{L^2(I,\mu)} + |I|^{1/2} (\max_{\bar{I}} \rho)^{1/2}$ . In particular,

$$\|\bar{y}\|_{L^2(I)} \leq \frac{1}{\lambda} \left( \frac{\max_{\bar{I}} \rho}{\min_{\bar{I}} \rho} \right)^{1/2} \|\tilde{f}\|_{L^2(I)} + \left( \frac{\max_{\bar{I}} \rho}{\min_{\bar{I}} \rho} \right)^{1/2} \tilde{K}_3(|p| \vee |q|),$$

where  $\tilde{K}_3 = \frac{1}{\lambda} \|b\|_{L^2(I)} |I|^{-1} + \frac{1}{\lambda} \|c\|_{L^2(I)} + |I|^{1/2}$ .

*Proof.* Let  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then, we can check that  $\|\ell\|_{L^\infty(I)} \leq |p| \vee |q|$  and  $\|\ell'\|_{L^\infty(I)} \leq |I|^{-1}|p-q|$ . Let  $y := \bar{y} - \ell$ . Then,  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  is a strong solution to (3.11) where  $\tilde{f}$  is replaced by  $\tilde{f} - b\ell' - c\ell$ . Then, we derive from (4.1) in Theorem 4.1 that  $y = \bar{y} - \ell$  and that

$$\begin{aligned} \|\bar{y} - \ell\|_{L^2(I,\mu)} &= \|y\|_{L^2(I,\mu)} \leq \frac{1}{\lambda} \|\tilde{f} - b\ell' - c\ell\|_{L^2(I,\mu)} \\ &\leq \frac{1}{\lambda} \|\tilde{f}\|_{L^2(I,\mu)} + \frac{1}{\lambda} \|b\|_{L^2(I,\mu)} \|\ell'\|_{L^\infty(I)} + \frac{1}{\lambda} \|c\|_{L^2(I,\mu)} \|\ell\|_{L^\infty(I)}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|\bar{y}\|_{L^2(I,\mu)} &\leq \|\bar{y} - \ell\|_{L^2(I,\mu)} + \|\ell\|_{L^2(I,\mu)} \leq \|\bar{y} - \ell\|_{L^2(I,\mu)} + |I|^{1/2} (\max_{\bar{I}} \rho)^{1/2} \|\ell\|_{L^\infty(I)} \\ &\leq \frac{1}{\lambda} \|\tilde{f}\|_{L^2(I,\mu)} + \frac{1}{\lambda} \|b\|_{L^2(I,\mu)} |I|^{-1} |p-q| + \left( \frac{1}{\lambda} \|c\|_{L^2(I,\mu)} + |I|^{1/2} (\max_{\bar{I}} \rho)^{1/2} \right) (|p| \vee |q|) \\ &\leq \frac{1}{\lambda} \|\tilde{f}\|_{L^2(I,\mu)} + \left( \frac{1}{\lambda} \|b\|_{L^2(I,\mu)} |I|^{-1} + \frac{1}{\lambda} \|c\|_{L^2(I,\mu)} + |I|^{1/2} (\max_{\bar{I}} \rho)^{1/2} \right) (|p| \vee |q|). \end{aligned}$$

The rest immediately follows from the above.  $\square$

**Theorem 4.3.** Let  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq \lambda$  on  $I$  for some constant  $\lambda > 0$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $\rho$  be a function defined in (3.3). Let  $\bar{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi \in H^{2,2}(I) \cap C^1(\bar{I})$ . Then,

$$\|\bar{y} - \Phi\|_{L^2(I)} \leq \left( \frac{\max_{\bar{I}} \rho}{\min_{\bar{I}} \rho} \right)^{1/2} \frac{1}{\lambda} \|\tilde{f} - \mathcal{L}[\Phi]\|_{L^2(I)} + \left( \frac{\max_{\bar{I}} \rho}{\min_{\bar{I}} \rho} \right)^{1/2} \tilde{K}_3 \left( |p - \Phi(x_1)| \vee |q - \Phi(x_2)| \right),$$

where  $\tilde{K}_3 > 0$  is a constant as in Theorem 4.2 and  $\mathcal{L}$  is a differential operator defined by (3.12).

*Proof.* Note that  $\mathcal{L}[\bar{y}] = -\bar{y}'' + b\bar{y}' + c\bar{y} = \tilde{f} \in L^2(I)$  and  $\mathcal{L}[\Phi] = -\varepsilon\Phi'' + b\Phi' + c\Phi \in L^2(I)$ . Let  $u = \bar{y} - \Phi$ . Then,  $u \in H^{2,2}(I) \cap C^1(\bar{I})$  and  $\mathcal{L}[u] = \tilde{f} - \mathcal{L}[\Phi] \in L^2(I)$ . Moreover, by Theorem 3.2  $u$  is a unique strong solution to the following problem:

$$\begin{cases} -\varepsilon u'' + bu' + cu = \tilde{f} - \mathcal{L}[\Phi] & \text{on } I, \\ u(x_1) = p - \Phi(x_1), \quad u(x_2) = q - \Phi(x_2). \end{cases}$$



By Theorem 4.2,

$$\|u\|_{L^2(I)} \leq \left(\frac{\max_{\bar{I}}\rho}{\min_{\bar{I}}\rho}\right)^{1/2} \frac{1}{\lambda} \|\tilde{f} - \mathcal{L}[\Phi]\|_{L^2(I)} + \left(\frac{\max_{\bar{I}}\rho}{\min_{\bar{I}}\rho}\right)^{1/2} \tilde{K}_3 \left( |p - \Phi(x_1)| \vee |q - \Phi(x_2)| \right),$$

as desired.  $\square$

**Theorem 4.4.** Let  $\varepsilon > 0$  be a constant,  $b \in L^2(I)$ ,  $c \in L^2(I)$  with  $c \geq \lambda$  on  $I$  for some constant  $\lambda > 0$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Let  $p, q \in \mathbb{R}$  be given. Let  $\rho$  be a function defined in (3.3). Let  $\tilde{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi_A \in H_0^{1,2}(I) \cap H^{2,2}(I) \cap C^1(\bar{I})$  and  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then,

$$\begin{aligned} \|\tilde{y} - (\Phi_A + \ell)\|_{L^2(I)} &\leq \left(\frac{\max_{\bar{I}}\rho}{\min_{\bar{I}}\rho}\right)^{1/2} \frac{1}{\lambda} \|\tilde{f} - \mathcal{L}[\Phi_A + \ell]\|_{L^2(I)} \\ &\leq \exp\left(\int_I \frac{1}{2\varepsilon} |b| dx\right) \frac{1}{\lambda} \|\tilde{f} - \mathcal{L}[\Phi_A + \ell]\|_{L^2(I)}, \end{aligned}$$

where  $\mathcal{L}$  is a differential operator defined by (3.12).

*Proof.* Since  $(\Phi_A + \ell)(x_1) = p$  and  $(\Phi_A + \ell)(x_2) = q$ , the assertion follows from Theorem 4.3, where  $\Phi$  there is replaced by  $\Phi_A + \ell$  here.  $\square$

#### 4.2. Numerical experiments for PINN with $L^2(I, \mu)$ -contraction estimates

Let us consider the conditions as in Section 3.2 and additionally assume that  $c \geq \lambda$  for some  $\lambda > 0$ . Then, by using Theorem 4.4 with the Monte Carlo integration as in (2.2), we discover that for sufficiently large  $n \in \mathbb{N}$

$$\begin{aligned} \underbrace{\frac{1}{n} \sum_{i=1}^n |\tilde{y}(X_i) - (\Phi_A + \ell)(X_i)|^2}_{=: \text{Error}} &\leq \frac{1}{\lambda^2} \left(\frac{\max_{\bar{I}}\rho}{\min_{\bar{I}}\rho}\right) \cdot \frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi_A + \ell](X_i))^2, \\ &\leq \frac{1}{\lambda^2} \exp\left(\int_I \frac{1}{\varepsilon} |b| dx\right) \cdot \underbrace{\frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi_A + \ell](X_i))^2}_{=: \text{Loss}}, \quad \text{very likely.} \end{aligned} \tag{4.2}$$

**Example 4.5.** Let  $I = (0, 1)$ . Assume that  $b = 2$ ,  $p = 0$ ,  $q = 1$  and  $\tilde{f} = 0$ . Let  $\lambda > 0$  be a constant and  $c := \lambda$ . Then, (3.10) is expressed as

$$\begin{cases} -\varepsilon \tilde{y}'' + 2\tilde{y}' + \lambda \tilde{y} = \tilde{f} & \text{on } I, \\ \tilde{y}(0) = 0, \quad \tilde{y}(1) = 1, \end{cases} \tag{4.3}$$

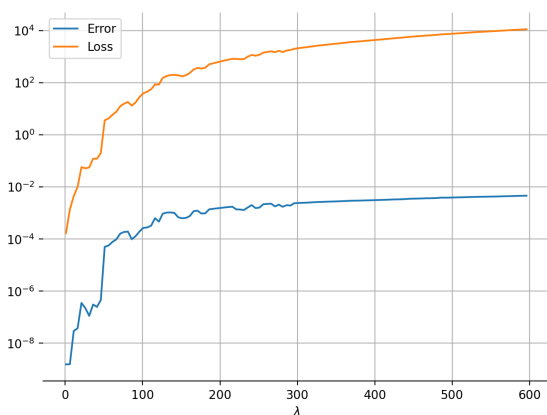
which has a unique solution by Theorem 3.4. All experimental details are identical to those described in Example 3.8.

(1) Let  $\varepsilon = 1$ . Then,

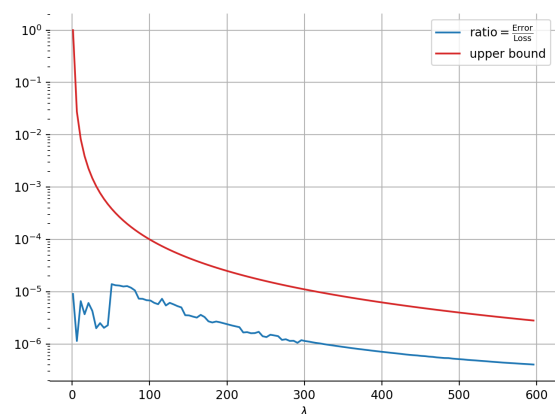
$$\tilde{y}(x) := \frac{1}{e^{1+\sqrt{1+\lambda}} - e^{1-\sqrt{1+\lambda}}} e^{(1+\sqrt{1+\lambda})x} + \frac{-1}{e^{1+\sqrt{1+\lambda}} - e^{1-\sqrt{1+\lambda}}} e^{(1-\sqrt{1+\lambda})x}, \quad x \in \mathbb{R}$$

is a unique solution to (4.3).

In Figure 5(a), we can observe that as  $\lambda$  increases, **Loss** increases whereas **Error** is relatively stable. The main reason for this is explained by Figure 5(b), where  $\frac{\text{Error}}{\text{Loss}}$  decreases with the upper bound  $\frac{1}{\lambda^2}$  as  $\lambda$  increases. In the error estimates (4.2), we derive the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  as  $\frac{1}{\lambda^2} \exp\left(\int_I \frac{1}{\varepsilon} |b| dx\right) = \frac{e^2}{\lambda^2}$ . However, by Figure 5(b), one can see that  $\frac{1}{\lambda^2}$  is a more suitable upper bound for  $\frac{\text{Error}}{\text{Loss}}$  than  $\frac{e^2}{\lambda^2}$ .



(a) Loss and Error after 500 epochs as  $\lambda$  increases



(b)  $\frac{\text{Error}}{\text{Loss}}$  bounded by the upper bound  $\frac{1}{\lambda^2}$  as  $\lambda$  increases

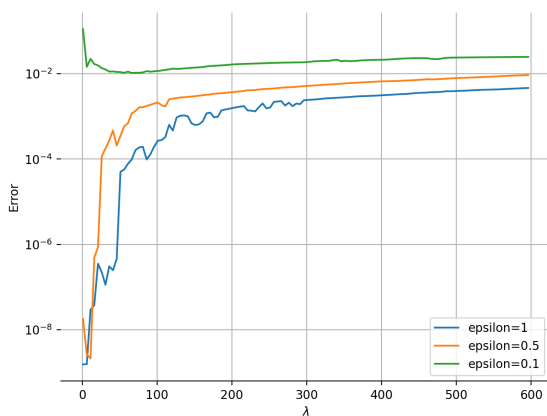
**Figure 5.** As  $\lambda$  increases, **Loss** increases but **Error** remains robust due to the decrease of  $\frac{\text{Error}}{\text{Loss}}$ .

(2) For general  $\varepsilon > 0$ , we obtain that

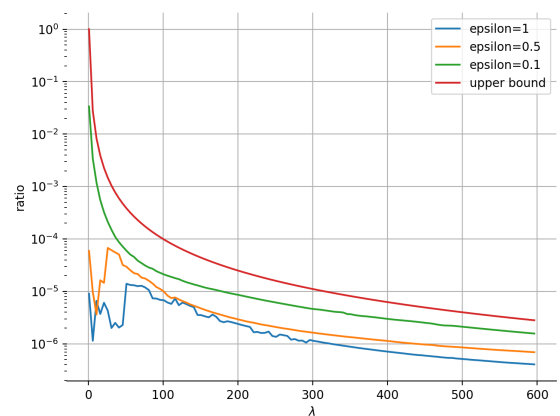
$$\tilde{y}(x) := \frac{1}{e^{\frac{1}{\varepsilon} + \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}} - e^{\frac{1}{\varepsilon} - \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}}} e^{\left(\frac{1}{\varepsilon} + \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}\right)x} + \frac{-1}{e^{\frac{1}{\varepsilon} + \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}} - e^{\frac{1}{\varepsilon} - \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}}} e^{\left(\frac{1}{\varepsilon} - \sqrt{\frac{1}{\varepsilon^2} + \frac{\lambda}{\varepsilon}}\right)x}, \quad x \in \mathbb{R}$$

is a unique solution to (4.3) by Theorem 3.4.

Figure 6(a) describes the behavior of **Error** for each case when  $\varepsilon = 1, 0.5, 0.1$ . In Figure 6(a), if  $\lambda$  is small, there are rapid changes of **Error**, but if  $\lambda$  is over 100, **Error** remains relatively stable. Figure 6(b) describes the behaviors of the ratios  $\left(\frac{\text{Error}}{\text{Loss}}\right)$  which rapidly decrease as  $\lambda$  increases. The rapid decrease of  $\frac{\text{Error}}{\text{Loss}}$  as the increase of  $\lambda$  in Figure 6(b) explains the reason **Error** remains relatively stable in Figure 6(a). Here, we point out that the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  is  $\frac{1}{\lambda^2}$  in Figure 6(b). Indeed, if  $\varepsilon = 0.1$ , then one can derive an upper bound of  $\frac{\text{Error}}{\text{Loss}}$  as  $\frac{1}{\lambda^2} \exp\left(\int_I \frac{1}{\varepsilon} |b| dx\right) = \frac{e^{20}}{\lambda^2}$  as in (4.2), which is much bigger than  $\frac{1}{\lambda^2}$ . Therefore, we have to derive sharper error estimates than (4.2), which will be presented in the next section.



(a) Comparison of Errors after 500 epochs as  $\lambda$  increase when  $\varepsilon = 1, 0.5, 0.1$



(b) Comparison of the ratios  $\left(\frac{\text{Error}}{\text{Loss}}\right)$  as  $\lambda$  increases when  $\varepsilon = 1, 0.5, 0.1$  and the upper bound  $\frac{1}{\lambda^2}$

**Figure 6.** The change of  $\varepsilon$  does not significantly affect Error behavior.

### 5. Error analysis via $L^2(I, dx)$ -contraction estimates

#### 5.1. $L^2(I, dx)$ -contraction estimates

**Theorem 5.1** ( $L^2(I, dx)$ -contraction estimates). *Let  $\varepsilon > 0$  be a constant,  $b \in H^{1,1}(I)$ ,  $c \in L^1(I)$  with  $c \geq 0$  on  $I$  and  $f \in L^2(I)$ , where  $I$  is a bounded open interval. Assume that there exists a constant  $\gamma > 0$  such that  $-\frac{1}{2}b' + c \geq \gamma$  on  $I$ . Let  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  be a unique strong solution to (3.6). Then,*

$$\|y\|_{L^2(I)} \leq \frac{1}{\gamma} \|f\|_{L^2(I)}. \tag{5.1}$$

*Proof.* Multiplying both sides of (3.6) by  $y$  and integrating over  $I$  with respect to  $dx$ , we have

$$\begin{aligned} \gamma \int_I y^2 dx &\leq \int_I \varepsilon |y'|^2 \rho dx + \int_I \left(-\frac{1}{2}b' + c\right) y^2 dx = \int_I \varepsilon |y'|^2 dx + \int_y b y' y dx + \int_I c y^2 dx \\ &= \int_I f y dx \leq \left(\int_I f^2 dx\right)^{1/2} \left(\int_I y^2 dx\right)^{1/2}, \end{aligned}$$

and hence the assertion follows. □

**Remark 5.2.** *The above contraction estimates can be extended to the case of the multidimensional linear elliptic equations. As a general form of  $L^2$ -contraction estimates, one can derive  $L^r$ -contraction estimates with  $r \in [1, \infty]$  by using the sub-Markovian property and the Riesz-Thorin interpolation with a duality argument (see [14]).*

**Theorem 5.3.** *Let  $\varepsilon > 0$  be a constant,  $b \in H^{1,1}(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^2(I)$ , where  $I$  is a bounded open interval. Assume that there exists a constant  $\gamma > 0$  such that  $-\frac{1}{2}b' + c \geq \gamma$  on  $I$ . Then, the unique strong solution  $\tilde{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  to (3.10) as in Theorem 3.4 satisfies*

$$\|\tilde{y}\|_{L^2(I)} \leq \frac{1}{\gamma} \|\tilde{f}\|_{L^2(I)} + \tilde{K}_4 (|p| \vee |q|),$$

where  $\tilde{K}_4 := \frac{1}{\gamma} \|b\|_{L^2(I)} |I|^{-1} + \frac{1}{\gamma} \|c\|_{L^2(I)} + |I|^{1/2}$ .

*Proof.* Let  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then, we can verify that  $\|\ell\|_{L^\infty(I)} \leq |p| \vee |q|$  and  $\|\ell'\|_{L^\infty(I)} \leq |I|^{-1}|p-q|$ . Let  $y := \tilde{y} - \ell$ . Then, by Theorem 3.4  $y \in H_0^{1,2}(I) \cap H^{2,1}(I) \cap C^1(\bar{I})$  is a unique strong solution to (3.11) where  $f$  is replaced by  $\tilde{f} - b\ell' - c\ell$ , and hence it follows from (5.1) in Theorem 5.1 that

$$\begin{aligned} \|\tilde{y} - \ell\|_{L^2(I)} &= \|y\|_{L^2(I)} \leq \frac{1}{\gamma} \|\tilde{f} - b\ell' - c\ell\|_{L^2(I)} \\ &\leq \frac{1}{\gamma} \|\tilde{f}\|_{L^2(I)} + \frac{1}{\gamma} \|b\|_{L^2(I)} \|\ell'\|_{L^\infty(I)} + \frac{1}{\gamma} \|c\|_{L^2(I)} \|\ell\|_{L^\infty(I)}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|\tilde{y}\|_{L^2(I)} &\leq \|\tilde{y} - \ell\|_{L^2(I)} + \|\ell\|_{L^2(I)} \leq \|\tilde{y} - \ell\|_{L^2(I,\mu)} + |I|^{1/2} \|\ell\|_{L^\infty(I)} \\ &\leq \frac{1}{\gamma} \|\tilde{f}\|_{L^2(I)} + \frac{1}{\gamma} \|b\|_{L^2(I)} |I|^{-1} |p-q| + \left( \frac{1}{\gamma} \|c\|_{L^2(I)} + |I|^{1/2} \right) (|p| \vee |q|) \\ &\leq \frac{1}{\gamma} \|\tilde{f}\|_{L^2(I)} + \left( \frac{1}{\gamma} \|b\|_{L^2(I)} |I|^{-1} + \frac{1}{\gamma} \|c\|_{L^2(I)} + |I|^{1/2} \right) (|p| \vee |q|), \end{aligned}$$

as desired.  $\square$

**Theorem 5.4.** Let  $\varepsilon > 0$  be a constant,  $b \in H^{1,1}(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Assume that there exists a constant  $\gamma > 0$  such that  $-\frac{1}{2}b' + c \geq \gamma$  on  $I$ . Let  $p, q \in \mathbb{R}$  be given and  $\tilde{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi \in H^{2,2}(I) \cap C^1(\bar{I})$ . Then,

$$\|\tilde{y} - \Phi\|_{L^2(I)} \leq \frac{1}{\gamma} \|\tilde{f} - \mathcal{L}[\Phi]\|_{L^2(I)} + \tilde{K}_4 \left( |p - \Phi(x_1)| \vee |q - \Phi(x_2)| \right),$$

where  $\tilde{K}_4 > 0$  is a constant in Theorem 5.3 and  $\mathcal{L}$  is a differential operator defined as in (3.12).

*Proof.* Let  $u := \tilde{y} - \Phi$ . Then,  $u \in H^{2,2}(I) \cap C^1(\bar{I})$  and  $\mathcal{L}[\Phi] \in L^2(I)$ . Moreover, by Theorem 3.4  $u$  is a (unique) strong solution to the following problem:

$$\begin{cases} -\varepsilon u'' + bu' + cu = \tilde{f} - \mathcal{L}[\Phi] & \text{on } I, \\ u(x_1) = p - \Phi(x_1), \quad u(x_2) = q - \Phi(x_2). \end{cases}$$

By Theorem 3.4, the result follows.  $\square$

**Theorem 5.5.** Let  $\varepsilon > 0$  be a constant,  $b \in H^{1,1}(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $\tilde{f} \in L^2(I)$ , where  $I = (x_1, x_2)$  is a bounded open interval. Assume that there exists a constant  $\gamma > 0$  such that  $-\frac{1}{2}b' + c \geq \gamma$  on  $I$ . Let  $p, q \in \mathbb{R}$  be given and  $\tilde{y} \in H^{2,1}(I) \cap C^1(\bar{I})$  be the unique strong solution to (3.10) as in Theorem 3.4. Let  $\Phi_A \in H_0^{1,2}(I) \cap H^{2,2}(I) \cap C^1(\bar{I})$  and  $\ell(x) := \frac{q-p}{x_2-x_1}(x-x_1) + p$ ,  $x \in \mathbb{R}$ . Then,

$$\|\tilde{y} - (\Phi_A + \ell)\|_{L^2(I)} \leq \frac{1}{\gamma} \|\tilde{f} - \mathcal{L}[\Phi_A + \ell]\|_{L^2(I)},$$

where  $\mathcal{L}$  is a differential operator defined by (3.12).

*Proof.* Since  $(\Phi_A + \ell)(x_1) = p$  and  $(\Phi_A + \ell)(x_2) = q$ , the assertion follows from Theorem 5.4 where  $\Phi$  there is replaced by  $\Phi_A + \ell$  here.  $\square$

5.2. Numerical experiments for PINN with  $L^2(I, dx)$ -contraction estimates

Let us consider the conditions in Section 3.2 and also assume that  $b \in H^{1,1}(I)$ ,  $c \in L^2(I)$  with  $c \geq 0$  on  $I$  and  $f \in L^2(I)$ , where  $I$  is a bounded open interval. Additionally, assume that

$$-\frac{1}{2}b' + c \geq \gamma, \quad \text{for some } \gamma > 0. \tag{5.2}$$

Then, by using Theorem 5.5 with the Monte Carlo integration as in (2.2), we obtain that for sufficiently large  $n \in \mathbb{N}$

$$\underbrace{\frac{1}{n} \sum_{i=1}^n |\bar{y}(X_i) - (\Phi_A + \ell)(X_i)|^2}_{=: \text{Error}} \leq \frac{1}{\gamma^2} \cdot \underbrace{\frac{1}{n} \sum_{i=1}^n (\tilde{f}(X_i) - \mathcal{L}[\Phi_A + \ell](X_i))^2}_{=: \text{Loss}}, \quad \text{very likely.} \tag{5.3}$$

**Example 5.6.** Let  $I = (0, 1)$ . Let  $\lambda, k \geq 0$  be constants. Let  $b(x) = -kx$ ,  $c(x) = \lambda$  and  $\bar{y}(x) = x(1-x)e^x + x^2$ ,  $x \in \mathbb{R}$ . Then,

$$\bar{y}'(x) = (1-x-x^2)e^x + 2x, \quad \bar{y}''(x) = (-3x-x^2)e^x + 2, \quad \text{for all } x \in \mathbb{R}.$$

Let  $f(x) := -\varepsilon \bar{y}''(x) + b(x)\bar{y}'(x) + c(x)\bar{y}(x) = (kx^3 + (k-\lambda+\varepsilon)x^2 + (3\varepsilon-k+\lambda)x)e^x + \lambda x^2 - 2kx^2 - 2\varepsilon$ ,  $x \in \mathbb{R}$ . Then,  $\bar{y}$  is a unique strong solution to (3.10) with  $p = 0$  and  $q = 1$ . Moreover, we can choose

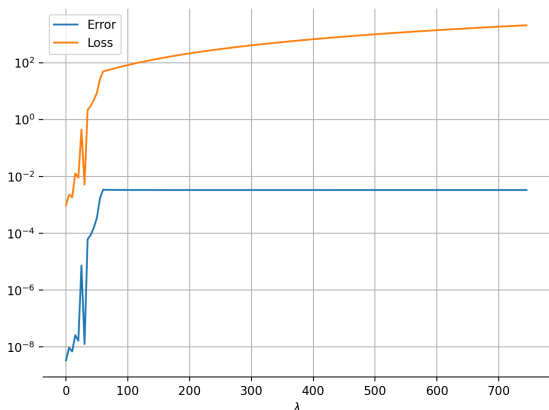
$$\gamma = \frac{1}{2}k + \lambda = -\frac{1}{2}b' + c.$$

Therefore, the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  is expressed as

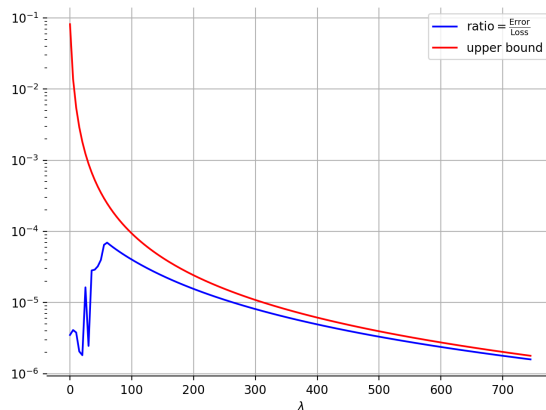
$$\frac{1}{\gamma^2} = \frac{4}{(k+2\lambda)^2}.$$

All experimental details are identical to those described in Example 3.8.

(1) Let  $\varepsilon = 1$  and  $k = 7$ .



(a) Loss and Error after 500 epochs as  $\lambda$  increases ( $\varepsilon = 1, k = 7$ )

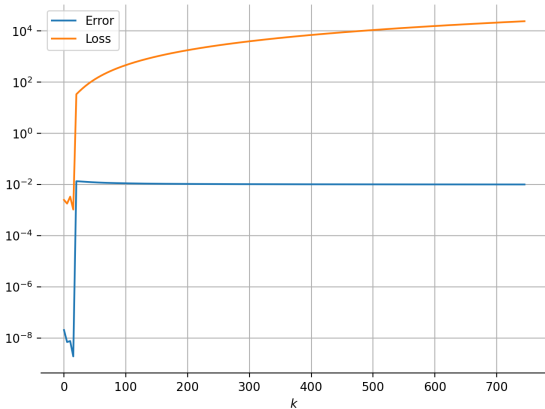


(b) Ratio  $\left(\frac{\text{Error}}{\text{Loss}}\right)$  bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(7+2\lambda)^2}$  ( $\varepsilon = 1, k = 7$ )

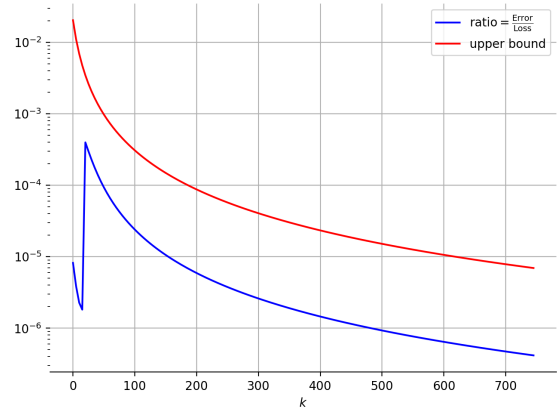
**Figure 7.** As  $\lambda$  increases, Loss increases but Error remains robust (Example 5.6).

In Figure 7(a), we can observe that as  $\lambda$  increases, Loss also increases but Error remains stable. This phenomenon is well-explained by Figure 7(b) where the ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) rapidly decreases with the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(7+2\lambda)^2}$  as  $\lambda$  increases. Furthermore, it is observed in Figure 7(b) that the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  is very sharp when  $\lambda$  is large enough.

(2) Let  $\varepsilon = 1$  and  $\lambda = 7$ .



(a) Loss and Error after 500 epochs as  $k$  increases ( $\varepsilon = 1, \lambda = 7$ )

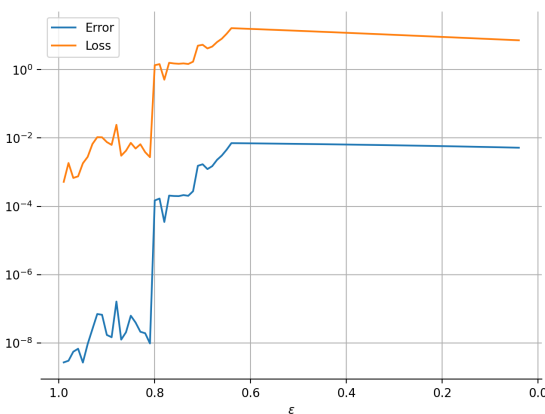


(b) Ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(k+14)^2}$  as  $k$  increases ( $\varepsilon = 1, \lambda = 7$ )

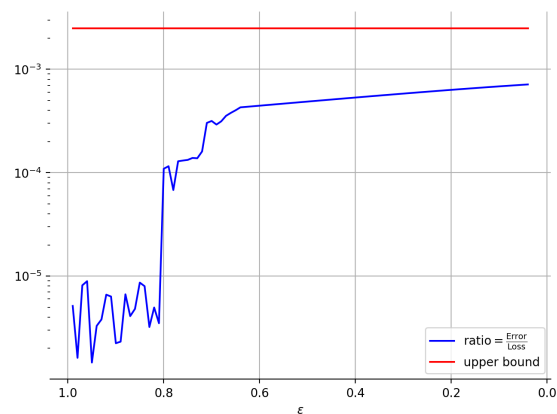
**Figure 8.** As  $k$  increases, Loss increases but Error remains robust (Example 5.6).

In Figure 8(a), we can observe that as  $\lambda$  increases, Loss also increases but Error remains stable. This phenomenon is well-explained by Figure 8(b) where the ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) rapidly decreases with the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(k+14)^2}$  as  $k$  increases.

(3) Let  $k = 10$  and  $\lambda = 15$ .



(a) Loss and Error after 500 epochs as  $\varepsilon$  decreases ( $k = 10$  and  $\lambda = 15$ )



(b) Ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{1}{400}$  as  $\varepsilon$  decreases ( $k = 10$  and  $\lambda = 15$ )

**Figure 9.** As  $\varepsilon$  decreases, Loss and Error remain stable (Example 5.6).

In Figure 9(a), we can observe that as  $\varepsilon$  decreases, Loss and Error remain robust. Since  $\varepsilon$  is a very small value, the stability of Loss is well-explained. Since  $k$  and  $\lambda$  are fixed as 10 and 15,

respectively and  $\varepsilon$  only decreases, we can expect that by our error estimates (5.3),  $\frac{\text{Error}}{\text{Loss}}$  does not change dramatically. This expectation is realized by Figure 9(b), where  $\frac{\text{Error}}{\text{Loss}}$  are bounded by  $\frac{1}{\gamma^2} = \frac{1}{400}$ , and hence the robustness of Error is well-explained.

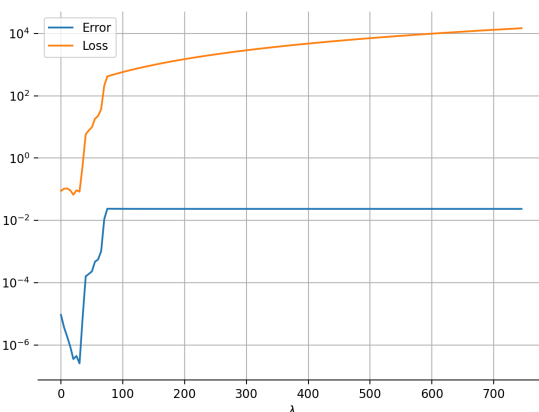
**Example 5.7.** Let  $I = (0, 1)$ . Let  $\lambda, k \geq 0$  be constants. Let  $b(x) = -kx$ ,  $c(x) = \lambda$  and  $\tilde{y}(x) = \sin \pi x + \cos \pi x$ ,  $x \in \mathbb{R}$ . Then,

$$\tilde{y}'(x) = \pi \cos \pi x - \pi \sin \pi x, \quad \tilde{y}''(x) = -\pi^2 \sin \pi x - \pi^2 \cos \pi x, \quad \text{for all } x \in \mathbb{R}.$$

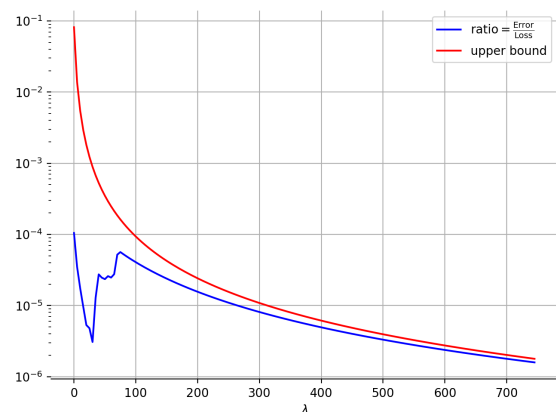
Let  $f(x) := -\varepsilon \tilde{y}''(x) + b(x)\tilde{y}'(x) + c\tilde{y}(x) = (\varepsilon\pi^2 + \pi kx + \lambda) \sin \pi x + (\varepsilon\pi^2 - \pi kx + \lambda) \cos \pi x$ ,  $x \in \mathbb{R}$ . Then,  $\tilde{y}$  is a unique strong solution to (3.10) with  $p = 1$ ,  $q = -1$ . As in Example 5.6, the upper bound of  $\frac{\text{Error}}{\text{Loss}}$  is expressed as  $\frac{1}{\gamma^2} = \frac{4}{(k+2\lambda)^2}$ .

All experimental details are identical to those described in Example 3.8.

(1) Let  $\varepsilon = 1$  and  $k = 7$ .



(a) Loss and Error after 500 epochs as  $\lambda$  increases ( $\varepsilon = 1$ ,  $k = 7$ )

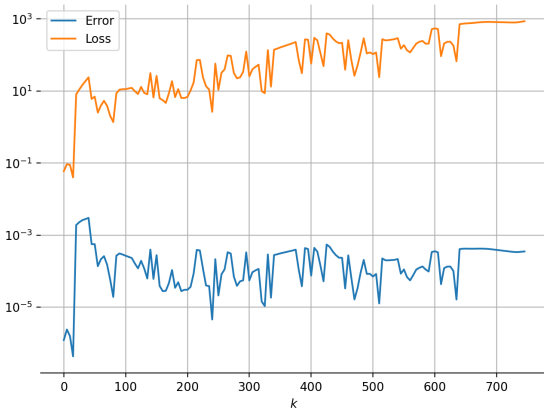


(b) Ratio  $\left(\frac{\text{Error}}{\text{Loss}}\right)$  bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(7+2\lambda)^2}$  as  $\lambda$  increases ( $\varepsilon = 1$ ,  $k = 7$ ).

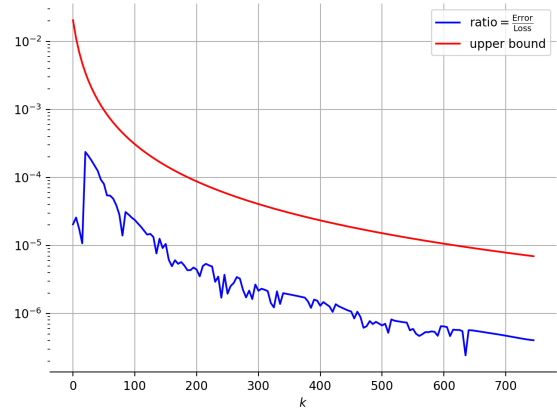
**Figure 10.** As  $\lambda$  increases, Loss increases but Error remains robust (Example 5.7)

In Figure 10(a), we can observe that as  $\lambda$  increases, Loss also increases but Error remains stable. This phenomenon is well-explained by Figure 10(b), where the ratio  $\left(\frac{\text{Error}}{\text{Loss}}\right)$  rapidly decreases with the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(7+2\lambda)^2}$  as  $\lambda$  increases.

(2) Let  $\varepsilon = 1$  and  $\lambda = 7$ .



(a) Loss and Error after 500 epochs as  $k$  increases ( $\varepsilon = 1, \lambda = 7$ ).

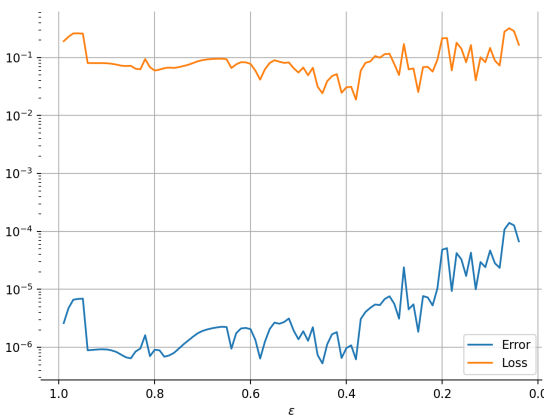


(b) Ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(k+14)^2}$  as  $k$  increases ( $\varepsilon = 1, \lambda = 7$ )

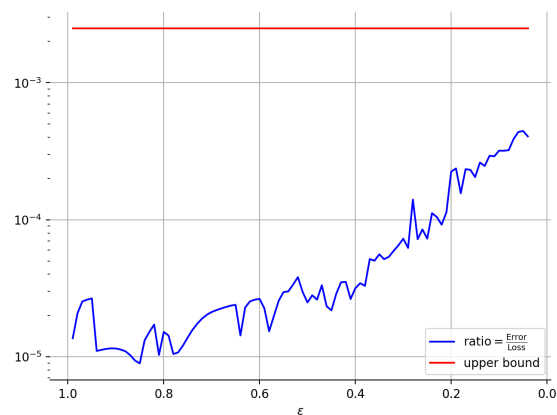
**Figure 11.** As  $k$  increases, Loss increases but Error remains robust (Example 5.7).

In Figure 11(a), we can observe that as  $\lambda$  increases, Loss also increases but Error remains stable. This phenomenon is well-explained by Figure 11(b), where the ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) rapidly decreases with the upper bound  $\frac{1}{\gamma^2} = \frac{4}{(k+14)^2}$  as  $k$  increases.

(3) Let  $k = 10$  and  $\lambda = 15$ .



(a) Loss and Error after 500 epochs as  $\varepsilon$  decreases ( $k = 10, \lambda = 15$ )



(b) Ratio ( $\frac{\text{Error}}{\text{Loss}}$ ) bounded by the upper bound  $\frac{1}{\gamma^2} = \frac{1}{400}$  as  $\varepsilon$  decreases ( $k = 10, \lambda = 15$ )

**Figure 12.** As  $\varepsilon$  decreases, Loss and Error remain stable (Example 5.7).

In Figure 12(a), we can observe that as  $\varepsilon$  decreases, Loss and Error remain stable. Since  $\varepsilon$  is a very small value, the stability of Loss is well-explained. Since  $k$  and  $\lambda$  are fixed as 10 and 15, respectively and  $\varepsilon$  only decreases, we can expect that by our error estimates (5.3),  $\frac{\text{Error}}{\text{Loss}}$  does not change dramatically. This expectation is realized by Figure 12(b), where  $\frac{\text{Error}}{\text{Loss}}$  are bounded by  $\frac{1}{\gamma^2} = \frac{1}{400}$ , and hence the robustness of Error is well-explained.



## 6. Conclusions and discussion

Based on the universal approximation theorem derived from the Stone-Weierstrass theorem, solutions to differential equations can be uniformly approximated by neural networks. However, due to the abstract nature of the Stone-Weierstrass theorem, finding explicit neural networks that approximate a solution to a differential equation would be very challenging. As an alternative, through a posteriori analysis, we have demonstrated that Error is controlled by Loss (see (3.13), (3.14), (4.2), (5.3)), and hence we can obtain effective error bounds in terms of Loss.

Our analysis is specifically based on the Sobolev space theory with a variational approach. Initially, we rigorously established the existence and uniqueness of solutions to one-dimensional boundary value problems for second-order linear elliptic equations with highly irregular coefficients. There are two approaches for selecting a trial function to approximate the solution to a differential equation (see Section 3.2): The first approach is to directly use a neural network itself as a trial function (PINN I), and the second approach is to use a function based on the neural network that exactly fulfills the given boundary conditions (PINN II). Experimentally, we confirmed that (PINN II) is more efficient in approximating the solution.

Ultimately, our main error estimate is (5.3) which demonstrates that under the assumption of (5.2),  $\frac{\text{Error}}{\text{Loss}}$  decreases rapidly with the upper bound  $\frac{1}{\gamma^2}$  as  $\gamma$  increases, so that Error is robustly maintained even though Loss increases. These results refute the conventional belief that numerical cost increases as the quantities of the coefficients increase. Our research indicates that PINN can be efficient numerical solvers for differential equations with large quantities of coefficients.

In the case of the upper bound of the ratio, Figure 7(b) and Figure 10(b) describe very sharp upper bounds for  $\frac{\text{Error}}{\text{Loss}}$ , while Figure 8(b) and Figure 11(b) present less sharp bounds, and hence further studies can be conducted to derive sharper upper bounds. Further discussion is required to determine under what additional conditions the rapid decrease in  $\frac{\text{Error}}{\text{Loss}}$  can be expected. Empirically, it is observed that the larger quantities of the first-order coefficients  $b$  and the zero-order coefficients  $c$  imply the small values of  $\frac{\text{Error}}{\text{Loss}}$ . In future research, a rigorous mathematical demonstration is required to present additional explicit conditions on the coefficients that guarantee a rapid decrease in  $\frac{\text{Error}}{\text{Loss}}$  beyond the condition (5.2).

The most innovative aspect of this paper is that we specifically calculated the constant arising in the  $L^2$ -estimate and, in particular, verified through  $L^2$ -contraction estimates that this constant decreases rapidly as the lower bound of the zero-order term increases. This is highly significant as it provides a concrete upper bound for the error-to-loss ratio in a posteriori error analysis of PINN, which enables us to estimate  $L^2$ -error by calculating  $L^2$ -training loss. Our research emphasizes the importance of precisely calculating constants in  $L^2$ -estimates in future studies. However, the current results are limited to the one-dimensional Dirichlet boundary value problem, and the condition on the coefficients for obtaining  $L^2$ -contraction estimates is somewhat restrictive because we need a constant  $\gamma > 0$  that is not small and satisfies  $-\frac{1}{2}b' + c \geq \gamma$  in  $I$ . Further research is necessary to overcome these limitations.

Through this paper, it was confirmed that robust error estimates are guaranteed regardless of the quantities of coefficients in one-dimensional boundary value problems for linear elliptic equations. Moreover, we could confirm the possibility of an efficient numerical solution through PINN for differential equations with very large quantities of coefficients, and we expect that this will be applied to more general multidimensional elliptic and parabolic differential equations with singular coefficients

---

and general boundary conditions.

### Author contributions

Both authors contributed equally to this manuscript. All authors have read and approved the final version of the manuscript for publication.

### Acknowledgments

We would like to thank the anonymous reviewers for their valuable comments and suggestions, which improve the content of this paper.

### Conflict of interest

All authors declare no conflicts of interest in this paper.

### References

1. A. Biswas, J. Tian, S. Ulusoy, Error estimates for deep learning methods in fluid dynamics, *Numer. Math.*, **151** (2022), 753–777. <https://dx.doi.org/10.1007/s00211-022-01294-z>
2. H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, New York: Springer, 2011. <https://dx.doi.org/10.1007/978-0-387-70914-7>
3. R. T. Q. Chen, Y. Rubanova, J. Bettencourt, D. K. Duvenaud, Neural ordinary differential equations, *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, Canada: Montréal, 2018, 6572–6583.
4. G. Cybenko, Approximation by superpositions of a sigmoidal function, *Math. Control Signals Syst.*, **2** (1989), 303–314. <https://dx.doi.org/10.1007/BF02551274>
5. T. De Ryck, S. Mishra, Error analysis for physics-informed neural networks (PINNs) approximating Kolmogorov PDEs, *Adv. Comput. Math.*, **48** (2022), 1–40. <https://dx.doi.org/10.1007/s10444-022-09985-9>
6. T. De Ryck, A. D. Jagtap, S. Mishra, Error estimates for physics-informed neural networks approximating the Navier-Stokes equations, *IMA J. Numer. Anal.*, **44** (2024), 83–119. <https://dx.doi.org/10.1093/imanum/drac085>
7. M. W. M. G. Dissanayake, N. Phan-Thien, Neural-network-based approximations for solving partial differential equations, *Commun. Numer. Methods Eng.*, **10** (1994), 195–201. <https://dx.doi.org/10.1002/cnm.1640100303>
8. P. A. Farrell, A. F. Hegarty, J. J. H. Miller, E. O’Riordan, G. I. Shishkin, Singularly perturbed convection-diffusion problems with boundary and weak interior layers, *J. Comput. Appl. Math.*, **166** (2004), 133–151. <https://dx.doi.org/10.1016/j.cam.2003.09.033>
9. K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Networks*, **2** (1989), 359–366. [https://dx.doi.org/10.1016/0893-6080\(89\)90020-8](https://dx.doi.org/10.1016/0893-6080(89)90020-8)

10. Y. Hong, C. Y. Jung, J. Laminie, Singularly perturbed reaction-diffusion equations in a circle with numerical applications, *Int. J. Comput. Math.*, **90** (2013), 2308–2325. <https://dx.doi.org/10.1080/00207160.2013.772987>
11. Y. Hong, S. Ko, J. Lee, Error analysis for finite element operator learning methods for solving parametric second-order elliptic PDEs, *arXiv*, 2024. <https://dx.doi.org/10.48550/arXiv.2404.17868>
12. I. E. Lagaris, A. Likas, D. I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Networks*, **9** (1998), 987–1000. <https://dx.doi.org/10.1109/72.712178>
13. H. Lee, I. Kang, Neural algorithm for solving differential equations, *J. Comput. Phys.*, **91** (1990), 110–131. [https://dx.doi.org/10.1016/0021-9991\(90\)90007-N](https://dx.doi.org/10.1016/0021-9991(90)90007-N)
14. H. Lee, On the contraction properties for weak solutions to linear elliptic equations with  $L^2$ -drifts of negative divergence, *Proc. Amer. Math. Soc.*, **152** (2024), 2051–2068, <https://dx.doi.org/10.1090/proc/16672>
15. S. Mishra, R. Molinaro, Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs, *IMA J. Numer. Anal.*, **42** (2022), 981–1022. <https://dx.doi.org/10.1093/imanum/drab032>
16. S. Mishra, R. Molinaro, Estimates on the generalization error of physics-informed neural networks for approximating PDEs, *IMA J. Numer. Anal.*, **43** (2023), 1–43. <https://dx.doi.org/10.1093/imanum/drab093>
17. J. Müller, M. Zeinhofer, Notes on exact boundary values in residual minimisation, *Proc. Math. Sci. Mach. Learn.*, **190** (2022), 231–240.
18. M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.*, **378** (2019), 686–707. <https://dx.doi.org/10.1016/j.jcp.2018.10.045>
19. H. G. Roos, M. Stynes, L. Tobiska, *Robust numerical methods for singularly perturbed differential equations: convection-diffusion-reaction and flow problems*, Heidelberg: Springer Berlin, 2008. <https://dx.doi.org/10.1007/978-3-540-34467-4>
20. W. Rudin, *Principles of mathematical analysis*, 3 Eds., New York: McGraw-Hill, 1976.
21. E. K. Ryu, *Infinitely large neural networks*, Lecture Notes in Mathematics, Research Institute of Mathematics, Number 58, 2023.
22. Y. Shin, J. Darbon, G. E. Karniadakis, On the convergence of physics informed neural networks for linear second-order elliptic and parabolic type PDEs, *Commun. Comput. Phys.*, **28** (2020), 2042–2074. <https://dx.doi.org/10.4208/cicp.OA-2020-0193>
23. Y. Shin, Z. Zhang, G. E. Karniadakis, Error estimates of residual minimization using neural networks for linear PDEs, *J. Mach. Learn. Model. Comput.*, **4** (2023), 73–101. <https://dx.doi.org/10.1615/JMachLearnModelComput.2023050411>
24. H. Son, J. Jang, W. Han, H. Hwang, Sobolev training for physics-informed neural networks, *Commun. Math. Sci.*, **21** (2023), 1679–1705. <https://dx.doi.org/10.4310/CMS.2023.v21.n6.a11>

25. N. Yadav, A. Yadav, M. Kumar, *An introduction to neural network methods for differential equations*, SpringerBriefs in Applied Sciences and Technology, Dordrecht: Springer, 2015. <https://dx.doi.org/10.1007/978-94-017-9816-7>
26. J. Yoo, J. Kim, M. Gim, H. Lee, Error estimates of Physics-Informed Neural Networks for initial value problems, *J. Korean Soc. Ind. Appl. Math.*, **28** (2024), 33-58. <https://dx.doi.org/10.12941/jksiam.2024.28.033>
27. M. Zeinhofer, R. Masri, K. A. Mardal, *A unified framework for the error analysis of physics-informed neural networks*, 2024. <https://doi.org/10.48550/arXiv.2311.00529>

### Supplementary

The code used for numerical experiments in this paper is available at <https://github.com/hahmYoo/Robust-error-estimates-of-PINN>.



AIMS Press

©2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)