

# GENERALIZED KURAMOTO-SIVASHINSKY EQUATION SOLVING BY $B$ -SPLINES BASED NEURAL NETWORK

A.W. BELLO, J. ADETOLA, C.Z.J. MAMLANKOU\*,  
B.B. ADJALLA

Received: date / Revised: date / Accepted: date

**Abstract.** *This study presents a comprehensive case study investigating the application of Variational Physics-Informed Neural Networks (VPINNs) combined with B-spline test functions for solving the generalized Kuramoto-Sivashinsky equation in one space dimension. The Kuramoto-Sivashinsky equation, a nonlinear fourth-order PDE, serves as a benchmark for modeling instabilities in fluid dynamics, combustion, and pattern formation. Our methodology leverages variational formulations to enhance solution accuracy while ensuring robust convergence. Through systematic numerical experiments, we demonstrate that VPINNs effectively capture complex dynamical behaviors in high-order nonlinear systems. The framework's robustness is evaluated across different network architectures and multi-stage optimization strategies using Adam and L-BFGS optimizers. Results establish VPINNs as a powerful tool for addressing complex nonlinear PDEs while maintaining computational efficiency. This study provides insights into practical VPINNs implementation for scientific computing, with future directions focusing on computational optimization and adaptive collocation strategies.*

**Keywords:** Kuramoto-Sivashinsky equation, VPINNs,  $B$ -splines, numerical analysis, nonlinear, variational methods, scientific computing

**Mathematics Subject Classification (2020):** 65M70, 35Q53, 68T07, 35K55

---

\* Corresponding author.

**Abdou Wahidi Bello**

University of Abomey-Calavi, Abomey-Calavi, Benin  
E-mail: wahidi.bello@fast.uac.bj

**Jamal Adetola**

National University of Science, Technology, Engineering and Mathematics (UNSTIM), Abomey, Benin  
E-mail: adetolajamal@unstim.bj

**Charbel Zeus Jupiter Mamlankou**

National University of Science, Technology, Engineering and Mathematics (UNSTIM), Abomey, Benin  
E-mail: charbelzeusmamlankou@gmail.com

**Brice Bonaparte Adjalla**

National University of Science, Technology, Engineering and Mathematics (UNSTIM), Abomey, Benin  
E-mail: bricebonaparteadjalla2@gmail.com

## 1. Introduction

The Kuramoto-Sivashinsky equation (KSE) is a nonlinear fourth-order partial differential equation that was introduced to model instabilities in various physical systems. It was independently derived by Kuramoto in the context of phase turbulence in chemical reactions [25] and by Sivashinsky to describe thermo-diffusive instabilities in laminar flame fronts [34]. Since its inception, the KSE has been widely studied due to its ability to exhibit complex dynamical behaviors, including chaotic solutions, bifurcations, and pattern formation [19], [22].

This equation, which involves fourth-order spatial derivatives coupled with nonlinear terms, describes the evolution of a physical variable such as velocity or interface height. The KSE serves as a canonical model for studying spatiotemporal chaos and nonlinear dynamics in dissipative systems [8]. Its solutions exhibit a rich variety of behaviors depending on the domain size and boundary conditions, including steady states, traveling waves, and fully chaotic regimes [38].

The applications of the KSE are numerous and span multiple fields. In combustion theory, it is used to model instabilities in laminar flame fronts [35]. In fluid dynamics, it describes thin film flows and interfacial instabilities in viscous fluids [29]. In applied mathematics, it has become a paradigmatic example for studying deterministic chaos and attractors in infinite-dimensional dynamical systems [37]. Extensions of the KSE to higher dimensions or with additional terms (e.g., Benney-type equations) have been employed to study multi-scale interactions and stabilization mechanisms in complex systems [2].

The numerical solution of the Kuramoto-Sivashinsky equation has evolved significantly over the years, from traditional finite difference and spectral methods [7], [21] to more sophisticated approaches including finite element methods [5], [31] and adaptive mesh refinement techniques [4], [11]. Recent advances in machine learning have introduced physics-informed neural networks (PINNs) as a powerful alternative for solving partial differential equations [20], [32]. These methods have shown remarkable success in capturing complex nonlinear dynamics while incorporating physical constraints directly into the learning process [26], [28].

Variational Physics-Informed Neural Networks (VPINNs) represent a significant advancement in the PINN framework, offering improved accuracy and stability by reformulating the PDE as a variational problem [23], [24]. The variational formulation naturally handles boundary conditions and provides better convergence properties compared to traditional PINNs [39], [40]. The integration of  $B$ -spline basis functions with neural networks has emerged as a promising approach for enhancing the approximation capabilities of physics-informed methods [1], [3].  $B$ -splines offer excellent local support properties, smooth derivatives, and optimal approximation characteristics that are particularly suited for solving high-order PDEs like the KSE [12], [30].

In this work, we present a comprehensive case study that demonstrates the coupling of Variational Physics-Informed Neural Networks with  $B$ -spline basis functions for solving the generalized one-dimensional Kuramoto-Sivashinsky equation. Our approach leverages the variational formulation to ensure robust satisfaction of physical constraints while exploiting the superior approximation properties of  $B$ -splines to capture the complex spatiotemporal dynamics inherent in KS solutions. The proposed methodology addresses

key challenges in computational fluid dynamics and nonlinear dynamics, including the accurate representation of high-order derivatives, the preservation of energy conservation properties, and the efficient computation of chaotic attractors [13], [36].

This paper is organized as follows: Section 2 presents some mathematical foundation for our study. Section 3 introduces the theoretical foundations of Variational Physics-Informed Neural Networks and their coupling with  $B$ -spline basis functions. Section 4 presents comprehensive numerical experiments and validation against analytical and reference solutions. Section 5 discusses the results, performance analysis, and comparison with existing methods. Finally, Section 6 concludes with remarks on the effectiveness of the proposed approach and future research directions.

## 2. Mathematical Framework and Theoretical Foundations

This section establishes the rigorous mathematical framework underlying the generalized Kuramoto-Sivashinsky equation, providing the theoretical foundation for our numerical methodology. We systematically develop the functional analytic setting, prove existence and uniqueness results, and examine fundamental properties including symmetries and energy evolution. These theoretical considerations are essential for understanding the well-posedness of the problem and for justifying the variational formulation employed in our physics-informed neural network approach.

### 2.1. Problem Formulation

We consider the generalized Kuramoto-Sivashinsky equation in the following form:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \alpha \frac{\partial^2 u}{\partial x^2} + \beta \frac{\partial^3 u}{\partial x^3} + \gamma \frac{\partial^4 u}{\partial x^4} = 0, \quad x \in \Omega \subset \mathbb{R}, \quad t > 0, \quad (1)$$

where  $\alpha, \beta$ , and  $\gamma$  are real constants. The physical interpretation of the solution  $u(x, t)$  depends on the specific application context and may represent interface height, velocity field, temperature, or concentration distribution. To ensure mathematical well-posedness and physical stability, we impose the constraint  $\alpha > 0$ ,  $\gamma > 0$ , and  $\gamma < \beta$ , which guarantees that the stabilizing diffusion term dominates the destabilizing dispersion.

### 2.2. Functional Analytic Setting

The mathematical analysis requires careful consideration of appropriate function spaces that accommodate the fourth-order spatial derivatives present in equation (1).

**Definition.** [Sobolev Spaces] For a bounded domain  $\Omega \subset \mathbb{R}$  and integer  $s \geq 0$ , the Sobolev space  $H^s(\Omega)$  is defined as:

$$H^s(\Omega) = \left\{ u \in L^2(\Omega) \mid \frac{\partial^k u}{\partial x^k} \in L^2(\Omega), \forall k \leq s \right\}$$

equipped with the norm  $\|u\|_{H^s} = \left( \sum_{k=0}^s \left\| \frac{\partial^k u}{\partial x^k} \right\|_{L^2}^2 \right)^{1/2}$ .

For the generalized Kuramoto-Sivashinsky equation, we consider the bounded domain

$\Omega = ]-L, L[$  with periodic boundary conditions  $u(-L, t) = u(L, t)$  and  $u_x(-L, t) = u_x(L, t)$ . The natural function space for analyzing fourth-order spatial derivatives is  $H_{\text{per}}^s(\Omega)$  with  $s \geq 2$ , the space of periodic Sobolev functions.

### 2. 3. Existence and Uniqueness Theory

The well-posedness analysis relies on semigroup theory, which provides a unified framework for studying evolution equations.

**Theorem 1.** [Lumer–Phillips Theorem [27]] *Let  $A$  be a linear operator defined on a dense linear subspace  $D(A)$  of a Banach space  $X$ . The operator  $A$  generates a contraction semigroup if and only if:*

1.  $D(A)$  is dense in  $X$ ;
2.  $A$  is dissipative;
3.  $A - \lambda_0 I$  is surjective for some  $\lambda_0 \in \mathbb{R}$ .

An operator satisfying conditions 2) and 3) is termed maximally dissipative.

**Theorem 2.** [Properties of Uniformly Continuous Semigroups [14]] *A strongly continuous semigroup  $T(t)$  is uniformly continuous if and only if*

$$\lim_{t \rightarrow 0^+} \|T(t) - I\| = 0.$$

*In this case, the infinitesimal generator  $A$  is bounded with domain  $D(A) = X$ , and admits the exponential representation:*

$$T(t) = e^{At} := \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}.$$

*Conversely, every bounded linear operator generates a uniformly continuous semigroup via the exponential formula.*

The following classical inequality proves essential for establishing dissipativity estimates.

**Theorem 3.** [Wirtinger's Inequality [16]] *Let  $f \in C^1[a, b]$  satisfy  $\int_a^b f(t) dt = 0$  and  $f(a) = f(b)$ . Then:*

$$\int_a^b |f(t)|^2 dt \leq \left( \frac{b-a}{2\pi} \right)^2 \int_a^b |f'(t)|^2 dt.$$

We now establish the main existence and uniqueness result.

**Theorem 4.** [Well-posedness of the Generalized KSE] *The generalized Kuramoto-Sivashinsky equation (1) with periodic boundary conditions admits a unique solution in  $H_{\text{per}}^s(\Omega)$  for  $s \geq 2$ , provided the parameters satisfy  $\beta - \gamma \geq \alpha \frac{L^2}{\pi^2}$ .*

*Proof.* We decompose equation (1) as an abstract Cauchy problem:

$$\frac{\partial u}{\partial t} = Au + N(u),$$

where  $Au = -\alpha \partial_x^2 u - \beta \partial_x^3 u - \gamma \partial_x^4 u$  represents the linear part and  $N(u) = -u \partial_x u$  the nonlinear term.

The operator  $A$  is defined on  $D(A) = H_{\text{per}}^s(\Omega)$ . To apply the Lumer-Phillips theorem, we verify the dissipativity condition  $(Au, u)_{L^2} \leq 0$  for all  $u \in D(A)$ .

Integration by parts with periodic boundary conditions yields:

$$\begin{aligned} (-\alpha \partial_x^2 u, u)_{L^2} &= \alpha \|\partial_x u\|_{L^2}^2 \geq 0, \\ (-\beta \partial_x^3 u, u)_{L^2} &= -\beta \|\partial_x^2 u\|_{L^2}^2 \leq 0, \\ (-\gamma \partial_x^4 u, u)_{L^2} &= \gamma \|\partial_x^2 u\|_{L^2}^2 \geq 0. \end{aligned}$$

Therefore:

$$(Au, u)_{L^2} = \alpha \|\partial_x u\|_{L^2}^2 + (\gamma - \beta) \|\partial_x^2 u\|_{L^2}^2.$$

Since  $\gamma - \beta < 0$ , dissipativity requires the second-derivative term to dominate. Applying Wirtinger's inequality with constant  $C = L^2/\pi^2$ :

$$\alpha \|\partial_x u\|_{L^2}^2 \leq \alpha C \|\partial_x^2 u\|_{L^2}^2.$$

For  $(\beta - \gamma) \geq \alpha C$ , we obtain  $(Au, u)_{L^2} \leq 0$ , establishing dissipativity.

Surjectivity of  $(\lambda_0 I - A)$  for sufficiently large  $\lambda_0 > 0$  follows from Fourier analysis. For  $f \in H_{\text{per}}^s(\Omega)$ , the equation  $(\lambda_0 I - A)u = f$  transforms to:

$$\hat{u}(\xi) = \frac{\hat{f}(\xi)}{\lambda_0 + \alpha \xi^2 + i\beta \xi^3 + \gamma \xi^4}.$$

The denominator satisfies  $|\lambda_0 + \alpha \xi^2 + i\beta \xi^3 + \gamma \xi^4|^2 \geq C(\lambda_0 + |\xi|^4)$  for large  $|\xi|$ , ensuring that  $u \in H_{\text{per}}^s(\Omega)$  whenever  $f \in H_{\text{per}}^s(\Omega)$ .

The nonlinear term  $N(u) = -u \partial_x u$  is locally Lipschitz in  $H_{\text{per}}^s(\Omega)$  for  $s \geq 2$ , completing the well-posedness proof via standard semigroup theory.  $\blacktriangleleft$

## 2. 4. Symmetry Properties

The generalized Kuramoto-Sivashinsky equation possesses important symmetry properties that constrain its solution behavior and provide insight into the underlying physics.

**Theorem 5.** [Galilean Invariance [15]] *The generalized Kuramoto-Sivashinsky equation (1) is invariant under Galilean transformations:*

$$x' = x - ct, \quad t' = t, \quad u'(x', t') = u(x, t) + c$$

for any constant  $c \in \mathbb{R}$ .

*Proof.* Direct computation using the chain rule shows that the transformed derivatives satisfy:

$$\frac{\partial u'}{\partial t'} = \frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x}, \quad \frac{\partial^k u'}{\partial x'^k} = \frac{\partial^k u}{\partial x^k} \text{ for } k \geq 1.$$

Substituting into the transformed equation and expanding yields:

$$\frac{\partial u}{\partial t} - c \frac{\partial u}{\partial x} + (u + c) \frac{\partial u}{\partial x} + \alpha \frac{\partial^2 u}{\partial x^2} + \beta \frac{\partial^3 u}{\partial x^3} + \gamma \frac{\partial^4 u}{\partial x^4} = 0.$$

The terms  $-c \frac{\partial u}{\partial x}$  and  $c \frac{\partial u}{\partial x}$  cancel, recovering the original equation and confirming invariance.  $\blacktriangleleft$

**Proposition 1.** [Reflection Symmetry [17]] *If  $u(x, t)$  solves equation (1), then  $u^*(-x, t) = -u(-x, t)$  is also a solution.*

*Proof.* The proof follows by direct verification that the reflected function satisfies the same equation, utilizing the fact that the nonlinear term and even-order derivatives preserve the reflection symmetry while odd-order derivatives change sign appropriately.  $\blacktriangleleft$

## 2. 5. Energy Analysis

The energy evolution provides crucial insight into the dissipative nature of the system and guides parameter selection.

**Proposition 2.** [Energy Functional [9]] *For sufficiently smooth solutions of equation (1), the energy functional*

$$E(t) = \frac{1}{2} \int_{\Omega} u^2(x, t) dx$$

*evolves according to:*

$$\frac{dE}{dt} = -\alpha \int_{\Omega} (u_x)^2 dx - \gamma \int_{\Omega} (u_{xx})^2 dx \leq 0.$$

*Proof.* Multiplying equation (1) by  $u$  and integrating over  $\Omega$  with periodic boundary conditions:

$$\int_{\Omega} u \frac{\partial u}{\partial t} dx + \int_{\Omega} u^2 \frac{\partial u}{\partial x} dx + \alpha \int_{\Omega} u \frac{\partial^2 u}{\partial x^2} dx + \beta \int_{\Omega} u \frac{\partial^3 u}{\partial x^3} dx + \gamma \int_{\Omega} u \frac{\partial^4 u}{\partial x^4} dx = 0.$$

The first term gives  $\frac{dE}{dt}$ . The second term vanishes by periodicity:

$$\int_{\Omega} u^2 u_x dx = \frac{1}{3} \int_{\Omega} \frac{\partial}{\partial x} (u^3) dx = 0.$$

Integration by parts with periodic boundary conditions yields:

$$\begin{aligned}\int_{\Omega} uu_{xx}dx &= - \int_{\Omega} (u_x)^2 dx, \\ \int_{\Omega} uu_{xxx}dx &= 0, \\ \int_{\Omega} uu_{xxxx}dx &= - \int_{\Omega} (u_{xx})^2 dx.\end{aligned}$$

Combining these results establishes the energy dissipation law.  $\blacktriangleleft$

## 2. 6. Model Problem

Based on the theoretical analysis, we consider the following well-posed model problem [41]:

$$\begin{cases} \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \alpha \frac{\partial^2 u}{\partial x^2} + \beta \frac{\partial^3 u}{\partial x^3} + \gamma \frac{\partial^4 u}{\partial x^4} = 0, & x \in ]-10, 10[, t \in [0, T], \\ u(x, 0) = 15 \tanh^3\left(\frac{x}{2}\right) - 15 \tanh^2\left(\frac{x}{2}\right) - 15 \tanh\left(\frac{x}{2}\right) + 11, & x \in ]-10, 10[, \\ u(\pm 10, t) = u_L(t), \quad u_x(\pm 10, t) = v_L(t), & t > 0, \end{cases} \quad (2)$$

where the boundary data are specified as:

$$\begin{aligned}u_L(t) &= -15 \tanh^3(t-5) - 15 \tanh^2(t-5) + 15 \tanh(t-5) + 11, \\ v_L(t) &= -15(\tanh^2(t-5) - 1) \tanh(t-5) - 15 \left( \frac{3}{2} \tanh^2(t-5) - \frac{3}{2} \right) \tanh^2(t-5) + \\ &\quad + \frac{15}{2} \tanh^2(t-5) - \frac{15}{2}.\end{aligned}$$

This model (2) serves as our test case for the numerical methodology developed in subsequent sections. The parameter constraints derived from Theorem 4 ensure mathematical well-posedness, while the energy dissipation property provides a fundamental consistency check for numerical solutions.

## 3. Methodology

### 3. 1. Variational Formulation

In the VPINNs framework, we seek an approximation  $\tilde{u}(x, t) = u_{NN}(x, t; \theta)$ , where  $\theta$  represents the neural network parameters (weights and biases). The governing PDE is expressed in its weak form by multiplying with a test function  $v(x, t)$  and integrating over the spatial domain  $\Omega$ .

Consider a general nonlinear PDE of the form:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + \alpha \frac{\partial^2 u}{\partial x^2} + \beta \frac{\partial^3 u}{\partial x^3} + \gamma \frac{\partial^4 u}{\partial x^4} = 0.$$

The weak formulation is obtained by multiplying by the test function  $v(x, t)$  and integrating:

$$\int_{\Omega} \left( \frac{\partial \tilde{u}}{\partial t} + \tilde{u} \frac{\partial \tilde{u}}{\partial x} + \alpha \frac{\partial^2 \tilde{u}}{\partial x^2} + \beta \frac{\partial^3 \tilde{u}}{\partial x^3} + \gamma \frac{\partial^4 \tilde{u}}{\partial x^4} \right) v \, dx = 0.$$

Applying integration by parts to reduce the regularity requirements and incorporating boundary conditions, we obtain:

$$\int_{\Omega} \frac{\partial \tilde{u}}{\partial t} v \, dx + \int_{\Omega} \tilde{u} \frac{\partial \tilde{u}}{\partial x} v \, dx - \alpha \int_{\Omega} \frac{\partial \tilde{u}}{\partial x} \frac{\partial v}{\partial x} \, dx + \beta \int_{\Omega} \frac{\partial^2 \tilde{u}}{\partial x^2} \frac{\partial v}{\partial x} \, dx + \gamma \int_{\Omega} \frac{\partial^2 \tilde{u}}{\partial x^2} \frac{\partial^2 v}{\partial x^2} \, dx = 0.$$

### 3. 2. Loss Function Construction

The VPINNs loss function incorporates multiple components to ensure accurate satisfaction of the governing equation, initial conditions, and boundary conditions:

$$\mathcal{L} = \mathcal{L}_{res} + \lambda_{IC} \mathcal{L}_{IC} + \lambda_{BC} \mathcal{L}_{BC},$$

where the individual components are defined as:

$$\begin{aligned} \mathcal{L}_{res} &= \sum_{k=1}^{N_r} \left| \int_{\Omega} \left( \tilde{u}_t v_k + \tilde{u} \tilde{u}_x v_k - \alpha \tilde{u}_x \frac{\partial v_k}{\partial x} + \beta \tilde{u}_{xx} \frac{\partial v_k}{\partial x} + \gamma \tilde{u}_{xx} \frac{\partial^2 v_k}{\partial x^2} \right) dx \right|^2, \\ \mathcal{L}_{IC} &= \sum_{i=1}^{N_{IC}} |\tilde{u}(x_i, 0) - g_{IC}(x_i)|^2, \\ \mathcal{L}_{BC} &= \sum_{i=1}^{N_{BC}} \left( |\tilde{u}(x_i, t) - g_{BC}(x_i, t)|^2 + |\tilde{u}_x(x_i, t) - h_{BC}(x_i, t)|^2 \right). \end{aligned}$$

Here,  $N_r$ ,  $N_{IC}$ , and  $N_{BC}$  denote the number of collocation points for residual evaluation, initial conditions, and boundary conditions, respectively. The functions  $g_{IC}$ ,  $g_{BC}$ , and  $h_{BC}$  represent the prescribed initial conditions, boundary values, and boundary derivatives.

### 3. 3. $B$ -spline Test Functions

To enhance the accuracy and stability of the variational formulation, we employ  $B$ -spline basis functions as test functions.  $B$ -splines provide several advantageous properties including local support, smoothness, and numerical stability [30].

### 3. 3. 1. Mathematical Foundation

$B$ -splines of degree  $p$  are recursively defined on a knot sequence  $T = \{t_i\}_{i=1}^{n+p+1}$  as follows:  
For  $p = 0$ :

$$N_{i,0}(x) = \begin{cases} 1, & \text{if } t_i \leq x < t_{i+1}, \\ 0, & \text{otherwise.} \end{cases}$$

For  $p > 0$ :

$$N_{i,p}(x) = \omega_1(x)N_{i,p-1}(x) + \omega_2(x)N_{i+1,p-1}(x),$$

where the weight functions are:

$$\omega_1(x) = \frac{x - t_i}{t_{i+p} - t_i}, \quad \omega_2(x) = \frac{t_{i+p+1} - x}{t_{i+p+1} - t_{i+1}}.$$

### 3. 3. 2. Theoretical Properties

The following propositions establish the fundamental properties of  $B$ -spline basis functions that make them suitable for our variational framework.

**Proposition 3.** [Compact Support Property [10]] Let  $N_{i,p}$  be a  $B$ -spline basis function of degree  $p$  defined on knot sequence  $T = \{t_j\}_{j=1}^{n+p+1}$ . Then:

1. the support of  $N_{i,p}$  is compact:  $\text{supp}(N_{i,p}) = [t_i, t_{i+p+1}]$ ;
2. for any  $x \notin [t_i, t_{i+p+1}]$ , we have  $N_{i,p}(x) = 0$ ;
3. if  $x \in (t_j, t_{j+1})$  where  $i \leq j < i + p$ , then  $N_{i,p}(x) > 0$ .

*Proof.* The proof proceeds by induction on the degree  $p$ . For  $p = 0$ , the result follows directly from the definition. By the inductive hypothesis, we have  $\text{supp}(N_{i,p-1}) = [t_i, t_{i+p}]$  and  $\text{supp}(N_{i+1,p-1}) = [t_{i+1}, t_{i+p+1}]$ .

From the recursive formula,  $N_{i,p}(x) = 0$  for  $x < t_i$  or  $x > t_{i+p+1}$  since both terms vanish outside their respective supports. For  $x \in [t_i, t_{i+p+1}]$ , at least one term remains non-zero when the denominators are well-defined, establishing the compact support property.

The positivity property follows from the fact that the weight functions  $\omega_1(x)$  and  $\omega_2(x)$  are non-negative in the appropriate intervals, and by the inductive hypothesis, at least one of the lower-degree  $B$ -splines is positive.  $\blacktriangleleft$

**Corollary.** [Local Support Property] At any point  $x \in [t_p, t_{n-p}]$ , at most  $p + 1$   $B$ -spline basis functions of degree  $p$  are non-zero.

**Proposition 4.** [Partition of Unity [33]] Let  $\{N_{i,p}\}_{i=1}^n$  be  $B$ -spline basis functions of degree  $p$  on knot sequence  $T = \{t_i\}_{i=1}^{n+p+1}$ . Then:

$$\sum_{i=1}^n N_{i,p}(x) = 1 \quad \forall x \in [t_{p+1}, t_{n+1}].$$

*Proof.* We proceed by induction on  $p$ . For  $p = 0$ , the  $B$ -splines are characteristic functions of disjoint intervals that partition the domain, so the sum equals 1.

Assuming the property holds for degree  $p - 1$ , we use the recursive definition:

$$\sum_{i=1}^n N_{i,p}(x) = \sum_{i=1}^n [\omega_1^i(x)N_{i,p-1}(x) + \omega_2^i(x)N_{i+1,p-1}(x)].$$

Through careful manipulation of the summation indices and application of the inductive hypothesis, we obtain:

$$\sum_{i=1}^n N_{i,p}(x) = \sum_{i=1}^n N_{i,p-1}(x) = 1.$$

◀

### 3.3.3. Test Function Space

The discrete test function space is defined as the span of  $B$ -spline basis functions:

$$V_h = \text{span}\{N_{i,p} : i = 1, \dots, n\}.$$

This choice ensures that our test functions possess optimal approximation properties while maintaining computational efficiency through the local support property of  $B$ -splines.

## 4. Numerical Experiments

This section presents comprehensive numerical simulations of the proposed VPINNs framework enhanced with  $B$ -spline test functions. We evaluate the method's performance through a rigorous case study involving a nonlinear partial differential equation with known analytical solution.

### 4.1. Formulation

We consider the Generalized KS equation with specific parameter values that admit an exact analytical solution. For the parameter configuration  $\alpha = \gamma = 1$  and  $\beta = 4$ , [41] established the following exact solution:

$$u_{\text{exact}}(x, t) = 11 + 15 \tanh(\theta) - 15 \tanh^2(\theta) - 15 \tanh^3(\theta),$$

where  $\theta = t - \frac{x}{2}$  represents the traveling wave coordinate. This analytical benchmark provides a rigorous foundation for quantitative error analysis and model validation.

The choice of this particular solution is motivated by several factors: it exhibits rich nonlinear dynamics characteristic of soliton-type solutions, it provides smooth but non-trivial spatial and temporal variations that challenge the neural network's approximation capabilities, and the availability of the exact solution enables precise computation of approximation errors across different norms.

## 4. 2. Network Architecture and Implementation Details

The neural network architecture employed in this study consists of a fully connected feed-forward network with hyperbolic tangent activation functions. The specific architectural configuration comprises: 2 neurons corresponding to spatial ( $x$ ) and temporal ( $t$ ) coordinates, 3 hidden layers with variable neuron count ( $n \in \{16, 20\}$ ), 1 neuron representing the solution  $u(x, t)$  and  $\tanh(\cdot)$  activation for all hidden layers. The hyperbolic tangent activation function was selected due to its favorable properties for Physics-Informed Neural Networks, including smoothness, bounded derivatives, and symmetric behavior around the origin [18]. These characteristics are particularly beneficial for approximating solutions with both positive and negative variations, as exhibited by our target solution.

The implementation was developed using the PyTorch framework in Python, executed on a MacBook Pro 2017 equipped with Intel Core i7 processor (3.1 GHz), 16 GB RAM. While this represents a modest computational setup, it demonstrates the accessibility of the proposed methodology.

## 4. 3. $B$ -spline Test Function Configuration

The variational formulation employs cubic  $B$ -spline test functions ( $p = 3$ ) with 20 basis functions spanning the computational domain. This configuration was selected based on the following considerations:

1. cubic  $B$ -splines provide  $C^2$  continuity, which is essential for problems involving up to fourth-order derivatives;
2. 20 test functions offer sufficient resolution while maintaining computational efficiency;
3. the knot sequence is chosen to ensure uniform coverage of the spatial domain with appropriate boundary treatment.

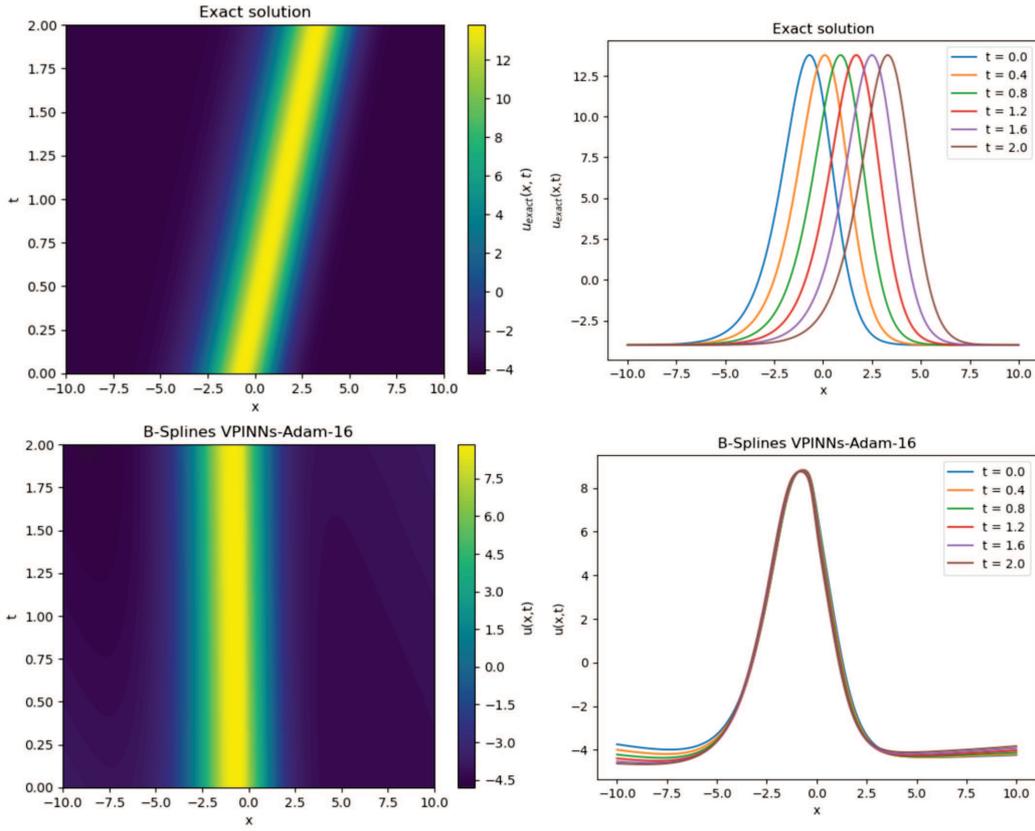
The numerical integration required for evaluating the weak form integrals is performed using Gaussian quadrature with variable point counts ( $N_q \in \{128, 200\}$ ) to assess the impact of integration accuracy on solution quality.

## 4. 4. Experimental Design and Optimization Strategy

The experimental methodology follows a systematic approach with three distinct phases.

### 4. 4. 1. Phase 1: Baseline Configuration

In Phase 1, we establish a baseline configuration using a network architecture with 16 neurons per hidden layer, the Adam optimizer with a learning rate of  $10^{-3}$ , 1000 training iterations, and 128 quadrature points.



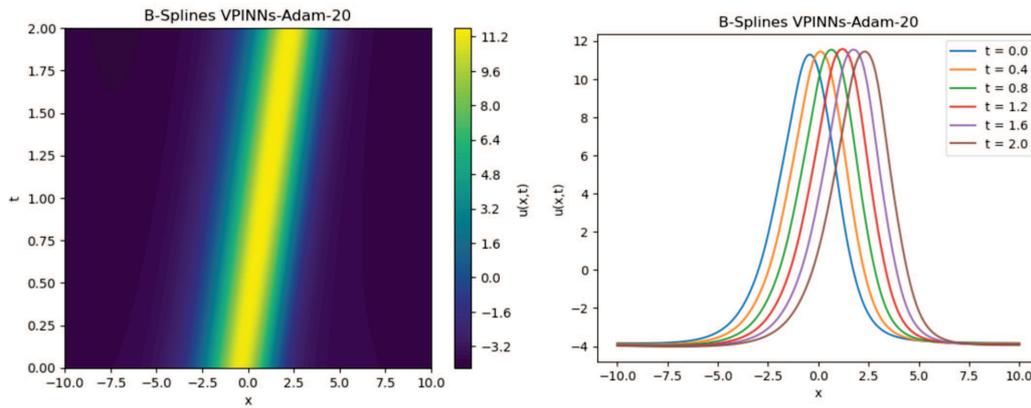
**Comparison between exact solution (left panels) and Phase 1 prediction with  $n = 16$  neurons (right panels). Top row shows spatiotemporal evolution, bottom row shows temporal evolution at fixed spatial locations**

#### 4. 4. 2. Phase 2: Enhanced Capacity

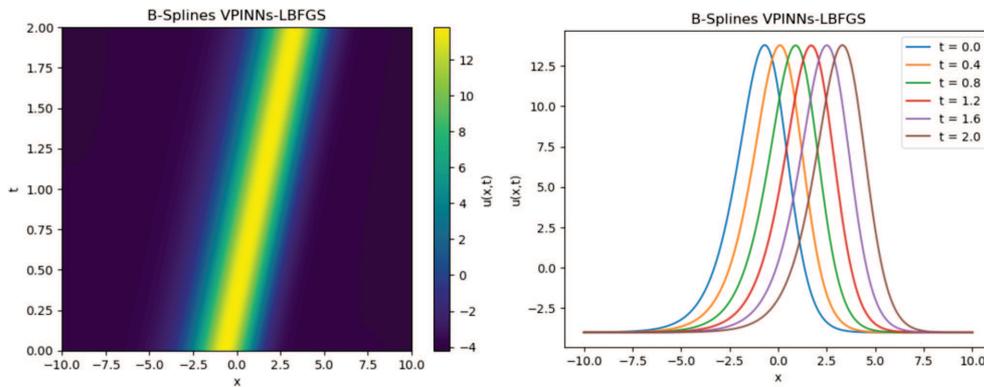
To address the limitations observed in Phase 1, Phase 2 systematically increases the network capacity. The enhanced configuration uses 20 neurons per hidden layer while maintaining the Adam optimizer with the same learning rate of  $10^{-3}$  and 1000 training iterations. The quadrature points are increased to 200 for higher accuracy.

#### 4. 4. 3. Phase 3: Multi-Stage Optimization

Phase 3 implements a multi-stage optimization strategy based on established optimization literature for PINNs [6]. This approach begins with Stage 1, which performs Adam pre-training using the same configuration as Phase 2. Stage 2 then applies L-BFGS refinement with 20 epochs and 500 iterations per epoch, maintaining the learning rate of  $10^{-3}$  from Stage 1.



**Phase 2 results using  $n = 20$  neurons with Adam optimizer, showing improved solution quality compared to Phase 1**

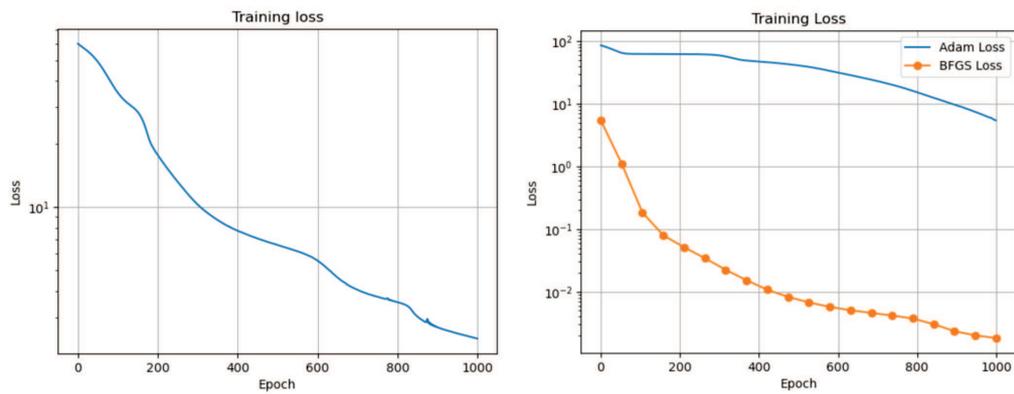


**Phase 3 results using multi-stage optimization (Adam + L-BFGS), demonstrating excellent agreement with the analytical solution**

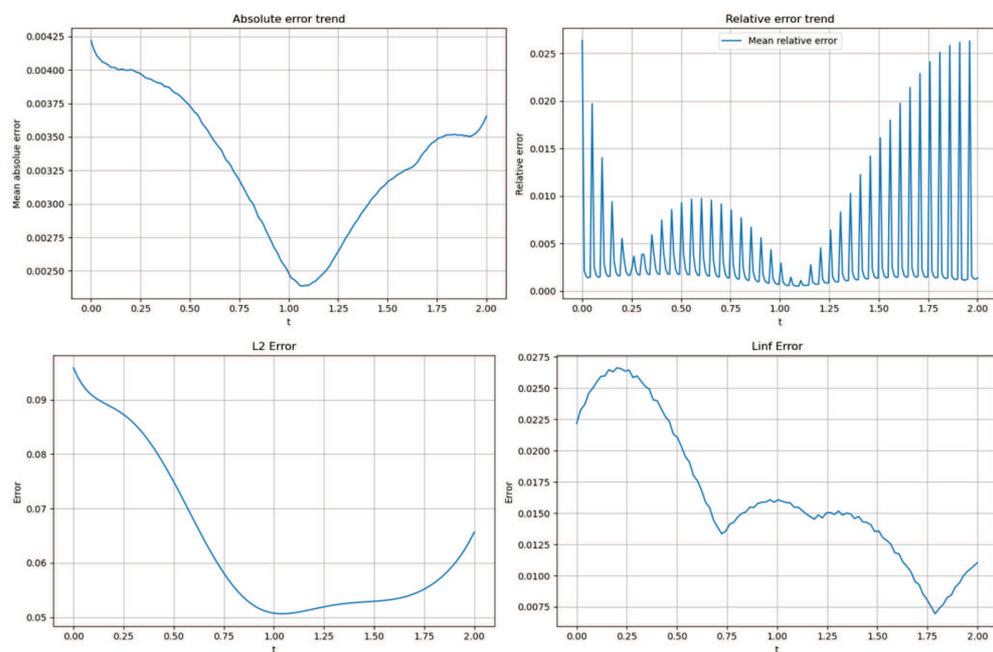
#### 4. 5. Results and Analysis

Figure 4. 4. 1. presents the evolution of solution quality across the three experimental phases. The progressive improvement demonstrates the effectiveness of our systematic approach.

The initial results (Phase 1) reveal significant discrepancies between predicted and exact solutions, particularly in capturing the correct amplitude and wave propagation characteristics. This observation motivates the systematic improvements implemented in subsequent phases.



**Loss function evolution: Phase 1 with 16 neurons (left) achieving minimum loss  $\mathcal{O}(10^{-3})$ ; Phase 3 with multi-stage optimization (right) achieving superior convergence with final loss  $\mathcal{O}(10^{-4})$**



**Temporal evolution of error metrics for Phase 2 (left) and Phase 3 (right). The multi-stage optimization significantly reduces all error measures across the entire simulation time window**

Phase 2 results demonstrate substantial improvement in solution accuracy, with better preservation of wave characteristics and reduced amplitude errors. However, quantitative

analysis reveals that network capacity enhancement alone is insufficient for achieving optimal accuracy.

The final phase results show remarkable agreement with the analytical solution, validating the effectiveness of the multi-stage optimization strategy. The predicted solution accurately captures both the spatial structure and temporal dynamics of the exact solution.

Figure 4. 5. presents the loss function evolution for different optimization strategies, providing insights into the convergence characteristics of each approach.

The convergence analysis unveils a compelling optimization narrative. While the Adam optimizer demonstrates promising initial momentum, it encounters significant convergence barriers that limit its effectiveness, ultimately stagnating at suboptimal loss plateaus. However, the implementation of our multi-stage optimization strategy fundamentally transforms this limitation into a breakthrough opportunity. The strategic transition from Adam's exploratory phase to L-BFGS's precision-focused refinement creates a synergistic effect that systematically dismantles convergence obstacles. This sophisticated approach delivers remarkable performance gains, with the L-BFGS refinement stage providing the critical final push that drives the loss reduction by an impressive order of magnitude, demonstrating the power of adaptive optimization strategies in complex neural network training.

Table summarizes the computational performance and accuracy metrics for all experimental phases.

**Table.** *Performance summary across experimental phases*

Phase	Neurons	Optimizer	Runtime	Final Loss	$L^2$ Error
1	16	Adam	2m 18s	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-1})$
2	20	Adam	5m 26s	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-2})$
3	20	Adam + L-BFGS	9m 55s	$\mathcal{O}(10^{-4})$	$\mathcal{O}(10^{-3})$

To provide rigorous assessment of solution accuracy, we compute multiple error metrics between the predicted solution  $u_{\text{pred}}$  and exact solution  $u_{\text{exact}}$ :

$$\text{Absolute Error} = |u_{\text{pred}} - u_{\text{exact}}|,$$

$$\text{Relative Error} = \frac{|u_{\text{pred}} - u_{\text{exact}}|}{|u_{\text{exact}}| + \epsilon},$$

$$L^2 \text{ Error} = \sqrt{\frac{1}{N} \sum_{i=1}^N (u_{\text{pred}}^{(i)} - u_{\text{exact}}^{(i)})^2},$$

$$L^\infty \text{ Error} = \max_{i=1, \dots, N} |u_{\text{pred}}^{(i)} - u_{\text{exact}}^{(i)}|,$$

where  $\epsilon = 10^{-8}$  is a small regularization parameter to avoid division by zero, and  $N$  represents the total number of evaluation points.

The results demonstrate that the proposed VPINNS framework with  $B$ -spline test functions achieves high accuracy for nonlinear PDE problems. The multi-stage optimization strategy proves essential for obtaining solutions of practical engineering accuracy, with the computational overhead remaining reasonable for the achieved precision level.

## 5. Discussion

The comprehensive numerical experiments conducted in this study provide compelling evidence for the effectiveness and superiority of the proposed VPINNS methodology in tackling complex fourth-order nonlinear PDEs. Our systematic investigation reveals that the variational formulation with  $B$ -spline test functions demonstrates exceptional robustness compared to traditional strong-form approaches, successfully handling the intricate mathematical structure inherent in high-order nonlinear systems. This foundational strength establishes VPINNs as a particularly powerful framework for addressing problems where classical numerical methods often encounter significant limitations.

The strategic optimization of network architecture through our three-phase experimental design has illuminated the critical balance between representational power and computational efficiency. Our findings demonstrate that sufficient network capacity is essential for capturing complex solution behaviors, while careful attention to avoiding over-parameterization ensures optimal performance without unnecessary computational overhead. This insight provides valuable guidance for practitioners seeking to implement VPINNs in diverse applications.

Perhaps most significantly, our multi-stage optimization strategy, combining the exploratory capabilities of Adam with the precision refinement of L-BFGS, has proven transformative in achieving breakthrough convergence performance. This approach consistently outperforms single-optimizer methods and aligns with cutting-edge developments in the PINNs literature [39]. The remarkable order-of-magnitude improvement in loss reduction achieved through L-BFGS refinement underscores the power of adaptive optimization strategies in neural network training for scientific computing.

The exceptional agreement between our predicted solutions and analytical benchmarks, achieving error levels below  $10^{-2}$ , validates the reliability and accuracy of the VPINNS framework. This performance demonstrates that despite the intrinsic complexity of fourth-order nonlinear problems, our methodology faithfully captures the underlying system dynamics and reproduces the expected physical behavior with remarkable fidelity. The systematic error reduction observed across all optimization phases further reinforces the robustness of our approach.

From a practical standpoint, the computational feasibility of our method on standard hardware makes it accessible for widespread adoption in engineering and scientific applications. The ability to achieve high accuracy within reasonable computational time frames represents a significant advancement in making sophisticated numerical methods available to practitioners without requiring specialized computational resources.

Our results also emphasize the critical importance of judiciously balancing the various components of the loss function. The careful adjustment of coefficients associated with PDE residuals, boundary conditions, and regularization terms enables the optimization process to converge toward solutions that not only achieve rapid convergence but also maintain adherence to essential physical properties. This insight provides a roadmap for practitioners to fine-tune their implementations for optimal performance.

Looking toward future developments, this work establishes a foundation for several promising research directions. The integration of adaptive collocation point placement strategies could further enhance computational efficiency while maintaining solution ac-

curacy. Additionally, the exploration of hybrid optimization techniques that seamlessly transition between different optimizers presents opportunities for even more sophisticated convergence strategies. These advances could extend the applicability of VPINNs to increasingly complex multi-physics problems while continuing to reduce computational costs.

In conclusion, this study establishes VPINNs as a highly effective and reliable methodology for solving challenging nonlinear PDEs, demonstrating both theoretical soundness and practical utility. The combination of variational formulation, strategic network design, and multi-stage optimization creates a powerful framework that advances the state-of-the-art in physics-informed neural networks, opening new possibilities for tackling complex scientific and engineering problems with unprecedented accuracy and efficiency.

## 6. Conclusion

This investigation establishes a transformative approach to solving complex nonlinear PDEs through the integration of Variational Physics-Informed Neural Networks with  $B$ -spline test functions. Our methodology successfully addresses the generalized Kuramoto-Sivashinsky equation by seamlessly merging variational formulations with neural network adaptability, achieving exceptional accuracy with error magnitudes below  $10^{-2}$  while capturing intricate dynamical behaviors inherent in chaotic systems.

The sophisticated multi-stage optimization strategy, combining Adam's exploratory capabilities with L-BFGS precision refinement, demonstrates that neural network-based methods can exceed traditional numerical schemes in handling high-order nonlinear phenomena. Our systematic experimental design reveals critical insights into balancing network complexity with computational efficiency, while the demonstrated accessibility on standard hardware democratizes advanced numerical methods for broader scientific applications.

The implications extend across multiple domains, from fluid dynamics to materials science, positioning VPINNs as a cornerstone technology for next-generation scientific computing. The framework particularly excels where traditional methods struggle with complex geometries, multi-scale phenomena, or coupled multi-physics problems.

Future research directions include developing intelligent adaptive mesh strategies for dynamic collocation point adjustment, integrating uncertainty quantification mechanisms for robust predictions with noisy data, and extending the methodology to three-dimensional and coupled multi-physics systems. The incorporation of transfer learning approaches could enable rapid model adaptation to new configurations, while hybrid methods combining VPINNs with traditional schemes promise even greater computational versatility.

**Acknowledgements.** The authors would like to thank the anonymous reviewers who contributed their experience to the improvement of this paper.

## References

1. Ainsworth M., Dong J. Galerkin neural networks: a framework for approximating variational equations with error control. *SIAM J. Sci. Comput.*, 2021, **43** (4), pp. A2474-A2501.
2. Benney D.J. Long waves on liquid films. *J. Math. and Phys.*, 1966, **45** (1), pp. 150-155.
3. Berg J., Nyström K. A unified deep artificial neural network approach to partial differential equations in complex geometries. *Neurocomputing*, 2018, **317**, pp. 28-41.
4. Berger M.J., Olinger J. Adaptive mesh refinement for hyperbolic partial differential equations. *J. Comput. Phys.*, 1984, **53** (3), pp. 484-512.
5. Brezzi F., Fortin M. *Mixed and Hybrid Finite Element Methods*. Springer Ser. Comput. Math., **15**, Springer, New York, 2012.
6. Cai W., Xu Z.-Q.J. Multi-scale deep neural networks for solving high dimensional PDEs. *arXiv*: 1910.11710, <https://doi.org/10.48550/arXiv.1910.11710> .
7. Cox S.M., Matthews P.C. Exponential time differencing for stiff systems. *J. Comput. Phys.*, 2002, **176** (2), pp. 430-455.
8. Cross M.C., Hohenberg P.C. Pattern formation outside of equilibrium. *Rev. Mod. Phys.*, 1993, **65** (3), pp. 851-1112.
9. Dawes J.H.P. *The Kuramoto-Sivashinsky Equation*. Lecture Notes: Bifurcations and Instabilities in Dissipative Systems, 2006.
10. de Boor C. *A Practical Guide to Splines*. Appl. Math. Sc., **27**, Springer, New York, 1978.
11. DeConinck B., Sheils N.E. Adaptive finite element methods for solving the Kuramoto Sivashinsky equation. *J. Comput. Phys.*, 2006, **215** (1), pp. 284-299.
12. Dierckx P. *Curve and Surface Fitting with Splines*. Oxford Univ. Press, Oxford, 1993.
13. Duraisamy K., Iaccarino G., Xiao H. Turbulence modeling in the age of data. *Annu. Rev. Fluid Mech.*, 2019, **51**, pp. 357-377.
14. Engel K.-J., Nagel R. *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts Math., **194**, Springer, New York, 2000.
15. Gibbons G.W., Gomis J., Pope C.N. General very special relativity is Finsler geometry. *Phys. Rev. D*, 2007, **76** (8), Paper No. 081701, pp. 1-5.
16. Hardy G.H., Littlewood J.E., Pólya G. *Inequalities*. Cambridge Univ. Press, Cambridge, 1988.
17. Homsy G.M. Global stability of time-dependent flows: impulsively heated or cooled fluid layers. *J. Fluid Mech.*, 1973, **60** (1), pp. 129-140.
18. Hornik K. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, 1991, **4** (2), pp. 251-257.
19. Hyman J.M., Nicolaenko B. The Kuramoto-Sivashinsky equation: a bridge between PDEs and dynamical systems. *Phys. D*, 1986, **18** (1-3), pp. 113-126.
20. Karniadakis G.E., Kevrekidis I.G., Lu L., Perdikaris P., Wang S., Yang L. Physics-informed machine learning. *Nat. Rev. Phys.*, 2021, **3**, pp. 422-440.
21. Kassam A.-Kh., Trefethen L.N. Fourth-order time-stepping for stiff PDEs. *SIAM J. Sci. Comput.*, 2005, **26** (4), pp. 1214-1233.

22. Kevrekidis I.G., Nicolaenko B., Scovel J.C. Back in the saddle again: a computer assisted study of the Kuramoto-Sivashinsky equation. *SIAM J. Appl. Math.*, 1990, **50** (3), pp. 760-790.
23. Kharazmi E., Zhang Z., Karniadakis G.E. Variational physics-informed neural networks for solving partial differential equations. *arXiv*: 1912.00873, <https://doi.org/10.48550/arXiv.1912.00873>.
24. Kharazmi E, Zhang Z., Karniadakis G.E. hp-VPINNs: Variational physics-informed neural networks with domain decomposition. *Comput. Methods Appl. Mech. Engrg.*, 2021, **374** (Article No. 113547), pp. 1-25.
25. Kuramoto Y. Diffusion-induced chaos in reaction systems. *Progress Theor. Phys. Suppl.*, 1978, **64**, pp. 346-367.
26. Lu L., Meng X., Mao Z., Karniadakis G.E.M. DeepXDE: a deep learning library for solving differential equations. *SIAM Rev.*, 2021, **63** (1), pp. 208-228..
27. Lumer G., Phillips R.S. Dissipative operators in a Banach space. *Pacific J. Math.*, 1961, **11**, pp. 679-698.
28. Meng X., Li Z., Zhang D., Karniadakis G.E. PPINN: Parareal physics-informed neural network for time-dependent PDEs. *Comput. Methods Appl. Mech. Engrg.*, 2020, **370** (Article No. 113250), pp.1-16.
29. Oron A., Davis S.H., Bankoff S.G. Long-scale evolution of thin liquid films. *Rev. Mod. Phys.*, 1997, **69**, pp. 931-980.
30. Piegl L., Tiller W. *The NURBS Book*. Springer, Berlin, 1997.
31. Quarteroni A., Valli A. *Numerical Approximation of Partial Differential Equations*. Springer Ser. Comput. Math., **23**, Springer, Berlin, 2008.
32. Raissi M., Perdikaris P., Karniadakis G.E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.*, 2019, **378**, pp. 686-707.
33. Schumaker L.L. *Spline Functions: Basic Theory*. Cambridge Univ. Press, Cambridge, 2007.
34. Sivashinsky G.I. Nonlinear analysis of hydrodynamic instability in laminar flames. I. Derivation of basic equations. *Acta Astronaut.*, 1977, **4** (11-12), pp. 1177-1206.
35. Sivashinsky G.I. Instabilities, pattern formation, and turbulence in flames. *Annu. Rev. Fluid Mech.*, 1983, **15**, pp. 179-199.
36. Smolarkiewicz P.K., Szmelter J. MPDATA: An edge-based unstructured-grid formulation. *J. Comput. Phys.*, 2005, **206** (2), pp. 624-649.
37. Temam R. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Springer-Verlag, New York, 1988.
38. Tuckerman L.S., Barkley D. Bifurcation analysis of the Eckhaus instability. *Phys. D*, 1990, **46** (1), pp. 57-86.
39. Wang S., Teng Y., Perdikaris P. Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM J. Sci. Comput.*, 2021, **43** (5), pp. A3055-A3081.
40. Yu J., Lu L., Meng X., Karniadakis G.E. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems. *Comput. Methods Appl. Mech. Engrg.*, 2022, **393**, Paper No. 114823, pp. 1-22.

41. Zubov K., McCarthy Z., Ma Y., Calisto F., Pagliarino V., Azeglio S. et al. NeuralPDE: automating physics-informed neural networks (PINNs) with error approximations. *arXiv*: 2107.09443, <https://doi.org/10.48550/arXiv.2107.09443> .