RESEARCH STATEMENT

JUNHO CHOI
junho_choi@kaist.ac.kr

Eureka moment. That is one of the fundamental motivations for studying mathematics. Specifically, we can encounter eureka moments by identifying unknown solutions of partial differential equations (PDEs) from known data. Accordingly, my research topics are: 1) machine learning methodology for fast and general-purpose PDE solvers; and 2) the mathematical modeling and simulation of ocean wave dynamics.

My first research topic is dedicated to developing unsupervised operator learning networks for generating the rapid and accurate solution of parametric PDEs, including the 3D Navier-Stokes equations (see Sec.1). Once trained by the structure of a target PDE, rather than by reference solutions to the PDE, the operator learning networks are capable of assigning the corresponding solutions to given PDE data in nearly real-time. Accordingly, this capability is useful for generating massive datasets to train physical artificial intelligences within a digital twin environment. Furthermore, this framework allows for real-time ensemble computing across various scenarios, enabling efficient, probabilistic forecasting in weather and climate modeling.

The second topic involves the mathematical modeling and simulation of ocean wave dynamics, specifically focusing on highly non-linear phenomena like rogue waves (see Sec.2). Building upon the Euler equations, I have introduced approximate PDE hierarchy and performed simulations using the approximate PDEs to realize rogue wave dynamics. This computational work is expected to serves as a key clue for justifying the relations among ocean wave models, and constructing robust, real-time forecasting and warning systems.

1. **Operator learning networks for solving parametric partial differential equations**

1.1. **Background on machine learning framework as a PDE solver.** Conventionally, numerical schemes, for example FEM, FVM, FDM, have long been the standard for calculating PDE solutions. While the demand has recently soared for generating large-scale solutions, such as both for building digital twins, and for conducting resource-intensive ensemble computing, conventional schemes have suffered a bottleneck to perform these computations. Accordingly, it has been driving the need for alternatives like operator learning networks. An operator learning network is a machine learning framework designed to emulate a PDE operator which assigns the solutions to given PDE data such as initial conditions, boundary

Date: December 14, 2025.

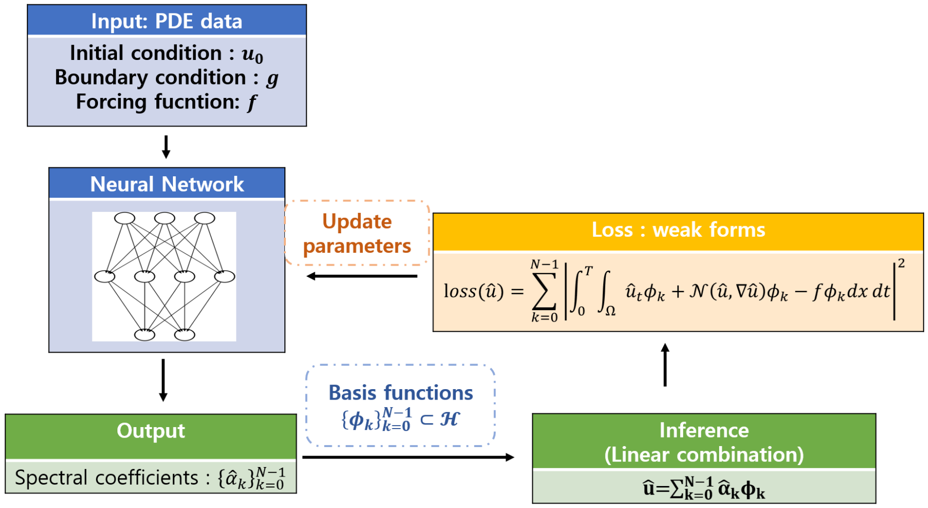

FIGURE 1. **Schematic diagram upon the training procedure of SON for emulating a PDE operator** SON is designed to map given PDE data to a set of spectral coefficients. These coefficients are transformed into an inference solution defined by $\widehat{u} := \sum_{k=0}^{N-1} \widehat{\alpha}_k \phi_k$ with a set of basis functions in H. While the loss function derived from the weak form of (1.1) is minimized, the network is trained, driving the inference toward the true solution.

conditions, or forcing functions. Thanks to rapid advances in GPU performance for parallel tensor computation and foundational theories like the Universal Approximation Theorem (UAT) for operators, frameworks of operator learning network have been developed as a powerful alternative PDE solver for immediately generating numerous solutions. Specifically, the Fourier Neural Operator (FNO) [10] and Deep Operator Network (DeepONet) [11] are widely adopted as operator learning networks. Once trained, an operator learning network can generate multiple solutions in a fraction of the time. However, their accuracy is contingent upon having a sufficient number of reference solutions for training, which necessitates substantial initial computational costs outside of the network's training and inference process. This dependency on large, high-fidelity reference solutions remains a critical obstacle to their application. In contrast, Physics-Informed Neural Networks (PINNs) [13] have emerged to address this data dependency by using the residual of a governing PDE, thereby eliminating reliance on reference solutions for training. The key limitation of PINNs, however, is their inability to solve parametric problems efficiently since the network is designed to solve only a single instance of a PDE at a time, it must be fully retrained for every new set of PDE data.

1.2. **Research achievements.** To overcome these limitations, I have advanced a novel unsupervised operator learning network based on spectral element methods, called Spectral Operator Network (SON)[4, 6, 7].

(1.1)

$$\begin{cases} f = u_t + \mathcal{N}(u, \nabla u) & \text{in } (0,T]\times\Omega, \\ u_0 = u & \text{for } t = 0, \Longrightarrow \int_0^T \int_\Omega u_t \phi dx dt + \int_0^T \int_\Omega \mathcal{N}(u, \nabla u)\phi dx dt = \int_0^T \int_\Omega f \phi dx dt. \\ g = u & \text{on } \partial\Omega \end{cases}$$

The training procedure starts from defining a weak formulation for a target PDE on a proper vector space denoted by H as in (1.1). As illustrated in Fig.1, the network is then designed to map given an input function, such as a forcing function f, an initial condition u_0, and a boundary condition g, to a set of spectral coefficients, $\{\widehat{\alpha}_k\}_{k=0}^{N-1}$. These coefficients are converted into an inference solution defined by $\widehat{u} := \sum_{k=0}^{N-1} \widehat{\alpha}_k \phi_k$ with a set of basis functions, $\{\widehat{\phi}_k\}_{k=0}^{N-1}$ in H. Besides, the loss function is derived from the residual of the weak formulation to which the inference solution is substituted. Consequently, while the loss function decreases, the network is trained, driving the inference toward the true solution.

The core contribution of my research is listed as follows:

(1) SON integrates the benefits of both an operator learning framework and the spectral element method. It can assign multiple inference solutions in near real-time as an operator learning network. In addition, thanks to adopting the spectral method, it can achieve lower accuracy with a reduced number of nodal points. Moreover, it does not need an additional network to infer a set of basis functions, thereby resulting in a simpler network architecture with fewer training parameters than existing operator learning networks.
(2) It does not require reference solutions for training because it adopts the loss function derived from a weak formulation of a target PDE. Furthermore, by inherently learning the underlying structure of a target PDE, the trained network demonstrates enhanced robustness for a wider range of unseen input data.
(3) This network is flexible to various types of inputs, including initial conditions, boundary conditions, and forcing functions.
(4) This approach is, to my knowledge, the sole unsupervised operator learning network designed for the velocity-pressure form of the incompressible 3D Navier-Stokes equations with no-slip boundary conditions [4]. Besides, SON is also adapted to infer solutions of various types of elliptic equations [6]; Burgers equation, Boundary layer equation, 2D Kuramoto-Sivashinsky equation [7].

1.3. **Future research plan.** Despite fast computing performance of current machine learning (ML) methodologies, conventional numerical schemes are still regarded as more reliable standard for computing PDE solutions. This is because there is few theories to support ML methodologies as a PDE solver. Additionally, since ML methodologies serve as regression models, trained networks struggle to generalize accurately to out-of-sample input data, in contrast to conventional numerical schemes. Thus, to establish SON as a trustworthy and general-purpose PDE solver, my future research will focus on three key areas as follows.

1.3.1. ***Theory of SON for efficient approximation.*** My future research will focus on founding the theory upon how to efficiently represent PDE operators using the SON architecture. In fact, UAT only proves the existence of an operator network capable of approximating a target PDE operator but fails to provide guidelines for efficiently approximating it. This theoretical deficiency necessitates countless trial-and-error for designing a working operator network. In addition, it is typically preferable for a network where the number of parameters exceeds the number of training data, resulting in an underdetermined regression model. Thereby, optimizing the parameters of the network is analogous to solving an underdetermined system to find a solution among infinitely many solutions by a gradient descent method. Nevertheless, it leads not only to significant inefficiency in design of the network, but also to overfitting the training data and failing to generalize robustly to out-of-sample data, unlike conventional numerical schemes.

To address these issues, my approach first involves identifying how to represent a target PDE operator (functional) in a proper function space. For example, the Riesz Representation Theorem ensures that a linear (or linearlized) PDE operator defined in a Hilbert space can be represented by an inner product between a Green function and a function provided by PDE data. Using this theorem, I will then investigate the Green function to express the target PDE operator by such an inner product in a proper function space. If it proven, I will then proceed to effectively approximate the Green function using neural networks in the function space. In addition, analyzing the properties of the function space - for example, orthogonality between basis functions - is crucial for optimizing the network's representation and reducing the number of training parameters. The resulting theorems will then optimize the training procedure, thereby fundamentally improving both generalization and computational tractability.

1.3.2. ***Theory of SON for efficient training.*** I plan to research how to efficiently minimize a loss function, especially derived from PDE residuals. Typically, iterative methods using variations of Stochastic Gradient Descent (SGD) are employed to minimize a loss function. However, the iterative methods suffer from long training time because their rate of convergence may become very slow depending on an initial parameter guess and the complex shape of the loss function. Moreover, given that the loss in an unsupervised neural network is derived from an underlying PDE residual, the convergence rate slows down if the

	Euler equation	**Potential flow equation(PFE)**	**Benney-Luke equation(BLE)**	**Kadomtsev-Petviashvili equation(KPE)**	**Korteweg-de Varies (KdV) equation**
Dimension of domain	3D	3D	2D	2D	1D
Maximum height vs. Individual soliton height a_i	N/A	N/A	N/A	$9a_i$	$4a_i$
Approximation relation	The original	$>$	$>$	$>$	The simplest

Table 1. Approximation hierarchy among ocean wave models [8]

residual has a large condition number. Therefore, to accelerate minimizing a loss function, I will investigate the operator network where the optimal number of training parameters is determined by the properties of the function space found in Sec.1.3.1. In addition, in order to reduce the number of training iterations, I will adapt classical iterative methods. For example, given that the Generalized Minimal Residual method and the Conjugate Gradient method can solve a linear system exactly in fewer iterations than the system's rank, the concept of the methods is likely to be adapted into our framework.

1.3.3. ***Application to real-world problems.*** I intend to apply the SON framework to real-world problems, collaborating with experts in the relevant research topic (as discussed in Sec.2). The immediate focus is simulating and forecasting rogue waves, a challenge where the computational demands of multi-instance simulations on 3D large domain are expensive even with High-Performance Computing. Thus, the SON is ideally suited for rapidly generating multiple solutions of rogue waves, thereby enabling the development of instant high-wave warning systems.

The framework is also directly applicable to weather forecasting. Modern weather prediction employs ensemble forecasting which computes multiple weather scenarios, requiring massive computation. To reduce the computation, machine learning methods have been employed [1, 12]; however, their reliance on labeled solutions creates an extra computational burden for training data collection. Given that many atmospheric models are governed by the Navier-Stokes equations, SON provides an essential advantage: it can be applied directly without reliance on expensive labeled solutions, offering a fast, data-efficient alternative for ensemble forecasting.

Finally, SON is ideally suited for constructing a digital twin for the purpose of training physics-driven AI models. Training these advanced models necessitates generating massive volumes of high-fidelity data governed by physical laws within a virtual environment. SON will address this necessity by rapidly and accurately generating this physically-governed simulation data.

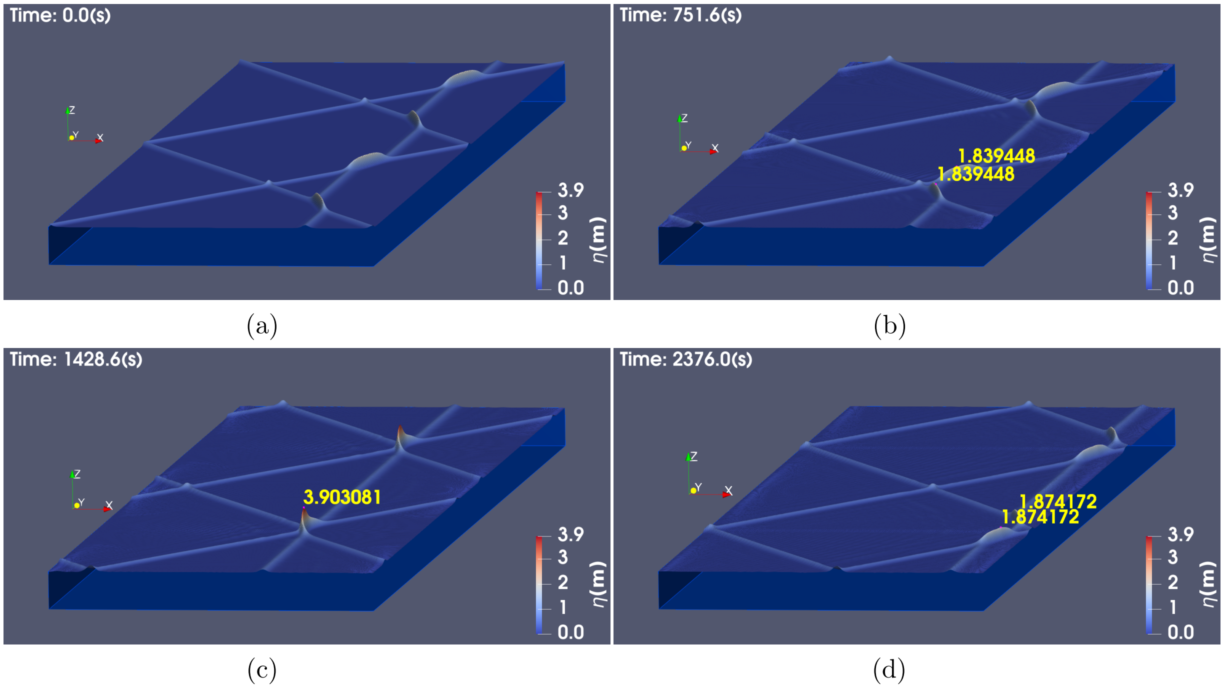

FIGURE 2. **A numerical simulation of BLE.** The simulation depicts the interaction among three oblique line solitons. (a): Initially, three line solitons obliquely cross each other, each having a height of 0.5 m. (b) and (c): As these solitons obliquely travel, they gather at a single point, reaching a maximum height of circa 3.9 m. This peak height is more than 7.8 times higher than the height of the individual constituent solitons. (d): After this peak wave occurs, the three line solitons then separate and continue travelling independently. (Refer to the relevant movie linked in the footnote)[1]

2. Rogue waves

2.1. **A rogue wave theory: interaction among three oblique line solitons.** My second research subject involves the mathematical modeling and simulation of ocean waves, specifically focusing on highly non-linear phenomena like rogue waves. Rogue waves are defined as waves exceeding twice the significant wave height, characterized by their sudden, unpredicted occurrence. A theory for their genesis is rooted in Kadomtsev–Petviashvili equation (KPE) [9]. It posits that the interaction between three oblique line-solitons, described by KPE, can generate a significantly higher wave than the individual constituent solitons. Because it is derived as an approximation from the leading terms of the Euler equation's perturbation expansion, KPE serves as a ocean wave model to explain this

[1] Kalogirou, A. Benney–Luke simulation of three obliquely interacting solitary waves (2022). https://youtu.be/Dx8-vhXwJSU

mechanism. However, two challenges remain: proving the existence of a corresponding Euler solution that behaves like the KPE three-soliton interaction, and then justifying the approximate relationship between the KPE solution and the Euler solution.

2.2. **Research achievements.** My research goal is to computationally and theoretically validate the three-line soliton model as an explanation for rogue wave dynamics. To achieve this, I have introduced the approximate relationship of ocean wave models, formally deriving the Potential Flow Equation (PFE), the Benney-Luke Equation (BLE), and the KPE from the full Euler equations as in table 1 and [8]. I then conducted a comprehensive analysis of the KPE solution that describes the interaction of three oblique line solitons. Using this exact KPE solution as a seeding initial condition, I conducted numerical simulations of the intermediate ocean wave models, BLE and PFE to computationally assess whether the rogue wave behavior persists at the higher ocean wave model. This computational work is expected to offer theoretical justification for the existence of the corresponding solutions to the ocean wave models.

The main contribution of my research is stated as follows.

(1) I have comprehensively analyzed KPE solutions that models the interaction of three oblique line solitons [3]. This solution demonstrates an extreme event that the amplitude of wave crest reaches up to nine times that of the constituent individual line solitons. Crucially, I have derived the condition that guarantees this nine-fold maximum amplification within the KPE framework. This finding is pivotal, as it provides the necessary parameters to design initial conditions for simulation of PFE and BLE, which are more physically meaningful ocean wave models.
(2) I have designed a stable numerical scheme to realize the interaction of three oblique line solitons in BLE [3] and PFE [5]. Specifically, I adapted the KPE solution which exhibits a maximum ninefold wave amplification to seed an initial condition of BLE and PFE. For a numerical scheme, I designed FEM variational formulation coupled with a second-order modified-midpoint temporal scheme. In fact, this scheme allowed to prevent artificial dissipation of the initial wave energy over time, leading to the conserved wave energy concentrating at the peak interaction point. Consequently, the simulations of BLE and PFE demonstrated stable energy evolution. In addition, as shown in Fig.2, they successfully replicated the three-soliton interaction observed in the KPE solution, achieving a maximum wave amplification of up to 7.8-fold. Furthermore, this crucial computational evidence strongly suggests the existence of a deep hierarchical relationship among the KPE, BLE, and PFE models concerning the interaction of three oblique line solitons.
(3) Utilizing PFE, I have developed a robust numerical algorithm specifically designed to simulate real-world problems upon ocean wave propagation [2]. This algorithm is capable of simulating waves generated by a wave maker in a wave basin. In addition, the algorithm realizes wave-absorption at a shallow beach by simulating

the transition of the potential-flow dynamics in deeper waters to shallow-water dynamics at the beach. Since these simulations are designed to replicate real-world physical environments, the resulting high-fidelity models are directly applicable for the construction and validation of practical ocean wave forecasting and warning systems.

2.3. Future research plan.

2.3.1. ***Analysis on the hierarchy.*** Building upon the computational evidence, I am now undertaking a rigorous analysis on the formal approximate relation among the PFE, BLE, and KPE. The initial focus is on establishing the relationship between their respective one-dimensional forms. Given that KPE possesses known traveling wave solutions, the primary analytical goal is to identify the existence of equivalent traveling wave solutions for BLE and PFE that exhibit similar behavior. By applying the structure of the KPE solution to BLE and PFE, each equation is transformed into a system of ordinary differential equations (ODEs). The transformed ODE system will be solved precisely using the Newton method. After that, I will analyze the relationship between their respective solutions. Then, this analytical justification will be the first step in constructing a comprehensive theoretical foundation for the entire model hierarchy.

2.3.2. ***Application of the ocean wave models for addressing real-world problems.*** I intend to extend the ocean wave framework to apply it real-world hazard prediction, including various wave situations. Specifically, I will focus on the complex non-linear interaction between surface waves and various seabed topographies in coastal regions. Given that PFE is a deep-water model, I will devise extended ocean wave models which couple deep-water dynamics with shallow-water dynamics based on PFE. Subsequently, I will develop efficient numerical schemes to realize the dynamic transition between these two regimes under various seabed topographies. Thereby, the ultimate objective is to accurately forecast the impact and destructive potential of rogue waves near shorelines, which will provide essential, high-fidelity data for designing effective coastal defense strategies and enhancing early warning systems.

2.3.3. ***Application of machine learning framework for accelerating computation.*** For the final topic, I will research the collaborative application of the ocean wave framework and machine learning models. Simulating and forecasting rogue waves presents a significant computational challenge, as multi-instance simulations over large, three-dimensional domains remain prohibitive even with HPC. For example, a single instance of the PFE simulation for the three-line-soliton interaction, conducted using FEM, required up to 56 hours [5]. Whereas such conventional methods yield accurate results, they are impractical for instant, operational forecasting. Given my prior work developing a machine learning framework for solving the 3D Navier-Stokes equations in nearly real-time, this collaborative application would significantly reduce the computational time, thereby helping to develop robust real-time high-wave warning systems.

References

[1] Kaifeng Bi, Lingxi Xie, Hengheng Zhang, Xin Chen, Xiaotao Gu, and Qi Tian. Accurate medium-range global weather forecasting with 3d neural networks. *Nature*, 619(7970):533–538, 2023.

[2] Onno Bokhove, Yang Lu, Mark Kelmanson, Junho Choi, and Floriane Gidel. Variational numerical-modelling strategies for the simulation of driven free-surface waves. *EarthArXiv eprints*, page X5PR06, 2025.

[3] Junho Choi, Onno Bokhove, Anna Kalogirou, and MA Kelmanson. Numerical experiments on extreme waves through oblique–soliton interactions. *Water Waves*, 4(2):139–179, 2022.

[4] Junho Choi, Teng-Yuan Chang, Namjung Kim, and Youngjoon Hong. A data free neural operator enabling fast inference of 2d and 3d navier stokes equations. *arXiv preprint arXiv:2510.23936*, 2025.

[5] Junho Choi, Anna Kalogirou, Yang Lu, Onno Bokhove, and Mark Kelmanson. A study of extreme water waves using a hierarchy of models based on potential-flow theory. *Water Waves*, 6(2):225–277, 2024.

[6] Junho Choi, Namjung Kim, and Youngjoon Hong. Unsupervised legendre–galerkin neural network for solving partial differential equations. *IEEE Access*, 11:23433–23446, 2023.

[7] Junho Choi, Taehyun Yun, Namjung Kim, and Youngjoon Hong. Spectral operator learning for parametric pdes without data reliance. *Computer Methods in Applied Mechanics and Engineering*, 420:116678, 2024.

[8] Floriane Gidel, Onno Bokhove, and Anna Kalogirou. Variational modelling of extreme waves through oblique interaction of solitary waves: application to mach reflection. *Nonlinear Processes in Geophysics*, 24(1):43–60, 2017.

[9] Yuji Kodama. Kp solitons in shallow water. *Journal of Physics A: Mathematical and Theoretical*, 43(43):434004, 2010.

[10] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

[11] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021.

[12] Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, et al. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators. *arXiv preprint arXiv:2202.11214*, 2022.

[13] Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378:686–707, 2019.