RSFNO: A Fourier Neural Operator–Based Solver for High-Dimensional Neutron Transport Equations

Authors: WU, Dr. Jiajun, HUANG, Miss Yifei, Wu, Dr. Zhen, SHANG, Dr. Jiahe, Zhou, Mr. Yang, HAO, Dr. Yisheng, HU, Dr. Ankang, QIU, Dr. RUI 邱睿, Zhang, Dr. Hui, LI, Prof. JUNLI, LI, Prof. JUNLI

Date: 2026-02-03T23:32:03+00:00

Abstract

This work establishes the theoretical foundations and practical implementation of a Fourier Neural Operator-based solver for high-dimensional neutron transport equations. We leverage the operator-learning framework to approximate the solution operator of the Boltzmann equation, allowing for efficient representations of reactor-scale neutron flux. RSFNO integrates stochastic geometric parameterization with frequency-domain learning to establish a direct end-to-end mapping from shielding configurations to flux distributions. Evaluated on a reactor shielding benchmark, the model demonstrates remarkable efficiency: despite being trained on just 800 samples, it reduces the full-domain mean squared error by 58.95% compared to a strong Unet3D baseline. In deep-penetration regions, its accuracy rivals or exceeds that of a 10^6-particle Monte Carlo (MC) simulation. Crucially, RSFNO completes a full-domain prediction in 11.79~s on a single NVIDIA RTX~4090, achieving an effective $\sim 2795\times$ speedup over the MC baseline. While not yet a replacement for high-fidelity safety analysis, RSFNO offers a reliable, ultra-fast surrogate for the computationally expensive deterministic steps in hybrid variance-reduction workflows (e.g., CADIS). These results highlight the promise of Fourier-operator-based solvers in accelerating large-scale transport analysis and enabling scalable, data-driven reactor shielding design.

Full Text

Preamble

RSFNO: A Fourier Neural Operator–Based Solver for High-Dimensional Neutron Transport Equations∗ Jiajun WU,1, 2 Yifei HUANG,3 Zhen WU,1, 2,

4 Jiahe SHANG,1, 2 Yang ZHOU,1, 2 Yisheng HAO,1, 5 Ankang HU,1, 2 Rui QIU,1, 2 Hui ZHANG,1, 2, † and Junli LI1, 2, ‡ 1Department of Engineering Physics, Tsinghua University, Beijing 100084, China 2Key Laboratory of Particle and Radiation Imaging of Ministry of Education, Beijing 100084, China 3School of Physics, Sichuan University, Chengdu 610065, Sichuan, China 4Nuctech Company Limited, Beijing 100084, China 5China Research and Development Academy of Machinery Equipment, Beijing 100089, China This work establishes the theoretical foundations and practical implementation of a Fourier Neural Operator-based solver for high-dimensional neutron transport equations. We leverage the operator- learning framework to approximate the solution operator of the Boltzmann equation, allowing for ef- ficient representations of reactor-scale neutron flux. RSFNO integrates stochastic geometric param- eterization with frequency-domain learning to establish a direct end-to-end mapping from shielding configurations to flux distributions. Evaluated on a reactor shielding benchmark, the model demon- strates remarkable efficiency: despite being trained on just 800 samples, it reduces the full-domain mean squared error by 58.95% compared to a strong Unet3D baseline. In deep-penetration regions, its accuracy rivals or exceeds that of a 106-particle Monte Carlo (MC) simulation. Crucially, RSFNO completes a full-domain prediction in 11.79 s on a single NVIDIA RTX 4090, achieving an effective 2795× speedup over the MC baseline. While not yet a replacement for high-fidelity safety anal- ysis, RSFNO offers a reliable, ultra-fast surrogate for the computationally expensive deterministic steps in hybrid variance-reduction workflows (e.g., CADIS). These results highlight the promise of Fourier-operator-based solvers in accelerating large-scale transport analysis and enabling scalable, data-driven reactor shielding design.

Keywords: Fourier neural operator; Monte Carlo simulation; Global radiation field; Neural network; MC- Shield

INTRODUCTION

The exponential growth of artificial intelligence (AI) has created an urgent demand for reliable, clean en- ergy [1]. To meet this challenge, advanced nuclear systems—including sodium-cooled fast reactors (SFRs), small modular reactors (SMRs), and space nuclear power units—are being rapidly developed [2, 3]. However, the complex geometries and deep-penetration shielding characteristic of these next-generation reactors im- pose strict radiation protection standards under the ALARA principle. Consequently, high-fidelity calcula- tion of the global neutron flux distribution is essential for shielding optimization, structural integrity assessment, and personnel safety [6]. This requirement presents a formidable computational challenge: the governing linear Boltzmann transport equation operates in a high- dimensional phase space —coupling spatial, angular, and energy variables—rendering full-domain solutions com- putationally prohibitive [4]. ∗ Supported by the National Natural Science Foundation of China (General Program) (125B2113, U23B2067, U2167209, 12375312, 12175114), the National Key R&D Program of China (2022YFC2402304) and the Open-end Fund Projects of China Institute for Radiation Protection Scientific Research Platform (CIRP-HYYFZH-2023ZD001) † Co-corresponding author, zhanghui34@tsinghua.edu.cn ‡ Corresponding author, lijunli@mail.tsinghua.edu.cn Conventional numerical strategies typically employ de- terministic or Monte Carlo (MC) approaches [5, 8, 9].

Deterministic solvers, such as the discrete ordinates (SN) method [7, 10], offer efficiency on structured grids but incur discretization errors, such as ray effects, when ap- plied to the irregular geometries typical of engineering applications. Conversely, MC methods provide precise geometric fidelity through constructive solid geometry (CSG) and serve as the benchmark for accuracy. How- ever, they are constrained by a slow statistical conver- gence rate (proportional to $N -1/2$), which makes global field reconstruction extremely time-consuming, particu- larly in deep-shielding scenarios with low particle sur- vival rates. While variance reduction techniques (e.g., CADIS [33], AIS [12, 38], DeGVR [13], PDMC [14], MAGIS-GPS [15], SsGVR [16], Weight Window [17]) al- leviate some computational burden, they often require complex parameter tuning and do not fundamentally al- ter the underlying algorithmic complexity.

The integration of AI into nuclear engineering offers a pathway to accelerate these calculations [18]. Early data-driven approaches utilized Convolutional Neural Networks (CNNs) and U-Nets to approximate radiation fields or expedite MC convergence [19–25]. Although effective in limited contexts, these models map between finite-dimensional vector spaces (i.e., fixed-resolution im- ages). Consequently, they fail to generalize across vary- ing geometries or capture the continuous nature of phys- ical fields, restricting their utility in iterative reactor de- sign.

To overcome these discretization limits, the field has advanced towards Operator Learning, with Deep- [27–29] and the Fourier Neural Operator (FNO) [30, 31] emerging as leading paradigms. While DeepONet utilizes a branch-trunk architecture to flex- ibly handle arbitrary unstructured grids, its point-wise evaluation mechanism can become computationally pro- hibitive when reconstructing dense, high-dimensional fields typical of reactor shielding ($>$ 106 voxels). In con- trast, FNO is architecturally optimized for such contin- uous, uniform domains. By parameterizing the integral kernel directly in the frequency domain, FNO achieves a global receptive field with quasi-linear complexity $O(N \log N)$. This spectral approach aligns intrinsically with the non-local physics of neutron transport—where particles stream across the entire domain—offering a distinct efficiency advantage over point-based architec- tures for full-core 3D reconstruction. Consequently, we adopt the FNO framework, benchmarking it against the industry-standard Unet3D to demonstrate the superior- ity of spectral operator learning.

However, applying FNOs to reactor physics is non- trivial. The non-periodic nature of reactor boundaries and the high-frequency features arising from sharp mate- rial interfaces frequently induce spectral leakage (Gibbs phenomenon) [34, 35]. Here, we introduce RSFNO, a specialized Fourier Neural Operator-based solver for high-dimensional neutron transport equations. By in- tegrating geometric parameterization with frequency- domain learning, RSFNO establishes a direct mapping from shielding configurations to 3D neutron flux distri- butions. Using high-fidelity training data generated by the MC code MCShield [11, 36], we demonstrate that RSFNO significantly accelerates global field prediction relative to MC simulations. Furthermore, it outper- forms conventional convolutional baselines (specifically Unet3D) in reconstruction accuracy, providing a robust surrogate model for scalable, data-driven shielding de- sign.

II. MATERIALS AND METHODS A. Computational Framework

1. Neutron Transport Physics

The steady-state distribution of neutrons within a reactor shield is governed by the linear Boltzmann transport equation, a standard formulation in reactor physics [39, 40].

It balances particle streaming, colli- sions, and external sources in a high-dimensional phase space. Let D (cid:26) R3 be the spatial domain, S2 the unit sphere of directions, and E the energy variable. The angular neutron flux ψ(r, E, Ω) satisfies: Ω (cid:1) rψ(r, E, Ω) + Σt(r, E)ψ(r, E, Ω) = Σs(r, E′ ! E, Ω′ ! Ω)ψ(r, E′, Ω′)dΩ′dE′ + S(r, E, Ω), where Σt is the macroscopic total cross-section, Σs is the differential scattering cross-section, and S represents the fixed external source (e.g., the reactor core emission).

The equation is closed with appropriate boundary con- ditions on ∂D (e.g., vacuum or reflective boundaries).

2. Operator Learning Formulation

To facilitate the application of neural operators, we reformulate the transport problem as a mapping between function spaces. The equation can be expressed in the abstract operator form: discontinuities at material interfaces. Assuming the problem is well-posed, there exists an inverse operator, or solution operator, G†, that maps the geometry field directly to the flux solution:

G† : a(r) 7! ψ(r, E, Ω) = (L(a)) −1 S. The objective of this work is to construct a data-driven approximation, denoted as Gθ (where θ are learnable pa- rameters), of this solution operator. Specifically, we fo- cus on the mapping from the 3D material configuration a(r) to a scalar quantity of interest ϕ(r) (e.g., the energy-integrated scalar flux or dose):

Laψ = S, ϕ(r) = R(E)ψ(r, E, Ω)dΩdE, linear transport whereLarepresents operator (streaming plus collision minus scattering). Here, the parameter a(r) encapsulates the heterogeneous material properties and geometry of the shielding configura- tion. In reactor shielding, a(r) is typically a piecewise constant function representing different material zones (e.g., steel, sodium, concrete), characterized by sharp where R(E) is a response function. By training on high-fidelity Monte Carlo data, the neural operator $G\theta$ learns to implicitly encode the complex integral inversion of the transport operator, capturing the global dependencies of the radiation field while handling the geometric irregu- larities inherent in a(r).

3. Fourier Neural Operator Architecture

a. Operator Formulation We formulate the reac- tor shielding problem as learning a mapping between two infinite-dimensional Banach spaces. Specifically, we seek to approximate the solution operator G† defined in Eq. (3). While the underlying transport physics involves high-dimensional phase space, for practical shielding de- sign, we target the marginal scalar field map:

G : a(x) 2 A 7! ϕ(x) 2 U , where A = L∞(D; Rda) represents the material parame- ter space on the domain D, and U = H s(D; R) denotes the resulting scalar flux field space. b. Spectral Convolution and the FNO Layer Build- ing upon the seminal framework introduced by Li et al. [30], the Fourier Neural Operator approximates G by parameterizing a surrogate $G\theta$. The architecture com- prises a stack of integral operator layers. Let vt(x) 2 Rdv be the feature representation at layer t. The update rule for the (t + 1)-th layer is defined as: vt+1(x) = σ (W vt(x) + (K(ϕ)vt)(x)) , x 2 D, where σ is a non-linear activation function (e.g., GELU).

This update rule consists of two parallel branches that are critical for transport problems:

1. Global Spectral Branch (K): This term imple-

ments a global convolution via the Fast Fourier Transform (FFT). (Kvt)(x) = F −1 [Rt(ξ) (cid:1) (Fvt)(ξ)] (x), where F denotes the FFT, and Rt(ξ) 2 Cdv×dv is a learnable spectral weight tensor. To ensure com- putational efficiency and regularization, we trun- cate the spectrum, retaining only the lowest kmax frequency modes. This branch effectively captures the global dependency of neutron streaming, anal- ogous to the integral transport operator.

2. Local Residual Branch (W): The term W vt(x) rep-

resents a pointwise linear transformation (or a local convolution). c. Addressing Transport-Specific Challenges To adapt the standard FNO architecture to the specific con- straints of reactor physics, we introduce two architec- tural modifications:

1. Mitigation of Gibbs Phenomena: Neutron trans-

port in heterogeneous reactors is characterized by piece- wise constant cross-sections, leading to fields with low regularity (sharp gradients) at material interfaces. Pure spectral truncation often results in non-physical oscil- lations (Gibbs phenomenon) near these discontinuities.

In Eq. (6), the local branch W acts as a high-frequency compensator. While the spectral branch K resolves the smooth, global transport modes, W preserves the local high-frequency information lost during spectral trunca- tion, thereby sharpening material interfaces and reduc- ing oscillatory artifacts.

2. Handling Non-Periodic Boundaries: Unlike stan-

dard FFT algorithms that assume periodic boundaries, which contradicts the vacuum or leakage boundaries of a reactor shield. We propose a domain padding strategy.

The input domain D is padded with zeros (or a specific boundary value) prior to the Fourier transform. This allows the convolution kernel to decay naturally at the boundaries, preventing the " wrapping around" of neu- tron flux from one side of the reactor to the other, thus physically mimicking non-periodic leakage. d. Approximation Error Analysis The total ap- proximation error of the trained operator Gθ relative to the true operator G can be decomposed as: kG (cid:0) Gθk (cid:20) kG (cid:0) PkGk Truncation Error + kPkG (cid:0) Gθk Optimization Error where Pk is the projection onto the first kmax Fourier modes. The truncation error depends on the Sobolev regularity of the target flux ϕ. By incorporating the local branch W and using a sufficiently expressive hid- den dimension dv, RSFNO effectively minimizes the op- timization error while the hybrid architecture mitigates the perceived truncation error near singularities. e. Full Architecture The complete RSFNO solver is constructed as a composition:

Gθ = Q (cid:14) LT (cid:14) (cid:1) (cid:1) (cid:1) (cid:14) L1 (cid:14) P, where P (Lifting) projects the input a(x) to the high- dimensional feature space dv, Lt are the FNO layers defined in Eq. (6), and Q (Projection) maps the final features back to the scalar flux ϕ(x).

B. Dataset Generation and Representation The efficacy of data-driven operator learning hinges on the representativeness of the training measure over the input function space. According to the universal approximation theorem for neural operators [27], mini- mizing the generalization error requires that the training samples form a dense cover of the compact subset of the input space A relevant to the physics. To achieve this, we employ a stochastic media generation strategy that synthesizes geometrically diverse and materially hetero- geneous shielding configurations.

1. Stochastic Geometry Construction

The computational domain is discretized into uniform cubic voxels. The input field is represented by a material descriptor tensor M (r) = (ρ(r), δ(r)), where ρ denotes the mass density and δ the discrete material type. To encompass the range of neutron interactions encountered in reactor engineering, materials are categorized into four functional groups:

1. Strong Absorbers ($\delta = 1$): Materials with high

capture cross-sections (e.g., Gd, B, Cd).

2. Moderators ($\delta = 2$): Low-atomic-mass materials

for neutron thermalization (e.g., H2O, Graphite).

3. Structural Materials ($\delta = 3$): Medium-weight nu-

clei providing scattering and integrity (e.g., Fe, high statistical precision, the generation of each sam- ple was computationally intensive, requiring an average of 1,042 seconds on 100 parallel CPU cores. The dataset comprises 18 representative reactor materials, with detailed elemental compositions and reference den- sities listed in Table 1 . A visualization of the generated material identifiers and corresponding density fields is presented in Fig. 1 [Figure 1: see original paper].

Implementation Details and Training Protocol

1. Network Architecture and Implementation

(e.g., U, Pu). To generate shielding configurations that are both ge- ometrically diverse and physically realistic, we propose a novel hybrid stochastic generation strategy that governs material topology and density distribution separately: a. Geometric Topology via Wavelet Transformation.

To ensure structural continuity and introduce realistic spatial correlations (rather than random noise), the dis- crete material type distribution δ(r) is generated us- ing a wavelet-based method. A latent random field is first initialized in the frequency domain and projected onto the spatial domain via an Inverse Discrete Wavelet Transform (IDWT). The resulting scalar field W (r) is then mapped to discrete material types using amplitude thresholding: δ(r) = T (W (r)), where T is a step function defined by specific thresh- olds for the four material groups. This process creates continuous, multi-scale material regions. b. Continuous Density Variation. Within the de- fined material regions, the mass density is not uniform but exhibits continuous spatial variations. This hetero- geneity is modeled using a Gaussian Radial Basis Func- tion (RBF). For a given material type δ, the density field follows a Gaussian decay centered at a randomly sam- pled position rc: ρ(r) = hδ(r) exp (cid:18) (cid:19) (cid:0) kr (cid:0) rck2 where hδ(r) is the reference peak density corresponding to the material type at location r, and σ controls the gradient spread. This dual-layer generation strategy—combining wavelet-based topology with Gaussian-based density gradients—ensures that the dataset captures both the complex boundaries of reactor components and the continuous physics of material heterogeneity.

2. Monte Carlo Simulation

Ground-truth radiation fields were computed using the high-fidelity transport code MCShield [36]. To ensure The proposed RSFNO is implemented as a specialized 3D Fourier Neural Operator. As illustrated in Fig. 2 [Figure 2: see original paper], the architecture transforms the input material field into the target flux field through a sequence of spectral opera- tions. The network comprises three functional modules:

1. Feature Lifting: A shallow neural network projects

the input material tensor (density and material type channels) into a high-dimensional latent space v0(x) 2 Rdv .

2. Iterative Spectral Processing: The core processing

unit consists of four cascaded FNO blocks. In each block, the global spectral convolution is computed via the Fast Fourier Transform (FFT). To strictly enforce the non-periodic boundary conditions re- quired by reactor physics, we employ a padding strategy: the spatial domain is padded by p = 6 voxels prior to the FFT operation. This prevents spectral leakage and the " wrapping" artifacts typ- ical of periodic Fourier bases.

3. Projection and Reconstruction: The final latent

representation is projected back to the scalar domain via a two-layer Multi-Layer Perceptron (MLP) to recover the neutron flux ψ(x).

Weight parameters in the spectral domain are initial- ized using Xavier normal initialization with a gain of 0.01 to ensure gradient stability during the early train- ing phase.

2. Data Preprocessing

Given the physical characteristics of neutron trans- port, raw data requires rigorous preprocessing to ensure numerical stability: • Input Standardization: The material density fea- tures x are standardized using Gaussian normal- ization: xnorm = x (cid:0) µx Fig. 1. Visualization of a representative training sample: (a) discrete material identifiers defining geometry; (b) continuous material density distributions generated via Gaussian superposition.

TABLE 1. Elemental composition and reference density (hδ) of the 18 representative materials used in the dataset generation.

Material Elemental Composition (wt%) U235: 1.85, U238: 62.75, O: 18.10, Zr: 15.14, H: 1.18, Others: 0.61 O: 88.88, H: 11.11, B: 0.01 Fe: 97.90, Mn: 1.30, Ni: 0.55, C: 0.25 Fe: 68.46, Cr: 18.40, Ni: 9.92, Mn: 1.98, O: 1.23 Fe: 69.00, Cr: 19.00, Ni: 10.00, Mn: 2.00 Fe: 65.80, Cr: 18.12, Ni: 9.54, O: 4.13, Mn: 1.90, H: 0.52 U–Zr Fuel Alloy Borated Water Carbon Steel Porous Steel Light Steel Low-H Steel Concrete Simulant O: 50.50, Si: 34.09, Ca: 4.40, Fe: 3.82, Al: 3.43, Na: 1.62, K: 1.31, H: 0.51, Mg: 0.22, C: 0.10 Cr-Oxide Alloy O: 33.45, Cr: 31.27, Ni: 16.46, Fe: 11.35, H: 4.18, Mn: 3.29 N: 70.00, O: 30.00 SS-304 Type I Fe: 69.00, Cr: 19.00, Ni: 10.00, Mn: 2.00 SS-304 Type II Fe: 72.18, Cr: 19.09, Ni: 5.65, Mn: 2.01, O: 0.95, H: 0.12 Fe: 66.43, Cr: 18.29, Ni: 9.63, O: 3.30, Mn: 1.93, H: 0.41 SS-304 Type III Fe-Oxide Alloy I O: 58.26, Fe: 23.78, Cr: 6.55, Ni: 3.45, H: 7.28, Mn: 0.69 Fe-Oxide Alloy II O: 50.28, Fe: 29.97, Cr: 8.25, Ni: 4.34, H: 6.28, Mn: 0.87 Fe-Oxide Alloy III Fe: 62.81, Cr: 17.32, Ni: 9.12, O: 7.94, Mn: 1.82, H: 0.99 Fe-Oxide Alloy IV Fe: 56.01, Cr: 15.42, Ni: 8.12, O: 16.74, H: 2.09, Mn: 1.62 Fe-Oxide Alloy V O: 74.92, Fe: 10.84, H: 9.37, Cr: 2.99, Ni: 1.57, Mn: 0.31 Fe-Oxide Alloy VI Fe: 55.44, Cr: 15.27, Ni: 8.03, O: 17.47, H: 2.18, Mn: 1.61 hδ [g/cm3] where µx and σx are the global mean and standard deviation of the dataset. To prevent division-by- zero errors in void regions, σx is lower-bounded at 10−8. • Log-Transformed Output: The neutron flux ψ typ- ically spans 5-10 orders of magnitude (10−5 to 10−10). To compress this dynamic range, we model the logarithm of the flux. The target variable y is transformed as: ytarget = ln(ψ + ϵ) (cid:0) µln where ϵ is the minimum positive flux value in the dataset (handling vacuum regions), and (µln, σln) are statistics of the log-transformed field. This log- arithmic formulation serves a dual purpose: it com- presses the dynamic range of the neutron flux and, upon inverse transformation, intrinsically enforces the non-negativity constraint of the physical field.

Consequently, the solver requires no ad-hoc post- processing to maintain physical realizability.

3. Training Protocol

a. Loss Function The network is trained to mini- mize the Mean Squared Error (MSE) in the logarithmic domain, which effectively corresponds to optimizing the relative error in the physical domain. The loss function Fig. 2. Schematic of the proposed RSFNO network architecture, detailing the lifting layer, spectral convolution blocks with boundary padding, and the final projection head. is defined as: kGθ(a(i)) (cid:0) y(i) target spectrum. The material specifications for each region are detailed in Table 2 .

IV. RESULTS AND DISCUSSION where N is the batch size. b. Optimization Setup Training is performed using the Adam optimizer with a weight decay of 1 (cid:2) 10−4 (L2 regularization). We employ a StepLR scheduler to de- cay the learning rate progressively. To mitigate gradient explosion —common in deep operator learning—gradient clipping is applied, and updates with abnormal gradient norms are filtered. c. Evaluation Metrics In addition to the training loss, model performance is evaluated using three physics- aware metrics: (1) MSE on the recovered physical scale, (2) mean relative error (MRE), and (3) relative L2 error norm. An early stopping mechanism halts training if the validation loss stagnates for 50 epochs. All experiments are conducted on a cluster equipped with 3 NVIDIA GeForce RTX 4090 GPUs. Data loading is parallelized with memory pinning, and mixed-precision training is selectively disabled at critical spectral layers to preserve numerical precision.

V. EXPERIMENTAL RESULTS To evaluate the predictive fidelity of RSFNO in real- istic engineering scenarios, we constructed the Reactor Pressure Vessel (RPV) benchmark. This case is designed to simulate the neutron flux distribution within the crit- ical shielding components of a pressurized water reactor.

The RPV model consists of four concentric cylindri- cal regions: the reactor core, cavity, reflector, and the pressure vessel itself (see Fig. 3 [Figure 3: see original paper]). The neutron source is located within the core region and follows a Watt fission A. Model Convergence and Generalization Dynamics The training dynamics of RSFNO demonstrate supe- rior stability and convergence compared to conventional convolutional architectures. As illustrated in Fig. 4 [Figure 4: see original paper], the training loss (Log-Space MSE) of RSFNO decreases monotonically over 921 epochs, dropping from an initial 0.630 to 0.0024 —a reduction of over 99.6%. In contrast, the baseline Unet3D [26, 32, 37], despite achieving a com- parable minimum training loss (0.0027), exhibits severe overfitting. While the validation loss of Unet3D diverges significantly from a minimum of 0.0038 to a final value of 0.032, RSFNO maintains a consistent descent, reaching a final test loss of 0.0026. This indicates that RSFNO ef- fectively learns the underlying physical operator without memorizing noise in the training data.

A quantitative analysis of the generalization gap (de- fined as jLtrain (cid:0) Ltestj) further highlights this advan- tage (Fig. 5 [Figure 5: see original paper]). RSFNO achieves an exceptionally narrow generalization gap of 1.40 (cid:2) 10−4 at the end of train- ing. Averaged over the final 50 epochs, the mean gap is 1.47 (cid:2) 10−4 with a negligible standard deviation of 1.0 (cid:2) 10−5. Conversely, Unet3D exhibits a mean gap of 2.51(cid:2)10−2 with a large standard deviation (3.14(cid:2)10−2), reflecting unstable convergence. These results confirm that the spectral regularization inherent in FNO pre- vents overfitting, a common pitfall in high-dimensional regression tasks.

Computational efficiency during training also favors the proposed architecture. The per-epoch training time for RSFNO is stable at 191.7 (cid:6) 11.6 s, approximately 37.5% faster than Unet3D (307.0 (cid:6) 5.4 s). This speedup, Fig. 3. Cross-sectional geometry of the RPV benchmark model TABLE 2. Material composition and reference densities for the RPV benchmark components.

Elemental Composition (wt%) Component Reactor Core U238: 64.15, O: 17.05, Zr: 15.80, U235: 1.80, H: 1.14, Fe: 0.06 Cavity Reflector Pressure Vessel Fe: 97.89, Mn: 1.31, Ni: 0.55, C: 0.25 N: 76.70, O: 23.30 O: 88.83, H: 11.13, B: 0.05 hδ [g/cm3] combined with the elimination of early stopping require- ments, makes RSFNO highly suitable for large-scale 3D operator learning.

Crucially, the log-Gaussian normalization scheme en- sures predictive consistency across the large dynamic range of neutron flux. As shown in Fig. 6 [Figure 6: see original paper], while the logarithmic MSE stabilizes at (cid:25) 2.6 (cid:2) 10−3, the error in the physical space reaches as low as 2.27 (cid:2) 10−14. This precision surpasses the Unet3D baseline (2.78 (cid:2) 10−13), ensuring reliable predictions across orders-of-magnitude variations.

B. Performance on Reactor Shielding Benchmark To assess performance in deep-penetration scenarios, we deployed the trained models on the Reactor Pressure Vessel (RPV) benchmark. We evaluated prediction ac- curacy relative to high-fidelity Monte Carlo (MC) sim- ulations (108 particles, statistical uncertainty $<$ 5% in the shielding region) and analyzed the preservation of flux gradients.

1. Global Field Reconstruction

As shown in Fig. 8 [Figure 8: see original paper], RSFNO successfully reconstructs the global flux distribution. The predicted flux range covers [1.00 (cid:2) 10−10, 1.82 (cid:2) 10−5] cm−2, closely matching the reference range of [1.00 (cid:2) 10−10, 1.84 (cid:2) 10−5] cm−2, with a peak flux deviation of only 1.09%. Visually, the RSFNO prediction is indistinguishable from the high-fidelity reference in the primary shielding regions.

It is observed that in the extreme corners of the do- main, the RSFNO prediction exhibits values below the visualization threshold ($<$ 5 (cid:2) 10−9 cm−2), appearing as " blank" regions in the flux contours. This behavior is directly correlated with the stochastic properties of the training data.

As evidenced by the statistical error map of the refer- ence simulation (Fig. 7 [Figure 7: see original paper]), these corner regions correspond to areas of significant statistical uncertainty, with rela- tive errors exceeding 10% (depicted in red) even with 108 particle histories. In these deep-penetration zones, the signal-to-noise ratio of the Monte Carlo "ground truth" is inherently low. Consequently, the neural oper- ator, driven by the global MSE loss, tends to suppress these high-variance fluctuations. While under-prediction in deep penetration regions presents a theoretical non- conservative risk, we observe that these " blank" regions strictly correlate with the noise-dominated zones of the reference Monte Carlo data (statistical error $>$ 10%).

Crucially, the model preserves the correct flux gradients and global topology. As will be quantitatively substan- tiated in the subsequent gradient analysis (Section 4.2.2, where a contour matching score of 0.969 is reported), this topological fidelity ensures that when used to gen- erate biasing parameters for hybrid variance reduction Fig. 4. Comparative convergence history: RSFNO (orange) demonstrates sustained decrease in both training and test loss, whereas Unet3D (red) exhibits significant overfitting after epoch 100. (e.g., CADIS workflows), RSFNO provides accurate di- rectional guidance for particle transport, leaving the fi- nal rigorous safety quantification to the subsequent high- fidelity Monte Carlo simulation. where M∗ denotes the set of optimal point matches be- tween c1 and c2 that minimizes the total transport cost, and k = min(jc1j, jc2j) is the number of matched pairs.

The final matching score is defined as:

2. Gradient Field and Iso-Flux Analysis

In shielding design, accurately capturing the gradi- ent of flux decay is as critical as the absolute values.

We analyzed the iso-flux contours extracted from the gradient fields (Fig. 9 [Figure 9: see original paper]). Five contours were generated at geometrically spaced intervals from 2.00 (cid:2) 10−8 to 4.80 (cid:2) 10−6 cm−2.

To quantify topological similarity, we define a compos- ite metric combining the Hausdorff distance (DH) and Earth Mover’ s Distance (DEMD):

Score = 1 + 0.5 (cid:1) D(cref, ci) Quantitative analysis reveals that RSFNO achieves a contour matching degree of 0.969, surpassing both the 106-particle MC simulation (0.944) and Unet3D (0.959).

This confirms that RSFNO’ s spectral convolutions ef- fectively preserve the smooth, continuous nature of the radiation field, filtering out the stochastic noise inherent in low-count MC simulations while maintaining sharp gradients at material interfaces.

D(c1, c2) = wH DH (c1, c2) DEMD(c1, c2)

3. Accuracy in Shielding Region and Efficiency

where Lmax = 181 (max Euclidean distance), wH = 0.7, and wE = 0.3.

The Hausdorff distance DH measures the worst-case local mismatch between the two contour point sets: (cid:18) DH (c1, c2) = max kp (cid:0) qk2, max kp (cid:0) qk2 Complementarily, DEMD captures the global distribu- tional shift.

It is calculated as the average Euclidean distance of the optimal point-wise assignment obtained via the Hungarian algorithm:

DEMD(c1, c2) = (p,q) M∗ kp (cid:0) qk2, (cid:19) We performed a rigorous error analysis within the des- ignated shielding target region (2200(cid:2)2200(cid:2)3400 mm3).

Table 3 summarizes the results. RSFNO demonstrates a superior error distribution: 46.63% of voxels have a relative deviation < 5%, and 70.83% are within < 15% error. In contrast, Unet3D only achieves < 5% error in 20.54% of voxels. Notably, with only 800 training sam- ples, RSFNO reduces the full-domain Mean Squared Er- ror (MSE) by 58.95% compared to Unet3D (3.60(cid:2)10−14 vs. 8.77 (cid:2) 10−14).

This accuracy gain is attributed to the global recep- tive field of the Fourier operator, which is essential for modeling the deep penetration of neutrons where local interactions are coupled over long distances. Convolu- tional networks like Unet3D struggle to capture these Fig. 5. Evolution of the generalization gap (|Ltrain − Ltest|) during training. RSFNO maintains a negligible gap, indicating robust generalization. non-local dependencies without an excessively deep hierarchy.

In terms of speed, RSFNO completes a full-domain prediction in 11.79 s (inference time 1.21 s) on a single NVIDIA RTX 4090 GPU. While slightly slower than Unet3D (10.87 s) due to the O(N log N) complexity of 3D FFT operations and wider channel dimensions, this overhead is negligible compared to the accuracy gains.

More importantly, compared to the 106-particle MC sim- ulation (3382.83 s on CPU), RSFNO achieves a speedup factor of approximately 2795(cid:2). This demonstrates that RSFNO can serve as a near real-time surrogate solver, providing accuracy comparable to or better than moderate-fidelity MC simulations at a fraction of the computational cost.

V. CONCLUSION We have introduced RSFNO, a Fourier Neural Operator-based framework designed to circumvent the computational bottlenecks of high-fidelity Monte Carlo (MC) simulations and the geometric rigidity of de- terministic solvers. By reformulating neutron trans- port as a mapping between infinite-dimensional func- tion spaces, our approach integrates stochastic geometric parameterization with frequency-domain operator learn- ing. This methodology effectively overcomes the recep- tive field limitations of traditional Convolutional Neural Networks, enabling the capture of long-range dependen- cies inherent in particle transport physics.

Quantitative evaluation on the Reactor Pressure Ves- sel (RPV) benchmark validates the efficacy of this ar- Fig. 6. Distribution of MSE across the test set evaluated in (a) logarithmic space and (b) the original physical space.

Method

TABLE 3. Performance comparison in the RPV shielding region.

Total Time Inference Time (s) Full-Space Proportion of Cells with Relative Error < 5% < 10% MC (106) 3819.4 ± 39.6 3382.8 ± 97.9 1.46e-13 24.21% 44.53% Unet3D 10.87 ± 1.12 1.19 ± 0.15 20.54% 34.69% 8.77e-14 1.21 ± 0.10 RSFNO

11.79 ± 0.58 3.60e-14 46.63% 63.53% < 25% 59.31% 62.08% 70.83% Note: Bold (x) indicates the best performance; underline (x) indicates the second best.

The computational acceleration achieved is transfor- mative. A full-core 3D shielding prediction requires only 11.79 s (with 1.21 s inference time) on a single GPU.

This represents a (cid:24) 2795(cid:2) speedup over a standard 106- particle single-threaded Monte Carlo simulation, reduc- ing time-to-solution from hours to seconds while main- taining engineering-grade accuracy in key regions.

While RSFNO is not intended to replace high-particle- count MC simulations for final safety licensing, it offers immediate utility as a high-speed surrogate within hy- brid variance reduction workflows, such as Consistent Adjoint Driven Importance Sampling (CADIS). Conven- tionally, generating global adjoint or forward flux dis- tributions for weight windows requires expensive deter- ministic calculations. RSFNO effectively substitutes this step, providing instantaneous estimations of global flux topology. This capability enables the rapid construction of biasing parameters, substantially accelerating subse- quent high-fidelity Monte Carlo simulations without the overhead of traditional SN solvers.

Future work will extend the solver' s universality by ex- panding the training domain to encompass more diverse reactor geometries, material compositions, and source spectra. Crucially, we aim to extend the RSFNO output from scalar fluxes to multi-group energy distributions.

By treating energy groups as high-dimensional feature channels, the solver will support spectrum-dependent applications, such as activation analysis and radiation damage calculation. Furthermore, we aim to adapt the frequency-domain formulation to multi-physics cou- pling scenarios, particularly thermal-hydraulic feedback, Fig. 7. Spatial distribution of the relative statistical error for the reference MC simulation (108 particles), confirming reliability (< 5%) in the shielding region. chitecture. Trained on a limited dataset of 800 samples, RSFNO reduced the full-domain Mean Squared Error (MSE) by approximately 58.95% relative to the widely used Unet3D baseline. In critical shielding regions, the model demonstrated superior reliability, with 46.63% of voxels exhibiting relative deviations within 5%. More- over, gradient-field analysis confirms that RSFNO recon- structs physically continuous flux contours, achieving a topological matching score of 0.969 against high-fidelity references.

Fig. 8. Cross-sectional comparison of neutron flux distributions at Y = 0: (a) Ground Truth (108 MC); (b) Low-fidelity MC (106); (c) Unet3D Prediction; (d) RSFNO Prediction. RSFNO reproduces the smoothness of the high-fidelity reference. where rapid resolution of the Boltzmann transport equa- tion is a prerequisite for holistic digital twin simula- tions. RSFNO thus represents a viable pathway toward real-time, data-driven reactor physics analysis, offering a scalable solution to the " curse of dimensionality" in advanced shielding design.

VI. BIBLIOGRAPHY [1] China Nuclear Energy Association. (2025). Annual re- port on China' s nuclear energy development 2025. Bei- jing: CNIA. [2] D.H. Zhang, S.P. Wang, Z.W. Dai, Review on Develop- ment of Sodium-Cooled Fast Reactor Worldwide, Atomic Energy Science and Technology 58, no. S2:265-70 (2024) doi: 10.7538/yzk.2024.youxian.0501 [3] Y. Wang, W. Chen, L. Zhang et al. , Small Modular Re- actors: An Overview of Modeling, Control, Simulation, and Applications, in IEEE Access, vol. 12, pp. 39628- 39650, (2024)doi: 10.1109/ACCESS.2024.3351220 [4] Zhang, TF., Li, ZP. Variational nodal methods for neu- tron transport: 40 years in review. Nucl. Eng. Technol. 54(9), 3181–3204 (2022). 10.1016/j.net.2022.04.012 [5] Deng, L., Li, G. Monte Carlo Simulation Methods and Applications for Particle Transport Problems. Beijing:

Sci. Press (2019). (in Chinese) [6] Hirayama, H., Nakashima, H., Morishima, M., Uematsu, M., Sato, O. Progress and prospects of calculation meth- ods for radiation shielding. J. Nucl. Sci. Technol. 52(11), 1339–1361 (2015). 10.1080/00223131.2015.1021283 [7] Yang, SL., Wei, JX., Hong, ZY., Liu, H. The SN algo- rithms for neutron transport and their applications. Sci.

Sin. Inf. 46, 1510–1526 (2016). 10.1360/N112016-00078 [8] Song, CX., Kawai, R. Monte Carlo and variance reduc- tion methods for structural reliability analysis: A com- prehensive review. Probabilist. Eng. Mech. 73, 103479 (2023). 10.1016/j.probengmech.2023.103479 [9] Li, R., Zhang, LY., Shi, DF., Deng, L. Crit- soluble boron iteration in MC icality search of Fig. 9. Flux gradient field analysis and iso-flux contours at X = 0. RSFNO (d) generates smooth contours that closely align with the high-fidelity reference (a), unlike the noisy contours from low-particle MC (b). code JMCT. Energy Procedia 127, 329–334 (2017). 10.1016/j.egypro.2017.08.118 [10] Herring, N.F., Beyer, K. et al. The general collision- count source method for ray effects mitigation in SN problems using Monte Carlo coupling. Prog. Nucl. En- ergy 176, 105373 (2024). 10.1016/j.pnucene.2024.105373 [11] Wu, J., Zhang, H., Hu, A. et al. DALB-MC: Dynamic adaptive load balancing for efficient Monte Carlo particle transport on heterogeneous architectures. J. Nucl. Sci.

Technol. 1–12 (2025). 10.1080/00223131.2025.2526784 [12] Li, J.L. et al. An auto-importance sampling method for deep penetration problems. Prog. Nucl. Sci. Technol. 2, 732–737 (2011).doi: 10.15669/pnst.2.732. [13] Pan, Q., Wang, L., Cai, Y. et al. Density-extrapolation global variance reduction (DeGVR) method for large- scale radiation field calculation. Comput. Math. Appl. 143, 10–22 (2023). 10.1016/j.camwa.2023.04.024 [14] Pan, Q., Lv, H., Tang, S. et al. Pointing probability driven semi-analytic Monte Carlo method (PDMC) – Part I: Global variance reduction for large-scale radia- tion transport analysis. Comput. Phys. Commun. 291, 108850 (2023). 10.1016/j.cpc.2023.108850 [15] Pan, Q., Lv, H., Tang, S. et al. MAGIC-GPS global variance reduction method for large-scale shield- ing calculation. Nucl. Eng. Des. 414, 112581 (2023). 10.1016/j.nucengdes.2023.112581 [16] Lei, JC., Chen, ZP., Peng, ZQ. et al. Space self-adaption global variance reduction (SsGVR) method for large- scale radiation shielding calculation. Ann. Nucl. Energy 223, 111622 (2025). 10.1016/j.anucene.2025.111622 [17] X.R. Zhang, H.N. Liu, D.C. Zhang, Development and Validation of an Automated Weight Window Generator for Reactor Monte Carlo Programs. Proceedings of the 2024 7th International Conference on Mathematics and Statistics, 30 -36(2024). doi: 10.1145/3686592.3686598 [18] Zhou, W., Sun, G., Yang, Z. et al. BP neural net- work based reconstruction method for radiation field applications. Nucl. Eng. Des. 380, 111228 (2021). 10.1016/j.nucengdes.2021.111228 [19] Li, M., Liu, Y., Peng, M. et al. A fast simulation method for radiation maps using interpolation in a virtual envi- ronment. J. Radiol. Prot. 38, 892 (2018). 10.1088/1361- 6498/aac392 [20] Hao, YS., Wu, Z., Pu, YH. et al. Research on inversion method for complex source-term distributions based on deep neural networks. NUCL SCI TECH 34, 195 (2023). 10.1007/s41365-023-01327-8 [21] Cao, P., Gan, Q., Song, J., Long, P., Wang, F., Hu, L., Wu, Y. An artificial neural network based neutron field reconstruction method for reactor. Ann. Nucl. Energy 138, 107195 (2020). 10.1016/j.anucene.2019.107195 [22] Bagheri, S., Khalafi, H., SMR. 3D source term sim- ulation for exact shielding design based on genetic algorithm. Ann. Nucl. Energy 191, 109915 (2023). 10.1016/j.anucene.2023.109915 [23] Song, Y., Mao, J., Zhang, Z. et al. A novel multi- objective shielding optimization method: DNN-PCA- NSGA- . Ann. Nucl. Energy 161, 108461 (2021). 10.1016/j.anucene.2021.108461 [24] Ma, R., Wang, X., Li, S. et al. An efficient computa- tional method for particle transport problems based on neural networks. Acta Phys. Sin. 73(07), 79-88 (2024). (in Chinese) 10.7498/aps.73.20231661 [25] Berry, J.J., Gil-Delgado, G.G., Osborne, A.G.S.

Classification of a multi- group collision probability model using machine learning. Ann. Nucl. Energy 160, 108367 (2021). 10.1016/j.anucene.2021.108367 group structures [26] Zhang, G., Song, Y., Zhang, Z., Yuan, W. Research and development of source term activity reconstruction system based on deep learning. Ann. Nucl. Energy 175, 109248 (2022). 10.1016/j.anucene.2022.109248 [27] Lu, L., Jin, P., Pang, G. et al. Learning nonlinear oper- ators via DeepONet based on the universal approxima- tion theorem of operators. Nat. Mach. Intell. 3, 218-229 (2021). 10.1038/s42256-021-00302-5 [28] Li, H., Miao, Y., Sharif Khodaei, Z., Aliabadi, M.H. An architectural analysis of DeepONet and a general extension of the physics-informed DeepONet model on solving nonlinear parametric partial differ- ential equations. Neurocomputing 611, 128675 (2025). 10.1016/j.neucom.2024.128675 [29] Yang, Y. DeepONet solving nonlinear partial differential equations with physics-informed training. arXiv:2410.04344 (2024). 10.48550/arXiv.2410.04344 [30] Li, Z., Kovachki, N.B., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A.M., Anandku- mar, A. Fourier neural operator for parametric par- tial differential equations. arXiv:2010.08895 (2020). 10.48550/arXiv.2010.08895 [31] Cao, L., Hong, H., Jiang, M. Fast solving partial differential equations via imitative Fourier neu- Int. Joint Conf. Neural Netw. ral operator. Proc. (IJCNN), Yokohama, Japan, 1–8 (2024). 10.1109/IJCNN60899.2024.10650596 [32] Qamar, S. et al. A variant form of 3D-UNet for infant brain segmentation. Future Gener. Comput. Syst. 108, 613–623 (2020).

[33] Peplow, D.E., Mosher, S.W., Evans, T.M. Consistent adjoint driven importance sampling using space, energy and angle. ORNL/TM-2012/7, Oak Ridge Natl. Lab., Oak Ridge, TN (2012). [34] Gottlieb, D., Shu, CW. On the Gibbs phenomenon and its resolution. SIAM Rev. 39(4), 644–668 (1997). 10.1137/1039095 [35] Gelb, A., Tadmor, E. Detection of edges in spectral data.

Appl. Comput. Harmon. Anal. 7(1), 101–135 (1999). 10.1006/acha.1999.0262 [36] Wu, J., Wu, Z., Hu, A. et al. Optimization progress of large-scale radiation shielding Monte Carlo simula- tion software based on AIS variance reduction tech- nique system: MCShield. EPJ Nucl. Sci. Technol. (2024). 10.1051/epjn/2024032 [37] Hao, Y., Wu, Z., Pu, Y. et al. Research on complex source term inversion and radiation field reconstruc- tion method based on deep neural networks. Proc. 11th Int. Symp. Radiation Safety and Detection Technol., Hanyang Univ., Seoul, Korea (2023). [38] Y.-S. Hao, Z. Wu, S.-S. Gao et al, Research on a Monte Carlo global variance reduction method based on an automatic importance sampling method. Nuclear Sci- ence and Techniques 2024 Vol. 35 Issue 5 Pages 86. doi: 10.57760/sciencedb.j00186.00389 [39] G. I. Bell and S. Glasstone, Nuclear Reactor Theory, Van Nostrand Reinhold, New York, 1970. [40] E. E. Lewis and W. F. Miller Jr., Computational Meth- ods of Neutron Transport, Wiley, New York, 1984.

Note: Figure translations are in progress. See original paper for figures.