# A Neural Network-Accelerated Source Iteration Method for Neutron Transport Simulations with One-Step Convergence

**Authors:** Zhou, Mr. Nan, Li, Prof. Qing, Yu, Prof. Yingrui, Chen, Prof. Zhang, Gong, Prof. Helin, Lou, Dr. Lei, Shen, Dr. Pengfei, Ji, Mr. Wenhao, Guo, Dr. Xia, Zhou, Mr. Nan

**Date:** 2026-02-14T21:59:01+00:00

## Abstract

A neural network framework is proposed to achieve one-step convergence in source iteration, a fundamental method for solving transport equations. By predicting the fully converged source distribution from either early-stage source distributions or intrinsic physical parameters, this framework effectively shortcuts the conventional iterative pathway. A residual convolutional network trained on 2 million randomly generated reactor configurations achieves prediction errors between 1.02% and 3.31% on validation data. When integrated into the iteration loop, the predictor induces an immediate transition to a high-accuracy regime, achieving in five iterations the accuracy that typically requires one hundred conventional steps-a greater than 20-fold acceleration, as validated on 1,000 distinct reactor cores using OpenMC. Further investigations into iteration-step flexibility and spatial decoupling demonstrate the framework ' s potential for generalization, with central-region predictions achieving accuracy comparable to that obtained after extensive iteration (e.g., ~129 steps). This work proposes a data-driven strategy in which high-fidelity predictions guide the iteration directly to convergence.

## Full Text

## Preamble

A Neural Network-Accelerated Source Iteration Method for Neutron Transport Simulations with One-Step Convergence Nan Zhou1, *, Qing Li1,* , Yingrui Yu1, Zhang Chen1, Helin Gong1,2, Lei Lou1, Pengfei Shen3, Wenhao Ji1 & Xia Guo4
[1] Nuclear Power Institute of China, Chengdu, Sichuan, China.

**4 Sichuan University, Chengdu, Sichuan, China.**

*e-mail: zhounan@npic.ac.cn; liqing@npic.ac.cn

## Abstract

A neural network framework is proposed to achieve one-step convergence in source iteration, a fundamental method for solving transport equations. By predicting the fully converged source distribution from either early-stage source distributions or intrinsic physical parameters, this framework effectively shortcuts the conventional iterative pathway. A residual convolutional network trained on 2 million randomly generated reactor configurations achieves prediction errors between 1.02% and 3.31% on validation data. When integrated into the iteration loop, the predictor induces an immediate transition to a high-accuracy regime, achieving in five iterations the accuracy that typically requires one hundred conventional steps —a greater than 20-fold acceleration, as validated on 1,000 distinct reactor cores using OpenMC. Further investigations into iteration-step flexibility and spatial decoupling demonstrate the framework ' s potential for generalization, with central-region predictions achieving accuracy comparable to that obtained after extensive iteration (e.g., ~129 steps). This work proposes a data-driven strategy in which high-fidelity predictions guide the iteration directly to convergence.

## 1. Introduction

Source iteration is a fundamental method in nuclear engineering for solving neutron transport and diffusion equations. Whether employing Monte Carlo methods for core simulations[1] or deterministic approaches such as finite difference[2], nodal methods[3], or the method of characteristics[4], the underlying iterative process remains source iteration.

In this paper, we present a novel framework that leverages early source distributions or intrinsic physical parameters to predict the converged solution with a residual convolutional neural network[5]. When this high-fidelity prediction is reintroduced into the iteration loop, it induces an immediate, one-step core convergence —effectively bypassing the traditional iterative trajectory. This approach achieves significant acceleration (>20×) while rigorously preserving physical consistency, as quantitatively validated on 1,000 distinct, randomly generated reactor core models.

Although this framework requires training for specific geometries, it holds significant value for nuclear industry practices, where the demand for repeated calculations on fixed geometries is widespread. When it comes to a specific reactor core in practical engineering, its geometry lattice is fixed. Whether for engineering design, fuel reload design, or rapid in-core detector diagnostics during operation, a vast number of high-fidelity solutions for this fixed lattice are required.

In the latter part of this paper, we explore iteration-step decoupling and spatial-decoupling strategies, demonstrating the framework' s potential for broader applicability.

## 2.1 Source Iteration Sample Generation and Analysis

To investigate the convergence dynamics of source iteration in particle transport, we developed a Stochastic Homogenized Assembly Core Generator (SHACG). SHACG generates statistically representative samples of the source iteration process by producing a multitude of randomized reactor core configurations. Each configuration is an 8$×$8 lattice comprising 3 to 7 unique fuel assembly types, with stochastic U-235 enrichment ranging from 1.5% to 7.5%. For each generated core, a modified version of the OpenMC [6] Monte Carlo code was employed to simulate the source iteration process, which consists of the following cyclic steps: 1) Source Sampling: Initial neutron histories are emitted by sampling from the spatial fission rate distribution determined in the preceding generation. 2) Transport Simulation: Neutron histories are tracked through the geometry until absorption or leakage. 3) Source Convergence: Fission events occurring during the generation are tallied to define the updated fission source distribution for the next cycle.

This iterative procedure drives the fission rate distribution toward convergence with the statistically precise solution obtained from subsequent active generations. SHACG recorded the evolving fission rate distribution over the first 100 inactive generations, along with the final converged reference distribution and the core' s enrichment map (as an intrinsic physical parameter), as one complete sample.

The convergence mechanism can be understood through eigenvalue decomposition. The , is a superposition of eigenvectors fission rate distribution at generation N, corresponding to different eigenvalues Assuming 1l is the dominant eigenvalue (| 1l | > | kl | for k > 1), and considering normalization, the influence of (the eigenvector associated with 1l ) becomes increasingly pronounced as N increases, while contributions from other eigenvectors decay.

Consequently, after sufficient generations (N » 1), the distribution converges to: c k R where is the converged solution.

## 2.2 Analysis of Source Iteration Convergence Process

To analyze the convergence dynamics, we employed SHACG to generate 3,000 samples of the source iteration process from randomized reactor cores. Macroscopically, the mean absolute error (MAE) relative to the reference solution decreased steadily over generations, reaching approximately 2.62% by generation 100 (Fig. 1 [FIGURE:1]). However, a microscopic examination of individual spatial points revealed more complex behaviors. While many points exhibited monotonic convergence (positive drift), a substantial and persistent fraction

(~50% in later generations) displayed non-monotonic oscillations around the final value, indicative of negative drift (Fig. 2 [FIGURE:2]). Notably, the proportion of points exhibiting continuous positive drift from the outset decayed rapidly: only about 24.5% persisted for the first 10 generations, dropping to nearly zero (~0.1%) by generation 50, as detailed in Table 1 .

Figure 1 Monte Carlo source iteration convergence Figure 2 Negative drift analysis in Monte Carlo source iteration Table 1 Proportion of data points exhibiting continuous positive drift Source Iteration Index First 10 First 20 First 30 First 40 First 50 Continuous Positive Drift Count Percentage 24.5%

## 2.3 Data Dynamics Analogy and Derivation of Sufficient Conditions

The observed micro-scale disorder (specifically, oscillations that persist amidst macro-convergence) necessitates a conceptual model. We propose an analogy to Newtonian dynamics: the source distribution at generation N, solution by a dominant "elastic force" elasticF , is driven toward the converged proportional to their difference. elastic The non-monotonic behavior arises from a secondary "neighbor influence force" neighbor . This force depends on local differences within the current distribution ( 1)NR - and intrinsic physical parameters such as the macroscopic cross-section Σ which is strongly correlated with enrichment. neighbor Consequently, the evolution from ( 1)NR + can be viewed as the result of an unknown mapping M acting on the current state and the combined forces: elastic neighbor Therefore, an unknown mapping H relates the current state and intrinsic physical parameters directly to the converged solution:

Beginning with an average source distribution, the fission rate distribution in each generation and the intrinsic parameters are recursively determined. This recursive application implies a fundamental relationship: the converged solution is a complex function of any intermediate distribution distribution Σ: (at an arbitrary generation N) or the intrinsic parameter This analysis yields a key insight: neural networks are inherently capable of approximating such complex mappings. Therefore, the sufficient conditions for predicting the converged solution with a neural network are the knowledge of either (1) an early-intermediate source distribution, or (2) intrinsic physical parameters strongly coupled to the macroscopic cross-section (e.g., enrichment). This derivation provides the theoretical foundation for the proposed accelerated convergence framework.

## 3.1 Neural Network Framework and Training

To validate the theoretical sufficiency conditions established in Section 2.3, we implemented a 12-layer residual convolutional neural network (Res-CNN) with 64 filters per layer (Fig. 3 [FIGURE:3]). The model accepts 8$×$8 matrices representing either enrichment distribution or early-generation fission rate distributions as input and outputs predicted converged fission rate distribution.

Residual connections ensure stable gradient propagation during the training of this deep architecture.

Figure 3 Schematic of Res-CNN architecture Training utilized a dataset of 2 million statistically unique reactor cores generated by SHACG. The combinatorial space exceeding 10^{69} guarantees zero duplicates. A separate validation set of 4,000 independent cores (0.2% of the total data) monitored generalization performance. The mean absolute error (MAE) between predictions and the reference converged solutions served as the loss function:

L MAE R Where N is the number of samples in a training batch, and the superscript k indexes the training samples.

Model optimization was performed using the Adam algorithm with an initial learning rate of 1$×$10-3, incorporating adaptive learning rate scheduling and gradient clipping (threshold: 1.0). The training process was accelerated via Horovod-based distributed computation.

Seven distinct input schemes were evaluated to probe the prediction capability:

Model 0: Enrichment distribution. Models 1–6: Fission rate distributions from generations 1, 6, 11, 16, 21, and 26, respectively.

All models achieved high predictive accuracy, with validation MAE ranging from 1.02% to 3.31% (Table 2 ). The enrichment-based Model 0 (MAE: 1.02%) outperformed the distribution-based models. This is attributed to the inherent stochastic noise in Monte Carlo fission tallies, which is absent in the deterministic enrichment parameter. Notably, distribution-based models maintained an MAE $\leq$ 3.31% even when trained on late-generation inputs, confirming the sufficiency of early-stage distributions for convergence prediction.

Minimal overfitting was observed, with a divergence of less than 0.5% between training and validation loss.

Table 2 Model performance across different training input schemes Model model0 model1 model2 model3 model4 model5 model6 Training Input Enrichment 1st Generation Distribution 6th Generation Distribution 11th Generation Distribution 16th Generation Distribution 21st Generation Distribution 26th Generation Distribution Validation MAE 1.02% 2.51% 3.31% 3.29% 3.10% 3.03% 3.16% These trained models—particularly Model 0 (highest accuracy) and Model 1 (earliest injection point)—were subsequently integrated into the source iteration process to trigger the predicted one-step convergence, as detailed in the following section.

## 3.2 Validation of One-step Convergence

Building upon the trained models described in Section 3.1, we integrated Model 0 (enrichment-based) and Model 1 (first-generation distribution-based) into the source iteration workflow within OpenMC. The validation was conducted on a set of 1,000 unseen reactor cores generated by SHACG following this procedure:

Figure 6

Figure 1: Figure 6

1) A standard source iteration was executed for the first inactive generation. 2) The neural network prediction was substituted for the source distribution of the second generation. 3) Convergence was monitored via the MAE relative to the reference solution. a) Results for Model 0 b) Results for Model 1 Figure 4 [FIGURE:4]: MAE evolution with neural network integration.

Critical Findings: Convergence Phase Transition: The injection of the neural network prediction caused an immediate discontinuity in the MAE trajectory (Fig. 4), bypassing the incremental refinement characteristic of standard iteration. The system transitioned directly to a stabilized convergence regime, achieving effective convergence in one step. Subsequent iterations serve only to marginally refine the solution.

Acceleration Factor: Model 0 maintained an MAE below 2.62% starting from generation

## 5. This performance is equivalent to that achieved after 100 traditional iterations, representing

an acceleration factor exceeding 20×. Model 1 achieved a sustained MAE below 2.62% from generation 16 onward.

## 3.3 Extending Applicability: Decoupled Prediction Scenarios

Although practical engineering often involves repeated simulations on fixed geometries, we explored two extensions to demonstrate the framework' s broader applicability: 1) Iteration-Step Flexibility, and 2) Spatial Decoupling.

Iteration-Step Flexibility To enable predictions at arbitrary stages (e.g., for checkpoint restarts), models trained on specific distributions (e.g., from generations 1 & 16) were tested with inputs from other, non-matching iterations. As shown in Fig. 5 [FIGURE:5], prediction accuracy degraded when inputs deviated from the trained generation. To address this, we implemented a decoupled training strategy where input data was randomly selected from defined generation windows (e.g., generations 14–21 or 5–25) during training. The retrained models showed improved accuracy across early generations (Fig. 6

), demonstrating their potential iteration-agnostic application. (Batch Number corresponds to Source Iteration Index) Figure 5 Accuracy degradation with input-generation mismatch. (Batch Number corresponds to Source Iteration Index) Figure 6 Performance of models retrained with decoupled (window-based) input sampling Spatial Decoupling Monte Carlo source iteration

inherently simulates neutron transport across geometric boundaries. To assess the potential for localized prediction in larger systems, models trained on 8$\times$8 *cores were applied to larger* 16$\times$16 geometries via a structured workflow:

Input Preprocessing: The full 16$\times$16 *fission rate distribution was cropped to an* 8$\times$8 *matrix by removing peripher*
*The neural network predicted the converged fission rate distribution for the cropped* 8$\times$8 *region.* 3) *Output Post—processing : The outer layers of the* 8$\times$8 *prediction were stripped away, retaining only the central* 4$\times$4 *or* 2$\times$2 core regions to ensure spatial fidelity. 4) Validation: The retained central regions were renormalized and compared with the reference solution to assess accuracy.

The localized prediction workflow is illustrated in Fig. 7 [FIGURE:7]. The convergence behavior of 500 full 16 $\times$16 cores served as the baseline (Fig. 8 [FIGURE:8]). For the most central (2 $\times$2) regions, Model 0 achieved a prediction accuracy equivalent to that obtained after 129 traditional iterations (Table 3 ), suggesting promising potential for applying the framework to localized prediction within larger systems.

Figure 7 Schematic workflow for localized prediction in larger (16$\times$16) geometries.

Figure 8 Source iteration convergence baseline for 16$\times$16 cores Table 3 Equivalent iteration steps achieved by localized predictions for central regions Res-CNN Model Central Region Average MAE Equivalent Iteration Step Model 0 Model 0 Model 1 Model 1 42.81% 17.39% 46.80% 20.29%

## 4. Conclusion and Open Questions

This study presents a neural network-driven framework that achieves one-step convergence in source iteration —a cornerstone algorithm for transport simulations. By establishing sufficient conditions for convergence prediction through a data-dynamics analogy and training residual convolutional networks on 2 million unique reactor configurations, we demonstrate that either early source distributions or intrinsic parameters (e.g., enrichment) enable accurate prediction of converged solutions, with validation errors ranging from 1.02% to 3.31%. Integrating these predictions into the iteration loop induces an immediate convergence transition, bypassing incremental refinement and yielding an acceleration factor exceeding 20 $\times$ while preserving physical consistency, as demonstrated on 1,000 distinct cores.

The framework' s potential for broader applicability is explored through two decoupling strategies: iteration-step decoupling, in which a model trained over a range of steps maintains accuracy across them; and spatial decoupling, which demonstrates promising capability for generating accurate localized predictions within larger geometries.

Several open questions remain for future investigation. Extending the framework to deterministic transport methods, which operate on distinct mathematical principles, represents a logical next step to broaden its impact. Additionally,

the development of more lightweight network architectures that retain the one-step convergence capability while reducing model complexity and training costs warrants exploration. Although our preliminary studies on step and spatial decoupling suggest pathways toward generalization, a universal, scenario-agnostic acceleration solution remains a long-term goal. Nevertheless, this work proposes a data-driven approach for optimizing iterative solvers, offering a potential pathway to enhance computational efficiency in scientific disciplines reliant on source iteration processes.

Author contributions N.Z. conceived and led the study, developed the core methodology and framework, conducted the major experiments, and drafted the manuscript.

Q.L. supervised the research and provided strategic direction on integrating deep learning with transport theory.

Y.Y. contributed to the physical interpretation and theoretical consistency of the work.

Z.C., H.G., and L.L. contributed to code development and optimization for data generation and training.

P.S., W.J., and X.G. assisted with data analysis, visualization, and manuscript preparation.

All authors reviewed the results and approved the final manuscript.

Reference [1]. Wang K, Liu Z, An N, et al. The Reactor Monte Carlo code RMC: The state-of-the-art technologies, advancements, applications, and next[J]. EPJ Nuclear Sciences & Technologies, 2024, 10: 24. [2]. Ferguson DR, Derstine KL. Optimized iteration strategies and data management considerations for fast reactor finite difference diffusion theory codes. Nuclear Science and Engineering 1977;64(2):593-604. [3]. An P, Ma YQ, Xiao P, Guo FC, Lu W, Chai XM. Development and validation of reactor nuclear design code CORCA-3D. Nuclear Engineering and Technology 2019;51(7):1721-1728. [4]. Eklund M, Alamaniotis M, Hernandez H, Jevremovic T. Method of characteristics –A review with applications to science and nuclear engineering computation. Progress in Nuclear Energy 2015;85:548-567. [5]. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In:

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016. p. 770-778. [6]. Romano PK, Horelik NE, Herman BR, Nelson AG, Forget B, Smith K. OpenMC: A state-of-the-art Monte Carlo code for research and development. Annals of Nuclear Energy 2015;82:90-97.