

# 基于神经网络的粒子输运问题高效计算方法\*

马锐堉<sup>1)</sup> 王鑫<sup>1)2)</sup> 李树<sup>1)</sup> 勇珩<sup>1)</sup> 上官丹骅<sup>1)†</sup>

1) (北京应用物理与计算数学研究所, 北京 100094)

2) (中物院高性能数值模拟软件中心, 北京 100088)

(2023年10月17日收到; 2024年2月2日收到修改稿)

蒙特卡罗方法是求解粒子输运问题的有力工具之一, 其局限性在于为达到精度要求需模拟大量粒子, 计算耗时长, 这阻碍了该方法的进一步应用, 尤其在需快速响应的情形. 本文结合神经网络和若干蒙特卡罗方法基本原理发展了一种计算方法, 能够实现源分布可变, 几何、材料和目标计数不变的中子输运问题的快速准确求解. 首先, 为高效生成用于神经网络训练的数据, 利用重要性原理实现在同样模拟次数基础上有效扩充训练数据集容量, 在一定程度上克服了使用蒙特卡罗计算获取训练数据耗时的缺点. 进而, 基于目标计数是源分布与重要性函数乘积积分的事实, 设计了利用神经网络实现快速输运计算的策略. 该网络的输入是中子源项, 输出是目标计数, 在几何、材料和目标计数固定的情况下, 该神经网络可重复使用, 根据新的源项快速准确得到目标计数. 本文所提出方法的原理和框架同样适用于其他种类粒子的同类型输运问题. 基于若干基准模型的验证表明, 训练得到的神经网络能在不到 1 s 的时间内得到目标计数, 且与蒙特卡罗大样本模拟得到基准结果的平均相对偏差均低于 5%.

**关键词:** 蒙特卡罗方法, 神经网络, 粒子输运, 重要性原理

**PACS:** 28.20.-v, 24.10.Lx, 02.70.Uu

**DOI:** 10.7498/aps.73.20231661

## 1 引言

蒙特卡罗 (Monte Carlo, MC) 方法和确定论方法是求解粒子输运问题的两类主要方法<sup>[1]</sup>. 随着核技术的发展, 各类新型核设施具有更复杂的几何结构和反应机制, 各向异性更强, 同时对计算精度的要求也在不断提高. 为解决这类复杂粒子输运问题, 在计算机硬件性能不断提高的基础上, MC 方法由于能够对物理过程和几何构型进行精确建模且并行效率高, 因此得到广泛的发展和应用<sup>[2]</sup>. 同时, pin-by-pin 反应堆模型输运<sup>[3]</sup>、复杂多物理场耦合计算<sup>[4]</sup> 等大规模精细模型快速计算要求也在不断推动 MC 方法相关研究的开展<sup>[5]</sup>. 由于 MC

方法求解目标计数的相对误差与样本数平方根成反比, 而计算时间与样本数成正比, 为使误差达到要求, 其模拟过程非常耗时<sup>[3,6-8]</sup>. 国内外研究团队已发展多种减方差方法<sup>[9-19]</sup>, 目标是以较少的样本数达到更低的统计误差. 权窗算法是其中重要的一类, 一般认为比几何重要性等算法效率更高, 其算法参数可由重要性原理产生. 重要性原理最早由 Booth 和 Hendricks<sup>[14]</sup> 提出, 网格的重要性定义为进入该网格的单位权重的粒子及其后代对计数器的总贡献. 虽然存在大量研究, 但 MC 输运的绝对计算时间仍然相当可观, 且各种减方差方法的适应条件不同, 在应用上有一定的局限性.

深度学习中的神经网络是一种高度并行的信息处理系统, 具有很强的自适应学习能力, 能处理

\* 国家自然科学基金 (批准号:12305173, 12375164, 12331010)、国家自然科学基金委员会-中国工程物理研究院 NSAF 联合基金 (批准号: U2230208) 和核数据重点实验室基金 (批准号: JCKY2022201C155) 资助的课题.

† 通信作者. E-mail: [sgdh@iapcm.ac.cn](mailto:sgdh@iapcm.ac.cn)

复杂的多输入、多输出非线性系统<sup>[20-23]</sup>, 训练完成的网络可根据输入快速输出目标值<sup>[24-27]</sup>, 在诸多领域有广泛应用<sup>[28-32]</sup>. 近年来, 在核相关领域利用数据驱动的神经网络进行中子输运计算的趋势日益明显. Berry 等<sup>[33]</sup> 研究利用神经网络和随机森林分类器评估给定能群结构的适用性, 以准确预测轻水堆中的中子倍增因子. Cao 等<sup>[34]</sup> 提出基于神经网络的重构中子场方法, 用于重建反应堆压力容器和堆芯区域内的中子场. Zhang 等<sup>[35]</sup> 提出一种基于神经网络的源项重建方法, 实现复杂的三维源项快速重建. 张海明等<sup>[36]</sup> 建立了基于卷积神经网络的深度学习模型, 可用于直接预测反应堆堆芯有效增殖因子. 针对多组分中子屏蔽材料优化设计中 MC 模拟计算时间长的制约问题, 林海鹏等<sup>[37]</sup> 讨论了利用神经网络算法快速预测不同材料中子屏蔽效果的方法. Osborne 等<sup>[38]</sup> 通过上采样方法实现从少量粒子数 MC 模拟结果到精确结果的映射. Kim 等<sup>[39]</sup> 构建了一个以堆芯结构为输入的神经网络, 可快速优化堆芯结构设计. 与其他类型粒子输运结合的研究, 主要是在医学图像处理以及人体剂量学领域<sup>[40-42]</sup>.

在上述数据驱动的神经网络应用中, 网络通过分析所提供数据集的模式和输入输出之间的映射关系来学习和进行快速预测, 因此训练数据集的数量和质量对网络的鲁棒性和有效性有很大影响, 而训练数据的获取通常采用 MC 模拟或其他数值计算方法得到, 难以避免大量计算时间和计算资源的消耗. 对于数据驱动型的神经网络, 如何利用神经网络代替特定问题的传统计算, 以及如何利用传统方法高效获取训练数据都具有重要的研究意义.

本文的研究对象是源分布可变, 几何构型、材料和目标计数不变的问题, 目标是在结果合理的前提下提高计算速度. 这类问题在国防科技和国民经济领域存在多种应用场景. 例如, 中子输运是战略国防科技研究所必须的多物理耦合计算中最消耗计算资源的一环, 计算时间可以占总体时间的 90% 以上, 在装置变化缓慢阶段, 可认为每一步中子输运的源项都有变化而装置的几何、材料变化可忽略不计, 因此呈现为变源问题; 在民用经济方面, 反应堆高置信度数值模拟需要快速计算压力容器的中子注量率, 而反应堆构型复杂、尺度大的特点导致其求解非常耗时, 且工况复杂、燃料布置方案

多样, 以堆芯整体为源项进行模拟计算时是典型的变源问题, 面临着快速求解难题; 另外, 变源问题的快速实时求解也是为“脏弹”恐怖袭击的科学预防提供备选方案的重要技术手段.

本文的主要贡献和创新之处在于: 1) 以高效产生训练数据为目标, 从重要性的定义出发, 首次将其应用在等效 MC 计算的训练样本生成中, 发展出神经网络训练数据高效生成方法, 达到在同样模拟次数基础上有效扩充训练数据集容量的效果. 2) 基于目标计数是源分布与重要性函数乘积积分的事实, 构建可重复使用的高效神经网络, 以快速准确获得各种源分布对应的目标计数, 一定程度上突破了传统 MC 输运计算的效率瓶颈. 虽然本文的研究对象是中子, 但该方法可适用于其他种类粒子的同类型输运问题中.

## 2 方法介绍

本文建立了一种利用神经网络加速 MC 模拟中子输运问题的方法 (图 1 为技术路线图), 该神经网络所必需的训练数据集是通过相应的输运问题进行 MC 模拟而获得的. 对于新的中子源分布, 将其输入给训练好的神经网络就能快速输出相应的目标计数. 数据驱动的神经网络需要足够的高质量训练数据, 但通过传统的 MC 模拟来获取这些训练数据非常耗时. 针对这一问题, 将重要性原理作为一种高效生成训练数据的手段, 由此在一定程度上提高了神经网络的训练效率.

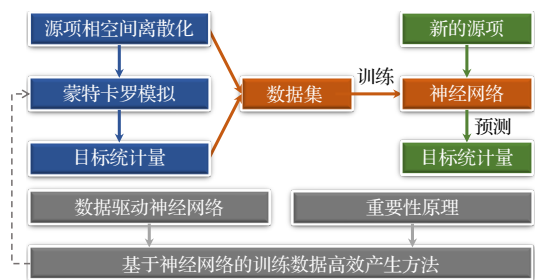


图 1 本文技术路线图

Fig. 1. Study framework of this paper.

### 2.1 理论基础

对于中子固定源问题, 统计体通量、能量沉积、剂量等物理量都与源项信息直接相关, 其中源项信息包括位置分布、能量分布、粒子发射方向等

多维信息. 中子的属性用  $\mathbf{p} = (\mathbf{r}, E, \mathbf{w})$  表示, 其中  $\mathbf{r}$  表示粒子空间位置,  $E$  表示能量,  $\mathbf{w}$  表示发射方向. 中子源分布满足归一条件:

$$\int_{\Omega} S(\mathbf{p}) d\mathbf{p} = 1. \quad (1)$$

MC 模拟中目标计数可以表示为

$$I = \int_{\Omega} S(\mathbf{p}) g(\mathbf{p}) d\mathbf{p}, \quad (2)$$

其中,  $S(\mathbf{p})$  为源分布,  $g(\mathbf{p})$  为重要性函数. 将相空间  $\Omega$  进行网格划分, 即  $\Omega = \sum_{i=1}^M \Omega_i$  ( $M$  足够大, 表示划分足够精细), 当源分布所在相空间划分足够精细的假设成立时, 相空间内单个网格的重要性函数可以用该网格内一个点的重要性函数表示, 则 (2) 式可被表示为

$$I = \sum_{i=1}^M \int_{\Omega_i} S(\mathbf{p}) g(\mathbf{p}) d\mathbf{p} \approx \sum_{i=1}^M g(\mathbf{p}_i^*) \int_{\Omega_i} S(\mathbf{p}) d\mathbf{p}. \quad (3)$$

$\mathbf{p}_i^*$  表示空间  $\Omega_i$  中一定点, 记  $\int_{\Omega_i} S(\mathbf{p}) d\mathbf{p} = q_i$ ,  $g(\mathbf{p}_i^*) = \omega_i$ , 则有

$$I \approx \sum_{i=1}^M q_i \omega_i. \quad (4)$$

对于固定源问题, 同一个模型 (包括几何及材料) 和目标量对应同一套  $\{\omega_i\}_{i=1}^M$ , 对于不同的源分布, 将  $S(\mathbf{p})$  离散化为  $\{q_i\}_{i=1}^M$  代入 (4) 式, 即可得到计数  $I$ .

本文以体平均通量作为目标计数 (相同方法可适用于任何计数). 对于不考虑时间变量的固定源问题, 定义

$$\phi(\mathbf{r}, E) = \int \Psi(\mathbf{r}, \mathbf{w}, E) d\mathbf{w}, \quad (5)$$

其中,  $\Psi(\mathbf{r}, \mathbf{w}, E)$  是中子角通量密度, 则体平均通量可以被表示为

$$\Phi_V = \frac{1}{V} \int_{E_{\min}}^{E_{\max}} dE \int dV \phi(\mathbf{r}, E). \quad (6)$$

$\Phi_V$  可以通过径迹长度估计法计算, 粒子一旦通过计数区域就会对目标计数造成贡献, 一般而言较其他估计法效率更高, 因此被多数 MC 输运程序采用.

## 2.2 基于神经网络的粒子输运问题高效计算方法

对于源分布多变, 几何、材料构型及目标计数稳定的一系列中子输运问题, MC 模拟可看作一个

黑箱, 其输入是源分布, 输出是固定的目标物理量, 这个黑箱通过精确度很高, 但是速度较慢的模拟将输入转换为输出. 通过设计合适的神经网络并以合理的形式表示输入端的源分布函数, 在获得充足的训练数据基础上, 训练得到的神经网络将以很快的响应速度得到新的中子源分布对应的目标物理量.

在实际应用中, 难以实现将连续形式的、不同分布的源项处理为神经网络的统一输入形式. 为了解决此问题, 本文首先对源分布进行  $N$  次采样, 从而产生  $N$  个源粒子. 然后将这些粒子的六维属性  $(\mathbf{r}, E, \mathbf{w})$  作为点分布输入网络. 最终网络产生  $N$  个输出的平均值被视为目标统计量的估计值.

为实现上述目标, 将源分布在相空间离散化处理, 进而以 (4) 式表达目标量后, 可用训练得到合适的神经网络模型代替 (4) 式中权值  $\{\omega_i\}_{i=1}^M$ , 方法和网络结构示意图如图 2 所示. 本文采用多层感知神经网络 (MLP) [21] 的模型结构. 神经网络中单层感知器的拟合能力有限, 而多层感知器的拟合能力更强, 可用于解决复杂问题, 因而 MLP 的优势在于它可以学习将任何输入映射到输出的权值. 最典型的 MLP 由三个网络结构组成: 输入层、隐藏层和输出层. 本文使用的网络包括一个输入层、多个全连接层和一个输出层. 在全连接层之间加入了整流线性单元 (ReLU) 激活函数, 激活函数为网络引入了非线性特征, 有助于网络学习输入和输出之间的复杂关系, 并能将神经元的输出限制在一定范围内. 常用的激活函数包括 Tanh, Sigmoid 和 ReLU, ReLU 的优点是能有效缓解梯度消失问题, 从而提高训练效率. 各层之间还引入了一个

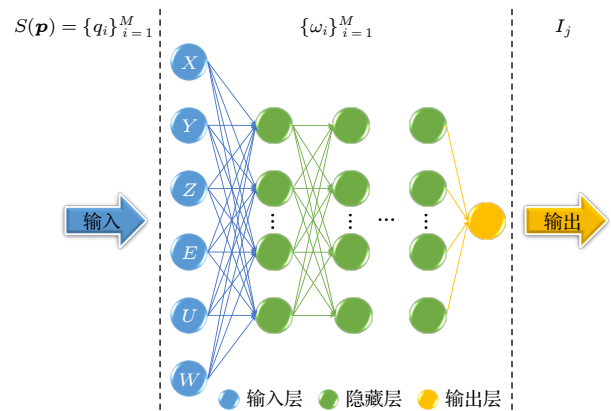


图 2 用于代理加速 MC 模拟的神经网络结构

Fig. 2. Data-driven neural network for Monte Carlo simulation.

dropout 层, 以减少过拟合, 提高神经网络的泛化能力, 使其更加稳健和精确.

使用 MC 模拟产生训练集, 每次模拟的源分布是  $\{S_j(\mathbf{p})\}_{j=1}^K$ , 得到对应的目标量结果记为  $I_j (j = 1, \dots, K)$ . 单个算例模拟足够粒子数使得  $\{I_j\}_{j=1}^N$  的统计误差满足精度要求, 由此得到高质量的训练样本. 对于一个新的源项  $S'(\mathbf{p})$ , 输入神经网络后即可输出对应的  $I'$ .

神经网络需要大量高质量的训练数据, 以减少过拟合和网络泛化性不足等问题. 然而通过 MC 模拟生成如此大量的训练样本必定成本高昂. 因此, 本文提出了一种基于重要性原理<sup>[14]</sup>的方法来高效获取足够的训练数据, 此方法可在单次 MC 模拟中获取尽可能多的训练数据. 重要性原理最早由 Booth 和 Hendricks<sup>[14]</sup> 提出, 用于运输计算中的权重窗参数生成算法. 在本文中, 重要性不再是作为生成权重窗参数的基础, 而是作为相应网格中的源对目标计数的最终贡献, 从而扩大了从单次算例模拟中获得的网络训练样本的数量. 具体说明如下: 将相空间划分为网格时, 网格  $i$  的重要性定义为通过该网格的粒子在单位权重下对目标计数的贡献.

$$\text{某栅元重要性} = \frac{\text{进入该栅元内的粒子及其后代对计数器的总贡献}}{\text{进入该栅元内的粒子总权重}}. \quad (7)$$

当  $\Omega$  空间内网格划分足够细时, 对  $\Omega_i$  内的点  $\mathbf{p}_i$  与其所在网格内定点  $\mathbf{p}_i^*$  距离足够近, 则有

$$\int_{\Omega_i} S(\mathbf{p})g(\mathbf{p})d\mathbf{p} = \int_{\Omega_i} \delta(\mathbf{p}-\mathbf{p}_i^*)g(\mathbf{p})d\mathbf{p} = g(\mathbf{p}_i^*), \quad (8)$$

其中,  $g(\mathbf{p}_i^*) \approx \omega_i$ . 因此, 一个网格的重要性可以用该网格中一个点的重要性来表示. 网格的重要性可视为该区域内点源产生的统计结果, 通过一个算例的 MC 模拟, 可根据 (7) 式统计出所有粒子所经过网格的重要性数据. 通过这种方法, 可以获得相空间内大量网格的重要性数据. 此外, 重要性结果的统计误差还可以作为判定是否可用于网络训练的标准. 在传统的 MC 方法中, 一个粒子的模拟只能产生一个用于网络训练的数据. 然而利用重要性原理, 每次模拟后都能获得多个网格的重要性数据. 符合统计误差要求的每个网格的重要性均可以视为该网格中点源产生的目标计数. 由此实现在同

样模拟次数的基础上有效扩展训练数据集.

上述方法在开源软件 OpenMC<sup>[43]</sup> 中得以实现. 对于每个网格, 定义了两个新的向量, 分别记录网格的总权重和总计数. 对于每个粒子, 都定义了一个新的向量来记录粒子在整个轨迹中每一步所穿过的网格以及当前的权重. 当前粒子被杀死时, 记录的网格和相应的权重会被累加到对应网格的总权重向量中. 此外, 如果粒子被判定已到达最终统计区域, 当前粒子的权重将被累加到它 (及其子代粒子) 穿过的所有网格的计数向量中. 当所有源粒子及其次级粒子的模拟结束后, 根据 (7) 式计算所有网格的重要性.

### 3 结果与讨论

#### 3.1 重要性原理产生数据的正确性验证

为验证重要性结果计算的正确性, 建立简单几何算例: 几何为  $10 \text{ cm} \times 10 \text{ cm} \times 10 \text{ cm}$ 、材料为  $^{56}\text{Fe}$  的正方体. 源项为位于坐标  $(0, 0, 0)$  的中子点源, 发射方向为沿  $X$  轴正方向的单一方向, 能量为  $10 \text{ MeV}$  的单一能量. 在几何相空间划分  $0.1 \text{ cm} \times 0.1 \text{ cm} \times 0.1 \text{ cm}$  的虚网格, 统计目标区域探测器 ( $1 \text{ cm} \times 1 \text{ cm} \times 1 \text{ cm}$ ) 的通量计数. 分别建立 10 个探测器, 其中心位于  $X$  轴的不同位置. 本数值实验基于相空间内一个点源的在进行 MC 模拟的同时统计重要性结果, 重要性结果取源项所在网格的重要性值, 即以该网格为源对目标区域产生的目标值贡献, 验证 MC 计算结果和重要性结果如图 3 所示.

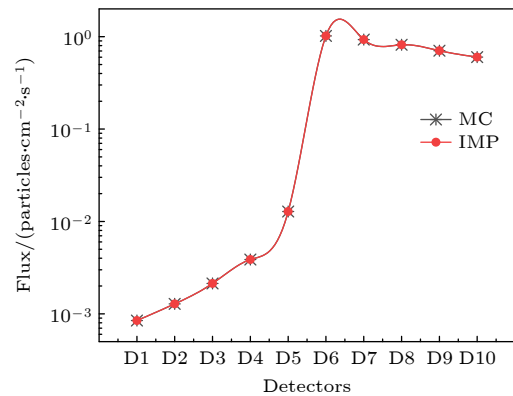


图 3 重要性结果和 MC 模拟结果对比 (MC-MC 计算结果, IMP-重要性原理计算结果)

Fig. 3. Comparison of results of MC simulation and importance data (MC- results of Monte Carlo simulation, IMP- results of importance data).

由重要性值的定义可知,本算例中重要性的统计过程为由该相空间网格内的此点源产生的样本对目标区域的贡献除以总粒子权重,而针对这一点源 MC 模拟同样为统计由此点源产生的目标结果除以总粒子数的平均值,因此对于此验证算例重要性与 MC 模拟结果完全一致,说明本研究在 OpenMC 中重要性数据产生方法实现的正确性.

### 3.2 神经网络结果验证

#### 3.2.1 简单模型

基于 3.1 节所设置简单算例对网络进行训练和测试. 相空间网格的划分参数如表 1 所列. 在相空间中随机生成 1000 个 MC 算例 (每个算例都是不同的点源), 每个算例模拟  $10^6$  个粒子, 以确保结果的最大统计误差低于 5%. 同时, 用重要性生成方法计算生成 1000 套重要性数据及统计误差, 以重要性结果的统计误差小于 2% 为标准选取重要性数据, 使用传统 MC 模拟计算 1000 个算例仅能产生 1000 个源项及其目标统计量作为神经网络的样本, 即训练集容量为 1000, 使用本文提出的基于重要性原理的高效训练数据产生方法后, 在 1000 个算例获得 MC 模拟结果的同时产生了 24285 个重要性数据, 均可以视为该网格中点源产生的目标统计量, 即将 1000 次模拟得到的训练集容量提升了 24 倍, 显著提升了单次模拟获得的训练样本数. 分别采用以下两组数据作为数据集对神经网络进行训练和测试: 1) 仅采用 1000 个算例 MC 模拟的结果; 2) 结合重要性数据和 MC 模拟结果. 预测通量与真值的相对偏差被用作损失, 并按 (9) 式计算:

$$L_{\text{relativeerr}} = |x - y|/x, \quad (9)$$

其中,  $x$  是作为真值的 MC 计算结果;  $y$  是网络预测的通量结果.

表 1 简单模型相空间划分参数

Table 1. Phase space meshing parameters of the Fe model.

变量	范围	网格尺寸	网格数目
$X/Y/Z$	-3—3 cm	0.2 cm	30
$E$	9—10 MeV	0.2 MeV	10
$U$	1.04—2.05 rad	0.2 rad	5
$W$	1.04—2.05 rad	0.2 rad	5

注:  $X/Y/Z$  为相空间网格空间维度在  $X/Y/Z$  方向坐标;  $E$  为相空间网格能量维度坐标;  $U/W$  为相空间网格角度维度在极角和方位角的坐标

神经网络的构建和计算使用 TensorFlow<sup>[44]</sup> 实现. 采用自适应学习率优化算法 Adam<sup>[45]</sup> 作为优化器, 学习率为  $10^{-4}$ . Adam 优化器的优势在于无需手动调整, 能根据每个参数的梯度和动量变化来调整其学习率, 可以纠正其学习率消失、收敛过慢或是高方差的参数更新导致损失函数波动较大等问题.

图 4 显示了网络训练过程中的损失函数曲线. 在每个网络的训练测试中, 80% 的数据集用于训练, 10% 用于验证, 10% 用于测试. 当基于验证数据的损失开始增加时, 即终止训练. 两个网络分别在 5000 和 10000 个算例时终止训练. 最终网络输出结果与 MC 计算给出的真值的相对偏差分别收敛到 19.75% 和 0.97%. 这表明, 采用重要性原理补充数据集可以使网络更好地收敛.

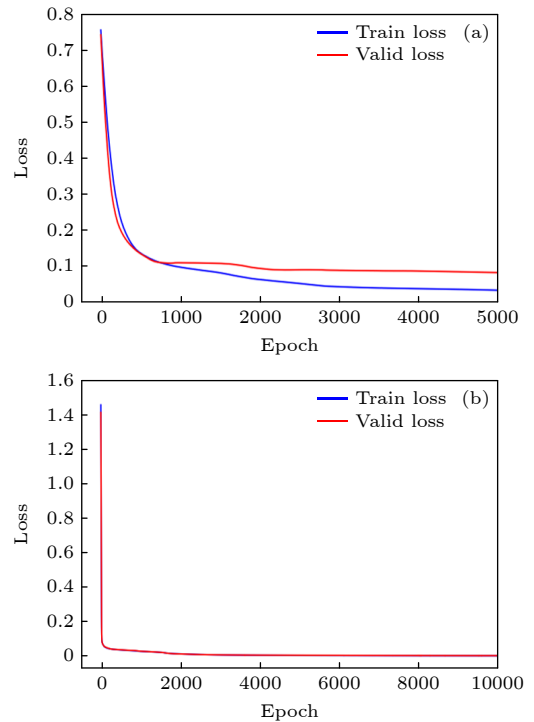


图 4 网络训练损失曲线 (a) 使用 1000 个 MC 模拟结果样本进行训练; (b) 结合重要性数据和 MC 结果进行训练

Fig. 4. Network training loss curves: (a) Trained with 1000 results samples of MC simulation; (b) trained with a combination of importance data and MC results.

对 MC 模拟和神经网络预测的时间耗费进行了比较. 在 MC 模拟中, 模拟  $10^6$  个粒子以确保统计误差保持在 5% 以下, 使用 1000 个实例的平均时间作为 MC 模拟用. 将 1000 个算例输入已训练好的神经网络, 并计算平均预测时间作为神经网络

用时. 结果表明, 神经网络将计算时间大约从 70 s 减少到 0.3 s, 这意味着计算用时减少了约 200 倍, 显著提高了计算速度.

### 3.2.2 Kobayashi 模型

使用 Kobayashi 模型<sup>[46]</sup>中的问题 1 测试了本文所提出的方法. 该模型包括三个区域: 源区域、真空区域和屏蔽区域 (图 5).  $X = 0$  和  $Y = 0$  平面是该模型的两个反射面. 该例题给定单一截面, 则仅在几何和角度维度进行相空间划分, 划分参数如表 2 所列.

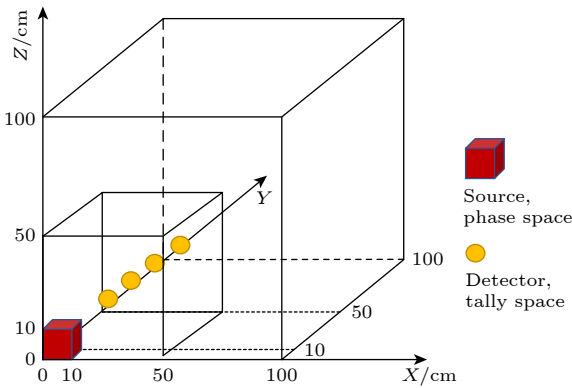


图 5 Kobayashi-1 模型几何示意图  
Fig. 5. Diagram of Kobayashi-1 benchmark.

表 2 Kobayashi 模型相空间划分参数

Table 2. Phase space division parameters of Kobayashi benchmark.

变量	范围	网格尺寸	网格数目
$X/Y/Z$	-10.1—10.1 cm	0.2 cm	100
$E$	—	—	—
$U$	0—1.57 rad	0.79 rad	11
$W$	0—1.57 rad	0.79 rad	11

该 Kobayash 问题一的基准包括两套不同截面设置: 在 Kobayash-1-i 中, 源区和屏蔽区的材料是纯吸收体; 在 Kobayash-1-ii 中, 源区和屏蔽区的材料设定为 50% 的散射截面. 这种材料的总截面为  $0.1 \text{ cm}^{-1}$ , 而空隙区域的总截面为  $10^{-4} \text{ cm}^{-1}$ . 对于每种截面, 分别设置四个探测器位于坐标 (15 cm, 15 cm, 15 cm), (25 cm, 25 cm, 25 cm), (35 cm, 35 cm, 35 cm) 和 (45 cm, 45 cm, 45 cm) 处. MC 模拟使用  $10^7$  个源粒子, 以确保所有探测器结果的最大统计误差保持在 5% 以下.

针对不同的探测器训练了 8 个网络, 每个网络的训练各基于 1000 个 MC 算例. 随机抽取 MC 和重要性数据的 80% 样本用作训练网络的训练集,

10% 样本作为用于优化网络参数的验证集, 10% 的样本作为最终测试训练好的网络预测效果的测试集合, 按照 (9) 式分别计算训练样本、验证样本、测试样本与真值的偏差作为训练偏差、验证偏差、测试偏差如表 3 所列, 测试偏差的分布如表 4 和图 6 所示, 由表 3 和表 4 可知所有情况的测试偏差均值均低于 5%, 小于 5% 的样本占总测试样本量的比例均超过 90%, 本数值实验中的神经网络输入是相空间内给定能量、位置、方向六维坐标的单个离散点, 对应到传统 MC 模拟中的单个样本, 而本方法利用神经网络进行目标物理量的统计为抽样所得源粒子对目标量贡献的均值, 因此预测均值偏差低于 5% 可以说明网络给出的结果精度符合要求. 偏差分布中最大偏差结果出现在源项距离探测器最远的位置, 分析原因是该位置的训练数据本身统计误差较大, 使得训练得到的网络对该位置单个点源给出的偏差较大, 这一特点与传统 MC 单个样本模拟结果的特点相同. 此外, 与原始 MC 模拟相比, 获得这些结果所需的时间几乎可以忽略不计, 用时从大约 5 min 缩短到不到 1 s, 说明神经网络对输运计算效率的提升效果显著. 作为参考, Igor 等<sup>[47]</sup>使用离散纵标法在不同硬件平台下计算完整 Kobayashi 基准题的时间花费在数小时量级.

表 3 Kobayashi 模型不同探测器的网络预测结果

Table 3. Prediction results for different detectors by networks of Kobayashi-1.

基准题	探测器位置/cm	训练偏差	验证偏差	测试偏差
Kobayashi-1-i	(15, 15, 15)	0.0424	0.0458	0.0408
	(25, 25, 25)	0.0426	0.0431	0.0442
	(35, 35, 35)	0.0473	0.0478	0.0458
	(45, 45, 45)	0.0473	0.0483	0.0452
Kobayashi-1-ii	(15, 15, 15)	0.0392	0.0401	0.0387
	(25, 25, 25)	0.0394	0.0404	0.0364
	(35, 35, 35)	0.0402	0.0456	0.0443
	(45, 45, 45)	0.0413	0.0455	0.0425

### 3.2.3 HBR2 反应堆基准题

H.B. Robinson-2 (HBR-2) 核电站是 Westinghouse 公司设计的商用压水式轻水反应堆, 于 1971 年开始运行, 归 Carolina Power and Light Company 所有. H.B. Robinson-2 压力容器测定基准题<sup>[48]</sup>主要测量 HBR2 核电站压力容器内部和外部的中子通量.

表 4 Kobayashi-1 模型不同探测器的网络预测偏差分布  
Table 4. Deviation distributions of predicted-results for different detectors by networks of Kobayashi-1.

基准题	探测器位置/cm	测试偏差			
		均值	最大值	标准差	< 0.05比例*/%
Kobayashi-1-i	(15, 15, 15)	0.0408	0.152	0.0078	95.5
	(25, 25, 25)	0.0442	0.178	0.0116	94.2
	(35, 35, 35)	0.0458	0.214	0.0256	92.6
	(45, 45, 45)	0.0452	0.220	0.0250	92.2
Kobayashi-1-ii	(15, 15, 15)	0.0387	0.126	0.0106	96.2
	(25, 25, 25)	0.0364	0.211	0.0230	95.7
	(35, 35, 35)	0.0443	0.245	0.0288	93.4
	(45, 45, 45)	0.0425	0.253	0.0298	93.3

注: \*预测结果与真值相对偏差小于5%的样本占总测试样本量的比例

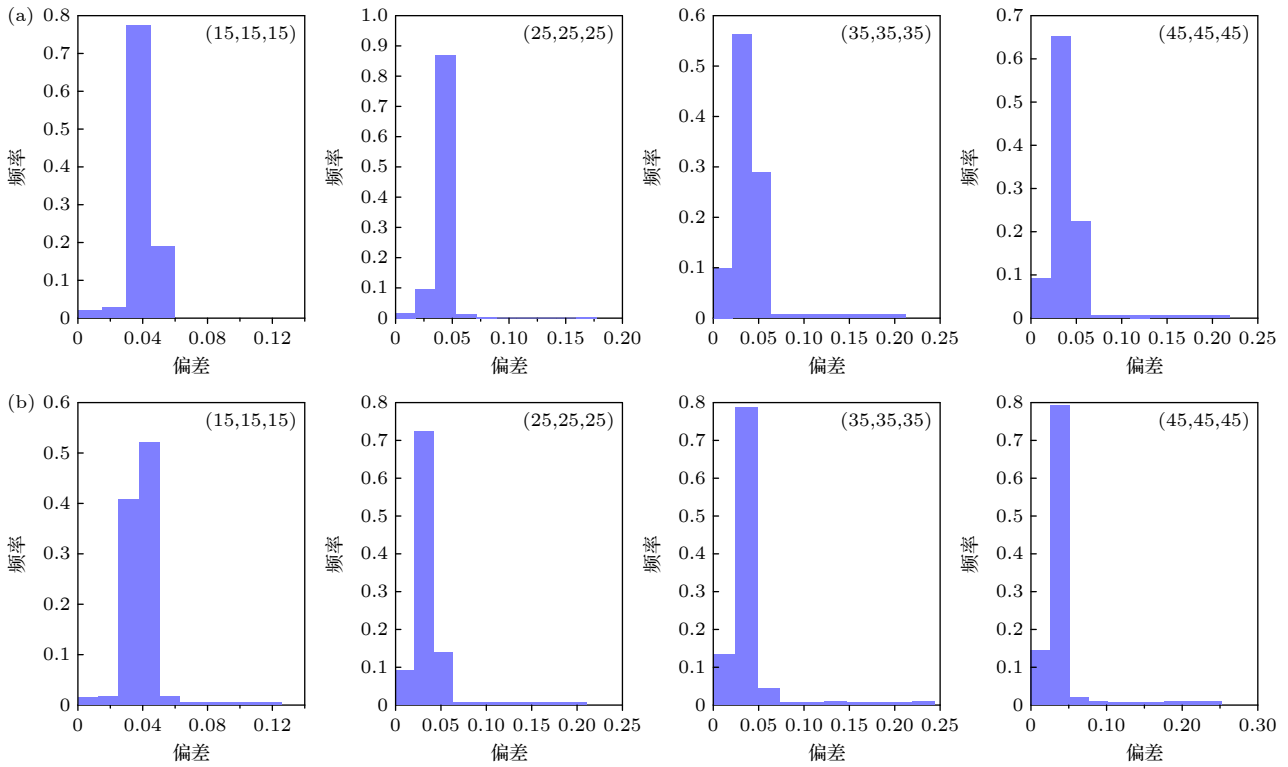


图 6 Kobayashi-1 模型不同探测器的网络预测偏差分布

Fig. 6. Deviation distributions of predicted-results for different detectors by networks of Kobayashi-1.

HBR-2 反应堆堆芯包含 157 个燃料组件、堆芯外包壳、堆芯吊篮、隔热罩、压力容器和生物屏蔽等部件. 在本研究中, 在 OpenMC 中构建了 HBR-2 反应堆结构. 统计以下两个探测器 (①和②) 的通量为目标统计量, 图 7 展示了 OpenMC 中 HBR2 基准题的几何结构及探测器示意图, 其中左半部分为水平面示意图, 右半部分为矢状面示意图. 在 OpenMC 中使用六层立方体来表示源项分布, 每层由 157 个立方体组成.

几何网格尺寸设定为 10.752 mm×10.752 mm×60.960 mm. 源粒子发射方向为各向同性, 能谱为

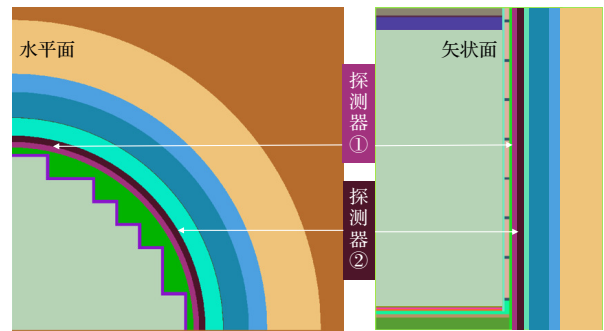


图 7 HBR-2 模型几何结构示意图

Fig. 7. Detector geometry diagram of HBR-2 benchmark.

瓦特谱. MC 模拟中模拟的粒子数量因源与探测器之间的距离而异, 在源与探测器距离较远的算例中, 有必要模拟更多的粒子数量以确保获得的通量统计误差达到要求, 因此每个算例都模拟了  $4.0 \times 10^6 - 6.0 \times 10^8$  个粒子, 以保证最大统计误差保持在 5% 以下. 最终利用 MC 模拟共建立了 942 个算例来生成真值数据集, 其中 754 个算例被用作为训练样本, 92 个算例被用作为验证样本, 92 个算例被用作为测试样本, 按照 (9) 式分别计算测试样本、验证样本、预测样本与真值的偏差作为训练偏差、验证偏差、测试偏差 (表 5).

表 5 HBR-2 模型不同探测器的网络预测结果  
Table 5. Prediction results for different location detector networks of HBR-2 benchmark.

探测器	训练偏差	验证偏差	测试偏差
①	0.0413	0.0402	0.0398
②	0.0425	0.0412	0.0442

表 5 列出了网络对这两个探测器的预测结果, 测试偏差的分布见表 6 和图 8. 网络预测结果的平均相对偏差分别为 3.98% 和 4.42%, 小于 5% 的样本占总测试样本量的比例均超过 90%, 同前算例一样, 最大偏差结果出现在靠近堆芯区域即源项距离探测器最远的位置, 由于该位置的训练数据本身统计误差相对较大而使得单个预测结果与真值偏差超过 5%. 考虑到目标物理量的统计过程为从源分布中多次抽样后输入神经网络并计算对目标量贡献的均值, 因此预测均值偏差低于 5% 表明网络输出结果精度符合要求. 在时间成本方面, 相对于 MC 模拟将计算时间从数小时缩短至约 0.3 s, 实现了显著提速. 作为参考, Roberto<sup>[49]</sup> 在不同硬件平台下使用 TORT 程序基于离散纵标法对 HBR-2 基准题进行计算, 单 CPU 时间花费在数十小时量级. 可见本文提出方法可以给出满足工程精度标准的预测结果, 而计算时间与原始 MC 模拟相比可以忽略不计, 从而获得显著的效率提升.

表 6 HBR-2 模型不同探测器的网络预测偏差分布  
Table 6. Deviation distributions of predicted-results for different detectors by networks of HBR-2.

探测器	测试偏差			
	均值	最大值	标准差	< 0.05 比例*
①	0.0398	0.0826	0.0096	92.3%
②	0.0442	0.0965	0.0132	90.2%

\*预测结果与真值相对偏差小于 5% 的样本占总测试样本量的比例

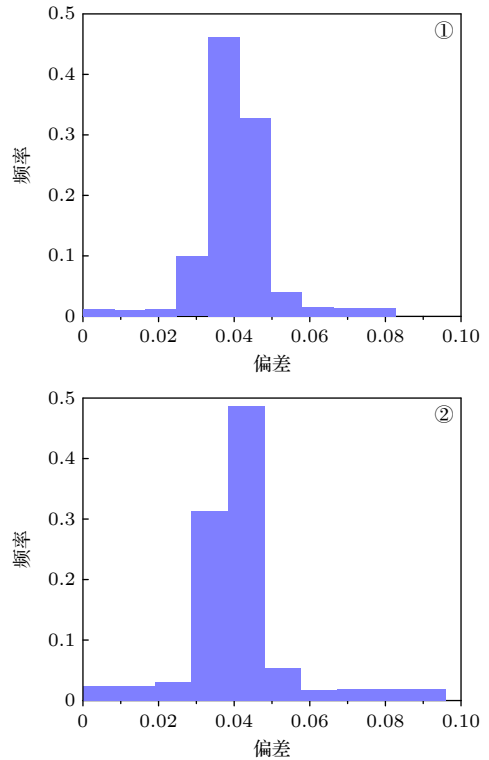


图 8 HBR-2 模型不同探测器的网络预测偏差分布

Fig. 8. Deviation distributions of predicted-results for different detectors by networks of HBR-2.

## 4 结 论

本文设计了用于加速中子输运问题 MC 模拟的神经网络, 只需一次性训练即可针对不同源项反复使用. 在训练该神经网络时, 使用重要性原理来高效获取足够多的高质量训练数据, 从而一定程度上克服了数据生成成本高的难题. 上述方法通过 Kobayashi 和 HBR-2 等基准问题进行了验证. 网络预测结果的相对偏差始终低于 5%. 而且与原始 MC 模拟相比, 获得这些结果所需的计算时间可忽略不计.

本文研究的一个局限性是只适用于源分布有变化, 而几何、材料及目标计数保持不变的情形. 如果几何、材料或目标计数发生变化, 则需要重新训练网络, 后续研究计划将几何、材料以及目标计数的变化也作为输入信息引入网络训练过程, 以增强网络的泛化能力, 并以迁移学习的训练方法提高训练效率.

感谢北京应用物理与计算数学研究所勇珩研究员牵头的“AI++”团队成员的讨论和帮助.

## 参考文献

- [1] Wu Y C 2017 *Fusion Neutronics* (Singapore: Springer Singapore) p21
- [2] Deng L, Li G 2019 *Monte Carlo Simulation Methods and Applications for Particle Transport Problems* (Beijing: Science Press) (in Chinese) [邓力, 李刚 2019 粒子输运问题的蒙特卡罗模拟方法与应用 (北京: 科学出版社)]
- [3] Shangguan D H, Yan W H, Wei J X, Gao Z M, Chen Y B, Ji Z C 2023 *Nucl. Sci. Tech.* **34** 58
- [4] Shangguan D H, Yan W H, Wei J X, Gao Z M, Chen Y B, Ji Z C 2022 *Acta. Phys. Sin.* **71** 090501 (in Chinese) [上官丹骅, 闫威华, 魏军侠, 高志明, 陈艺冰, 姬志成 2022 物理学报 **71** 090501]
- [5] Martin W R 2012 *Nucl. Eng. Technol.* **44** 151
- [6] Noh K, Lee D 2013 *Proceedings of the KNS Fall Meeting, Gyeongju, Korea, 2013, October 23–25*
- [7] Hassan M 2020 *Int. J. Sci. Res.* **9** 913
- [8] Gul A 2016 *Prog. Nucl. Energy* **92** 164
- [9] Huang Z P, Cao L Z, He Q M, Wu H C 2022 *J. Nucl. Sci. Technol.* **59** 1375
- [10] García-Pareja S, Lallena A M, Salvat F 2021 *Front. Phys.* **9** 633
- [11] Pooneh S, Mahdi S, Claudio T (Victor C Ed.) 2013 *Theory and Applications of Monte Carlo Simulations* (Rijeka: IntechOpen) pp152–172
- [12] Dubi A 1985 *Transport Theor. Stat.* **14** 167
- [13] Dubi A 1985 *Transport Theor. Stat.* **14** 195
- [14] Booth T E, Hendricks J S 1984 *Fusion Sci. Technol.* **5** 90
- [15] Davis A, Turner A 2011 *Fusion Eng. Des.* **86** 2698
- [16] Wagner J C, Haghighat A 1998 *Nucl. Sci. Eng.* **128** 186
- [17] Wagner J C, Peplow D E, Mosher S W 2014 *Nucl. Sci. Eng.* **176** 37
- [18] Sun A K, Chen Z P, Li L M, Liu C W, Yu T 2022 *Proceedings of the 23rd Pacific Basin Nuclear Conference Beijing, China, November 1–4, 2022* p821
- [19] Ramón J, Peña J 1995 *Comput. Phys. Commun.* **88** 76
- [20] LeCun Y, Bengio Y, Hinton G 2015 *Nature* **521** 436
- [21] Schmidhuber J 2015 *Neural Networks* **61** 85
- [22] Webb S 2018 *Nature* **554** 555
- [23] Tian S F, Li B 2023 *Acta Phys. Sin.* **72** 100202 (in Chinese) [田十方, 李彪 2023 物理学报 **72** 100202]
- [24] Xiao M J, Yu T C, Zhang Y S, Yong H 2023 *Comput. Fluids* **266** 106025
- [25] Liu L, Liu S P, Xie H, Xiong F S, Yu T C, Xiao M J, Liu L F, Yong H 2024 *J. Sci. Comput.* **98** 22
- [26] Xie H, Zhai C L, Liu L, Yong H 2022 arXiv: 2205.06658v1 [math. NA]
- [27] Hu Z H, Ying Y J, Yong H, Xu R R 2023 *At. Energy Sci. Technol.* **57** 812 (in Chinese) [胡泽华, 应阳君, 永珩, 续瑞瑞 2023 原子能科学技术 **57** 812]
- [28] Lu H Y, Li C H, Chen B B, Li W, Qi Y, Meng Z Y 2022 *Chin. Phys. Lett.* **39** 050701
- [29] Ma Y G, Pang L G, Wang R, Zhou K 2023 *Chin. Phys. Lett.* **40** 122101
- [30] Sun K W, Wang F 2023 *Chin. Phys. B* **32** 070705
- [31] Wu C C, Zhou P J, Wang J J, Li G, Hu S G, Yu Q, Liu Y 2022 *Acta Phys. Sin.* **71** 148401 (in Chinese) [武长春, 周甫钧, 王俊杰, 李国, 胡绍刚, 于奇, 刘洋 2022 物理学报 **71** 148401]
- [32] Yang Y, Cao H X 2023 *Acta Phys. Sin.* **72** 110301 (in Chinese) [杨莹, 曹怀信 2023 物理学报 **72** 110301]
- [33] Berry J J, Gil-Delgado G G, Osborne A G S 2021 *Ann. Nucl. Energy* **160** 108367
- [34] Cao P, Gan Q, Song J, Long P C, Wang F, Hu L Q, Wu Y C 2020 *Ann. Nucl. Energy* **138** 107
- [35] Zhang G M, Song Y M, Zhang Z H, Yuan W W 2022 *Ann. Nucl. Energy* **175** 109248
- [36] Zhang H M, Zhang H C 2022 *Mod. Appl. Phys.* **13** 020209 (in Chinese) [张海明, 张昊春 2022 现代应用物理 **13** 020209]
- [37] Lin H P, Li G D, Chen F G, Han Y, Liang R C 2020 *Radiat. Prot.* **40** 516 (in Chinese) [林海鹏, 李国栋, 陈法国, 韩毅, 梁润成 2020 辐射防护 **40** 516]
- [38] Osborne A, Dorville J, Romano P 2023 *Energy AI* **13** 100247
- [39] Kim S H, Shin S G, Han S, Kim M H, Pyeon C H 2020 *Prog. Nucl. Energy* **119** 103183
- [40] Zhou S K, Greenspan H, Davatzikos C, Duncan J S, Van Ginneken B, Madabhushi A, Prince J L, Rueckert D, Summers R M 2021 *Proc. IEEE* **109** 820
- [41] Peng Z, Shan H M, Liu T Y, Pei X, Wang G, Xu X G 2019 *IEEE Access* **7** 76680
- [42] Ma J H, Piao Z, Huang S, Duan X M, Qin G G, Zhou L H, Xu Y 2021 *Photonics Res.* **9** B45
- [43] Romano P K, Horelik N E, Herman B R, Nelson A G, Forget B, Smith K 2015 *Ann. Nucl. Energy* **82** 90
- [44] Abadi M, Agarwal A, Barham P, et al. 2016 arXiv: 1603.04467 [cs. DC]
- [45] Kingma D P, Ba J L 2014 arXiv: 1412.6980 [cs. LG]
- [46] Kobayashi K, Sugimura N, Nagaya Y 2001 *Prog. Nucl. Energy* **39** 119
- [47] Igor Z, Richard S 2001 *Prog. Nucl. Energy* **39** 207
- [48] Remec I, Kam F B K 1998 *H. B. Robinson-2 Pressure Vessel Benchmark Report* (United States) p2
- [49] Roberto O 2020 *Nucl. Eng. Technol.* **52** 2

# An efficient calculation method for particle transport problems based on neural network\*

Ma Rui-Yao<sup>1)</sup> Wang Xin<sup>1)2)</sup> Li Shu<sup>1)</sup> Yong Heng<sup>1)</sup> Shangguan Dan-Hua<sup>1)†</sup>

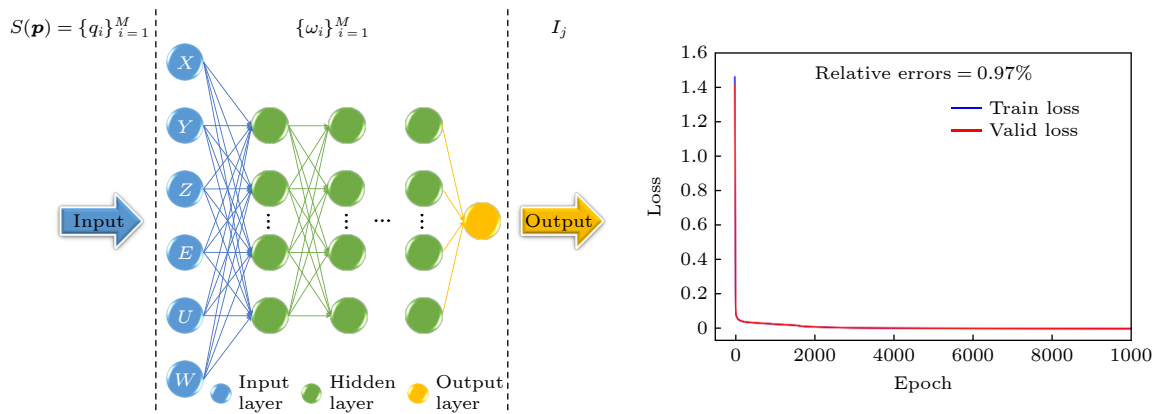
1) (*Institute of Applied Physics and Computational Mathematics, Beijing 100094, China*)

2) (*CAEP Software Center for High Performance Numerical Simulation, Beijing 100088, China*)

( Received 17 October 2023; revised manuscript received 2 February 2024 )

## Abstract

Monte Carlo (MC) method is a powerful tool for solving particle transport problems. However, it is extremely time-consuming to obtain results that meet the specified statistical error requirements, especially for large-scale refined models. This paper focuses on improving the computational efficiency of neutron transport simulations. Specifically, this study presents a novel method of efficiently calculating neutron fixed source problems, which has many applications. This type of particle transport problem aims at obtaining a fixed target tally corresponding to different source distributions for fixed geometry and material. First, an efficient simulation is achieved by treating the source distribution as the input to a neural network, with the estimated target tally as the output. This neural network is trained with data from MC simulations of diverse source distributions, ensuring its reusability. Second, since the data acquisition is time consuming, the importance principle of MC method is utilized to efficiently generate training data. This method has been tested on several benchmark models. The relative errors resulting from neural networks are less than 5% and the times needed to obtain these results are negligible compared with those for original Monte Carlo simulations. In conclusion, in this work we propose a method to train neural networks, with MC simulation results containing importance data and we also use this network to accelerate the computation of neutron fixed source problems.



**Keywords:** Monte Carlo method, neural network, particle transport, importance principle

**PACS:** 28.20.-v, 24.10.Lx, 02.70.Uu

**DOI:** 10.7498/aps.73.20231661

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 12305173, 12375164, 12331010), the Joint Fund of the National Natural Science Foundation of China (Grant No. U2230208), and the Key Laboratory of Nuclear Data Foundation, China (Grant No. JCKY2022201C155).

† Corresponding author. E-mail: [sgdh@iapcm.ac.cn](mailto:sgdh@iapcm.ac.cn)



## 基于神经网络的粒子输运问题高效计算方法

马锐焘 王鑫 李树 勇珩 上官丹骅

### An efficient calculation method for particle transport problems based on neural network

Ma Rui-Yao Wang Xin Li Shu Yong Heng Shangguan Dan-Hua

引用信息 Citation: *Acta Physica Sinica*, 73, 072802 (2024) DOI: 10.7498/aps.73.20231661

在线阅读 View online: <https://doi.org/10.7498/aps.73.20231661>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于物理信息神经网络的光斑质心计算

Calculation of spot centroid based on physical informed neural networks

物理学报. 2022, 71(20): 200601 <https://doi.org/10.7498/aps.71.20220670>

#### 多物理耦合计算中动态输运问题高效蒙特卡罗模拟方法

Efficient Monte Carlo algorithm of time-dependent particle transport problem in multi-physics coupling calculation

物理学报. 2022, 71(9): 090501 <https://doi.org/10.7498/aps.71.20211474>

#### 识别Z玻色子喷注的卷积神经网络方法

Tag Z boson jets via convolutional neural networks

物理学报. 2021, 70(6): 061301 <https://doi.org/10.7498/aps.70.20201557>

#### 热辐射输运问题的高效蒙特卡罗模拟方法

An efficient Monte Carlo simulation method for thermal radiation transport

物理学报. 2020, 69(2): 029501 <https://doi.org/10.7498/aps.69.20191315>

#### 利用神经网络识别高分子链在表面的吸附相变

Recognition of adsorption phase transition of polymer on surface by neural network

物理学报. 2019, 68(20): 200701 <https://doi.org/10.7498/aps.68.20190643>

#### 基于深度残差网络的高精度自然过渡模拟方法

High-order natural transition simulation method based on deep residual network

物理学报. 2020, 69(20): 204701 <https://doi.org/10.7498/aps.69.20200563>