

基于强化学习的多尺度输运问题网格剖分方法研究

钟家浩^{1,2} 唐文林¹ 马瑞笛¹ 彭晓东¹

1(中国科学院国家空间科学中心 北京 100190)

2(中国科学院大学 北京 101408)

摘要 空间环境中的多尺度输运过程通常涉及大尺度背景输运与局部强梯度结构的共存，对数值模拟的空间分辨率和计算效率提出较高要求。传统自适应网格剖分技术虽能提高关键区域分辨率，但高度依赖人工误差指示器和经验阈值，在复杂非线性问题中泛化能力有限。本文提出一种基于多智能体强化学习的网格剖分方法，将网格剖分建模为马尔可夫决策过程，把静态网格优化转化为序列决策任务；同时构建基于图结构的网格特征表示，利用图注意力机制编码邻域物理信息，并在集中训练、分散执行框架下实现多网格单元的协同剖分决策。数值实验表明，本文方法在训练场景和多类泛化测试场景中均取得较优效率指标，相较同一场景下最优基线方法提升约 3%~7%，统计检验和消融实验结果进一步表明，本文方法相较现有方法具有显著优势，且关键模块均对性能提升具有贡献。该方法在降低网格更新频率的情况下仍能保持较好的误差控制能力，展现出良好的数值稳定性与环境适应性。

关键词 多尺度输运，自适应网格剖分，强化学习，图注意力网络，多智能体
中图分类号 (TP399)

Reinforcement Learning-Based Adaptive Mesh Refinement Method for Multiscale Transport Problems

Zhong Jiahao^{1,2} Tang Wenlin¹ Ma Ruidi¹ Peng Xiaodong¹

1 (National Space Science Center, Chinese Academy of Sciences, Beijing 100190)

2 (University of Chinese Academy of Sciences, Beijing 101408)

Abstract Multiscale transport processes in the space environment typically involve the coexistence of large-scale background transport and localized structures with strong gradients, placing high demands on the spatial resolution and computational efficiency of numerical simulations. Although traditional adaptive mesh refinement techniques can improve resolution in key regions, they rely heavily on manually defined error indicators and empirical thresholds, resulting in limited generalization capabilities for complex nonlinear problems. This paper proposes a mesh partitioning method based on multi-agent reinforcement learning, which models mesh partitioning as a Markov decision process and transforms static mesh optimization into a sequential decision-making task. Additionally, it constructs a graph-based representation of mesh features, utilizes a graph attention mechanism to encode neighborhood physical information, and implements collaborative partitioning decisions across multiple mesh cells within a centralized training and distributed execution framework. Numerical experiments show that the proposed method achieves superior efficiency metrics in both training scenarios and multi-class generalization test scenarios, outperforming the best baseline method by approximately 3% to 7% under the same conditions. Statistical tests and ablation experiments further demonstrate that the proposed method has significant advantages over existing methods, and that all key modules contribute to the performance improvement. This method maintains good error control even while reducing the mesh update frequency, demonstrating excellent numerical stability and environmental adaptability.

Keywords Multiscale transport, Adaptive Mesh Refinement, Reinforcement Learning, Graph Attention Network, Multi-agent

* 基金项目：中国科学院院重点部署项目 KGFZD-145-24-33

0 引言

多尺度输运问题广泛存在于空间物理、流体动力学和偏微分方程数值模拟等领域。以空间环境模拟为例, 许多物理过程都包含物质、能量或场量在空间中的迁移与演化, 其数值特征常表现为大尺度背景输运与局部强梯度结构并存^[1]。

在多尺度输运问题的数值模拟中, 数值解的空间分布往往具有明显的非均匀性和动态演化特征。一方面, 局部集中结构、移动前沿、多区域边界、方向性传播以及由耗散或非线性作用引起的形态变化, 通常需要更高的局部分辨率才能被准确捕捉^[2-5]; 另一方面, 大范围背景区域内物理量变化相对平缓, 若采用统一高分辨率网格则会造成大量不必要的计算开销。因此, 如何根据解结构的分布及其演化趋势动态调整网格自由度, 在误差控制与计算效率之间取得平衡, 是多尺度输运问题数值模拟的关键问题^[3,4,6]。尤其当局部结构的位置、形态、传播方向或拓扑关系随时间发生变化时, 网格剖分策略不仅需要识别当前的高误差区域, 还应具备一定的前瞻性和跨区域协调能力, 以避免剖分滞后、局部过度剖分或关键结构分辨率不足等问题。

自适应网格剖分 (Adaptive Mesh Refinement, AMR) 技术的出现为解决多尺度输运问题模拟的计算精度与计算效率的平衡提供了重要途径^[3,4]。通过在物理量剧变区域局部剖分、在平稳区域降低分辨率, AMR 实现了计算资源的按需动态分配。AMR 技术有多种组织与实现方式, 主要分为两类: 一类是基于补丁的 AMR, 如 AMReX^[7]、Chombo^[8]、SAMRAI^[9]和 Uintah^[10]; 另一类是基于树结构的 AMR, 如 PARAMESH^[11]和 p4est^[12]。尽管在模拟中取得一定成效, 但传统 AMR 仍高度依赖人工构造的误差指示器及预设阈值。这类方法计算性能对参数选择很敏感, 缺乏统一的构建原则, 且大多基于局部指标进行决策, 难以捕获全局输运特征^[13]。在处理多尺度的问题时, 传统 AMR 表现出明显的泛化能力不足, 限制了其模拟精度与可靠性。

近年来, 随着深度学习与强化学习方法的发展, 将 AMR 问题从规则驱动转化为数据驱动的决策问题成为一种新的研究方向。2021 年, Bohn 等人首次将深度学习方法引入 AMR, 他们使用

循环神经网络构建 AMR, 不依赖具体问题的误差估计实现对一大类偏微分方程 (Partial Differential Equation, PDE) 的最优网格剖分, 为 PDE 求解提供了通用的黑箱网格剖分工具。更重要的是, 该研究给出了最优网格剖分的定义, 并证明了深度学习方法能够达到最优的网格剖分^[14]。2022 年, Pan 等人率先将四边形网格生成问题作为马尔可夫决策过程 (Markov Decision Process, MDP) 问题, 引入强化学习 (Reinforcement Learning, RL) 对该问题的求解进行了研究。该研究面向的是几何区域的初始网格生成问题, 其核心思想是在给定边界条件下, 通过智能体逐步选择网格生成动作, 从而自动构造满足质量要求的四边形网格, 重点在于提高网格生成过程的自动化程度和复杂几何适应能力^[15]。同年, Foucart 等人在 Pan 等的研究基础上, 进一步将强化学习引入 AMR 问题, 将已有数值计算网格上的单元细化、粗化操作建模为局部序贯决策过程, 并采用深度强化学习学习网格调整策略。该方法的研究对象并非初始网格生成, 而是在偏微分方程数值求解过程中, 根据局部解特征和计算资源约束动态调整网格分辨率。此外, Foucart 等人的研究一大创新是将计算资源的使用情况加入损失函数, 使得决策更加动态、智能^[16]。Yang 等人在 2023 年提出与 Foucart 等人研究相类似的方法, 其利用 RL 实现面向长期误差优化的网格剖分策略, 突破了传统基于瞬时误差的启发式方法在全局最优性与泛化能力方面的局限^[17]。

尽管近年来 RL 方法在 AMR 中的应用取得了初步进展, 为突破传统基于局部误差指示器的启发式方法提供了新的思路, 但仍面临若干关键挑战。首先, AMR 过程中网格拓扑会随细化与合并动态变化, 网格单元之间存在局部邻接、尺度差异和跨区域耦合关系, 如何构建能够适应不同网格拓扑并有效表征邻域依赖的状态表示, 是学习型剖分策略需要解决的基础问题^[18]。其次, 网格剖分并非单个局部单元的独立决策, 而是大量网格单元共同参与的联合优化过程。每个单元的细化、合并或保持操作都会影响整体误差分布和计算资源占用, 因此如何在保持局部决策效率的同时, 使多单元动作在全局误差一成本目标下协调一致, 是 AMR 强化学习建模中的关键问题

[17]。已有方法虽然可以通过图网络对网格单元进行表征，或采用参数共享策略逐单元输出动作，但局部动作与全局优化目标之间仍缺少显式的协同机制。

针对上述问题，本文从多尺度输运问题中网格剖分的多单元协同决策特征出发，提出一种基于多智能体强化学习（Multi-agent Reinforcement Learning, MARL）的 AMR 方法，从 AMR 的多单元决策特征出发，将每个网格单元建模为一个智能体，并在集中训练、分散执行（Centralized Training with Decentralized Execution, CTDE）框架下结合图注意力状态编码与价值分解机制，使局部单元动作不仅依赖自身及邻域特征，而且受到全局误差控制和计算成本目标的约束。与现有 RL-AMR 方法相比，本文工作的区别主要体现在两点：第一，在建模方式上，已有方法通常将网格剖分视为单一策略下的逐单元动作预测问题，即由同一网络根据局部状态或图编码特征分别输出各网格单元的细化、粗化或保持动作；本文则将每个网格单元显式建模为一个智能体，将 AMR 过程表述为多智能体协同决策问题，更强调不同网格单元动作之间的相互影响与协同关系。第二，在局部动作与全局目标的关联方式上，已有方法多侧重依据局部特征直接输出剖分动作，而本文在 CTDE 框架下结合价值分解机制，使训练阶段的局部动作能够受到全局误差控制和计算成本目标的共同约束，执行阶段则保持各网格单元依据自身及邻域观测进行分散决策。

为验证所提方法的作用机制和适用范围，本文设置了消融实验和多类泛化测试场景，其中消融实验考察关键模块对性能的影响，泛化测试场景包括多区域结构、不同网格拓扑、非线性演化、耗散型演化和传播型演化等典型数值现象，验证模型的适应能力。通过上述实验，本文从误差控制、计算效率和跨场景适应性等方面评估所提方法的有效性与局限性，为多尺度输运类问题中的学习型 AMR 策略设计提供参考。

本文组织结构如下：第 0 节为引言，介绍多尺度输运问题中自适应网格剖分的研究背景、传统方法局限以及强化学习方法的发展现状，并阐明本文研究思路与主要贡献。第 1 节将 AMR 问题建模为马尔可夫决策过程，定义状态、动作、状态转移和奖励函数。第 2 节介绍基于强化学习的网格剖分方法，包括图结构网格表示、图注意力邻域信息编码以及基于 CTDE 和价值分解的多智能体协同决策机制。第 3 节通过基础训练场

景验证、关键模块消融、跨场景泛化实验对所提方法进行验证，并与传统 AMR 方法及强化学习基线方法进行对比分析，评估本文方法的特点和优势。第 4 节对全文工作进行总结。

1 剖分问题的马尔可夫决策建模

传统 AMR 方法通常依赖于局部误差估计与预设阈值进行决策，例如基于误差指标的标记策略，其本质是一种局部贪心策略。这类方法仅根据当前时刻的误差信息进行网格调整，缺乏对求解演化过程的全局理解，难以实现对未来误差变化的前瞻性预测。特别对于时变问题，这一局限尤为突出。由于物理信息（如波、特征结构）在空间中的传播具有方向性和时间延迟，仅依赖当前局部误差往往无法提前识别即将出现的高误差区域，从而导致网格调整滞后，影响整体计算精度。因此，理想的网格剖分策略应能够综合考虑当前状态与未来演化趋势，在时间维度上实现长期优化。

本文将 AMR 建模为 MDP，在该框架下，AMR 系统可用一个四元组表示：

$$M = \langle S, A, P, R \rangle \quad (1)$$

其中状态空间 $S = \{s_1, s_2, \dots, s_n\}$ 表示当前网格及其物理场状态，包括局部误差、梯度强度、层级深度、邻域分布、密度和速度场等；动作空间 $A = \{refine, coarsen, hold\}$ 定义每个网格单元可执行的操作，包括“细化”、“粗化”或“保持不变”；状态转移函数 $P(s'|s, a)$ 描述在执行剖分动作 a 后网格及物理场的演化规律，由数值求解器和物理过程共同决定，状态转移满足概率的基本性质：对所有可能的 s' ，有 $\sum_{s'} P(s'|s, a) = 1$ ，且 $0 \leq P(s'|s, a) \leq 1$ 。奖励函数 $R(s, a, s')$ 量化了在状态转换过程中获得的即时反馈，其设计直接决定了策略优化的方向与最终收敛结果。

本文在奖励函数中同时引入误差项与计算成本项，以引导智能体学习合理的网格分配策略。为更准确地评估单步决策对整体目标的贡献，采用基于增量的奖励形式，综合考虑误差降低与计算成本变化，用以引导策略在精度与效率之间取得平衡^[16]：

$$R = [accuracy] - \gamma_c [cost] \\ = \begin{cases} +\log \Delta u_h - \gamma_c R_{AC}, & \text{if } refine \\ -\log \Delta u_h - \gamma_c R_{AC}, & \text{if } coarsen \\ 0, & \text{if } hold \end{cases} \quad (2)$$

其中, Δu_h 为剖分前后的数值差, γ_c 为比例 因

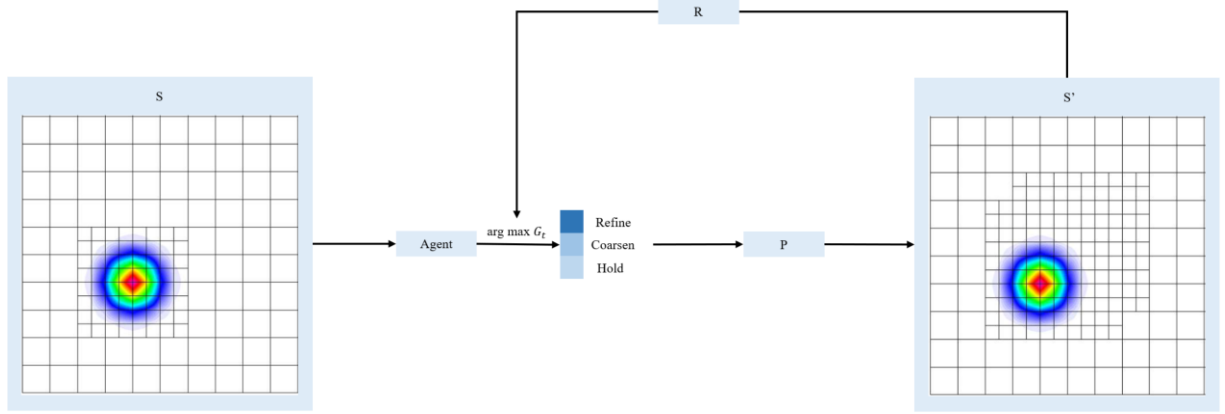


图 1 MDP 表述下的 AMR 问题。

Fig.1 AMR under an MDP Formulation.

子, 调和准确性与成本, 可作为超参数, $R_{\Delta C} = B(p') - B(p)$, 为决策前后的资源差, $p \in [0,1]$, 为资源占用率, $B: [0,1] \rightarrow [0,\infty)$, 为资源占用的映射函数。

智能体的目标是最大化累积奖励 G_t , 数学上, 累积奖励被定义为:

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots \\ &= \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \end{aligned} \quad (3)$$

其中, 折扣因子 γ 是一个介于 0 和 1 之间的参数, 控制对未来收益的关注程度, 以兼顾局部优化与长期目标。MDP 表述下的 AMR 问题决策、状态转移、奖励累计过程如图 1 所示。

2 基于强化学习的网格剖分方法设计

本章围绕 AMR 问题的强化学习建模与求解方法展开, 旨在构建一个兼具全局感知能力与局部决策效率的网格剖分框架。在前文将网格剖分问题形式化为马尔可夫决策过程的基础上, 本章进一步从状态表达、信息建模与决策机制三个层面展开方法设计。首先, 针对自适应网格在空间上具有动态拓扑结构的特点, 引入图结构对网格单元及其邻接关系进行统一建模, 并通过节点与边特征刻画局部物理状态及单元间耦合关系; 进一步地利用图注意力机制对邻域信息进行自适应加权聚合, 从而增强模型对跨尺度空间依赖的表达能力。随后, 面向多网格单元协同决策问题, 引入 MARL 框架, 在集中训练、分散执行机制下, 通过价值分解实现局部决策与全局优化目标之间的协调统一。

2.1 基于图的网格拓扑表示与邻域信息编码

多尺度输运过程具有显著的非局域性和空间各向异性。若仅依赖单一网格内的局部误差指示器进行剖分决策, 极易割裂跨尺度的物理联系, 导致网格分辨率错配与关键演化特征丢失。因此, 本文的 AMR 策略引入了整体拓扑结构与邻域信息交互机制, 将局部状态与跨网格的物理耦合效应隐式编码于图结构之中:

$$G = \{V, E\} \quad (4)$$

其中 $V = \{v_i\}_{i=1}^{n_t}$ 表示时间步 t 时刻的节点集合, 每一个节点 v_i 对应于一个网格, n_t 为当前网格数量, $E = \{e_{ij} = (v_i, v_j) | \Omega_i \cap \Omega_j \neq \emptyset\}$, Ω_i 为网格 i 对应的空间区域。

该图以网格为节点 V , 以网格间的空间邻接或物理耦合关系为边 E , 从而在结构层面建立起网格内部的信息通路。每个节点拥有内部信息 o_i , 用以描述其局部物理状态^[19]:

$$o_i = \left[\log(c_i^t), \text{one-hot}(\text{depth}(i)) \right] \quad (5)$$

其中 $c_i^t = \|u_i^t - \hat{u}_i^t\|_2$, u_i^t 为网格 i 在时间 t 的真实解, $\hat{u}_i^t$ 为网格 i 在时间 t 的数值解, 需要说明的是, 在本文基准实验中, 二维线性对流方程具有可构造参考解的特点, 因此局部误差项可通过解析解或高分辨率参考解计算, 并用于训练阶段的状态构造和性能评价。对于真实空间环境数值模拟等难以获得解析解的问题, 该误差项可由代理误差特征替代。one-hot(depth(i)) 为网格 i 细化层级的独热编码, 将每一层深度映射为一个仅对应位置为 1、其余为 0 的向量, 从而以离散且无序偏置的方式表达层级信息。

对于任意一对相邻网格 i, j , 构成一条有向边

e_{ij} , 其特征定义为^[17]:

$$e_{ij} = [\delta d_{ij}, \phi_{ij}] \quad (6)$$

$$\delta d_{ij} = \text{one-hot}(\text{depth}(i) - \text{depth}(j)) \quad (7)$$

$$\phi_{ij} = \left\langle \theta_j, \frac{x_i - x_j}{\|x_i - x_j\|_2} \right\rangle \quad (8)$$

其中, δd_{ij} 为相邻两个网格的层级差异的独热编码, θ_j 为单元 j 处的局部输运或特征传播方向向量, x_i, x_j 为两个单元的坐标, ϕ_{ij} 为相邻单元空间方向与局部输运方向之间的相对关系。需要说明的是, 式(8)中的坐标差归一化主要用于提取相邻单元之间的相对方向信息, 使该方向特征不受单元中心距离绝对大小的影响, 并不直接用于表示网格尺度。对于自适应剖分后可能出现的不同大小相邻单元, 本文通过 δd_{ij} 显式编码相邻单元的剖分层级差异, 从而反映局部网格尺度变化。同时, 图结构中的邻接关系根据当前实际网格拓扑构建, 单元中心坐标也由当前剖分后的真实几何位置计算得到。因此, 边特征同时包含相邻方向、局部传播方向以及剖分层级差异信息, 可在不引入额外尺度项的情况下描述不同大小网格单元之间的相对空间关系。

为有效建模网格单元之间的局部耦合关系, 本文进一步引入图注意力机制 (Graph Attention Network, GAT), 通过自适应加权的方式对邻域信息进行聚合, 其过程如图2所示。

在每一层图神经网络中, 节点特征通过与其邻居节点进行信息交互得到更新。具体而言, 对于节点 i , 首先通过注意力机制构造三组可学习参数矩阵:

$$W^q, W^k, W^v \quad (9)$$

分别对节点特征进行投影:

$$q_i = W^q v_i, k_j = W^k v_j, h_j = W^v v_j \quad (10)$$

其中, W^q, W^k, W^v 本质上是对原始节点特征进行线性嵌入, 将其映射到统一的特征空间, 以便后续进行相似性计算与信息融合。

随后, 通过点积计算节点 i 与邻居节点 $j \in \mathcal{N}_i$ 的相关性:

$$q_i = W^q v_i, k_j = W^k v_j, h_j = W^v v_j \quad (11)$$

并对所有邻居的相关性进行归一化, 得到注

意力权重:

$$\hat{a}_{ij} = \text{softmax}_j(\{a_{ik}\}_k) = \frac{\exp(a_{ij})}{\sum_{k \in \mathcal{N}_i} \exp(a_{ik})} \quad (12)$$

该权重刻画了邻居节点 j 对节点 i 的相对重要性。在此基础上, 结合节点特征与边特征构造消息项。与节点特征类似, 边特征通过线性映射矩阵 W^e 投影到同一特征空间:

$$b_{ij} = h_j + W^e e_{ij} \quad (13)$$

最终, 节点特征通过加权聚合邻域信息完成更新:

$$v_i' = \sum_{j \in \mathcal{N}_i} a_{ij} b_{ij} \quad (14)$$

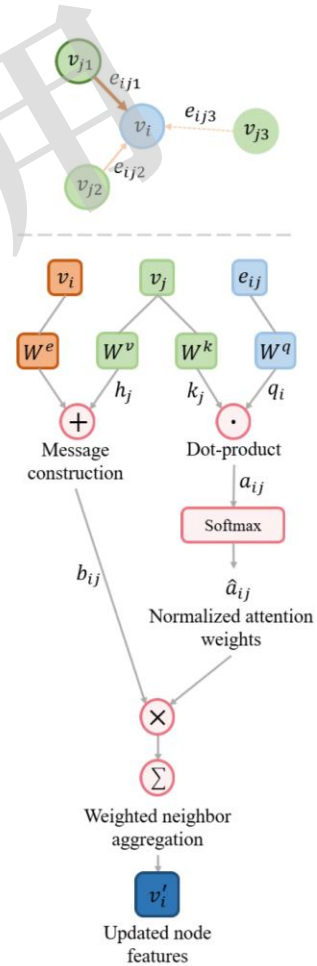


图2 基于图注意力机制的网格邻域信息聚合过程。

Fig.2 GAT-based aggregation of mesh neighborhood information.

2.2 基于CTDE的强化学习决策算法

在强化学习框架下, 本文采用时序差分方法

(Q-learning) 对网络参数进行优化。具体而言, 通过最小化当前估计值与目标值之间的均方误差, 不断逼近最优状态-动作价值函数。其更新目标定义为:

$$y_t = r_t + \gamma Q(s', a) \quad (15)$$

其中 r_t 为当前状态执行动作后的即时奖励, γ 为折扣因子, s' 为下一状态。通过这一迭代更新, 网络能够逐步学习到在保证局部物理精度的同时优化整体计算性能的最优动作策略。

多尺度输运问题中的局部结构通常具有明显的空间非均匀性和动态演化特征, 其形成与迁移不仅取决于当前单元内的局部状态, 还会受到邻域分布、传播方向、结构形态变化以及不同区域之间相互影响的共同作用。因此, 单纯依赖局部误差或局部梯度信息, 往往只能反映当前时刻、当前空间位置附近的变化情况, 难以充分描述结构迁移、边界演化或多区域协同变化对网格分辨率提出的要求。引入邻域信息和全局优化目标, 有助于模型从整体上刻画数值解的空间分布与演化特征, 使自适应网格剖分在局部指标响应之外进一步考虑误差控制与计算成本之间的整体平衡, 从而提高关键区域的分辨能力和数值模拟的稳定性。

为兼顾全局协同优化能力与大规模系统的可扩展性, 本文采用多智能体 CTDE 框架对 AMR 过程进行建模, 将每个网格单元视为一个智能体。各智能体基于自身局部状态及邻域信息输出剖分动作, 从而将全局网格剖分问题分解为多个局部单元的协同决策问题。在训练阶段, 所有智能体共享全局奖励信号 $Q(s, a)$, 并可利用全局状态信息对价值函数进行联合优化。通过集中式的价值函数估计, 可以将各个智能体的局部决策与全局性能指标 (如整体误差与计算成本) 建立联系, 从而实现对全局目标的直接优化。具体而言, 将全局价值函数 $Q(s, a)$ 表示为各个智能体局部价值函数的加性组合:

$$Q(s, a) = \sum_{i=1}^n Q_i(o_i, a_i) \quad (16)$$

其中, $Q_i(o_i, a_i)$ 表示第 i 个智能体在其局部观测 o_i 下采取动作 a_i 的个体价值函数。该分解方式将原本高维、耦合的联合优化问题转化为多个低维的局部优化问题, 从而显著降低了计算复杂度, 并提升了模型在大规模系统中的可扩展性。

在执行阶段, 每个智能体仅依赖自身的局部

观测 o_i 和对应的价值函数 $Q_i(o_i, a_i)$ 独立选择动作, 无需访问其他智能体的信息或全局状态。这种分布式决策机制使得模型能够适应 AMR 过程中智能体数量动态变化的特点, 同时显著降低了在线计算开销。

为保证局部最优决策能够一致地导出全局最优策略, 需要满足个体-全局最大一致性 (Individual-Global-Max, IGM) 性质^[19], 即如果全局价值函数可以表示为各个智能体局部价值函数的加性组合, 则其最优联合动作可由各智能体局部最优动作独立组合得到:

$$a^* = (a_1^*, \dots, a_n^*) \quad (17)$$

$$a_i^* = \arg \max_{a_i} Q_i(s_i, a_i) \quad (18)$$

其中, a^* 为全局最优联合动作, a_i^* 为第 i 个智能体的局部最优动作。

在满足 IGM 条件时, 全局最优策略的求解无需在指数规模的联合动作空间中进行搜索, 而可以通过各智能体独立求解局部最优动作实现, 从而显著降低决策复杂度。在采用加性价值分解的情况下, 若全局价值函数能够表示为各智能体局部价值函数之和, 则局部贪心动作组合可与全局贪心动作保持一致, 从而满足 IGM 条件^[19]。

为提升模型的表达能力与稳定性, 本文在图注意力机制中引入多头注意力结构。通过设置多个独立的注意力头, 对节点特征进行不同子空间的映射与信息聚合, 从而使模型能够从多个角度刻画节点间的相互作用关系。各注意力头的输出在特征维度上进行拼接, 并通过线性变换得到最终的节点表示, 有助于增强模型对复杂邻域结构的建模能力。

在此基础上, 采用多层图神经网络结构对特征进行逐层传播与更新。每一层均执行一次基于注意力的邻域聚合操作, 使节点能够逐步获取更远距离的邻域信息, 从而扩大感受野。同时, 引入残差连接与归一化操作以缓解深层网络中的梯度消失问题, 并提高训练稳定性。通过多头与多层结构的结合, 模型能够在非结构网格上实现多尺度信息建模。算法决策过程与结构如图 3 所示。

3 数值实验与结果分析

3.1 实验环境与设置

为验证所提出方法在输运类问题中的 AMR 效果, 本文基于二维对流方程构建数值实验环境。

需要说明的是，本文采用二维线性对流方程作为

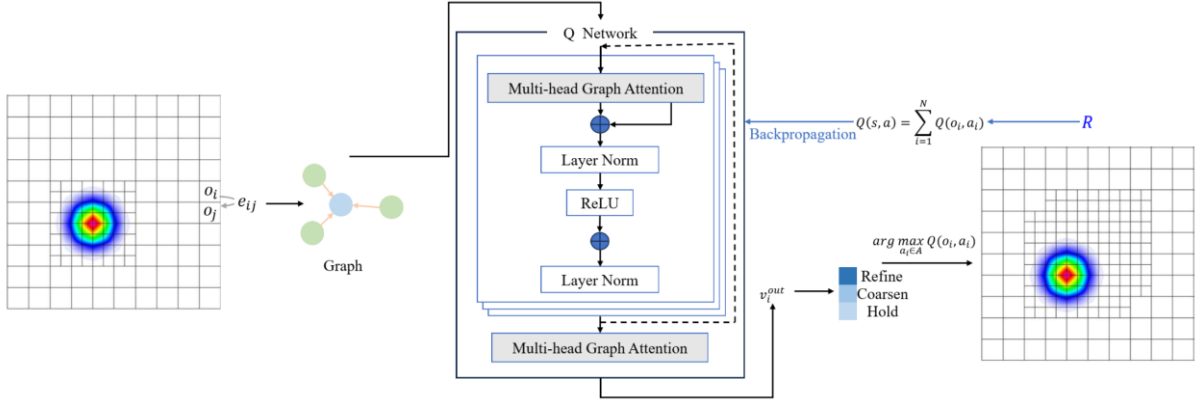


图3 强化学习框架下的动作决策与策略更新过程。

Fig.3 Action selection and policy learning process in the reinforcement learning framework.

验证对象，主要用于在简化模型下考察所提出强化学习网格剖分方法对细化区域选择、局部结构跟踪和误差一成本权衡的调控能力。该方程能够保留输运问题中的方向性传播和局部结构迁移特征，但尚未包含实际输运问题中的复杂物理过程。因此，本文结果应理解为方法层面的初步验证，而非完整过程的直接模拟。后续工作将进一步面向更复杂的多物理场耦合模型开展验证，以评估该方法在实际空间环境数值模拟中的推广能力。所有实验均在统一配置下进行训练与测试，以保证结果的可比性与可重复性。

训练阶段采用的二维线性对流方程^[20]：

$$u_t + \mathbf{v} \cdot \nabla u = 0 \quad (19)$$

其他训练阶段数值求解器参数见表 1，神经网络结构参数见表 2。

表 1 训练阶段数值求解器参数设置。

Tab.1 Solver parameters used in the training phase.

参数	记号	值
PDE 类型	–	Advection
时间步长	dt	0.002
网格剖分步长	t_{step}	125
终止时间	–	0.75
网格类型	–	quadrilateral
边界条件	–	periodic
最大剖分层级	–	1
是否允许粗化	–	是

强化学习算法的核心参数设置见表 3。其中，采用 ϵ -greedy 策略进行探索，并通过逐步衰减探索率实现从探索到利用的平滑过渡。折扣因子设为 0.99 以强调长期回报，从而引导模型学习前瞻性网格剖分策略。经验回放机制用于提高样本利

用效率，而优先经验回放进一步强化了对关键样本的学习能力。此外，引入 Double DQN 与 Dueling DQN 结构以缓解 Q 值过估计问题并提升训练稳定性。

表 2 神经网络结构参数设置。

Tab.2 Neural network architecture parameters.

参数	值
特征维度	64
注意力头数	2
注意力层数	2
递归传播次数	3
是否使用注意力	True
是否使用残差连接	True
是否使用 LayerNorm	True

表 3 强化学习算法关键参数设置。

Tab.3 Key hyperparameters of the reinforcement learning algorithm.

参数	值
Batch size	16
学习率	0.005
折扣因子	0.99
Buffer size	10000
ϵ -greedy (start/end)	0.5 / 0.01
ϵ 衰减步数	150000
训练轮数	400000
优先经验回放	True

为进一步综合评估误差与计算成本之间的权衡关系，本文引入效率指标 η ^[17]，用于刻画不同方法在精度与资源消耗之间的整体表现。该指标基于归一化误差与归一化计算成本定义为：

$$\eta = 1 - \sqrt{c^2 + d^2} \quad (20)$$

其中, $\tilde{c}$ 表示归一化全局误差, $\tilde{d}$ 表示归一化计算资源占用。该指标的取值范围为[0,1], 数值越大表示在相同计算成本下获得更低误差, 或在相同误差水平下消耗更少资源, 从而具有更高的整体效率。通过该指标, 可以将误差与计算成本统一到同一评价框架中, 为不同方法之间的性能比较提供直观依据。

本文还引入去中心化执行一致性指标 (Decentralized Execution Consistency, DECG), 以衡量中心化训练价值函数与去中心化逐单元执行策略之间的一致性。对于给定状态 s , 设中心化联合动作价值函数为 $Q_{tot}(s, \mathbf{a})$, 去中心化执行动作记为 $\mathbf{a}^{dec}$, 候选联合动作集合中的最大和最小中心化价值分别为 Q_{tot}^{max} 和 Q_{tot}^{min} , 则 DECG 定义为:

$$DECG = 1 - \frac{Q_{tot}(s, \mathbf{a}^{dec}) - Q_{tot}^{min}}{Q_{tot}^{max} - Q_{tot}^{min}} \quad (21)$$

DECG 越小, 说明去中心化执行动作越接近中心化训练目标下的最优动作; 当 DECG 接近 0 时, 表示中心化训练目标能够被局部执行策略较好地继承。由于联合动作空间规模会随网格单元数量呈指数增长, 直接枚举全部联合动作在实际 AMR 场景中不可行。因此, 本文在每个状态 s 下构造有限候选联合动作集合 $\mathcal{A}_{cand}(s)$, 用于近似评估中心化目标与去中心化执行动作之间的一致性。

在实验设计方面, 本文围绕基础训练场景验证、关键模块消融、跨场景泛化能力评估和综合效率评价等层面开展数值实验。首先, 基于训练

场景考察所提出方法的基本网格剖分能力, 重点分析模型能否根据数值解的空间分布特征及其演化趋势合理选择剖分区域, 并在有限计算资源条件下有效抑制全局误差增长。在此基础上, 为进一步明确各组成模块对整体性能的贡献, 本文设置消融实验, 分别考察图注意力机制、价值分解机制和多智能体集中训练框架的作用, 从而验证所提出方法相较于简化模型及参数共享单智能体策略的差异。随后, 为评估模型在未参与训练场景下的适应能力, 本文构造多类泛化测试场景, 覆盖多区域边界结构、复杂局部形态、非结构网格拓扑、非线性演化、耗散型演化和传播型演化等典型数值现象, 用以考察模型在不同解结构、网格连接关系和物理演化机制下的剖分稳定性。最后, 本文采用效率指标对不同方法在误差控制与计算资源消耗之间的综合表现进行统一评价, 并结合多次独立测试结果, 比较所提出方法与传统 AMR 方法及单智能体强化学习方法在不同实验场景下的性能差异。

本文的训练与实验环境为 Ubuntu 23.10 操作系统, 配备双路 Intel Xeon Platinum 8476C 处理器、256 GB 系统内存; 运行环境包括 Python 3.12 与 TensorFlow 2.16.1。

3.2 训练场景实验结果与分析

在初始网格为 16×16 、最大剖分深度为 1 的条件下, 实验结果如图 4 所示, 其中上排为强化学习方法得到的网格剖分结果, 下排为传统阈值方法在相同时间步下的网格剖分结果。

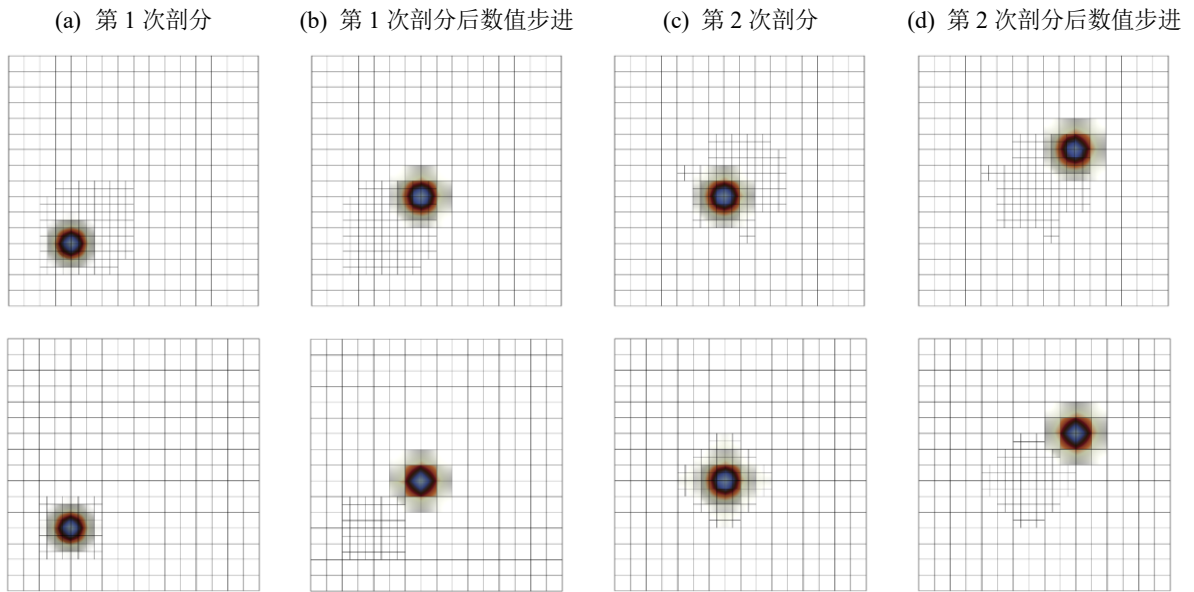


图4 强化学习方法与传统方法的网格剖分结果对比。

Fig.4 Comparison of mesh refinement results between the reinforcement learning method and the traditional method.

实验结果表明，强化学习方法能够根据数值解的空间分布动态调整网格分辨率，使剖分区域较好地跟随物理量高梯度区域的移动，并在解变化较为平缓的区域保持

较低分辨率，从而实现对计算资源的合理分配。

相比之下，传统阈值方法主要依据当前时刻的局部误差或阈值判断进行网格剖分，其剖分区域虽然能够在一定程度上覆盖主要分布区域，但在连续时间演化过程中容易出现剖分区域滞后、分布不够集中或局部过度剖分等现象。因此所提出的强化学习方法在网格分辨率动态剖分和剖分区域空间协调方面具有一定优势。

为定量分析不同网格剖分策略对数值误差演化的影响，本文统计了训练环境下不同方法的全局误差增长情况，结果如图 5 所示。

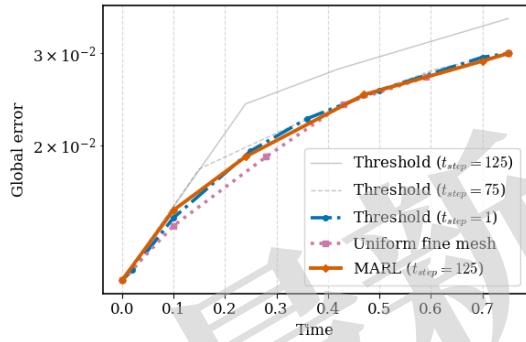


图 5 训练环境下强化学习方法与不同剖分步长阈值方法的全局误差增长曲线对比。

Fig.5 Comparison of global error growth curves between the reinforcement learning method and threshold-based methods with different refinement step sizes in the training environment.

图 5 中对比了不同剖分步长下的传统阈值方法、均匀细网格方法以及强化学习方法的全局误差变化曲线。可以看出，随着时间推进，各方法的全局误差均呈现逐渐增长趋势，但不同方法之间的误差增长幅度存在明显差异。当传统阈值方法采用较大的剖分步长时，其网格更新频率较低，难以及时跟随数值解中高梯度区域的运动变化，因此全局误差增长较快。随着剖分步长减小，阈值方法能够更频繁地调整网格结构，误差控制效果有所改善，但同时也会带来更高的网格更新开销和计算负担。相比之下，强化学习方法在相同剖分步长条件下能够获得更低的全局误差，其误差增长曲线与小剖分步长阈值方法以及均匀细网

格方法较为接近，说明该方法能够在较低网格调整频率下实现较好的误差控制效果。

这一结果表明，强化学习方法并非简单依赖频繁的网格更新来降低误差，而是通过学习数值解演化特征与网格剖分行为之间的关系，在有限的网格调整次数下更加合理地选择剖分区域。与传统阈值方法相比，所提出方法能够在保持较低计算开销的同时有效抑制全局误差增长，从而在计算精度与计算效率之间取得更好的平衡。

为进一步分析本文方法中关键模块的影响，本文在训练场景下设计了消融实验。消融实验采用与前文相同的二维线性对流方程、初始网格、最大剖分深度和训练参数设置，仅对模型结构中的关键模块进行替换或移除，以保证不同方法之间具有可比性。为降低随机初始条件对实验结论的影响，每种方法均在 30 组随机初始条件下重复实验，并统计效率指标与 DECG 指标的均值和标准误。表中括号内数值表示标准误，用于反映重复实验下均值估计的稳定性。

本文设计了四组对比模型。其中，Yang-RL 采用非严格价值分解的混合决策结构作为对照；MARL w/o GAT 去除了图注意力模块，用于考察邻域信息聚合能力对网格局部决策的影响；MARL w/o VD 去除了价值分解结构，用于分析价值分解机制对中心化训练目标与去中心化执行一致性的作用；MARL 表示本文提出的完整模型，包含图注意力网络、价值分解结构以及 CTDE 训练框架。实验结果如表 4 所示。

表 4 不同强化学习模型在 30 次随机初始条件下的消融实验结果。

Tab.4 Ablation results of different reinforcement learning models under 30 random initial conditions.

模型	效率指标	DECG
Yang-RL	0.83(0.02)	0.313(0.005)
MARL w/o GAT	0.85(0.03)	0.110(0.003)
MARL w/o VD	0.82(0.02)	N/A
MARL	0.89(0.02)	0.093(0.002)

由表 4 可以看出，完整 MARL 模型在效率指标上取得最优结果，其平均效率指标为 0.89，高于 Yang-RL 的 0.83、MARL w/o GAT 的 0.85 以及 MARL w/o VD 的 0.82。相较于 Yang-RL，完整 MARL 的效率指标提升约 7.2%；相较于去除图注意力模块的 MARL w/o GAT，提升约 4.7%；相较于去除价值分解结构的 MARL w/o

VD, 提升约 8.5%。这表明, 本文方法中图注意力网络、价值分解结构与 CTDE 训练框架的联合使用, 能够有效提升模型在误差控制与计算资源利用之间的综合表现。

从 DECG 指标来看, 完整 MARL 的 DECG 值为 0.093, 低于 Yang-RL 的 0.313 和 MARL w/o GAT 的 0.110, 说明完整模型在中心化训练目标与去中心化执行动作之间具有更好的一致性。Yang-RL 的 DECG 值较高, 表明其混合决策结构虽然能够生成网格剖分动作, 但全局训练目标与局部单元动作之间仍存在较明显偏差。MARL w/o GAT 的 DECG 值高于完整 MARL, 说明去除图注意力模块后, 模型对邻域误差分布和局部耦合关系的感知能力下降, 从而削弱了单元级动作与全局优化目标之间的一致性。对于 MARL w/o VD, 由于其去除了价值分解结构, 不再具备将中心化价值目标分解到局部智能体贡献的机制, 因此 DECG 指标无法计算, 记为 N/A。

消融实验结果表明, 图注意力网络主要提升了模型对邻域信息和误差空间结构的感知能力, 价值分解结构增强了中心化训练目标向局部动作决策的可解释分解能力, 而 CTDE 框架则使模型能够在训练阶段利用全局信息、在执行阶段保持单元级去中心化决策。三者共同作用, 使得完整 MARL 模型在效率指标和 DECG 指标上均取得最优结果。

3.3 泛化实验结果与分析

为了进一步验证所提出强化学习网格剖分策略的泛化能力, 本文在未参与训练的初始条件、网格结构、控制方程类型以及边界条件下开展泛化实验。泛化实验包括 5 个场景, 分别从初始分布复杂性、网格拓扑变化、方程非线性、物理机制差异和边界传播效应等方面对训练环境进行扩展。泛化场景 1 为线性对流方程下的对向双环形分布, 重点考察模型对多个分离局部结构、环形边界以及相反传播方向的协调剖分能力; 泛化场景 2 为非结构三角形网格下的线性对流方程, 重点考察基于图结构的状态表示对不规则单元形状和复杂邻接关系的适应能力; 泛化场景 3 为非线

性对流方程, 重点考察模型在传播速度和误差演化受解本身影响时的动态剖分能力; 泛化场景 4 为具有 Neumann 边界条件的扩散方程, 重点考察模型对耗散型演化、梯度平滑和边界条件变化的适应能力; 泛化场景 5 为具有固定壁边界条件的波动方程, 重点考察模型对波动传播、边界反射和局部振荡区域的跟踪能力。各泛化场景的方程类型、初始分布、网格结构和边界条件如表 5 所示, 与训练场景相比, 泛化实验中的物理量分布形式、网格连接关系、方程演化机制和边界约束均发生变化, 因此能够更充分地检验模型是否学习到了数值解演化与网格剖分之间的内在关系, 而不是仅仅记忆训练样本中的特定剖分模式。

实验结果如图 6 至图 10 所示, 图中 (a) 和 (c) 表示在当前时刻根据网格剖分策略得到的剖分结果, (b) 和 (d) 表示在相应剖分网格上完成数值步进后的计算结果。由结果图可以看出, 在泛化场景 1 中, 强化学习方法能够同时识别两个相互分离且传播方向相反的环形高梯度区域, 并在内外边界附近布置较高分辨率网格, 说明模型具有一定的多目标区域协调剖分能力。

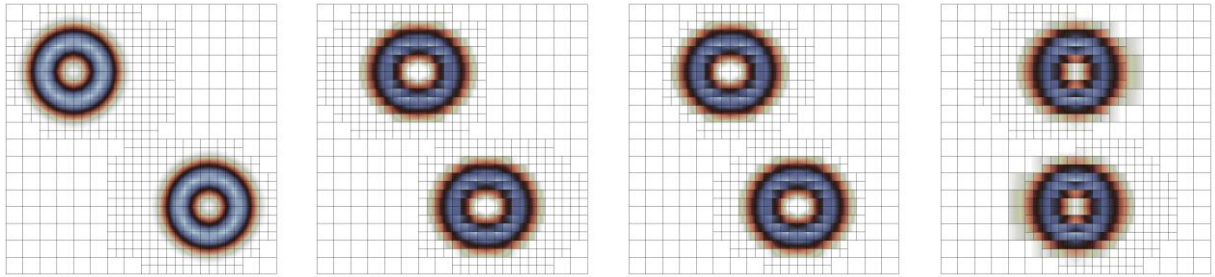
在泛化场景 2 中, 尽管网格单元形状和连接关系与训练阶段的规则网格存在明显差异, 模型仍能够将剖分区域集中于数值解变化剧烈的位置, 表明基于图结构的状态表示和邻域信息聚合机制有助于增强模型对不同网格拓扑的适应能力。在泛化场景 3 中, 数值解的传播速度和局部梯度变化与解本身相关, 误差增长区域相较线性对流问题更加复杂。实验结果表明, 所提出方法仍能够根据解的局部演化特征选择主要剖分区域, 并在物理量变化平缓的位置保持较低网格分辨率, 从而避免不必要的计算资源消耗。在泛化场景 4 中, 解的演化主要表现为局部峰值的平滑扩散和梯度逐渐减弱, 模型能够在梯度变化较明显的区域进行适度加密, 而不是对整个区域进行均匀剖分, 说明该方法能够根据误差分布变化动态调整剖分策略。在泛化场景 5 中, 数值解具有明显的传播和边界反射特征, 模型能够在波前、反射区域以及局部振荡较强的位置保持较高分辨率, 表明所

表 5 泛化实验场景设置。

Tab.5 Generalization test scenarios.

场景编号	方程类型	网格类型	初始分布	边界条件
1	线性对流方程	结构四边形网格	对向双环形分布	周期边界条件
2	线性对流方程	非结构三角形网格	高斯型局部分布	周期边界条件
3	非线性对流方程	结构四边形网格	高斯型局部分布	周期边界条件

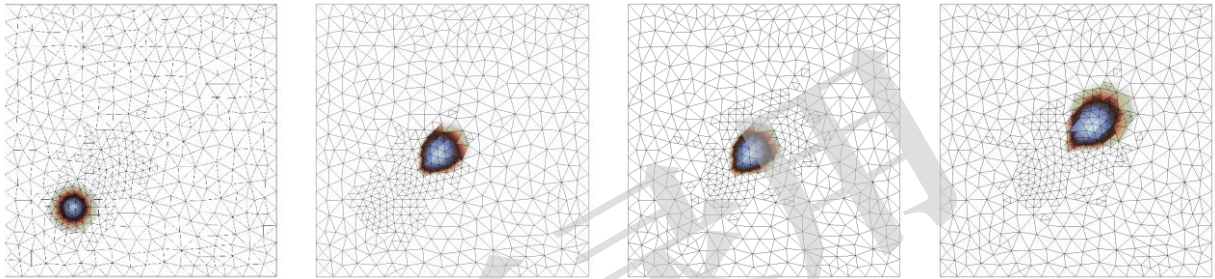
4	扩散方程	结构四边形网格	高斯型局部分布	Neumann 边界条件
5	波动方程	结构四边形网格	高斯型局部分布	固定壁边界条件



(a) 第 1 次剖分 (b) 第 1 次剖分后数值步进 (c) 第 2 次剖分 (d) 第 2 次剖分后数值步进

图 6 泛化测试场景 1 下强化学习方法的网格剖分结果。

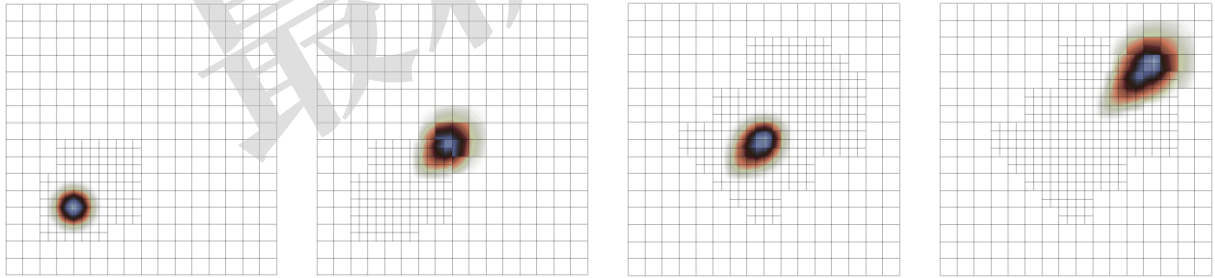
Fig.6 Mesh refinement results of the reinforcement learning method under generalization test scenario 1.



(a) 第 1 次剖分 (b) 第 1 次剖分后数值步进 (c) 第 2 次剖分 (d) 第 2 次剖分后数值步进

图 7 泛化测试场景 2 下强化学习方法的网格剖分结果。

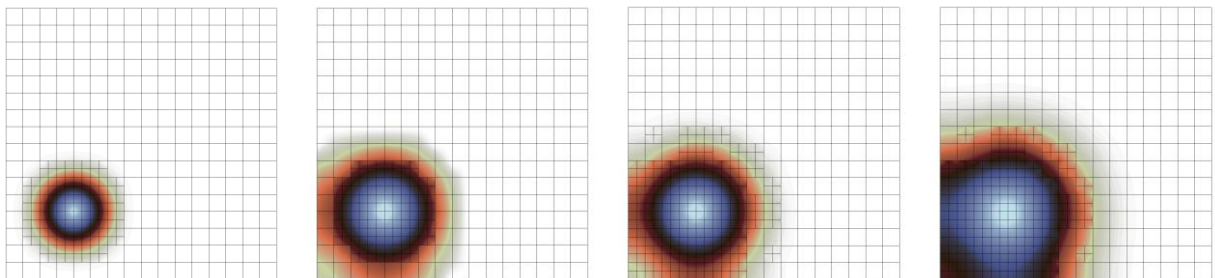
Fig.7 Mesh refinement results of the reinforcement learning method under generalization test scenario 2.



(a) 第 1 次剖分 (b) 第 1 次剖分后数值步进 (c) 第 2 次剖分 (d) 第 2 次剖分后数值步进

图 8 泛化测试场景 3 下强化学习方法的网格剖分结果。

Fig.8 Mesh refinement results of the reinforcement learning method under generalization test scenario 3.



(a) 第 1 次剖分 (b) 第 1 次剖分后数值步进 (c) 第 2 次剖分 (d) 第 2 次剖分后数值步进

图 9 泛化测试场景 4 下强化学习方法的网格剖分结果。

Fig.9 Mesh refinement results of the reinforcement learning method under generalization test scenario 4.

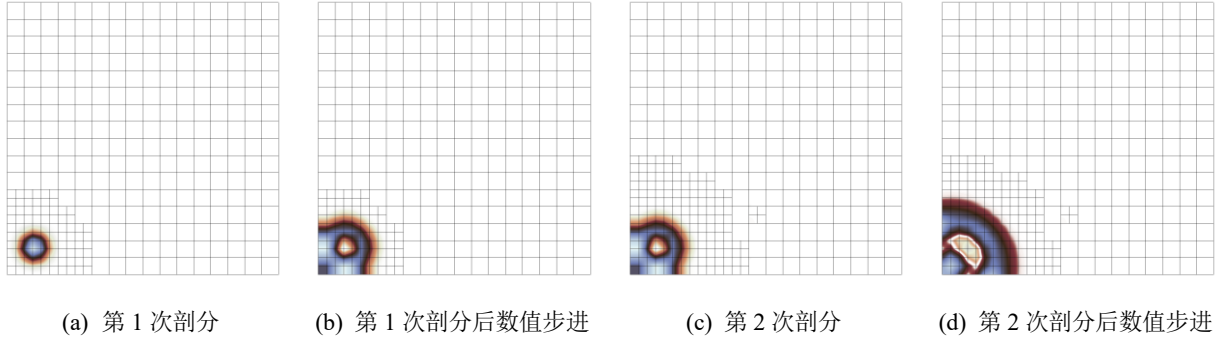


图 10 泛化测试场景 5 下强化学习方法的网格剖分结果。

Fig.10 Mesh refinement results of the reinforcement learning method under generalization test scenario 5.

提出策略对波动型问题中的动态误差传播也具有一定捕捉能力。

在此基础上,本文引入多类代表性基线方法,对所提出的 MARL 网格剖分方法进行综合评价。传统 AMR 基线包括基于梯度指示器的 Grad-based AMR^[21]、基于 Zienkiewicz–Zhu 型恢复误差指示器的 ZZ-type AMR^[22]、基于相邻网格解差的 Löhner-type AMR^[5]和 Kelly-type AMR^[23,24]。其中, Grad-based AMR 主要依据局部物理量梯度大小进行网格剖分,能够直接反映局部变化剧烈区域; ZZ-type AMR 则通过恢复型误差指示器估计局部误差,相较于单纯梯度指标具有更好的平滑性和稳定性; Löhner-type AMR 利用相邻网格单元的解值变化构造归一化差分指标,用于识别局部曲率较大或变化突出的区域; Kelly-type AMR 则依据相邻单元之间的解跳跃、梯度跳跃或通量跳跃衡量局部离散误差。强化学习基线采用 Yang et al.^[17]的方法,用以对比单智能体强化学习网格剖分策略与本文多智能体强化学习框架之间的差异。结果如表 6 所示,表中 SARL 表示单智能体强化学习; MARL / best baseline 表示本文方法与同一测试场景下最优基线方法效率指标的比值,所有结果均基于 30 次随机初始条件重复实验统计得到,表中数值采用“均值(标准误)”

形式表示。

由表 6 可以看出,在训练场景中,本文 MARL 的效率指标达到 0.89,高于各传统 AMR 方法和 Yang-RL。其中, Yang-RL 为该场景下的最优基线,效率指标为 0.83,本文方法相对提升约 7%,说明其能够更有效地平衡误差控制与计算资源消耗。

在泛化测试场景中,本文 MARL 仍保持较好性能。对于线性对流方程的对向双环形分布,本文方法效率指标为 0.75,相对最优基线 0.72 提升约 4%,说明该方法能够协调多个分离高梯度区域的剖分决策。对于非结构三角形网格场景,本文方法效率指标为 0.69,相对最优基线 0.67 略有提升,表明模型对网格拓扑变化具有一定适应能力。在非线性对流方程场景中,本文方法效率指标为 0.77,高于最优基线 0.73,说明其能够较好捕捉非线性输运中的误差演化。

在扩散方程 Neumann 边界条件和波动方程固定壁边界条件场景中,本文方法效率指标分别为 0.79 和 0.75,相对最优基线提升约 4%和 3%。虽然提升幅度较小,但在方程类型和边界条件变化的情况下仍保持最优表现。总体来看, MARL / best baseline 在所有场景中均大于 1,表明本文方法在训练场景和多类泛化场景下均具有稳定优势。

表 6 不同 AMR 方法在训练场景及泛化测试场景下的效率指标对比。

Table 6 Comparison of efficiency metrics for different AMR methods in the training and generalization test scenarios.

方法类别	方法	训练场景	泛化场景 1	泛化场景 2	泛化场景 3	泛化场景 4	泛化场景 5
传统 AMR	Grad-based	0.73 (0.01)	0.57 (0.01)	0.59 (0.01)	0.62 (0.01)	0.69 (0.01)	0.55 (0.02)
传统 AMR	ZZ-type	0.77 (0.01)	0.61 (0.01)	0.61 (0.02)	0.58 (0.01)	0.75 (0.01)	0.59 (0.01)
传统 AMR	Löhner-type	0.81 (0.01)	0.70 (0.02)	0.63 (0.01)	0.60 (0.02)	0.72 (0.01)	0.66 (0.01)
传统 AMR	Kelly-type	0.76 (0.01)	0.68 (0.01)	0.65 (0.01)	0.64 (0.01)	0.73 (0.01)	0.63 (0.01)
SARL	Yang-RL	0.83 (0.02)	0.72 (0.02)	0.67 (0.01)	0.73 (0.02)	0.76 (0.01)	0.73 (0.01)
MARL	MARL	0.89 (0.01)	0.75 (0.01)	0.69 (0.01)	0.77 (0.02)	0.79 (0.01)	0.75 (0.01)

MARL/best baseline	1.07	1.04	1.03	1.05	1.04	1.03
--------------------	------	------	------	------	------	------

为进一步验证本文 MARL 方法相较于 Yang-RL 方法的提升是否具有统计可靠性, 本文基于 30 次随机初始条件重复实验结果进行配对显著性检验, 采用双侧配对 t 检验计算 p 值。结果如表 7 所示, 表中数值采用“均值 (标准误)”形式表示。

表 7 MARL 与 Yang-RL 的配对显著性检验结果。

Tab.7 Paired significance test results between MARL and

Yang-RL.				
测试场景	Yang-RL	MARL	相对提升	p 值
训练场景	0.83 (0.02)	0.89 (0.01)	7.2%	0.002
泛化场景 1	0.72 (0.02)	0.75 (0.01)	4.2%	0.013
泛化场景 2	0.67 (0.01)	0.69 (0.01)	3.0%	0.035
泛化场景 3	0.73 (0.02)	0.77 (0.02)	5.5%	0.010
泛化场景 4	0.76 (0.01)	0.79 (0.01)	3.9%	0.015
泛化场景 5	0.73 (0.01)	0.75 (0.01)	2.7%	0.021

由表 7 可以看出, MARL 在训练场景及各泛化测试场景中的效率指标均高于 Yang-RL。各场景下的 p 值均小于 0.05, 说明 MARL 相较 Yang-RL 的提升在统计意义上显著; 同时, 不同泛化场景下的提升幅度存在差异, 说明 MARL 在整体上具有稳定优势, 但其增益大小会受到方程类型、网格结构和边界条件变化的影响。

4 结论

针对多尺度输运过程数值模拟中局部强梯度结构捕捉与计算效率难以兼顾的问题, 本文提出一种基于多智能体强化学习的 AMR 方法。该方法将网格剖分过程建模为马尔可夫决策过程, 以网格单元为智能体, 结合图结构状态表示、图注意力邻域信息聚合以及 CTDE 框架下的价值分解机制, 实现了局部剖分动作与全局误差控制、计算资源优化目标之间的协同。

实验结果表明, 本文方法能够根据数值解的空间分布和演化特征动态调整网格分辨率, 在较低网格更新频率下仍保持较好的误差控制能力。消融实验进一步表明, 图注意力机制、价值分解结构和 CTDE 框架均对模型性能提升具有贡献。泛化实验结果显示, 本文方法在复杂初始分布、非结构网格、非线性对流、扩散方程和波动方程等未参与训练的场景下仍能保持较稳定的效率表现, 相较传统 AMR 方法和 Yang-RL 方法具有一定优势。

需要说明的是, 本文目前仍主要面向典型输

运类模型开展方法验证。后续工作将进一步面向包含源项、耗散效应、非线性耦合和背景场作用的复杂物理模型开展测试, 以评估该方法在更接近实际空间环境数值模拟问题中的适用性。

参考文献 (References)

- [1] TÓTH G, VAN Der HOLST B, SOKOLOV I V, *et al.* Adaptive numerical algorithms in space weather modeling[J]. *Journal of Computational Physics*, 2012, **231**(3): 870-903
- [2] LEVEQUE R J. Finite Volume Methods for Hyperbolic Problems[M]. Cambridge: Cambridge University Press, 2002
- [3] BERGER M J, OLIGER J. Adaptive mesh refinement for hyperbolic partial differential equations[J]. *Journal of computational Physics*, 1984, **53**(3): 484-512
- [4] BERGER M J, COLELLA P. Local adaptive mesh refinement for shock hydrodynamics[J]. *Journal of Computational Physics*, 1989, **82**(1): 64-84. DOI: 10.1016/0021-9991(89)90035-1
- [5] LÖHNER R. An adaptive finite element scheme for transient problems in CFD[J]. *Computer Methods in Applied Mechanics and Engineering*, 1987, **61**(3): 323-338. DOI: 10.1016/0045-7825(87)90098-3
- [6] HUANG W Z, RUSSELL R D. Adaptive Moving Mesh Methods[M]. New York: Springer, 2010
- [7] ZHANG W Q, ALMGREN A, BECKNER V, *et al.* AMReX: a framework for block-structured adaptive mesh refinement[J]. *Journal of Open Source Software*, 2019, **4**(37): 1370
- [8] COLELLA P, GRAVES D T, LIGOCCI T J, *et al.* Chombo Software Package for AMR Applications Design Document[R]. Berkeley: Lawrence Berkeley National Laboratory, 2003
- [9] WISSINK A M, HORNUNG R D, KOHN S R, *et al.* Large scale parallel structured AMR calculations using the SAMRAI framework[C]//Proceedings of the 2001 ACM/IEEE Conference on Supercomputing. Denver: IEEE, 2001: 22-22
- [10] BERZINS M, LUITJENS J, MENG Q Y, *et al.* Uintah: a scalable framework for hazard analysis[C]//Proceedings of the 2010 TeraGrid Conference. Pittsburgh: ACM, 2010: 1-8
- [11] MACNEICE P, OLSON K M, MOBARRY C, *et al.* PARAMESH: a parallel adaptive mesh refinement community toolkit[J]. *Computer Physics Communications*, 2000, **126**(3): 330-354
- [12] BURSTEDDE C, WILCOX L C, GHATTAS O. *p4est*: scalable algorithms for parallel adaptive mesh refinement on forests of octrees[J]. *SIAM Journal on Scientific Computing*, 2011, **33**(3): 1103-1133
- [13] CAO W M, HUANG W Z, RUSSELL R D. A study of monitor functions for two-dimensional adaptive mesh generation[J]. *SIAM Journal on Scientific Computing*, 1999, **20**(6): 1978-1994
- [14] BOHN J, FEISCHL M. Recurrent neural networks as optimal mesh refinement strategies[J]. *Computers & Mathematics with Applications*, 2021, **97**: 61-76
- [15] PAN J, HUANG J W, CHENG G D, *et al.* Reinforcement learning for automatic quadrilateral mesh generation: a soft actor-critic approach[J]. *Neural Networks*, 2023, **157**: 288-304
- [16] FOUCART C, CHAROUS A, LERMUSIAUX P F J. Deep reinforcement learning for adaptive mesh refinement[J].

- Journal of Computational Physics*, 2023, **491**: 112381
- [17] YANG J C, DZANIC T, PETERSEN B K, *et al.* Reinforcement learning for adaptive mesh refinement[C]//International Conference on Artificial Intelligence and Statistics. Valencia: PMLR, 2023: 5997-6014
- [18] PELISSIER U, PARRET-FRÉAUD A, BORDEU F, *et al.* Graph neural networks for mesh generation and adaptation in structural and fluid mechanics[J]. *Mathematics*, 2024, **12**(18): 2933
- [19] RASHID T, SAMVELYAN M, DE WITT C S, *et al.* Monotonic value function factorisation for deep multi-agent reinforcement learning[J]. *The Journal of Machine Learning Research*, 2020, **21**(1): 178
- [20] CROWLEY W P. Numerical advection experiments[J]. *Monthly Weather Review*, 1968, **96**(1): 1-11
- [21] MOSHFEGH R, LI X D, NILSSON L. Gradient-based refinement indicators in adaptive finite element analysis with special reference to sheet metal forming[J]. *Engineering Computations*, 2000, **17**(8): 910-932
- [22] ZIENKIEWICZ O C, ZHU J Z. The superconvergent patch recovery and a *posteriori* error estimates. Part 1: the recovery technique[J]. *International Journal for Numerical Methods in Engineering*, 1992, **33**(7): 1331-1364. DOI: 10.1002/nme.1620330702
- [23] KELLY D W, DE S R GAGO J P, ZIENKIEWICZ O C, *et al.* A *posteriori* error analysis and adaptive processes in the finite element method: part I—error analysis[J]. *International Journal for Numerical Methods in Engineering*, 1983, **19**(11): 1593-1619

- [24] DE S R GAGO J P, KELLY D W, ZIENKIEWICZ O C, *et al.* A *posteriori* error analysis and adaptive processes in the finite element method: part II—adaptive mesh refinement[J]. *International Journal for Numerical Methods in Engineering*, 1983, **19**(11): 1621-1656

作者简介



钟家浩 男, 2000 年 11 月出生于山东省, 现为中国科学院国际空间科学中心研究生, 主要研究方向为复杂系统仿真。

E-mail: 1210614005@qq.com



唐文林 (通信作者) 男, 1986 年 3 月生于云南省, 现为中国科学院国家空间科学中心研究员, 博士生导师, 主要研究方向为空间引力波探测, 空间等离子体应用研究等。

E-mail: tangwenlin@nssc.ac.cn