



Adaptive Swarm Mesh Refinement Using Deep Reinforcement Learning with Local Rewards

Niklas Freymuth¹ · Philipp Dahlinger¹ · Tobias Würth² · Simon Reisch¹ · Luise Kärger² · Gerhard Neumann¹

Received: 21 May 2025 / Revised: 18 April 2026 / Accepted: 2 May 2026
© The Author(s) 2026

Abstract

Simulating physical systems is essential in engineering, but analytical solutions are limited to straightforward problems. Consequently, numerical methods like the Finite Element Method (FEM) are widely used. However, the FEM becomes computationally expensive as problem complexity and accuracy demands increase. Adaptive Mesh Refinement (AMR) improves the FEM by dynamically placing mesh elements on the domain, balancing computational speed and accuracy. Classical AMR depends on heuristics or expensive error estimators, which may lead to suboptimal performance for complex simulations. While AMR methods based on machine learning are promising, they currently only scale to simple problems. In this work, we formulate AMR as a system of collaborating, homogeneous agents that iteratively split into multiple new agents. This agent-wise perspective enables a spatial reward formulation focused on reducing the maximum mesh element error. Our approach, Adaptive Swarm Mesh Refinement++ (ASMR++), offers efficient, stable optimization and generates highly adaptive meshes at user-defined resolution at inference time. Extensive experiments demonstrate that ASMR++ outperforms heuristic approaches and learned baselines, matching the performance of expensive error-based oracle AMR strategies. ASMR++ additionally generalizes to different domains during inference, and produces meshes that simulate up to 2 orders of magnitude faster than uniform refinements in more demanding settings.

Keywords Multi-agent reinforcement learning · Swarm systems · Graph neural networks · Adaptive mesh refinement · Local rewards

Editor: Javier Garcia.

✉ Niklas Freymuth
freymuth@kit.edu

¹ Karlsruhe Institute of Technology (KIT), Autonomous Learning Robots, Karlsruhe, Germany

² Karlsruhe Institute of Technology (KIT), Institute of Vehicle System Technology, Karlsruhe, Germany



1 Introduction

The numerical simulation of fundamental physical principles like mass, momentum, and energy conservation, often expressed through complex Partial Differential Equations (PDEs), is a cornerstone of modern engineering. Since analytical solutions of such PDEs are limited to simple cases, numerical approximations, particularly the Finite Element Method (FEM) are commonly employed (Brenner & Scott, 2008; Reddy, 2019; Anderson et al., 2021). The FEM provides a framework to find approximate solutions by transforming the continuous PDEs into a discrete system of equations. It partitions the continuous problem domain into a mesh consisting of smaller, finite elements, allowing for an efficient numerical solution whose accuracy depends on the number used elements. However, as the physics becomes more complex, accurate simulations become significantly more expensive (Wan-ner & Hairer, 1996; Zimmerling et al., 2022; Brandstetter et al., 2022).

To bypass the high computational costs of the FEM, an alternative line of research focuses on "learned simulators" that approximate physical solutions directly from data. Early work on these learned simulators uses simple feed-forward (Um et al., 2018; Zimmerling et al., 2019) or convolutional neural networks (Zimmerling et al., 2022; Guo et al., 2016; Zimmerling et al., 2019). Due to the used network architectures, this line of work is ill-suited for irregular, unstructured mesh-based representations in the FEM. Here, Graph Network Simulators (GNSs) (Sanchez-Gonzalez et al., 2020; Pfaff et al., 2021) have emerged as an alternative architecture acting on a graph encoding of the simulation state. As an alternative to data-driven learning approach, physics-informed neural networks (Raissi 2019, Würth 2023) directly optimize a neural network to predict simulations that satisfy the governing equations of a physical system. While usually mesh-free, recent extensions combine these approaches with GNSs and meshes to facilitate generalization to new domains during inference (Würth et al., 2024). All of these learned simulators share a common goal of using recent advances in deep learning to speed up the simulation of complex physical systems. Yet, they do so by directly approximating the simulation, causing any prediction error of the learned model to directly affect the simulated quantities.

Instead of replacing the solver entirely, a more robust and risk-averse approach to speed up the FEM is Adaptive Mesh Refinement (AMR), which dynamically allocates more mesh elements to regions of high solution variability, striking a favorable balance between computational efficiency and accuracy (Plewa et al., 2005; Huang & Russell, 2010; Fidkowski & Darmofal, 2011). As AMR allows the creation of a simulation-specific mesh, it has become increasingly important for complex problems across domains like fluid dynamics (Berger & Colella, 1989; Baker, 1997; Zhang et al., 2020; Wallwork et al., 2022), structural mechanics (Ortiz & Quigley Iv, 1991; Stein, 2007; Gibert et al., 2019), and astrophysics (Cunningham et al., 2009; Bryan et al., 2014; Guillet et al., 2019). Yet, general-purpose AMR methods often rely on either simple heuristics (Zienkiewicz & Zhu, 1992) or computationally intensive error estimation techniques, such as goal-oriented adaptive finite element methods (Becker et al., 2023) and dual-weighted residual methods (Bangerth & Rannacher, 2013). While these approaches can achieve improved accuracy for specific applications, they are generally limited by high computational overhead for high-fidelity estimates or a lack of robustness across different problem domains (Mukherjee, 1996; Kita & Kamiya, 2001; Yano & Darmofal, 2012; Cervený et al., 2019; Wallwork, 2021). The above constraints complicate their effective use in practical scenarios.

Several recent methods apply supervised learning to improve existing or devise new AMR strategies. Here, methods usually predict intermediary metrics for specific aspects of AMR, such as predicting an error or refinement marking per mesh element for a subsequent refinement step (Zhang et al., 2020; Wallwork et al., 2022; Bohn & Feischl, 2021; Roth et al., 2022; Służalec et al., 2023), or predicting local mesh densities over the considered domain (Huang et al., 2021). These supervised approaches typically propose a greedy and often local next-step refinement focused on minimizing the error in the next refinement step. However, this short-term focus may come at the cost of long-term optimization, particularly in problems where local refinements influence error propagation globally or where the simulation involves multiple sequential steps (Yang et al., 2023).

Finally, an emerging body of work employs Reinforcement Learning (RL) (Sutton & Barto, 2018) to formulate AMR, and specifically *h-refinement* (Arnold et al., 2000; Stevenson, 2008), i.e., element sub-division and combination, as a sequential decision-making process. Multiple strategies for RL-AMR (RL-AMR) have been developed, all of which are motivated by the potential of RL to optimize non-differentiable, long-term rewards that are formulated to correspond to an error measure on the mesh. When considering *h-refinement*, the discrete mesh structure evolves at each step, resulting in a dynamic state space that requires the RL policy to handle a varying number of inputs and outputs. Some work circumvents this issue by learning a global, per-mesh quantity (Gillette et al., 2022; Pan et al., 2023), acting on local mesh elements (Huerger et al., 2024; Dzanic et al., 2024), or sampling a fixed number of mesh edges to refine per step (Wu et al., 2023). More general methods that consider the full mesh either do so by iteratively selecting an element to refine (Yang et al., 2023) or training on individual element's refinements and then inferring on the full mesh (Foucart et al., 2023). Another recent method uses a value decomposition network (Sunehag et al., 2017) to learn the credit assignment of the individual elements on the mesh quality by decomposing a shared Q-function (Yang et al., 2023). All of these approaches only scale to simple problems, either due to an expensive inference process (Yang et al., 2023), misaligned objectives and high variance in the state transitions (Foucart et al., 2023), or noisy reward signals during training (Yang et al., 2023).

In this work, we instead formulate AMR via *h-refinement* as a Swarm RL (Hüttenrauch et al. 2019, Šošić et al. 2017) problem, extending existing frameworks to a shared observation space and per-agent, spatial rewards. Crucially, we allow for agents to split into new agents over time to model element sub-division in the refinement process, mapping a per-agent reward signal quantifying the reduction in simulation error over refinement steps to assign credits for individual agents throughout an episode. We use a Message Passing Networks (MPNs) (Sanchez-Gonzalez et al., 2020), a type of Graph Neural Network (GNN) (Scarselli et al., 2009; Bronstein et al., 2021), for our policy due to their effectiveness in learning physical simulations (Brandstetter et al., 2022; Pfaff et al., 2021). Our method, Adaptive Swarm Mesh Refinement (ASMR), consistently produces highly efficient mesh refinements with thousands of elements and can be applied to arbitrary PDEs. ASMR can further generalize to previously unseen domains, system dynamics and material parameters when trained appropriately, allowing for robust and adaptive solutions across a wide range of complex scenarios.

This paper extends a previously published conference paper on ASMR (Freymuth et al., 2023). While the original ASMR policies produced meshes at a fixed granularity, we now introduce an adaptive element penalty. This scalar value is added to the reward function to

penalize over-refinement and is provided to the policy as context, allowing a single model to generate meshes of varying densities. Additional extensions include an improved mapping between agents over time that acts as a regularizer and leads to better results, an improved network architecture and an improved reward formulation based on the reduction of the maximum local error of the mesh. While this reward formulation was already introduced and ablated in our prior work with inconclusive results, we now present a large scale study that shows its improved performance compared to the previously used reward. We refer to the improved version as ASMR++, and to the previous version as *ASMR*. Experimentally, we find that the proposed changes substantially improve the method on more difficult task setups. We further provide additional visualizations of our method, including a schematic of the agent mapping and a qualitative error comparison to a uniform mesh. Finally, we add two challenging tasks that utilize Neumann boundary conditions and a 3-dimensional domain to showcase the applicability of our method to a wider range of applications. Figure 1 provides a schematic overview of a single ASMR++ refinement step.

We validate the effectiveness of ASMR++ on a wide range of PDEs that require complex refinement strategies. Our experiments use triangular and tetrahedral meshes for simplicity, although ASMR++ is agnostic to the element type of the underlying mesh. We employ conforming elements and corresponding h-adaptive refinements (Arnold et al., 2000; Stevenson, 2008), i.e., refinements via element subdivision, due to their prevalence in engineering (Ho-Le, 1988; Jones & Plassmann, 1997; Nagarajan & Soghrati, 2018). Compared to non-stationary settings from related work, which only consider a low maximum refinement depth or local refinements, the challenge here is to find multiple levels of precise refinement on the full mesh. We consider several recent RL-AMR methods as baselines (Yang et al., 2023, 2023; Foucart et al., 2023). As these methods have been shown to work for mesh refinement and coarsening with a relatively low depth on dynamic tasks, we adapt them to our setting of stationary meshes that require multiple precise refinement steps and thousands of elements. We further compare to a threshold-based refinement heuristic that uses the popular Zienkiewicz-Zhu Error Estimator (*ZZ Error*) estimate (Zienkiewicz & Zhu,

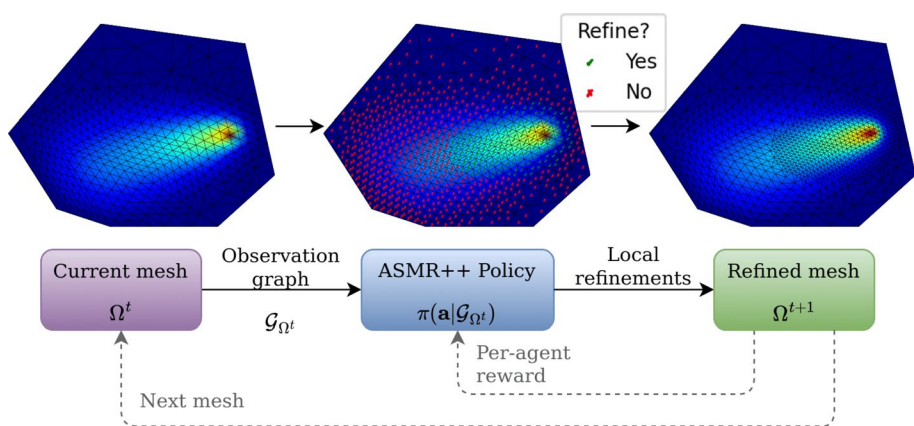


Fig. 1 A schematic ASMR++ refinement step. Given a current mesh, an observation graph encodes the elements as nodes and their neighborhood relationship as edges. A policy consumes this graph to decide on per-element markings that are given to a remeshing, which subsequently produces a refined mesh. Based on the quality of this finer mesh, per-agent rewards are calculated. The process is repeated until the mesh is fully refined

1992). Additionally, we consider different *oracle error estimates*, so called because they rely on privileged information from a high-resolution uniform mesh to approximate the true solution error, to mark elements for refinement.¹

2 Adaptive Swarm Mesh Refinement++

ASMR++ treats elements of a mesh as a swarm of homogeneous agents that collaborate to find an optimal refinement. For this, each agent's state and observation are defined through its topological position in the mesh and includes local, rotation- and translation-invariant features of both the mesh and the PDE. At each step, all agents decide whether to mark their respective element for refinement. Agents that refine their element receive a *local* reward based on how much their refinement has improved the quality of the underlying simulation. These local rewards are aggregated into a global term for the return to ensure that each mesh element optimizes its local region while contributing to the global and its total element countal solution quality. Crucially, each refinement subdivides the refined elements, introducing new agents in the process. We introduce a Adaptive Swarm Markov Decision Process (ASMDP) that features a mapping of agents over time, allowing us to optimize over multiple refinement steps by propagating reward from agents of later time steps back to related earlier agents. Figure 1 provides a schematic overview of ASMR++, and the following sections describe the individual aspects in more detail.

2.1 Adaptive Swarm Markov Decision Process

We view mesh refinement as a collaborative multi-agent RL problem with a changing number of homogeneous agents and agent-wise rewards. For this, we adopt a swarm RL view, adapting the SwarMDP framework (Hüttenrauch et al. 2019, Šošić et al. 2017) to vector-valued rewards and changing state, action and observation spaces. Formally, let an Adaptive Swarm Markov Decision Process (ASMDP) be a tuple $\langle \mathbb{S}, \mathbb{O}, \mathbb{A}, P, r, \xi, \phi \rangle$. The state space, observation space and action space are given by $\mathbb{S}$, $\mathbb{O}$, and $\mathbb{A}$ respectively. Since the number of agents changes over time, we denote subsets of the state, observation, and action spaces with N agents as $\mathcal{S}^N \subset \mathbb{S}$, $\mathcal{O}^N \subset \mathbb{O}$, $\mathcal{A}^N \subset \mathbb{A}$. The transition function $P : \mathcal{S}^N \times \mathcal{A}^N \rightarrow \mathcal{S}^M$ takes an action over N agents and leads to a new state with M agents. The reward function $r : \mathcal{S}^N \times \mathcal{A}^N \rightarrow \mathbb{R}^N$ takes the full state into account to produce a scalar reward for each agent. Similarly, the observation function $\xi : \mathcal{S}^N \rightarrow \mathcal{O}^N$ calculates local observations for each agent from the global state.

We consider a finite-horizon setting with a horizon of T steps. In the scalar reward case, i.e., with $r(s, a) \in \mathbb{R}$, an optimal RL policy $\pi : \mathbb{O} \times \mathbb{A} \rightarrow [0, 1]$ generally maximizes the return, i.e., the expected cumulative future reward

$$J^t := \mathbb{E}_{\pi(a|\xi(s))} \left[\sum_{k=0}^T r(s^{t+k}, a^{t+k}) \right]. \quad (1)$$

¹ We implement all tasks as OpenAI gym (Brockman et al., 2016) environments and publish our code, including all approaches and tasks presented in this paper, at <https://github.com/NiklasFreymuth/ASMRplusplus>.

A key challenge of using RL for AMR is the changing number of agents introduced by splitting agents that are marked to be refined. To adapt the SwarMDP framework and its underlying Markov Decision Process to changing numbers of agents within a single episode, we thus introduce an agent mapping $\phi^t \in \mathbb{R}^{N \times M}$. Intuitively, each entry ϕ_{ij}^t describes the responsibility of agent i at step t for agent j at step $t + 1$. The responsibilities of agents at step t for agents at step $t' > t$ can then be computed via the matrix multiplication

$$\phi^{t,t'} := \phi^t \phi^{t+1} \dots \phi^{t'} = \Pi_{\tau=t}^{t'} \phi^\tau.$$

Given this mapping, we can propagate the rewards of all future agents for which agent i is responsible for at step t to calculate a return

$$J_i^t := \mathbb{E}_{\pi(a|\xi(s))} \left[\sum_{t'=t}^T \gamma^{k} (\phi^{t,t'} r(s^{t'}, a^{t'}))_i \right]. \quad (2)$$

For training, e.g., a value function $V(s)$, this leads to a TD error (Sutton & Barto, 2018)

$$\delta^t = r(s^t, a^t)_i + \sum_j \phi_{ij}^t V_j(s^{t+1}) - V_i(s^t) \quad (3)$$

of actions $a \sim \pi(\xi(s))$. We derive the targets for Q -functions analogously.

The ASMDP framework allows for a collaborative optimization of changing numbers of agents, each of which is equipped with its own reward. Compared to, e.g., a centralized partially observable Markov Decision Process (MDP) (Yang et al., 2023; Foucart et al., 2023), it provides a more efficient learning environment, since every environment step provides a learning signal for each agent. Similarly, the ASMDP supports a multi-agent setting with changing numbers of homogeneous agents, circumventing the posthumous credit assignment problem (Cohen et al., 2021) without having to resort to dummy states and learned value decompositions (Yang et al., 2023). The agent mapping additionally allows for an optimization of the return, i.e., the reward over time, which would not be possible with a purely local per-agent view. While this work focuses on AMR, the ASMDP framework can readily be applied to other scenarios that are characterized by localized decisions of changing numbers of entities.

2.2 Actions and Agent Mappings

In the context of AMR, a system state $s \in \mathbb{S}$ is defined by the problem domain Ω , a mesh discretization Ω , and a PDE formulated on Ω to be solved under specific conditions. ASMR++ views every element Ω_i^t of a mesh $\Omega^t := \{\Omega_i^t \subseteq \Omega \mid \bigcup_i \Omega_i^t = \Omega\}$ as one of many homogeneous agents. Each agent has a binary action space, deciding on whether it wants to mark its element for refinement or not. These markings are provided to a remesher, yielding a finer mesh $\Omega^{t+1} = \{\Omega_j^{t+1}\}_j$. The remesher subdivides each refined element Ω_i^t into a set of smaller elements $\{\Omega_j^{t+1}\}_j$, such that the disjoint union of all Ω_j^{t+1} reconstructs Ω_i^t , i.e., $\bigcup_j \Omega_j^{t+1} = \Omega_i^t$. It may also refine unmarked elements to close the mesh and thus assert a

conforming solution (Arnold et al., 2000), i.e., to make sure that elements of the mesh align with each other at the boundaries to ensure continuity of solution variables between adjacent elements.

To map between elements and their successors, *ASMR* uses the indicator function $\phi_{ij}^t := \mathbb{I}(\Omega_j^{t+1} \subseteq \Omega_i^t)$. This mapping assigns each element to all elements that it subdivides into, or equivalently, lets each agent be responsible for all new agents that it spawns. Here, we extend the mapping to include a normalization factor

$$\phi_{ij}^t := \frac{|\Omega_j^t|}{|\Omega^{t+1}|} \mathbb{I}(\Omega_j^{t+1} \subseteq \Omega_i^t). \quad (4)$$

Compared to *ASMR*, the normalization factor $\frac{|\Omega_j^t|}{|\Omega^{t+1}|}$ scales the responsibilities such that

$\sum \phi^t = \sum \phi^{t'}$. This procedure intuitively ensures that the total ‘mass’ of agents is preserved over time. We find that this factor regularizes the mapping of rewards within the agent swarm, akin to, e.g., batch normalization (Ioffe & Szegedy, 2015).

Conceptually, we may represent each mesh in a series of refinements as a layer in a hierarchical graph, with nodes corresponding to mesh elements. In our case, this graph is a simple tree. Here, each node connects to its predecessor in the prior layer and to all successors in the subsequent layer. Applying an agent mapping equates to navigating through the respective level of hierarchy of this graph. An example refinement procedure and its corresponding refinement graph for an initial mesh with a single element and 2 refinement steps can be seen in Fig. 2

This work focuses on adaptive mesh refinement via iterative mesh subdivision. For mesh coarsening, the above mapping can be extended to respect element coarsening as, e.g.,

$$\phi_{ij}^t := \frac{|\Omega_j^t|}{|\Omega^{t+1}|} \left[\mathbb{I}(\Omega_j^{t+1} \subseteq \Omega_i^t) + \left(\mathbb{I}(\Omega_i^t \subsetneq \Omega_j^{t+1}) / \left(\sum_k \mathbb{I}(\Omega_k^t \subsetneq \Omega_j^{t+1}) \right) \right) \right].$$

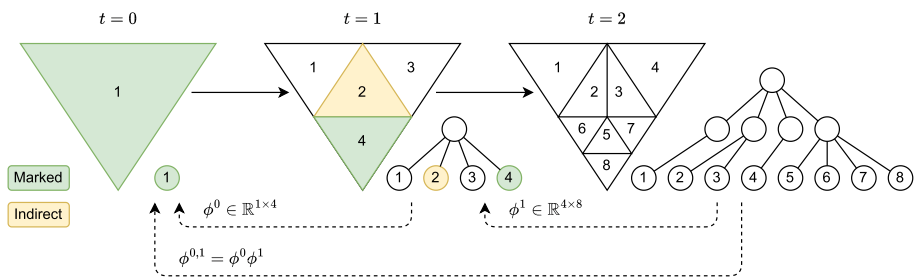


Fig. 2 Refinement procedure and responsibility mapping of ASMR++. Left: An initial mesh has its single element marked for refinement. Middle: The mesh element is subdivided into 4 new elements. The correspondence between the old and new mesh is represented as a directed acyclic graph. Constructing the matrix ϕ^0 from this graph allows us to map the new elements back to the old one. For the next step, Element 4 is marked for refinement, requiring an indirect refinement of Element 2 to ensure a conforming mesh. Right: After the second refinement, the resulting mesh consists of 8 elements. Using the refinement graph to construct the mapping ϕ^1 allows us to compute responsibilities of elements in the previous mesh for this mesh. Chaining these mappings as $\phi^{0,1} = \phi^0 \phi^1$ allows us to directly compute responsibilities between the initial mesh and the mesh after two refinement steps

In this case, the hierarchical graph is no longer a tree but instead a directed acyclic graph, since multiple elements can coarsen into the same new element. Similar modifications can be made to adapt ASMR++ to other AMR and adaptive mesh generation methods.

2.3 Observations and Policy Architecture

Given a mesh Ω^t and its corresponding system of equations, we construct an observation as the bidirectional graph $\mathcal{G}_{\Omega^t} = \mathcal{G} = (\mathcal{V}, \mathcal{E}, X_{\mathcal{V}}, X_{\mathcal{E}}) \in \mathcal{O}$ with mesh elements as nodes $\mathcal{V}$ and their neighborhood relation as edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Mesh- and task-dependent node and edge features of dimensions $d_{\mathcal{V}}$ and $d_{\mathcal{E}}$ are provided as $X_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}^{d_{\mathcal{V}}}$ and $X_{\mathcal{E}} : \mathcal{E} \rightarrow \mathbb{R}^{d_{\mathcal{E}}}$.

This observation graph is consumed by an MPN policy. MPNs (Pfaff et al., 2021; Linkerhägner et al., 2023) are a popular GNN architecture for mesh-based physical simulation (Pfaff et al., 2021; Linkerhägner et al., 2023; Würth et al., 2024) as they encompass the function class of several classical PDE solvers (Brandstetter et al., 2022). Given an input graph, they iteratively update high-dimensional latent representations on this graph for L consecutive message passing steps. Each step l receives the output of the previous step and updates the latent states $X_{\mathcal{V}}, X_{\mathcal{E}}$ for all nodes $v \in \mathcal{V}$ and edges $e \in \mathcal{E}$. Using linear embeddings x_v^0 and x_e^0 of the initial node and edge features, the l -th step is given as

$$x_e^{l+1} = f_{\mathcal{E}}^l(x_v^l, x_u^l, x_e^l), \text{ with } e = (u, v),$$

$$x_v^{l+1} = f_{\mathcal{V}}^l(x_v^l, \bigoplus_{e=(v,u) \in \mathcal{E}} x_e^{l+1}).$$

The operator $\oplus$ is a permutation-invariant aggregation such as a sum, mean or maximum over all aggregated elements. Each $f_{\cdot}^l$ is a learned function, usually parameterized as a simple Multilayer Perceptron (MLP). The MPN's final output is a learned representation x_v^L for each node $v \in \mathcal{V}$. This learned representation can be fed into a policy head to compute an action per agent, yielding a joint action vector $\pi(a|\mathcal{G}_{\Omega^t})$, or into a value head to compute a per-element value function estimate. Both are parameterized as MLPs that are shared between the different agents. Since MPNs are permutation-equivariant by design (Bronstein et al., 2021), this policy architecture ensures that the amount and ordering of mesh elements does not matter, i.e., that the elements are only defined by their features and local position in the mesh.

We note recent trends in surrogate modelling for engineering applications (Alkin et al., 2024, 2025) and RL for continuous control (Nauman et al., 2024; Lee et al., 2024) increasingly favor heavily regularized large-scale architectures, such as transformers (Vaswani et al., 2017). While our MPN architecture offers several desirable inductive biases, these larger and more complicated policy backbones are a promising extension that may enable ASMR++ to scale to larger domains and highly non-linear PDEs.

2.4 Reward Definition

Adaptive Mesh Refinement (AMR) aims to generate meshes with enhanced local resolution in areas that are relevant for the underlying system of equations. Its objective is to achieve

an optimal balance between the solution's accuracy on the mesh Ω^t , and its total element count $|\Omega_i^t \in \Omega^t|$.

Since closed-form solutions are generally not available for most systems of equations, we compute an approximate ground truth solution u_{Ω^*} using a fine-grained uniform reference mesh Ω^* (Yang et al., 2023). While calculating such a reference solution is expensive, it only needs to be computed once per training environment and is not required during inference. Given u_{Ω^*} , a refined mesh Ω^t and its solution u_{Ω^t} , we define the *error per element* as the maximum difference between the solutions on this element. We use the maximum difference over solutions as a measure of simulation quality that is independent of, e.g., the size of the used elements, simplifying the evaluation compared to, e.g., integrated $L2$ error computations. Given reference elements Ω_m^* with midpoints $p_{\Omega_m^*} \in \Omega_m^*$ as sampling points to query the solutions on, this yields an error

$$\text{err}(\Omega_i^t) \approx \max_{\Omega_m^* \subseteq \Omega_i^t} \left| u_{\Omega^*}(p_{\Omega_m^*}) - u_{\Omega^t}(p_{\Omega_m^*}) \right|. \quad (5)$$

We efficiently calculate the assignment $\Omega_m^* \in \Omega_i^t$ required for this reward using a k -d tree (Bentley, 1975). If a queried point $p_{\Omega_m^*}$ exactly lies on the edge between two elements in Ω^t , which is possible for some refinement algorithms, we assign it to both of these elements with a weight of 0.5. We normalize this error estimate with the initial mesh error to get the relative improvement of the mesh-error, i.e., $\text{err}(\Omega_i^t) = \text{err}(\Omega_i^t) / \sum_{\Omega_j \in \Omega^0} \text{err}(\Omega_j^0)$.

This estimate is consistent across geometries and process conditions. As the size of the solution u_{Ω^*} and subsequently that of the error $\text{err}(\Omega_i^t)$ may differ significantly between systems of equations due to varied process conditions, this normalization prevents systems of equations with large solution quantities from dominating the learning process. Given the normalized error estimate, we next construct an element-wise local reward function that estimates the benefit of a refinement and compares it to the cost of adding additional elements. In systems of equations with multiple quantities of interest, such as, e.g., elasticity problems where both the element displacement and stresses are important, the mesh must be optimized for all the quantities of interest. Here, we calculate rewards independently for each normalized solution dimension and use some problem-dependent average or norm to compute a scalar reward per element.

To define our reward, we introduce a variable element penalty α that penalizes each added element, and compare this to the decrease in error within each element when it is refined, yielding

$$r(\Omega_i^t) := \left(\text{err}(\Omega_i^t) - \max_j \phi_{ij}^t \text{err}(\Omega_j^{t+1}) \right) - \alpha \left(\sum_j \mathbb{I}(\Omega_j^{t+1} \subseteq \Omega_i^t) - 1 \right). \quad (6)$$

The local reward for each element is 0 if no refinement is made. If an element is refined, its reward is positive if and only if the refinement decreases the highest error in this element by more than the cost of adding new elements, which is scaled with the element penalty α . In comparison, *ASMR* (Freymuth et al., 2023) rewards the integrated average change in error for each mesh element instead of its maximum change in error. This average change in error is integrated over the element volume, meaning that small elements, i.e., those that

have been previously refined, quickly become comparatively insignificant for the reward. To counteract this, *ASMR* scales each reward by the inverse of the respective element's volume, leading to widely varying and potentially unbound reward values that do not directly match any optimization criteria on the full mesh. We instead directly reward the change in maximum error within the volume of each element. This maximum reward variant has been proposed as an ablation study in (Freymuth et al., 2023), but only been explored for a single task. We find that the maximum reward results in more well-behaved optimization, presumably because it directly targets refining elements where the reduction in error justifies adding new elements. We provide a more detailed comparison to the *ASMR* reward in B.2. Since Eq. 6 is based on an element's decrease in maximum error, it is independent of the element size, ensuring that rewards of elements of different sizes are on the same scale without the need for explicit scaling.

Equation 6 rewards each agent for its local refinement decision for each time step. While this direct and explicit local credit assignment leads to a temporally and spatially dense reward, it disregards non-local effects of mesh refinement, which are common for elliptical PDEs. To encourage global optimization, we therefore extend the return J_i^t of Eq. 2 to include a global term as

$$J_i^{t'} = \frac{1}{2} J_i^t + \frac{1}{2} J^t, \quad (7)$$

where J^t is the global return of Eq. 1 using the average reward $r = \frac{1}{N} \sum_j r_j$.

2.5 Element Penalty

We add the element penalty α in Eq. 6 as an input to the policy to inform it about the penalty of adding new elements. During training, we draw α from a log-uniform distribution at the start of each episode, allowing for environment samples on a wide range of different penalties. Similar to Value Decomposition Graph Networks (*VDGN*) (Yang et al., 2023), we then specify a value of α during inference to control the refinement level of the created mesh. Compared to *ASMR*, which trains a policy on a fixed element penalty value, we thus condition the policy on α for each rollout. This conditional policy allows for different mesh granularities during inference, whereas *ASMR* needs to train a new policy when the desired number of mesh elements changes.

Balancing the extra computational cost of adding mesh elements against the benefit of reducing simulation errors can be seen as an instance of multi-objective RL (Hayes et al. 2022, Van Moffaert and Nowé 2014). The cost of adding new elements increases linearly with a penalty factor α , and there exists some approximate ranking of the effectiveness of different mesh refinements at reducing simulation errors. This setup creates a roughly convex relationship between α and the policy's behavior. In other words, lowering α generally leads to a finer mesh that includes all refinements made with a higher α , while raising α results in fewer refinements throughout the mesh. Figure 3 provides an example of resulting *ASMR++* meshes for the same trained policy and Neumann boundary task for different values of the element penalty α . Regardless of the penalty, the parts of the mesh with the highest potential error reduction, in this case the areas near the load function and the sinusoidal Neumann boundaries, are refined the most.

2.6 Pseudocode

Algorithms 1 and 2 provide pseudocode for the training and inference procedures of ASMR++. ASMR++ iteratively interacts with a set of systems of equations during training, obtaining a reward by comparing the solution quality of its refined meshes to a fine-grained reference Ω^* . After training has finished, it can be applied to unseen problems without requiring access to such a fine-grained reference, providing accurate refinements from only the current coarse-mesh solution.

Input: Training environment $\mathcal{E}$ (Sampler for PDE, Ω^0 , Ω^*)
Input: Refinement steps T , Penalty range $[\alpha_{\min}, \alpha_{\max}]$
Input: Initial policy π_θ
Output: Optimized policy π_θ

```

1 while not converged do
2   Data Generation:
3   Initialize buffer  $\mathcal{B} \leftarrow \{\}$ 
4   for each rollout do
5     Sample  $(\text{PDE}, \Omega^0, \Omega^*) \sim \mathcal{E}$ 
6     Sample penalty  $\alpha \sim \text{LogUniform}(\alpha_{\min}, \alpha_{\max})$ 
7     for  $t = 0 \dots T - 1$  do
8       Get solution  $\mathbf{u}_{\Omega_t}$  from PDE and mesh  $\Omega^t$ 
9       Construct graph  $\mathcal{G}^t$  from  $\mathbf{u}_{\Omega_t}$  and  $\Omega^t$  (Section 2.3)
10      Sample refinement decisions  $\mathbf{a}_t \sim \pi_\theta(\mathcal{G}^t, \alpha)$ 
11       $\Omega^{t+1}, \phi_t \leftarrow \text{Refine}(\Omega^t, \mathbf{a}_t)$ 
12      Compute local rewards  $\mathbf{r}_t$  using  $\Omega^*$  and penalty  $\alpha$  (Equation 6)
13       $\mathcal{B} \leftarrow \mathcal{B} \cup \{(\mathcal{G}^t, \mathbf{a}_t, \mathbf{r}_t, \phi_t)\}$ 
14    end
15  end
16  Policy Update:
17  Compute global returns  $J^t$  (Equation 7) from rewards in  $\mathcal{B}$ 
18  Calculate loss and update parameters  $\theta$ 
19 end
20 return  $\pi_\theta$ 

```

Algorithm 1 ASMR++ Training

Input: Coarse mesh Ω^0 , Trained policy π_θ
Input: Refinement steps T , Element penalty α
Output: Final refined mesh Ω^T

```

1 for  $t = 0 \dots T - 1$  do
2   Get solution  $\mathbf{u}_{\Omega_t}$  from PDE and mesh  $\Omega^t$ 
3   Construct graph  $\mathcal{G}^t$  from  $\mathbf{u}_{\Omega_t}$  and  $\Omega^t$  (Section 2.3)
4   Select refinement decisions  $\mathbf{a}_t \leftarrow \arg\max \pi_\theta(\mathcal{G}^t, \alpha)$  // Deterministic
5    $\Omega^{t+1}, \_ \leftarrow \text{Refine}(\Omega^t, \mathbf{a}_t)$ 
6 end
7 return  $\Omega^T$ 

```

Algorithm 2 ASMR++ Inference (Deployment)

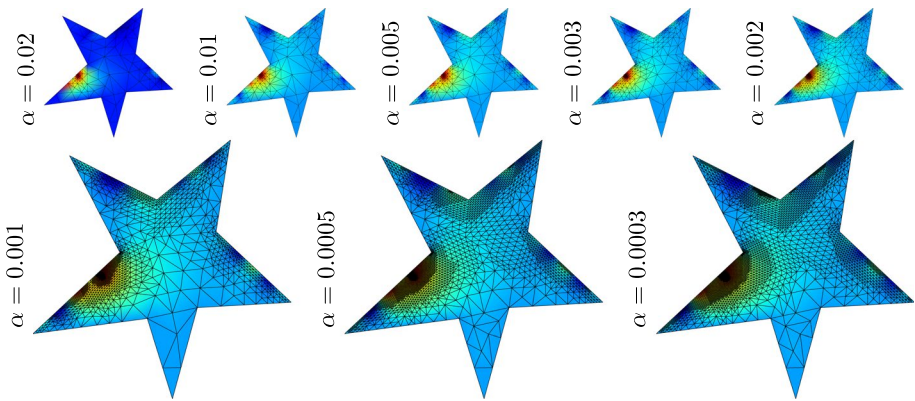


Fig. 3 ASMR++ refinements for a Poisson problem with sinusoidal Neumann boundary conditions for different inputs of the element penalty α . All refinements focus on the relevant parts of the problem, and lower element penalties lead to more fine-grained meshes

3 Experiments

3.1 Systems of Equations

The physical behavior of a technical system can be described by a set of PDEs. Simulations predict the physical states of a system by solving those PDEs. For non-trivial problems, the solving method is usually based on numerical approximation methods, such as the FEM, which considers the weak formulation of the PDEs. Multiplying the strong PDE formulation with a test function $v(x) \in V$ and integrating by parts yields the weak formulation (Larson & Bengzon, 2013), which we denote as $\Phi(u, v)$. We denote V as the space spanned by a set of basis functions $\phi_v(x) \in V$ and the sought solution as $u(x)$. In the Galerkin FEM, the solution $u(x, t)$ is approximated by a weighted sum of these basis functions $u(x) = \sum_{v=1}^{N_\varphi} u_v(t) \phi_v(x)$. Additionally, the same basis function are used to define the test function $v(x) = \sum_{v=1}^{N_\varphi} \phi_v(x)$. In this work we consider linear elliptical or parabolic PDEs. For elliptical PDEs, we seek for a space-dependent solution $u = u(x)$. This results in N_φ equations for N_φ unknown degrees of freedom u_v . Due to the linearity of the PDEs and the discretization of the FEM, we can assemble a linear system of equations $Au = b$ out of the weak formulation $\Phi(\phi_u, u_v, \phi_v)$. The time-dependent parabolic PDEs requires a discretization in time, e.g., using forward or backward Euler schemes. Subsequently, a system of equations can be assembled for each time step. Depending on the scheme, the resulting system is either solved with numerical solvers or evaluated directly. For each PDE, we create a task that consists of the weak formulations, a set of domains, and varying process conditions. The goal is to generate meshes that minimize the error of the FEM, relative to a fine-grained reference mesh Ω^* , which approximates the true solution of the system of equations, while keeping the number of elements in the generated meshes as low as possible.

Figure 4 shows examples for each task and visualizes final meshes created using an ASMR++ policy. Each task is associated with a family of domains, including L-shapes, rectangles with a square hole or multiple rhomboid holes, convex polygons, star shapes with different numbers of peaks, and a 3-dimensional plate. From left to right and top to

bottom, these tasks are briefly characterized as follows. *Laplace's equation* task considers a rectangular heat source at the inner boundary and features high solution gradients near this source. *Poisson's equation* uses a multi-modal load function comprised of a Gaussian mixture model, which in turn causes multiple distinct regions of interest in the solution. The *Stokes flow* task solves for the velocity of a fluid streaming into the domain from inlet on the left side. It uses more complex shape functions and requires high precision near the corners of the rhomboid obstacles. The *linear elasticity* task simulates the deformation and the resulting stress of a metal plate. The *heat diffusion* task simulates a time-dependent heat source and is evaluated at the final simulation step. Adding sinusoidal *Neumann Boundaries* to Poisson's Equation models scenarios where the boundary flux is specified, resulting in high solution variability near the domain edges. The *3d Poisson* task considers a 3-dimensional variant of Poisson's equation, using tetrahedral meshes and a different underlying refinement algorithm.

We implement the FEM, all systems of equations, meshes and refinements in *scikit-fem* (Gustafsson & Mcbain, 2020). All tasks are provided as OpenAI gym (Brockman et al., 2016) environments. While ASMR++ is designed to work with arbitrary element types, our experiments focus on conforming triangular and tetrahedral meshes. We use linear elements unless mentioned otherwise. Appendix A specifies the systems of equations and their process conditions.

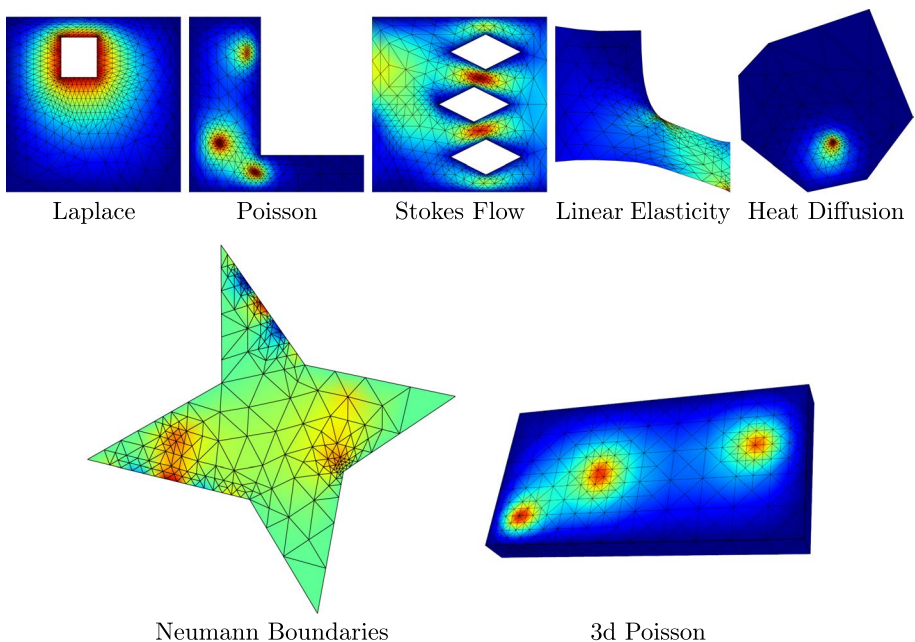


Fig. 4 Final ASMR++ meshes and corresponding FEM solution for different tasks. We visualize the velocity norm for Stokes Flow and a weighted average of the displacement norm and Von-Mises stress for Linear Elasticity. For the other tasks, we visualize their scalar solution quantity as a heatmap. ASMR++ provides highly adaptive meshes across a wide variety of PDEs by accurately optimizing task-specific regions of interests. It additionally supports Neumann boundary conditions and volumetric 3d meshes

3.2 Adaptive Mesh Refinement

For each system of equations, we generate an initial coarse mesh Ω^0 , which is then refined during an episode through iterative element subdivision. For simplicity, we use the default mesh refinement methods implemented in *scikit-fem* (Gustafsson & McBain, 2020). In 2-dimensional domains, we employ the red-green-blue refinement method (Carstensen, 2004), subdividing each marked element into 4 smaller elements and then closing the mesh, as illustrated in Fig. 2. For 3-dimensional meshes, we apply longest edge bisection (Rivara, 1984; Suárez et al., 2005) to halve each marked element. We emphasize that ASMR++ is compatible with arbitrary refinement algorithms, as it learns to mark elements for refinement based on the resulting error reduction.

In 2 dimensions, we uniformly refine the initial coarse mesh 6 times to create the reference mesh Ω^* , and use the same amount of 6 refinement steps per episode. We additionally evaluate a simpler task setup with 4 refinements for the reference and AMR algorithm to emulate the task complexity of existing work (Yang et al., 2023, 2023; Fortunato et al., 2022). In both cases, always refining all elements precisely reconstructs the reference mesh. In 3-dimensional domains, we uniformly refine three times, resulting in 8^3 times more elements in Ω^* compared to Ω^0 and employ 7 longest edge bisection AMR steps, yielding refinements that are topologically different from the reference mesh. The number of refinements was chosen to ensure that Ω^* is a sufficiently accurate proxy for the exact solution that allows for a challenging refinement task, while remaining computationally feasible for efficient experimentation.

3.3 Training and Evaluation

We train all RL-AMR methods on 100 randomly generated systems of equations, including randomized domains and process conditions. This choice reduces the number of expensive reference meshes Ω^* that must be computed during training, and is further explored in C.3. We evaluate each converged RL-AMR policy and the heuristics on 100 randomly sampled but fixed evaluation PDEs, training each RL-AMR algorithm for 10 random seeds. The systems of equations used for evaluation are disjoint from those used for training.

For all methods, we experiment with a range of target mesh resolutions to produce meshes of different refinement levels, as detailed in E.3. B.1 provides additional details on the experimental setup and additionally discusses the required computational budget. We train the RL-AMR methods using both Proximal Policy Optimization (PPO) (Schulman et al., 2017) with discrete actions as an on-policy algorithm and Deep Q-Network (DQN) (Mnih et al., 2013, 2015) as an off-policy variant.

We evaluate mesh quality by calculating the squared error between the solution on this mesh and the solution on the fine-grained reference Ω^* . This squared error over the domain is numerically approximated over evaluations at the midpoints $p_{\Omega_m^*}$ of the reference mesh as

$$\sum_{\Omega_m^* \in \Omega^*} \text{Volume}(\Omega_m^*) \left(u_{\Omega^*}(p_{\Omega_m^*}) - u_{\Omega^t}(p_{\Omega_m^*}) \right)^2. \quad (8)$$

We normalize the resulting value by that of the initial mesh for comparability across PDEs, and call the resulting metric the remaining squared *mesh error*. This squared error metric

captures the overall error of the solution on the refined mesh, while punishing outliers. It differs from our optimization objective in Eq. 6, which instead rewards minimizing the maximum error of the solution on the refined mesh. C.4 discusses and evaluates a mean error and an approximation of the maximum error as alternate metrics. We list further algorithm and network hyperparameters in E.1.

3.4 RL-AMR Baselines

We compare to *ASMR* (Freymuth et al., 2023) and several recent non-stationary RL-AMR methods (Yang et al., 2023, 2023; Foucart et al., 2023) that have been adapted to our application, focusing on stationary refinements. Following *ASMR*, we use the rewards as described in the respective papers for each method, employing the absolute integrated difference in solution compared to the reference mesh, i.e., Eq. 8 without the square as an error estimate.

Single Agent (Yang et al., 2023) uses the difference in error norm as a global reward and marks a single element for refinement in each step via a categorical action over the full mesh. *Sweep* (Foucart et al., 2023) randomly samples a single mesh element and decides on its refinement based on local features and a global resource budget during training. The method uses the logarithm of the change in error between consecutive steps as a reward. During inference, each timestep considers all mesh elements in parallel. *VDGN* (Yang et al., 2023) employs value decomposition networks (Sunhag et al., 2017) to estimate a global Q-function as the sum of element-wise local Q-functions. For a *VDGN*-like PPO variant, we instead analogously decompose the value function. As our method is an extension of *ASMR* (Freymuth et al., 2023), this baseline largely follows our method. However, it instead uses the volume-scaled reward in Eq. B.1, omits the normalization factor in the agent mapping in Eq. 4, considers an infinite horizon RL setting, and uses less regularization in the neural network architecture.

We use an MPN policy for all RL-AMR methods except for *Sweep*, which utilizes a simple MLP. The RL-AMR methods handle the trade-off between mesh resolution and solution accuracy slightly differently. *Single Agent* simply refines for a number of steps, *Sweep* uses a target number of elements, and the other methods penalize each added element. Here, *ASMR* trains a new policy for each penalty, while only *ASMR++* and the *VDGN*-like baseline use a conditional policy that can produce meshes of different resolutions during inference. E.2 lists all baseline-specific hyperparameters, while E.3 provides details on the mesh resolution parameters for all methods.

3.5 Heuristic Baselines

In addition to the RL-AMR methods above, we compare to a traditional error-based heuristics acting on a refinement threshold θ (Foucart et al., 2023; Binev et al., 2004; Bangerth et al., 2012). Given an error estimate $\text{err}(\Omega_i^t)$ for element i , the heuristic marks all elements for refinement for which $\text{err}(\Omega_i^t) > \theta \cdot \max_j \text{err}(\Omega_j^t)$. We consider an *Oracle Error Heuristic* that assumes knowledge about the fine-grained reference solution Ω^* and uses the absolute integrated difference to this solution as its error metric. Similarly, we compare to a *Maximum Oracle Error Heuristic*, for which the error estimate of Eq. 5 is used. Both variants require the reference mesh Ω^* to estimate an error, which is expensive to compute and thus may be unavailable during inference. We additionally consider the commonly used

ZZ Error, which instead uses a superconvergent patch recovery process for its error estimate (Zienkiewicz & Zhu, 1992) and does not require access to Ω^* . As the recovery process averages over neighboring mesh elements, the *ZZ Error* generally produces smooth error estimates. This property can lead to more coherent refinements when compared to the oracle heuristics, especially when considering the closing operations needed to ensure a conforming mesh after every refinement step.

All heuristics greedily refine based on local error estimates, in the sense that they are based purely on the current error rather than the potential reduction in error. Thus, they implicitly assume that refining elements with high error is the best way to minimize the error on the mesh. This assumption does not always hold, as, e.g., elliptical PDEs often feature globally propagating errors (Foucart et al., 2023; Strauss, 2007). The RL-AMR methods instead learn to optimize an objective function, which rewards the error reduction over time. This difference between greedily marking elements with a high local error, and learning to anticipate the long-term error reduction for any given refinement allows the RL-AMR methods to find more global strategies that take multiple refinement steps and non-local information of the mesh elements into account.

3.6 Observation Graph

As node features for the observation graph of Sect. 2.3, we use the mean and standard deviation of the FEM solution on the element's vertices, the element volume and the current environment timestep. For ASMR++ and the *VDGN*-like baseline, we also add the current element penalty to condition the refinement process of the policy on a given target resolution. Some tasks employ additional node features to encode their process conditions, which we describe in A. We use the Euclidean distance between the element midpoints as the single edge feature. Compared to using either a distance vector between element midpoints, or providing absolute positions, this ensures that the predicted sizing fields are invariant under Euclidean group operations (Pfaff et al., 2021; Bronstein et al., 2021).

4 Results

4.1 Qualitative Results

Figure 5 visualizes a full rollout of ASMR++ on an instance of the heat diffusion task, including the per-element refinement markings after every policy step. For this and other qualitative figures, the background color represents the magnitude of the problem-dependent scalar solution. ASMR++ is trained to produce a sequence of refinements that improve the solution accuracy enough to justify the additional mesh elements, which here corresponds to more refinements along the path of the heat source. Figure 3 shows final refined meshes for the same Poisson problem for a fixed ASMR++ policy that is conditioned on different values of the element penalty α . Here, decreasing the element penalty results in more mesh elements and thus a more accurate solution. Our method is able to focus on the interesting parts of the mesh regardless of the final mesh resolution, thus providing a favorable trade-off between solution accuracy and computational cost. Figure 4 provides final refinements of ASMR++ on exemplary systems of equations for all considered tasks. The refinement strategy indi-

vidually adapts to each task, choosing to refine elements that decrease the simulation error regardless of the underlying system of equations. We further explore this error reduction property in Fig. 6. The visualization shows that our method achieves a much more uniform distribution of errors across the mesh when compared to simple uniform refinement, while also acquiring a much smaller simulation error for the same number of mesh elements.

Appendix F visualizes all methods on the Stokes flow task to showcase the different refinement behaviors for different target mesh resolution parameters. Appendix G provides additional ASMR++ visualizations for the remaining tasks, including different camera angles for the 3-dimensional variant of Poisson's equation that uses tetrahedral meshes.

4.2 Quantitative Results

Quantitatively, we measure the mesh quality using a Pareto plot of the number of elements and the normalized squared error of the mesh, as defined in Eq. 8. We evaluate each RL-AMR policy on 100 evaluation environments per mesh resolution parameter. We then calculate the interquartile mean (Agarwal et al., 2021) of the resulting 100 used elements and normalized errors of the final produced meshes, and plot the resulting values. This process is done for 10 independently trained policy seeds and repeated across a range of mesh resolution parameters. For methods that require a fixed target mesh resolution, we use 10 different parameters, for the adaptive methods we instead use 20 parameters, as detailed in E.3. We further provide a log-log quadratic regression over the aggregated results of each method as a general trend-line. To improve clarity and focus on representative behavior, we omit data points with abnormally high element counts, which are sometimes produced as outliers by the *VDGN*-like baseline.

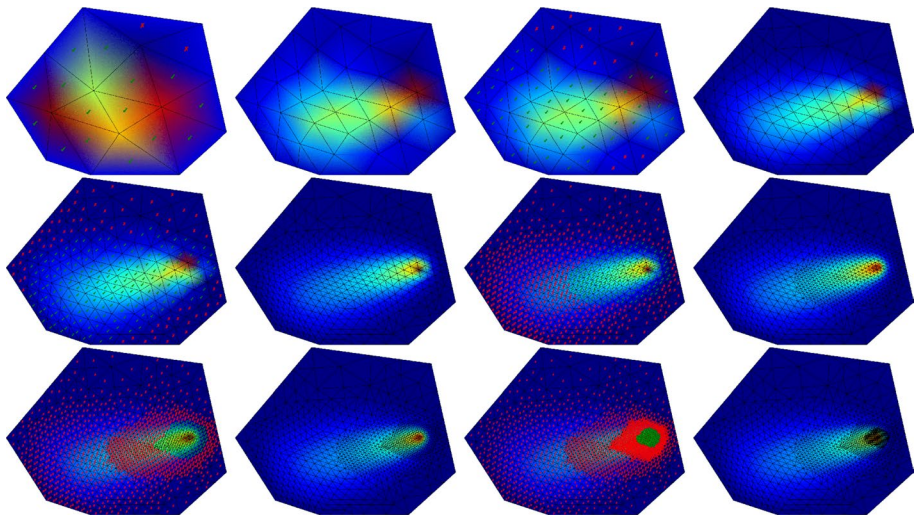


Fig. 5 Visualization of a full ASMR++ rollout for a heat diffusion example, including the policy action that marks individual elements for refinement. The mesh's background color denotes solution magnitude. The odd columns show which elements the policy marks for refinement (designated by a green tick), while the even columns show the resulting refined meshes. ASMR++ produces a sequence of refinements that improves solution accuracy while using few total mesh elements

We experiment with PPO as an on-policy algorithm and DQN as an off-policy variant of the RL backbone for all RL-AMR methods on the Stokes flow task in C.2. All RL-AMR methods, including ASMR++, perform better with PPO, indicating that an on-policy algorithm is favorable when dealing with action and observation spaces of varying size. Further, we compare our MPNs to the Graph Attention Network (GAT)-like (Velickovic et al., 2018) architecture proposed by *VDGN* (Yang et al., 2023) in C.3.1. For both our method and the *VDGN*-like baseline, the MPN shows minor but consistent benefits in performance. We thus use PPO and MPNs in all other experiments.

C compares the performance of the *ZZ Error Heuristic* when applied directly to the initial mesh to variants that instead uniformly refine the mesh once or twice before using the estimator. We find that the method benefits from 2 uniform initial refinements, and use this variant for all experiments. This tuning of the initial meshes is not required by the RL-AMR methods and prevents coarse refinements.

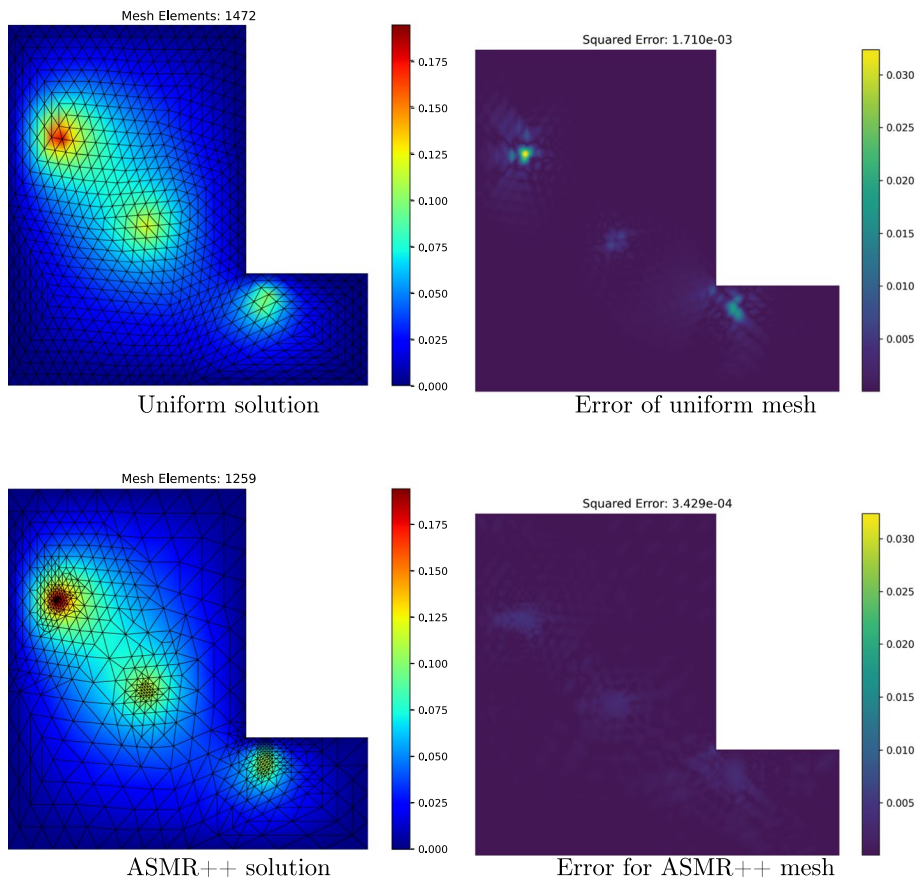


Fig. 6 Comparison of exemplary ASMR++ refinement with a uniform mesh on Poisson's equation. (Left) Mesh and PDE solution for a uniform mesh (Top) and our approach (Bottom). Both meshes have a similar number of total elements, but different spatial resolutions. (Right) Normalized squared difference in solution to the ground truth reference Ω^* , evaluated at each reference element's midpoint. ASMR++ has less than a third of the error of the uniform mesh, and a much more even distribution of errors across the mesh

Using these initial results, we compare the different approaches on Laplace's equation. On the left of Fig. 7, we evaluate a simplified setup that uses 4 refinement steps for both the reference mesh and the AMR algorithms, except for *Single Agent*, which instead refines for up to 100 time steps. All methods clearly outperform uniform refinements. The *ZZ Error Heuristic* performs well for larger meshes, but can not produce meshes with few elements as it starts with 2 uniform initial refinements.

The right of Fig. 7 increases the complexity to 6 refinement steps, and 400 time steps for *Single Agent*. Similarly, Fig. 8 provides results on the remaining 6 tasks. Across all tasks, only *ASMR* and *ASMR++* can handle larger meshes while the other RL-AMR methods fail to provide better-than-uniform meshes or exhibit high variance in the quality of the proposed refinements. Our method additionally outperforms all heuristics, including the *Oracle Error* and *Maximum Oracle Error Heuristics*, in several instances. These results indicate that our optimization method facilitates learning non-local long-horizon refinement strategies that minimize the mesh error instead of simple targeting elements that currently have a high error. *ASMR++* further improves over *ASMR* in most settings, especially in the complex Stokes flow task where it compares well to the different *Heuristics*. In the tasks where *ASMR++* does not perform better than *ASMR*, both perform similar to or better than all heuristics, potentially indicating that both methods offer near-optimal refinements for these tasks. On the 3d Poisson's Equation, the different methods quickly converge to a relatively high final mesh error, likely due to the difference in mesh topology between the refined meshes and the uniform reference that is caused by the longest edge bisection refinement strategy.

C.4 presents additional results using a mean error and an adapted maximum error metric. On both metrics, *ASMR++* is generally competitive with or better than the *Heuristics* and significantly outperforms the *Single Agent*, *Sweep* and *VDGN*-like baselines. Comparing the different metrics, the *Maximum Oracle Error Heuristic* improves over the *Oracle Error Heuristic* on the maximum error metric and vice versa. Likewise, *ASMR*, which uses a scaled variant of the mean error instead of the maximum objective of Eq. 6, outperforms *ASMR++* on half of the tasks in terms of the remaining mean error, while *ASMR++* is superior for every setup when evaluating the maximum error metric.

4.3 Ablations and Additional Experiments

We conduct several ablations and additional experiments on the challenging Stokes flow task to determine what makes our approach uniquely effective. These experiments individually change a single aspect of *ASMR++* while leaving the rest of the approach unchanged.

4.3.1 Reward

We compare the local maximum error reward of Eq. 6 to different local and global variants. For the global variants, we obtain a scalar reward function by averaging over the local reward functions. Here, we compute the return using Eq. 1, i.e., without mapping between agents over time. In addition to the maximum reward of Eq. 6, we compare to the volume-scaled reward proposed by *ASMR* in Eq. B.1, and a variant that simple rewards a decrease in integrated error, i.e., Eq. B.1 without the volume scaling term.

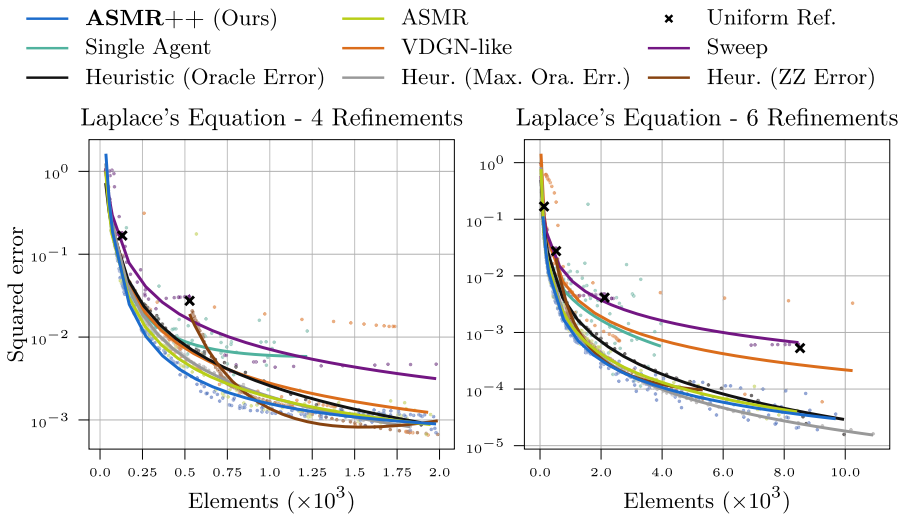


Fig. 7 Pareto plot of normalized squared errors and number of final mesh elements on Laplace's equation. (Left) All RL-AMR methods produce better-than-uniform meshes for shallow refinements and reference meshes with only 4 refinement steps. (Right) When increasing the complexity to 6 refinement steps, most RL-AMR methods exhibit high variance in mesh quality and sometimes fail to provide useful meshes. ASMR++ consistently produces high-quality meshes for both task complexities, performing on par with or better than the heuristics and providing a slight but consistent benefit over *ASMR*

Figure 9 shows the results. Using any type of global reward function leads to unstable refinements, especially for large meshes, likely because the credit for each reward cannot be properly assigned to the large number of agents in the system. For the local reward variants, the local maximum reward performs best, closely followed by the volume-scaled variant. Simply integrating the error of each mesh element and rewarding a reduction in this integrated error performs poorly for finer meshes. This result is likely the case because smaller elements generally contain less integration points of the reference mesh, and thus have less integrated error to reduce. Optimizing this reward produces meshes that have an even distribution of integrated element errors, rather than refinements that effectively decrease the average error on the mesh.

4.3.2 Return and Agent Mapping

Equation 4 enables the computation of a TD error across agents at successive time steps by mapping each agent at time step t to every agent it creates at time step $t + 1$. In this setup, each spawned agent is fully attributed to its predecessor, meaning the originating agent is considered wholly responsible for any agent it generates. Compared to *ASMR*, we additionally apply a regularization to this mapping through the ratio Ω^t / Ω^{t+1} . An alternative approach would be to instead limit the total responsibility per agent to 1, achieved by averaging the mapping as

$$\phi_{ij}^{tt} := \frac{\mathbb{I}(\Omega_j^{t+1} \subseteq \Omega_i^t)}{\sum_{j'} \mathbb{I}(\Omega_{j'}^{t+1} \subseteq \Omega_i^t)}.$$

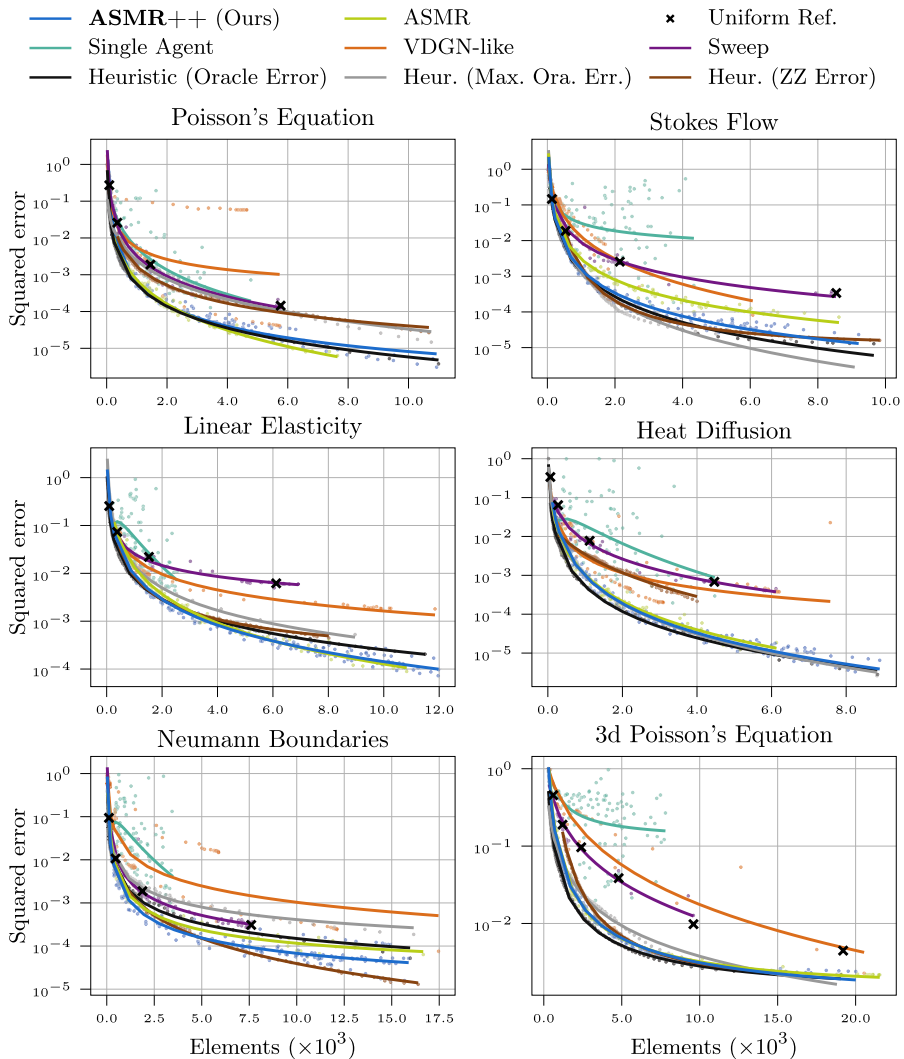


Fig. 8 Pareto plot of normalized squared errors and number of final mesh elements across different tasks. On all tasks, the error decreases in a log-linear relation to the number of used elements. ASMR++ gracefully scales to meshes of more than ten thousand elements, significantly outperforming the learned *Single Agent*, *VDGN-like* and *Sweep* in terms of consistency and mesh quality. ASMR++ also slightly improves over *ASMR* on most tasks and generally compares favorably to the *Oracle*, *Maximum Oracle* and *ZZ ErrorHeuristics*

We assess these two methods in Fig. 10, finding that the normalization factor $\frac{\Omega^t}{\Omega^{t+1}}$ significantly boosts performance, likely due to its role as a regularizer during training. Furthermore, using a sum mapping, i.e., assigning each new agent a weight of 1 from its creator, proves more effective than the mean mapping, which averages the total mapping weights per old agent. We additionally explore optimizing a return that omits the global term in Eq. 7, i.e., that only optimizes the local return of each agent. Here, we find that the partially global objective improves performance, likely because it aids in global decision making.

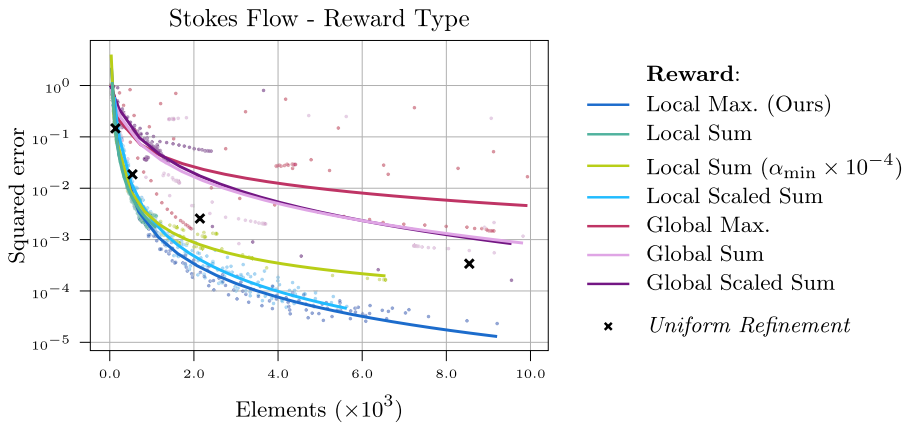


Fig. 9 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different reward functions for ASMR++. In general, a using a local reward function, i.e., individual rewards for each agent, is crucial for the performance of the method. Here, optimizing the reduction in total error per mesh element leads to sub-optimal meshes for large numbers of elements, even when significantly decreasing the minimum sampled element penalty $\alpha_{\min}$. Scaling this optimization with the inverse area of each mesh element as done by *ASMR* offsets this issue, but performs worse than the simpler maximum reward of ASMR++

4.3.3 Target Mesh Resolutions

Each RL-AMR method uses a single parameter to control the number of target elements of the final refined mesh. ASMR++ and *VDGN* condition their policy on an element penalty α , while *ASMR* specifies a fixed element penalty as a hyperparameter. Similarly, *Sweep* uses a fixed budget $N_{\max}$ per policy, and *Single Agent* trains on a specified number of rollout steps T . Fig. 22 evaluates the impact of these parameters for ASMR++ and the *VDGN*-like baseline. ASMR++ remains stable for a wide range of penalty parameters. The penalty can be tuned by identifying the order of magnitude that corresponds to the desired number of mesh elements. This process allows users to easily control the trade-off between solution accuracy and computational cost without requiring sensitive hyperparameter tuning. Additional details are provided in E.3 and Table 2.

Figure 11 compares training ASMR++ on a fixed element penalty, as done in *ASMR* (Frey-muth et al., 2023), to training on an adaptive penalty that the policy is conditioned on for each rollout. It additionally evaluates sampling the penalty uniformly instead of log-uniformly during training. Interestingly, training ASMR++ on a fixed element penalty does not improve performance, suggesting that the ASMR++ policy is able to learn to condition on a concrete α value. These results imply that ASMR++ can successfully learn optimal tradeoffs between element cost and error reduction and apply it during inference. Thus, we can train a single policy for a range of mesh granularities, omitting the need for *ASMR*'s expensive re-training for each target granularity. If we train on uniformly instead of log-uniformly sampled values of α , the method predominantly provides meshes with low numbers of elements, likely because there is a logarithmic relationship between the element penalty and the number of mesh elements.

Figure 12 visualizes the methods' performance for different target resolutions. Here, ASMR++ provides meshes with consistent numbers of elements for a given target resolu-

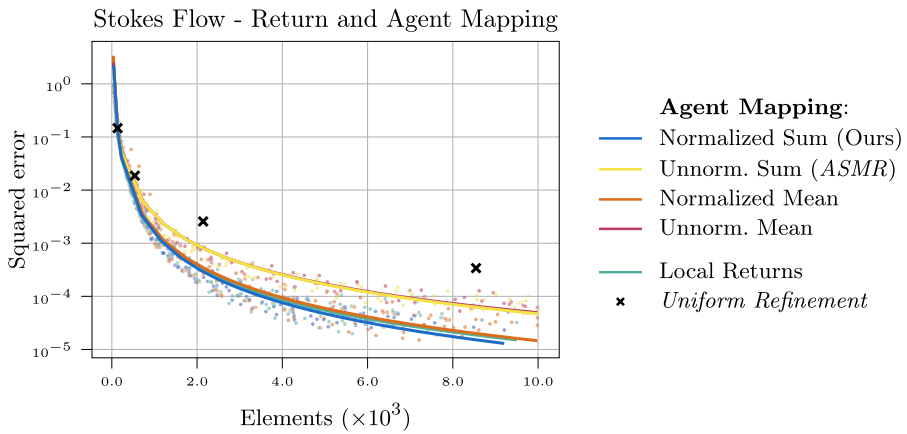


Fig. 10 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different variants of the agent mapping ϕ and local returns of ASMR++. The normalization factor in the mapping of Eq. 4 greatly improves performance. Using a sum instead of a mean mapping additionally improves performance. Adding a global term to the returns, as done in Eq. 7 further improves performance, indicating that a partially global objective improves global decision making of the individual agents

tion, while the other RL-AMR baselines produce poorly refined meshes or inconsistent refinement granularities over target resolutions.

4.3.4 Additional Ablations

C.3 provides additional ablations over changes in the network architecture, varying numbers of training PDEs and different scales of the element penalty parameter in the reward. The results indicate that as few as 10 training PDEs are sufficient to provide accurate refinements, with performance increasing for up to 100 PDEs. In terms of architecture, edge dropout (Rong et al., 2019) of 0.1, i.e., randomly omitting 10 % of edges in the mesh topology during training, slightly improves performance. Linearly increasing or decreasing the range of the element penalty α leads to meshes with fewer or more elements, respectively, yet ASMR++ always provides stable and efficient refinements for the respective element penalty ranges.

4.4 Runtime Comparison

Figure 13 compares the runtime of ASMR++ with directly computing the uniform reference Ω^* across all tasks. Our method's timing includes the creation of the initial mesh, and iteratively solving the system of equations on the mesh, creating an observation graph, querying the policy for a refinement strategy, and finally refining a total of 6 times. For the uniform mesh, we simply measure the time it takes to refine the initial mesh 6 times and to subsequently solve the problem on the resulting mesh. We use a single 8-Core AMD Ryzen 7 3700X Processor for all measurements.

Our approach is always significantly faster than computing the fine-grained reference solution Ω^* despite the computational overhead of computing the observation graph and

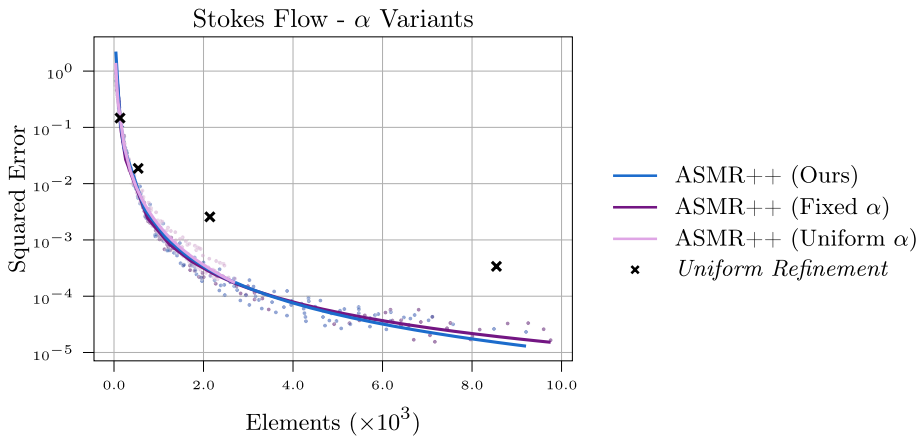


Fig. 11 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different configurations of the element penalty α for ASMR++. Training a new policy for each value of α , as done by *ASMR* (Freymuth et al., 2023), does not significantly improve performance, which indicates the utility and benefit of training a conditional policy on a range of α values. Sampling α uniformly instead of log-uniformly during training leads to meshes with significantly fewer elements, as those values that would create large meshes are sampled less often

executing the policy. Notably, higher squared error tolerances lead to reduced computation time, showing that ASMR++ can adapt to task-specific computational budgets. The Stokes flow task uses comparatively expensive P_2/P_1 Taylor-Hood-elements (John et al., 2016), causing our method to be more than 30 times faster than Ω^* even for highly refined and accurate final meshes.

ASMR uses the same inference procedure as ASMR++ and thus shows the same general runtime behavior as ASMR++ when controlling for network size. Similarly, the other RL-AMR methods also feature an iterative refinement procedure and thus comparable runtime, but produce meshes of worse quality than ASMR++. In contrast, the *Oracle Error Heuristic* and *Maximum Oracle Error Heuristic* produce high-quality meshes, but require computing Ω^* for an error estimate. As such, these heuristics are impractical in terms of runtime, as they depend on the fine-grained solution that AMR aims to approximate.

The *ZZ Error Heuristic*, which relies on element gradient information, is comparatively fast. Figure 14 compares it to ASMR++ for both a fixed number of elements and for achieving a target error. While ASMR++ is slightly slower per element due to the additional cost of generating an observation graph and executing a trained policy, its superior refinement quality leads to comparable or better runtimes for achieving a given target error on the Linear Elasticity and Heat Diffusion tasks. Additionally, the *ZZ Error Heuristic* relies on initial mesh tuning (compare Fig. 19) and exhibits significant variability in solution quality across tasks. While it can be faster than ASMR++ on tasks where its heuristic happens to align well, it lacks the principled, error-focused optimization objective provided by ASMR++ and its underlying RL framework. Prior work (Yang et al., 2023) has also demonstrated that even simpler RL-based AMR approaches outperform it on various problem classes, particularly those involving dynamic or time-varying behavior.

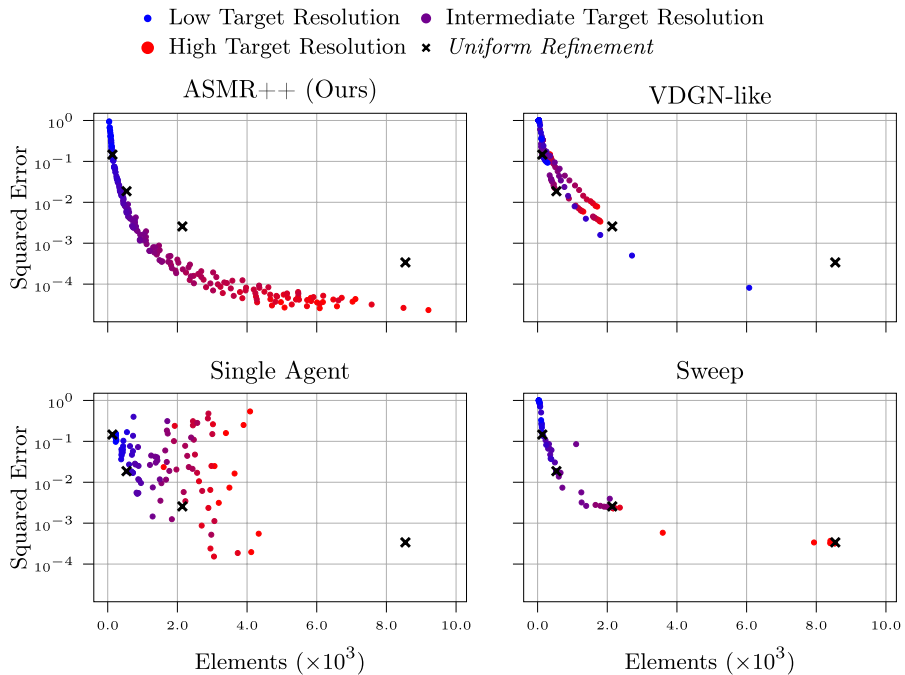


Fig. 12 Pareto plot of normalized squared errors and number of final mesh elements for the Stokes flow task for all RL-AMR methods. Each plot represents a single method for different evaluations of its respective mesh resolution parameters, as detailed in Appendix E.3 and Table 2. Small blue dots indicate meshes target with few elements, while large red dots correspond to finer target meshes. While the different RL-AMR baselines provide poor-performing or inconsistent policies, ASMR++ yields accurate refinements and consistent numbers of final mesh elements when the policy is conditioned on any given target resolution

4.5 Generalization Capabilities

We further experiment with the generalization capabilities of ASMR++ to unseen and larger domains during inference. Appendix D.1 shows the generalization capabilities of our approach when evaluated for different domains and load functions for Poisson's equation, indicating that it strongly generalizes to arbitrary domains and process conditions during inference.

4.5.1 Generalization to Larger Domains

We additionally train on a variant of Poisson's equation that may sample load functions modes outside of the considered domain and randomizes solution values on the domain's boundaries. Appendix D.2 provides training details. These changes effectively modify the training PDEs to mimic small patches of a larger mesh, acting as a data augmentation technique that facilitates up-scaling to larger domains and meshes during inference. An exemplary refinement on a 20×20 from a policy trained on 1×1 domains can be seen in Fig. 15. The refined mesh has more than 50 000 elements, which is significantly larger than

All Tasks - Runtime Comparison

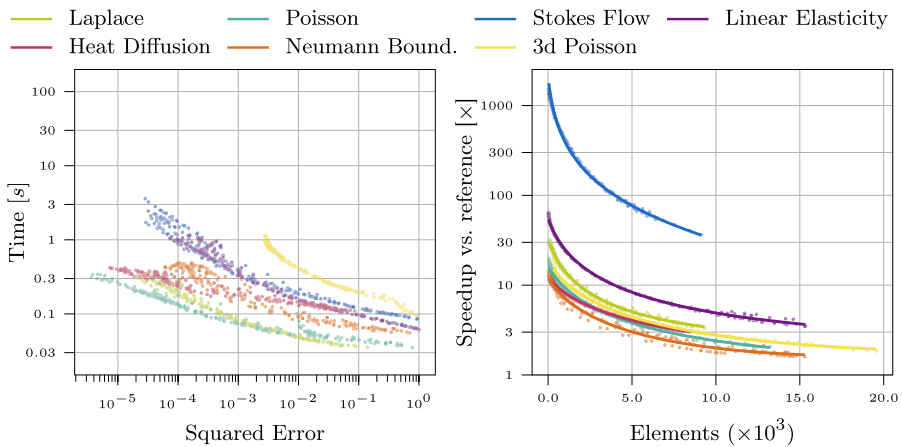


Fig. 13 Wallclock-time evaluation of ASMR++ for all tasks. Each color corresponds to a different task. (Left) Runtime of ASMR++ and the squared error when compared to the uniform solution Ω^* , which has an error of 0 by definition. (Right) Speed-up of ASMR++ compared to calculating the uniform solution Ω^* , which contains roughly 10^5 elements, with concrete numbers varying depending on the domain. Our approach is able to quickly produce high-quality meshes, achieving speedups of 1 to 2 orders of magnitude, depending on the task and target resolution of the mesh

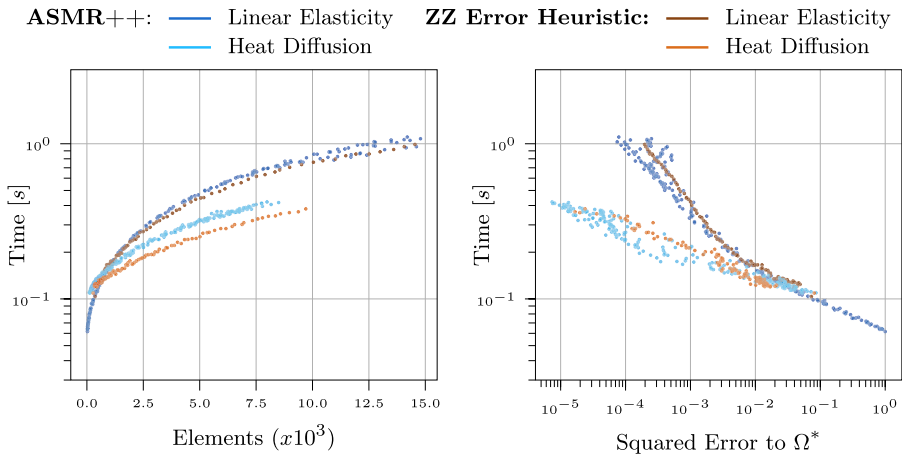


Fig. 14 Wall-clock time comparison of ASMR++ and the *ZZ ErrorHeuristic* for the Linear Elasticity and Heat Diffusion tasks. For the *ZZ ErrorHeuristic*, we measure the time required to solve the system of equations, execute the heuristic, and refine the mesh. For ASMR++, the timing also accounts for generating the observation graph and executing the trained policy. (Left) ASMR++ is slower for the same number of mesh elements due to additional computational steps. (Right) For a given target error, ASMR++ achieves comparable or better runtime on these tasks by requiring fewer mesh elements, thanks to its higher refinement quality

any meshes seen during training, yet accurately captures the interesting parts of the solution. The full refinement procedure is roughly 100 times faster than solving the fine-grained reference Ω^* .

Figure 16 shows the mesh error and inference speed for ASMR++ policies trained on the augmented setup for Poisson's equation when evaluated on domains of different sizes. For the 9×9 domain we used 10 instead of 100 evaluation PDEs due to their increased runtime and memory requirements. While the number of elements for uniform refinements is linear in the domain's volume, meshes created by our method grow less quickly to achieve the same error threshold. These results suggest that there are fewer elements with a significant error in larger domains. Accordingly, evaluating the reference mesh Ω^* quickly becomes expensive as the domain size gets larger, while ASMR++ provides fast and efficient refinements. For large meshes, our method is about 100 times faster than evaluating Ω^* while maintaining a mesh error of less than 0.001. Appendix D.2 shows additional visualizations for adapted training PDE and final refined meshes for different domain sizes.

The generalization capabilities of our method are likely a result of the agent-wise optimization and the utilized MPN, which facilitate refinement strategies that are largely based on local element neighborhoods. Especially the ability of ASMR++ to efficiently up-scale to larger meshes during inference opens up promising applications for practical engineering applications. A policy can be trained on small and cheap environments, and then used in larger and more challenging setups during inference.

4.5.2 Generalization to Different Materials

An essential aspect of AMR strategies is their ability to generalize across a wide range of systems of equations. To evaluate this for ASMR, we introduce varying material parameters to the Heat Diffusion and Linear Elasticity tasks. We define three levels of variability for both tasks, namely *fixed materials* that keep the constant parameters of the main experiments, *small variation* where the material parameters are varied within a narrow range, and *large variation* where the parameters span a wide range. Table 1 summarizes the specific parameter ranges for each task and variation level.

For training and evaluation, we follow the methodology and setup of the main experiments. Material parameters are sampled log-uniformly for diffusivity and Young's modulus, and uniformly for the Poisson ratio. The log of diffusivity is added as an input feature to the heat diffusion ASMR++ policy, while the log of Young's modulus and the Poisson ratio are added for linear elasticity. We evaluate policies trained on each of the three variability levels on all levels, and also compare to the *ZZ Error Heuristic* on all setups. Figure 17 presents the quantitative results. Training ASMR++ on a limited range of parameter variations enables it to generalize effectively to a broader set of out-of-distribution parameters, while still maintaining high refinement accuracy on the fixed material parameters used in the main experiments. This generalization ability comes at no additional computational cost, since the difference between these policies is the variation in material parameters in their training data. Although ASMR++ consistently outperforms the *ZZ Error Heuristic*, refinement quality declines slightly as parameter variability increases. This decline is likely due to the limited training data of only 100 systems of equations, and the comparatively limited capacity of the MPN backbone used for the policy. For even more complex PDEs and generalization settings, we thus expect ASMR++ to require a larger number of training

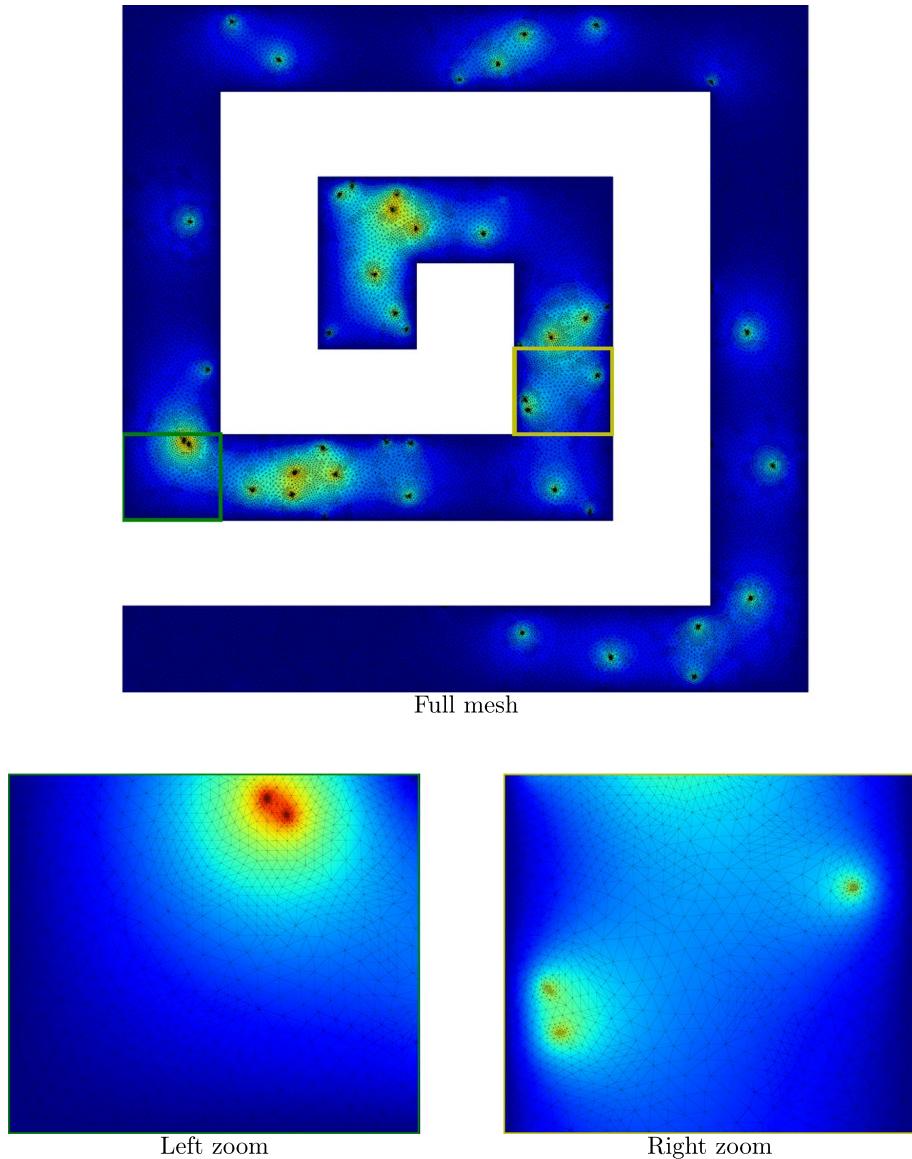


Fig. 15 (Top): Visualization of a final ASMR++ mesh and corresponding solution magnitude on a 20×20 spiral domain. The mesh consists of 52 223 elements and the full refinement procedure takes about 12.2 seconds on a regular CPU. In contrast, computing the uniform refinement Ω^* takes roughly 20 minutes. (Bottom): Close-ups of the full mesh

problems and an increased latent dimensionality of the policy architecture to ensure robust and accurate refinements.

Figure 18 provides final refinements of ASMR++ on exemplary systems of equations for different material parameters. For each task, we use a single policy trained on the “large variation” environments to generate all refinements. We find that training on a range of

Poisson's Equation - Domain Size Generalization

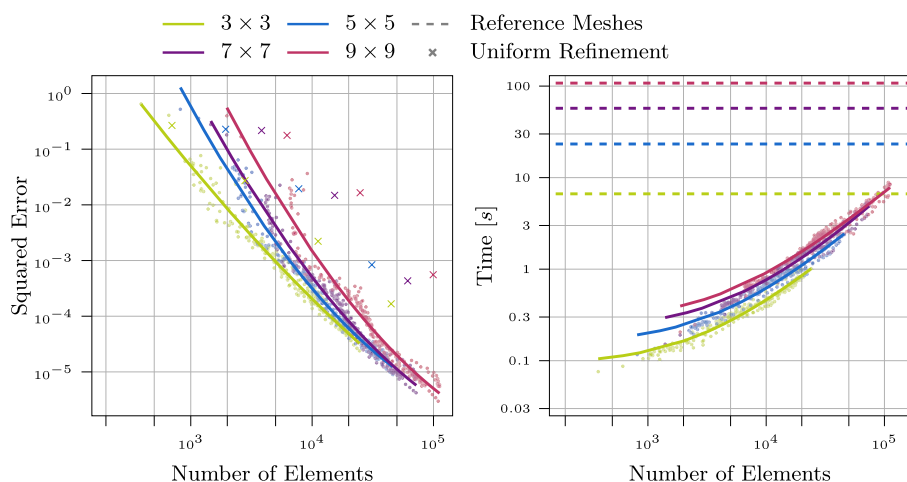


Fig. 16 ASMR++ mesh error and inference speed over different sizes $N \times N$ of generalization domains after training on augmented 1×1 domains. Each evaluated domain size is denoted by a different color. (Left) Pareto plot of squared errors and number of final mesh elements for ASMR++ and uniform refinements. Our method efficiently up-scales to larger domains during inference. (Right) Wallclock-time in seconds of ASMR++ for different numbers of elements compared to the uniform reference Ω^* (dashed lines, color corresponding to domain size). While evaluating the uniform reference quickly becomes expensive, ASMR++ provides efficient and comparatively cheap refinements even for larger domains

Table 1 Parameter ranges for the material generalization experiments by task

Task	Parameter (Unit)	Level	Value
Heat diffusion	Diffusivity (m^2/s)	Fixed materials	$1.0e-3$
		Small variation	$[1.0e-4, 1.0e-2]$
		Large variation	$[1.0e-5, 1.0e-1]$
Linear elasticity	Poisson Ratio	Fixed materials	0.3
		Small variation	$[0.2, 0.4]$
		Large variation	$[0.0, 0.499]$
	Young's Modulus (GPa)	Fixed materials	1.0
		Small variation	$[0.3, 3.0]$
		Large variation	$[0.1, 10.0]$

Parameter units are specified in parentheses. Meshes are normalized to unit length, so physical units correspond to this scale

material parameters allows ASMR++ to provide tailored refinements on a wide range of material parameters. For example, ASMR++ picks up on the high solution gradient in the lower left corner of the L-shape for Polypropylene, while finding that refinement is not necessary for the more evenly deforming Concrete. Additionally, since the model is trained to minimize global error, it may occasionally keep secondary refinements that have a negligible impact on the total computational budget.

5 Conclusion

We introduce Adaptive Swarm Mesh Refinement++ (ASMR++), an adaptive mesh refinement method leveraging swarm reinforcement learning that increases the efficiency of numerical simulations with the finite element method. Our method iteratively evaluates a system of equations on a mesh, generates an observation graph from the mesh and solution,

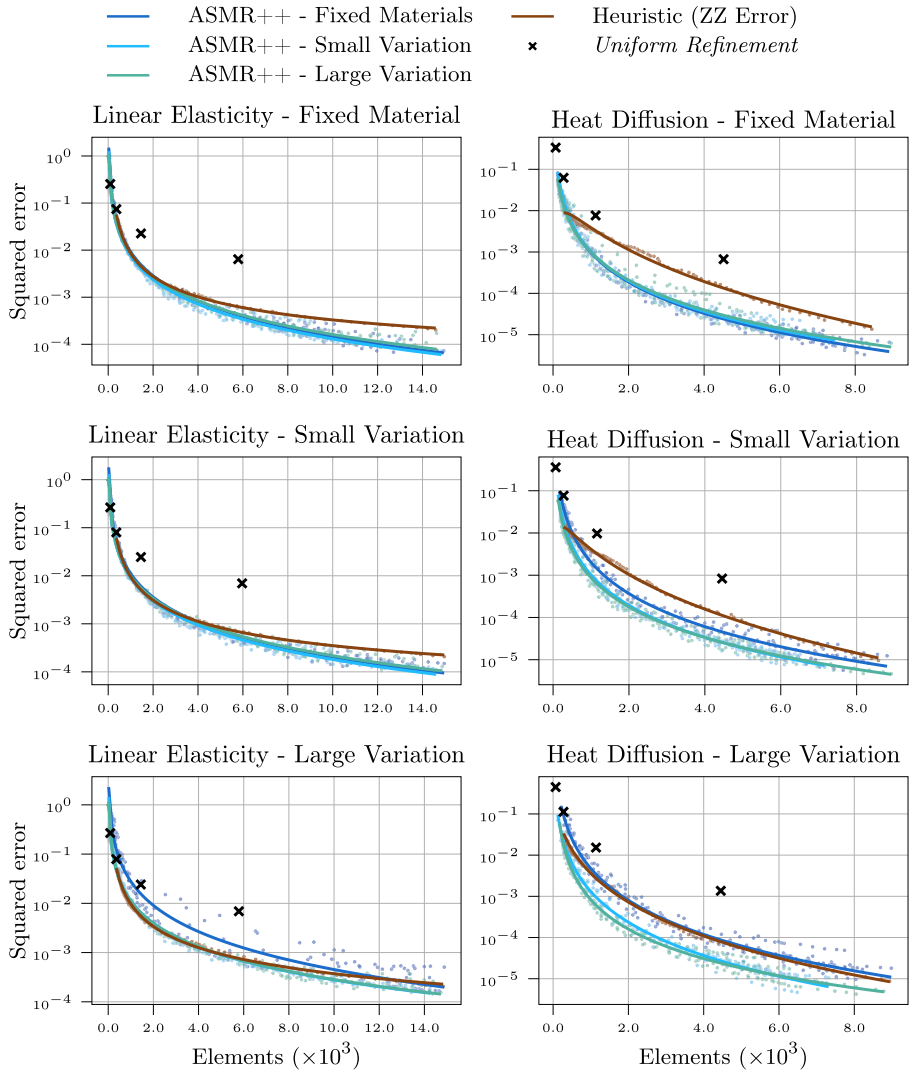


Fig. 17 Pareto plot of normalized squared errors and number of final mesh elements for the (Left) Linear Elasticity and (Right) Heat Diffusion tasks. Each sub-plot compares ASMR++ trained on different material parameters and the *ZZ Error* Heuristic on one set of evaluation environments. More precisely, we evaluate (Top) fixed parameters, (Middle) small variations and (Bottom) large variations in parameters. While ASMR++ trained on fixed parameters struggles to generalize to unseen parameter ranges during evaluation, training on a modest range of parameters enables generalization to unseen setups while maintaining refinement accuracy on the previously seen parameter ranges

processes this graph with a message passing graph neural network, and subsequently marks mesh elements for subdivision. We treat each mesh as a homogeneous agent in a collaborative multi-agent system, training all agents with a shared policy. Crucially, we provide a per-agent reward that balances the local reduction in simulation error against the cost of adding additional mesh elements. This cost can be tuned during inference to generate meshes of different resolution using the same policy. Accommodating different mesh sizes within an episode due to subdivisions, we propose a location-based mapping that assigns each new mesh element to the element in the previous mesh that created it.

Experimental results on stationary, conforming 2d and 3d meshes demonstrate that this agent-wise approach creates high-quality refinements for meshes with thousands of elements and without requiring any expensive error estimate during inference. ASMR++ surpasses both existing reinforcement learning-based methods and traditional refinement techniques, achieving a mesh quality comparable to expensive oracle-based error heuristics. We conduct additional experiments that highlight the benefits of coupling our per-agent reward function with our temporal agent mapping. Further, the method directly generalizes to different process conditions and domains during inference. When trained on a range of material parameters, it generalizes well to new, out-of-distribution material parameters, and when combined with simulation-specific data augmentation during training, it scales to significantly larger domains and meshes during inference. For the considered systems of equations, our method is up to 30 times faster than using a uniform refinement in small domains and can exceed a speed-up of factor 100 when scaling to larger evaluation setups during inference.

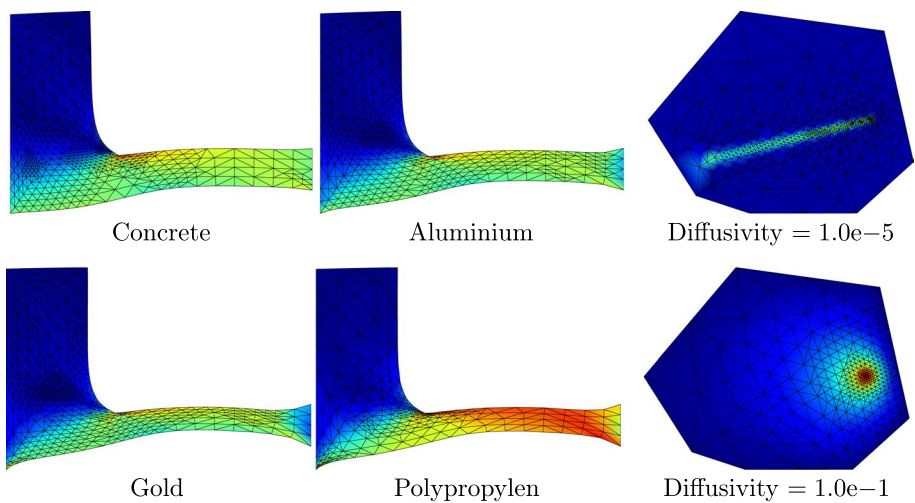


Fig. 18 Final ASMR++ meshes and corresponding FEM solutions for the same Linear Elasticity and Stokes Flow problems with varying material parameters. Left, Middle: Training on systems of equations with randomly sampled Poisson's ratio and Young's modulus enables the same ASMR++ policy to deliver precise refinements for simulations on several different real-world materials, such as concrete (Young's modulus 32 GPa, Poisson ratio 0.15), gold (77.2 GPa, 0.43), aluminum (68 GPa, 0.32), and polypropylene (1.2 GPa, 0.42). Note that the displacements in the figures have been intentionally chosen to be extreme to highlight different refinement behaviors. Right: Similarly, training with diverse diffusivity parameters allows for refinements for diffusivities spanning several orders of magnitude

5.1 Limitations and Future Work

Our method requires solving the system of equations on the mesh following each refinement, which quickly becomes computationally expensive for finer meshes. To mitigate this, we aim to integrate auxiliary physics-based losses or policy distillation methods to implicitly learn to predict which regions to refine from just the problem's process conditions. Additionally, our reward function is currently based on an expensive reference mesh and solution. While this reference is only required during training and ASMR++ generalizes well from as few as 10 different training systems of equations, the maximum mesh resolution is limited by this reference. Another promising direction is to apply ASMR++ to other numerical discretization methods such as the Finite Volume Method. Lastly, we aim to extend our method from stationary to time-dependent refinement strategies and thus include both mesh refinement and coarsening operations, as well as scale ASMR++ to more complex non-linear problems, such as anisotropic diffusion.

Systems of Equations

The following briefly describes the system of equations used for our experiments, including weak PDEs, domains and process conditions. Each task uses an underlying system of equations and varies either the process conditions, the used domain, or both. For some tasks, we add additional information about the task as node features to the observation graph of the MPN policy.

Laplace's Equation

Consider a domain Ω bounded internally by $\partial\Omega_{\text{in}}$ and externally by $\partial\Omega_{\text{out}}$. We seek a function $u(\mathbf{x})$ satisfying Laplace's Equation

$$\int_{\Omega} \nabla u \cdot \nabla v \, d\mathbf{x} = 0 \quad \forall v \in V,$$

where $v(\mathbf{x})$ denotes the test function and the solution $u(\mathbf{x})$ has to satisfy the Dirichlet boundary conditions

$$u(\mathbf{x}) = 0, \mathbf{x} \in \partial\Omega_{\text{in}} \quad \text{and} \quad u(\mathbf{x}) = 1, \mathbf{x} \in \partial\Omega_{\text{out}}.$$

The domain is a unit square $(0, 1)^2$ with a squared hole defined by the inner boundary $\partial\Omega_{\text{in}}$. For each system of equations, the hole's size is uniformly sampled from $U(0.05, 0.25)$ in both directions, while its center is placed randomly in $U(0.2, 0.8)^2$. The minimum distance to $\partial\Omega_{\text{in}}$ of each face midpoint is included as a node feature.

Poisson's Equation

We address Poisson's Equation

$$\int_{\Omega} \nabla u \cdot \nabla v d\mathbf{x} = \int_{\Omega} f v d\mathbf{x} \quad \forall v \in V,$$

where $f(\mathbf{x})$ is the load function, and $v(\mathbf{x})$ is the test function. The solution $u(\mathbf{x})$ must satisfy $u(\mathbf{x}) = 0$ on $\partial\Omega$. We use L-shaped domain Ω , defined as $(0, 1)^2 \setminus (p_0 \times (1, 1))$, where the lower left corner p_0 is sampled from $U(0.2, 0.95)$ in x and y direction.

We employ a Gaussian Mixture Model with three components for each problem's load function. The mean of each component is drawn from $U(0.1, 0.9)^2$, with rejection sampling ensuring all means reside within Ω . We then independently draw diagonal covariances from a log-uniform distribution $\exp(U(\log(0.0001, 0.001)))$ and randomly rotate them to yield a full covariance matrix. Component weights are generated from $\exp(N(0, 1)) + 1$ and subsequently normalized. The load function evaluation at each face's midpoint is used as a node feature.

Stokes Flow

Let $\mathbf{u}(\mathbf{x})$ and $p(\mathbf{x})$ represent the velocity and pressure fields, respectively, in a channel flow. We analyze the Stokes flow, aiming to solve for u and p without a forcing term, i.e.,

$$\nu \int_{\Omega} \nabla \mathbf{v} \cdot \nabla \mathbf{u} d\mathbf{x} - \int_{\Omega} (\nabla \cdot \mathbf{v}) p d\mathbf{x} = 0 \quad \forall \mathbf{v} \in \mathbf{V},$$

$$\int_{\Omega} (\nabla \cdot \mathbf{u}) q d\mathbf{x} = 0 \quad \forall q \in V,$$

with test functions $\mathbf{v}(\mathbf{x})$ and $q(\mathbf{x})$ (Quarтерoni & Quarтерoni, 2009). The velocity field at the inlet is defined as

$$\mathbf{u}(x=0, y) = u_P y(1-y) + \sin(\varphi + 2\pi y).$$

At the outlet, we impose $\nabla \mathbf{u}(x=1, y) = \mathbf{0}$, and assume a no-slip condition $\mathbf{u} = \mathbf{0}$ at all other boundaries. For numerical stability, we utilize P_2/P_1 Taylor-Hood elements, with quadratic velocity and linear pressure shape functions (John et al., 2016).

The inlet profile parameter u_P is sampled from a log-uniform distribution $\exp(U(\log(0.5, 2)))$. The domain is structured as a unit square with three rhomboid holes of length 0.4 and height 0.2. The holes are centered at $y \in \{0.2, 0.5, 0.8\}$ and we randomly sample their x -coordinate from $U(0.3, 0.7)$. We optimize the meshes w.r.t. the velocity vector and present a scalar error as the norm of the vectorized velocity error.

Linear Elasticity

We investigate the steady-state deformation of a solid under stress caused by displacements at the boundary $\partial\Omega$ of the domain Ω . Let $\mathbf{u}(\mathbf{x})$ be the displacement field, $\mathbf{v}(\mathbf{x})$ the test function, and

$$\varepsilon(\mathbf{u}) = \frac{1}{2}(\nabla \mathbf{u} + (\nabla \mathbf{u})^\top)$$

the strain tensor. The linear-elastic and isotropic stress tensor is given as

$$\boldsymbol{\sigma}(\boldsymbol{\varepsilon}) = 2\mu\boldsymbol{\varepsilon} + \lambda\text{tr}(\boldsymbol{\varepsilon})\mathbf{I},$$

using Lamé parameters

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \text{and} \quad \mu = \frac{E}{2(1+\nu)}$$

with a problem specific Young's modulus $E = 1$ and Poisson ratio $\nu = 0.3$. Without body forces, the problem is given as (Zienkiewicz et al., 2005)

$$\int_{\Omega} \boldsymbol{\sigma}(\boldsymbol{\varepsilon}(\mathbf{u})) : \boldsymbol{\varepsilon}(\mathbf{v}) \, d\mathbf{x} = 0 \quad \forall \mathbf{v} \in \mathbf{V}.$$

We use the same class of L-shaped domains as in the Poisson problem in Sect. A.2. We fix the displacement of the left boundary to 0, i.e., $\mathbf{u}(x=0, y) = 0$, and randomly sample a displacement direction from $U[0, \pi]$ and magnitude from $U(0.2, 0.8)$ to displace the right boundary as $\mathbf{u}(x=1, y) = \mathbf{u}_P$. The stress $\boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{0}$ is zero normal to the boundary at both the top and bottom of the part. The task-dependent displacement vector $\mathbf{u}_P$ is added as a globally shared node feature to all elements. Our objective includes the norm of the displacement field $\mathbf{u}$ and the resulting Von-Mises stress, using an equal weighting between the two.

Non-stationary Heat Diffusion

Using temperature $u(\mathbf{x})$, a test function $v(\mathbf{x})$, and a thermal diffusivity constant $a = 0.001$, we address a non-stationary thermal diffusion problem defined by

$$\int_{\Omega} \frac{\partial u}{\partial t} \, d\mathbf{x} + \int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} \quad \forall v \in V.$$

Here, $f(\mathbf{x}, y)$ is a position-dependent heat distribution

$$f(\mathbf{x}, y) = 1000 \exp(-100((x - x_p(\tau)) + (y - y_p(\tau)))).$$

The position of the heat source's path $\mathbf{p}_{\tau}(\tau) = (x_p(\tau), y_p(\tau))$ is linearly interpolated over time as

$$\mathbf{p}_{\tau} = \mathbf{p}_0 + \frac{\tau}{\tau_{\max}}(\mathbf{p}_{\tau_{\max}} - \mathbf{p}_0).$$

The start and goal positions $\mathbf{p}_0$ and $\mathbf{p}_{\tau_{\max}}$ are randomly drawn from the entire domain. The temperature $u \in \partial\Omega$ is set to zero on all boundaries. We use a total of $\tau_{\max} = 20$ uniform time steps in $\{0.5, \dots, 10\}$ and an implicit Euler method for the time-integration. Domain geometry is derived from 10 equidistant points on a circle centered at $(0.5, 0.5)$ with a

radius of 0.4. Each point is randomly distorted by a value $U(-0.2, 0.2)^2$ before we proceed to calculate the points convex hull as our domain. We finally normalize the convex hull to lie in $(0, 1)^2$. The result is a family of convex polygons with up to 10 vertices. We provide and measure the error and solution of the final simulation step, and provide the distance to the start and end position of the heat source as additional node features for each element.

Neumann Boundaries

We adapt Poisson's Equation in Sect. A.2 to include Neumann boundary conditions. Using star-shaped domains, we randomly select half the line segments $\partial\Omega_l$ that make up the boundary of the star as Neumann boundaries, and use zero Dirichlet boundary conditions for the remaining segments. Each selected line segment is assigned a random sinusoidal Neumann boundary condition

$$\nabla u \cdot \mathbf{n} = 10 \cdot A \cdot p_x \cdot (1 - p_x) \cdot \sin(\pi \nu p_x), \quad \mathbf{x} \in \partial\Omega_l,$$

where $\mathbf{n}(\mathbf{x})$ denotes the normal vector of the boundary segment, $p_x \in [0, 1]$ is the normalized position along the line segment, $A \in U(1, 3)$ the amplitude and $\nu \in \{3, 5, 7\}$ the frequency.

To create the star-shaped domains, we define an outer radius of 0.5 and an inner radius of 0.2 from a point centered at (0.5, 0.5). We then sample a number of star points in $U(3, 5)$ and equidistantly interleave them between the two boundaries. We randomly distort each point by $U(-0.05, 0.05)^2$ before applying a random rotation and normalizing the full domain to be within $(0, 1)^2$. Initial meshes are created with a target volume of 0.02 instead of the 0.05 used for the other tasks to compensate for the lower total volume of the star-shaped domains, and the diagonal covariances for the Gaussian mixture model load function are drawn from $U(0.00005, 0.0005)$. Each element gets the distance to the closest Neumann and Dirichlet boundaries as additional node features.

3-Dimensional Poisson's Equation

Finally, we extend Poisson's Equation in Sect. A.2 to a 3-dimensional domain. We use a rectangular plate of length $l_x = 1$, width $l_y = 0.5$ and height $l_z = 0.1$ as our domain, and create an initial mesh from by dividing it into 0.1^3 cubes and triangulating each cube. We assign Dirichlet boundary conditions

$$u(x = 0, y, z) = u(x = 1, y, z) = u(x, y = 0, z) = u(x = 1, y = 0.6, z) = 0$$

on the sides of the plate and define natural boundary conditions on the lower $z = 0$ and upper side $z = 0.1$ of the plate. We use tetrahedral elements for this task, allowing for a fully volumetric mesh, and solve with an iterative instead of a direct solver to accommodate the larger number of elements. The remesher uses longest edge bisection (Rivara, 1984; Suárez et al., 2005) to halve marked elements, whereas the 2-dimensional domains employ the red-green-blue refinement method (Carstensen, 2004).

Further Experiments

Experiment Setup

We repeat each experiment for 10 different random seeds, randomizing the parameters of the neural network as well as the PDEs, domains and process conditions. We normalize all domains to unit size unless mentioned otherwise. We create initial meshes with meshpy² with a target element size of 0.05 for all tasks except for the Neumann boundary task, which uses a target element size of 0.02. For all experiments, we terminate an episode with a large negative reward of $-1\,000$ when a threshold of 20 000 elements, respectively 50 000 elements for the 3d Poisson task, is exceeded. Unless mentioned otherwise, we train all policies on 100 randomized systems of equations, but evaluate on the same set of 100 fixed systems of equations across different methods and seeds for better comparability. All experiments are run for up to 3 days on 8 cores of an Intel Xeon Platinum 8358 CPU, but usually terminate within a single day.

We train a total of 5 RL-AMR methods on 7 separate tasks for 10 repetitions each. Since 3 of these methods train on fixed number of target mesh elements, this leads to $(2 + (3 \cdot 10)) \cdot 7 \cdot 10 = 2240$ core experiments. A similar combined number of experiments is needed for the heuristics, ablations and preliminary experiments.

Volume Reward

Eq. 6 rewards the reduction in maximum element error for a given refinement. While this has been proposed as an ablation by *ASMR* (Freymuth et al., 2023), it instead uses a different reward for all experiments. This reward considers a numerically integrated error per element, i.e., replaces Eq. 5 with

$$\text{err}(\Omega_i^t) \approx \sum_{\Omega_m^* \subseteq \Omega_i^t} \text{Volume}(\Omega_m^*) |u_{\Omega^*}(p_{\Omega_m^*}) - u_{\Omega^t}(p_{\Omega_m^*})|$$

and changes Eq. 6 to a volume-scaled reduction in this integrated error, i.e.,

$$r(\Omega_i^t) := \frac{1}{\text{Volume}(\Omega_i^t)} \left(\text{err}(\Omega_i^t) - \sum_j M_{ij}^t \text{err}(\Omega_j^{t+1}) \right) - \alpha \left(\sum_j M_{ij}^t - 1 \right). \quad (\text{B.1})$$

Both reward formulations evaluates if for any refinement, the reduction in error is greater than the cost of adding additional elements. The volume scaling component in Eq. B.1 encourages the policy to give priority to refining smaller mesh elements, roughly canceling out the volume of the integration points in Eq. 5, aiming to reduce the average error per unit volume for each element. Optimizing this reward lowers the errors for each mesh element, adjusted by the inverse of its volume, which promotes a uniform distribution of error in rela-

²<https://github.com/inducer/meshpy>

tion to the size of each element across the mesh. Opposed to this, the maximum reward of Eq. 6 minimizes the maximum error across any integration point in the mesh, posing a more direct and clear objective. We compare to this reward in Sect. C.3 and show the difference in performance in Fig. 9, finding that both variants perform similarly.

Extended Results

The following sections evaluate various algorithmic design choices for the different RL-AMR methods and the *ZZ Error* heuristic. In particular, we evaluate different RL backbones and GNN architectures to ensure a fair comparison between the different approaches. We also conduct experiments on various small changes in the setup and training of ASMR++ to find and explain what makes it effective. All experiments are conducted on the Stokes flow task unless mentioned otherwise, as the task requires challenging refinements at various parts of the domain.

Initial Meshes for the ZZ Error Heuristic

Figure 19 demonstrates the performance of the *ZZ Error Heuristic* applied to the initial mesh, in comparison to its performance when the process begins with either one or two uniform mesh refinements. The results indicate that initial uniform refinements significantly improve the heuristic's effectiveness, likely because the uniform refinements make it easier to identify gradients in key areas, which may be overlooked if the mesh is too coarse. Consequently we opt for the twice-refined mesh approach for all experiments. In comparison, the RL-AMR methods do not need to tune their initial mesh, as the methods are trained to produce optimal refinement sequences.

Proximal Policy Optimization and Deep Q-Networks

Figure 20 shows results on the Stokes flow task for Proximal Policy Optimization (PPO) (Schulman et al., 2017) and Deep Q-Network (DQN) (Mnih et al., 2013, 2015) as the RL backbone for all RL-AMR algorithms. For the PPO version of the *VDGN*-like baseline, we apply the value decomposition (Sunehag et al., 2017) to the value function instead of the *Q*-function, i.e., we define the value function of the full mesh as the sum of value functions of the individual mesh elements. Further, we use a mean instead of a sum for the agent mapping of the TD error in Eq. 3 for training the *Q*-value of the DQN experiments of ASMR++, as this experimentally increases training stability. PPO generally outperforms DQN, suggesting that on-policy optimization is favorable for the changing observation and action spaces of AMR. We thus choose PPO for all further experiments as it leads to better performance for all methods.

Design Choices and Ablations

Graph Attention Networks

Figure 21 compares the MPNs of Sect. 2.3 with the Graph Attention Network (GAT) architecture proposed by *VDGN* (Yang et al., 2023). The MPN slightly improves over the GAT for both methods, leading us to choose this architecture for all further experiments.

Element Penalty

Figure 22 shows the effect of different ranges for the element penalty α for ASMR++ and the *VDGN*-like baseline. Here, we scale the minimum and maximum sampled penalty by a constant factor during training and inference. While the *VDGN*-like baseline works well for the chosen penalty, it is generally unstable and fails to produce consistent refinements for finer meshes, regardless of whether the range of α is scaled by a factor of 2. In comparison, ASMR++ produces high-quality meshes for different element penalty ranges. When decreasing all penalties by a factor of 10, ASMR++ produces finer meshes with significantly more elements compared to the regular penalty. Similarly, increasing the penalty makes refinements less attractive, reducing the number of produced elements.

Number of Training PDEs

We generally train ASMR++ on 100 systems of equations, each of which consists of a randomly sampled domain and random process conditions. Figure 23 explores how fewer or more systems of equations during training affect performance, finding that performance improves for up to 100 points of data. Beyond 100 training systems of equations, performance gains diminish. During training, the reward function requires solving a fine-grained reference mesh Ω^* for each PDEs, which is computationally expensive. We thus use 100

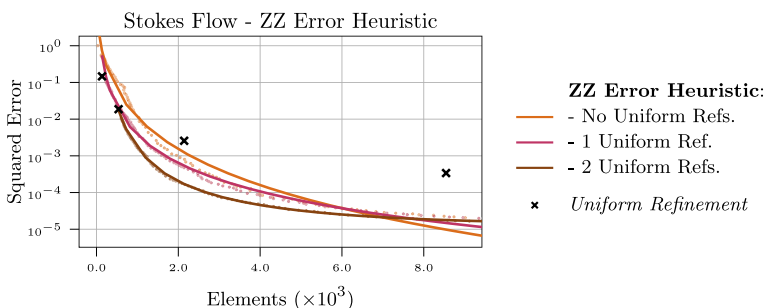


Fig. 19 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for the Zienkiewicz-Zhu Error Estimator (*ZZ Error*) Heuristic for 0, 1 and 2 initial uniform mesh refinements. A finer initial mesh leads to improved performance for the *ZZ ErrorHeuristic* but prevents the creation of very coarse meshes, indicating that the initial element size should be tuned for the *ZZ ErrorHeuristic*

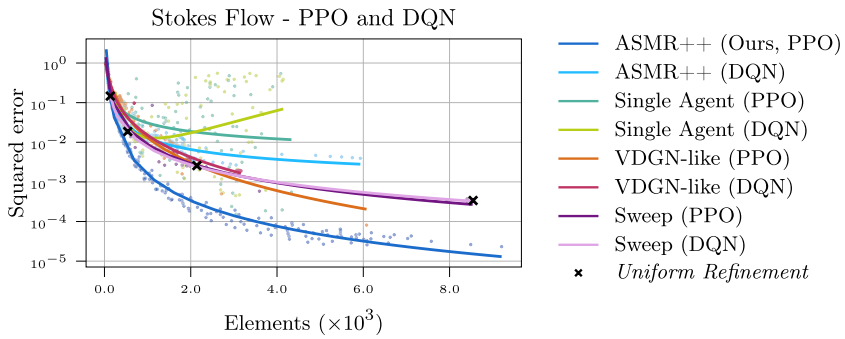


Fig. 20 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for the different RL-AMR methods using PPO and DQN backends. The on-policy PPO outperforms the off-policy DQN for all methods. Only ASMR++ consistently performs better than a uniform mesh

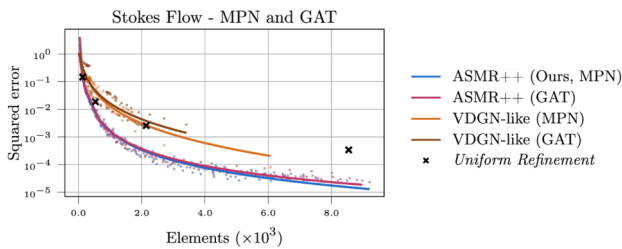


Fig. 21 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for ASMR++ and the *VDGN*-like baseline for MPN and GAT network architectures. For both methods, the MPN shows slightly better performance, though the impact of the architecture is comparatively minor

training systems of equations for our main experiments. Albeit noisy and inconsistent, ASMR++ performs better than uniform refinements on the unseen evaluation set when trained on a single system of equations, likely due to the mostly local optimization objective and network architecture.

Network Architecture

We ablate our choice of network architecture on the right side of Fig. 24. Using Graph Attention Networks (Velickovic et al., 2018) with edge features instead of our MPNs does not significantly impact performance. ASMR++ uses Edge Dropout (Rong et al., 2019) of 0.1, i.e., it randomly removes every tenth edge of the observation graph, regularizing the training data and incentivizing the model to learn general patterns instead of spurious correlations. Omitting this dropout leads to slightly worse refinements.

Mean and Maximum Error Metrics

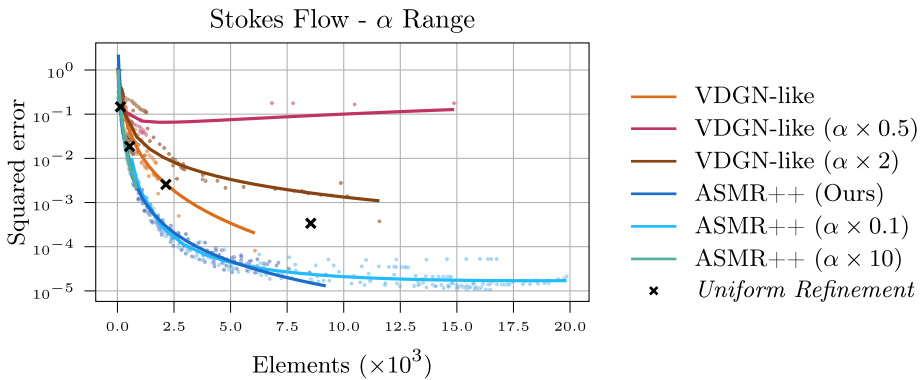


Fig. 22 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different ranges of the element penalty α for ASMR++ and the *VDGN*-like baseline. Increasing or decreasing the element penalty of ASMR++ by a factor of 10 leads to finer meshes for larger penalties and vice versa. Regardless of the scale of the element penalty, ASMR++ produces high-quality meshes. The *VDGN*-like baseline is comparatively unstable, yielding worse results when adapting the element penalty

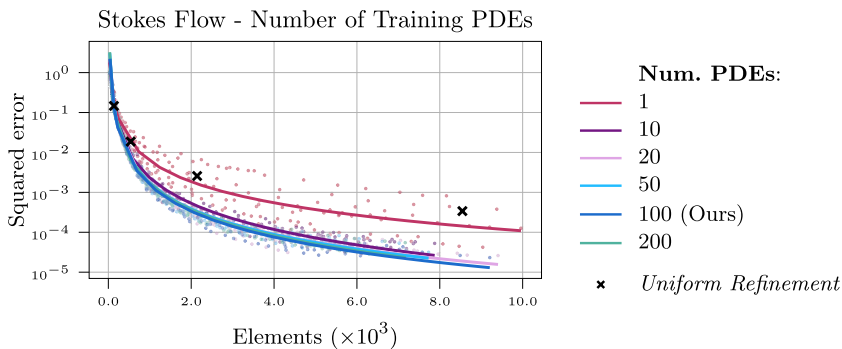


Fig. 23 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different numbers of training systems of equations. ASMR++ performs better than uniform on novel evaluation settings when trained on a single system of equations, and improves for up to 100 systems of equations in the training set

In addition to the squared error evaluated in Sect. 4, we present results for a mean and maximum mesh error metric. We approximate the maximum error as the average of the Top 0.1 % of errors of all integration points $p_{\Omega_m^*}$ to make it robust to outliers. We normalize both metrics by the error of the initial mesh for each PDE.

Figure 25 shows pareto plots for all tasks for a normalized mean mesh error. The *Oracle Error Heuristic* outperforms its maximum error counterpart, likely because it selects elements with a high integrated error rather than elements with a high maximum error for refinement, thus targeting areas with a high mean error. ASMR++ performs on par with or better than *ASMR* on average, even though it explicitly optimizes a decrease in maximum mesh error rather than a decrease in its mean error, likely due to its local optimization objective. Since the mean mesh error is less sensitive to outliers, more uniform refinements such as those produced by *Sweep* perform better than on other metrics. The mean error directly

quantifies the difference between the solution calculated on the fine ground truth mesh and that produced by the different methods. For most tasks, ASMR++ produces meshes that reduce the error between the initial mesh and the ground truth by more than 99 % with only a few thousand elements.

Figure 26 visualizes pareto plots for all tasks for the Top 0.1 % mesh error. Here, the *Maximum Oracle Error Heuristic* excels when compared to the regular *Oracle Error Heuristic*, likely because it refines elements that have the highest maximum error. Similarly, ASMR++ outperforms *ASMR* on all tasks, as it directly minimizes the maximum error instead of a scaled version of the average mesh error. Methods that produce relatively uniform meshes, such as *Sweep*, perform considerable worse on this approximated maximum error than on the other metrics, likely because uniform refinements produce a lot of elements on mesh areas that do not participate in the maximum mesh error. Notably, the maximum error for the 3d Poisson's equation is bound at around 0.1. This is likely the case because we use longest edge bisection for the adaptive refinements, which produces refined meshes with local elements that significantly differ from that of the fine-grained uniform ground truth reference. As bounding the maximum prediction error is important in many applications, these uniform meshes can potentially waste a lot of computational resources or lead to worse results than the highly adaptive meshes produced by, e.g. ASMR++. Overall, both metrics are consistent with the results in Figs. 7 and 8.

Inference-Time Generalization

Same-Size Generalization Capabilities

Figure 27 visualizes final ASMR++ meshes on Poisson's equation on different 1×1 domains and load functions that are not seen during training. We experiment on the 5 different domain classes that are used for the different tasks, plus a rectangular $\Omega = (0, 1)^2$ domain. For each class, we randomly sample 3 domains and apply a Gaussian mixture

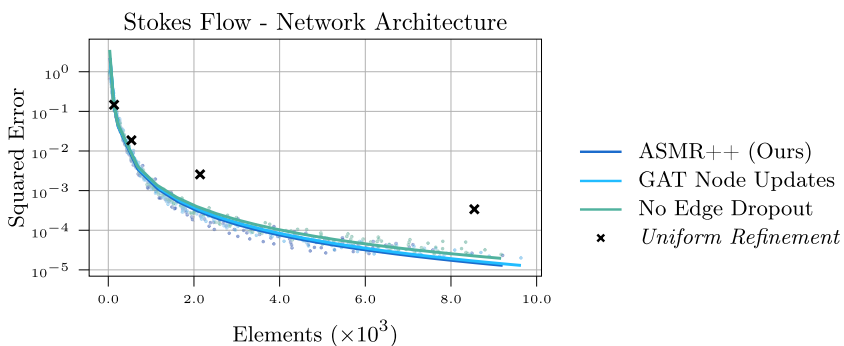


Fig. 24 Pareto plot of normalized squared errors and number of final mesh elements on the Stokes flow task for different network architectures. Performing message passing updates with Graph Attention Networks with edge features instead of MPNs does not significantly impact performance. Omitting Edge Dropout leads to slightly worse refinements

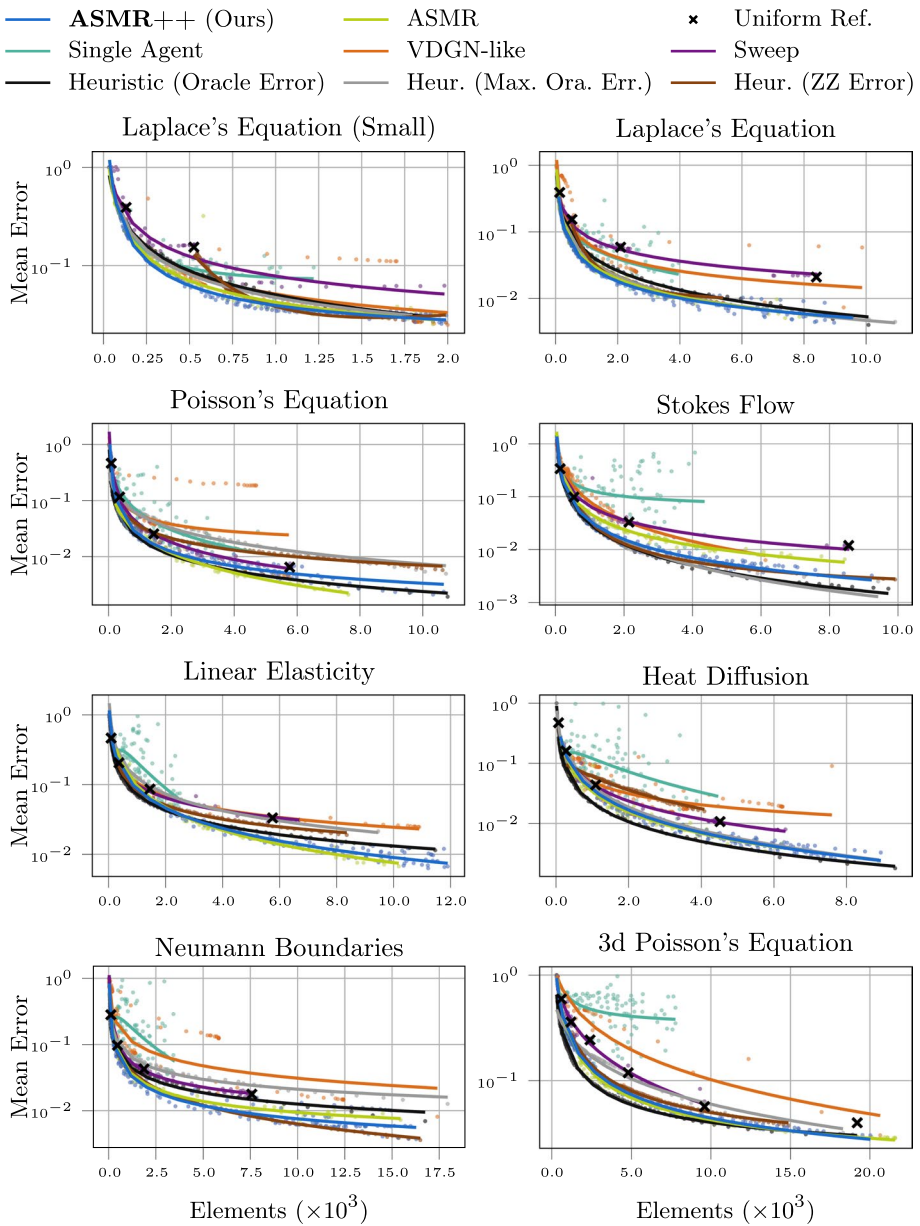


Fig. 25 Pareto plot of normalized mean errors and number of final mesh elements for all tasks. ASMR++ performs on par with or better than ASMR on all tasks, and both methods significantly outperform all RL-AMR baselines

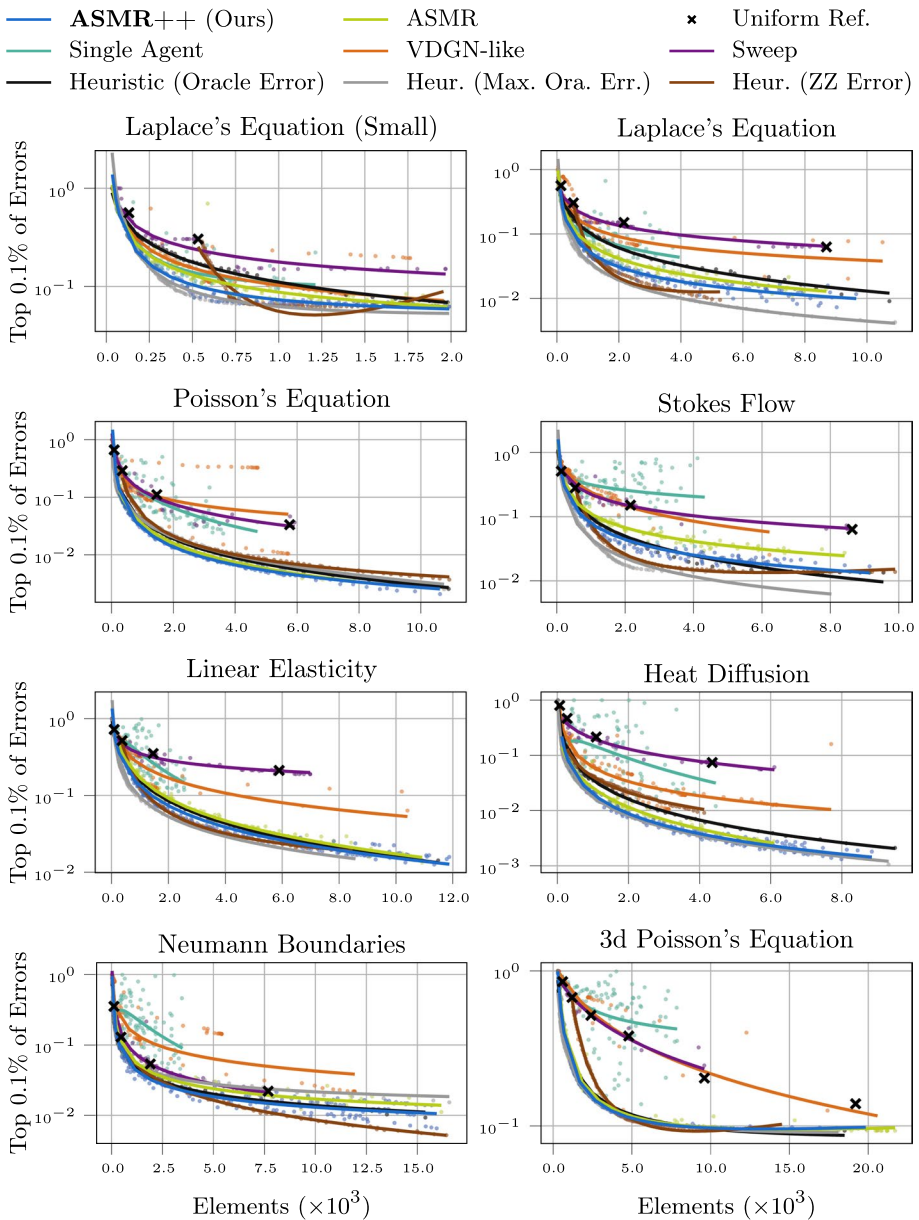


Fig. 26 Pareto plot of normalized maximum element error and number of final mesh elements for all tasks. ASMR++ directly minimizes the error of the maximum element during training. In contrast, ASMR optimizes a scaled variant of the average error. Consequently, ASMR++ surpasses ASMR on all tasks. Both methods outperform all other RL-AMR baselines

model load function with 1, 3, and 5 components, respectively. We find that our approach generalizes well to novel domains and load functions, likely due to the strong local inductive bias of the Swarm RL objective and the MPN policy.

Domain Size Up-Scaling

We additionally experiment with the abilities of *ASMR* to up-scale to larger domains during inference. This up-scaling is crucial for practical scenarios, as it allows applying policies that are trained on small and comparatively cheap domains to apply to complex, large-scale domains where training would otherwise be too expensive.

Using Poisson's equation as an example, we augment the training PDEs to cover a larger space of potential simulations and mimic larger mesh segments by altering boundary conditions and load functions. We train our policies on the square hole domains of, e.g., Fig. 27. We sample the means of the load function from a centered Gaussian with a standard deviation of 0.2, allowing components outside the mesh, and additionally apply random Gaussian loads to selected boundary edges to emulate a larger domain outside of the simulated mesh. As these changes lead to training PDEs of varying complexity, we use $10 \cdot \text{err}(\Omega)$ instead of $\text{err}(\Omega)$ in Eq. 6, i.e., we omit the per-domain normalization, and use an unlimited number of training PDEs to further increase variety. These changes only affect the training environments, and do not infer with our training algorithm. Figure 28 visualizes exemplary training PDEs and policy refinements.

Figure 29 compares *ASMR++* trained on these augmented training PDEs with the setup used throughout the paper. The results show that the augmented training setup leads to a slight decrease in performance relative to *ASMR++* trained under standard conditions. This decrease in performance is likely caused by the more challenging optimization problem. Yet, the augmented training yields high-quality refinements.

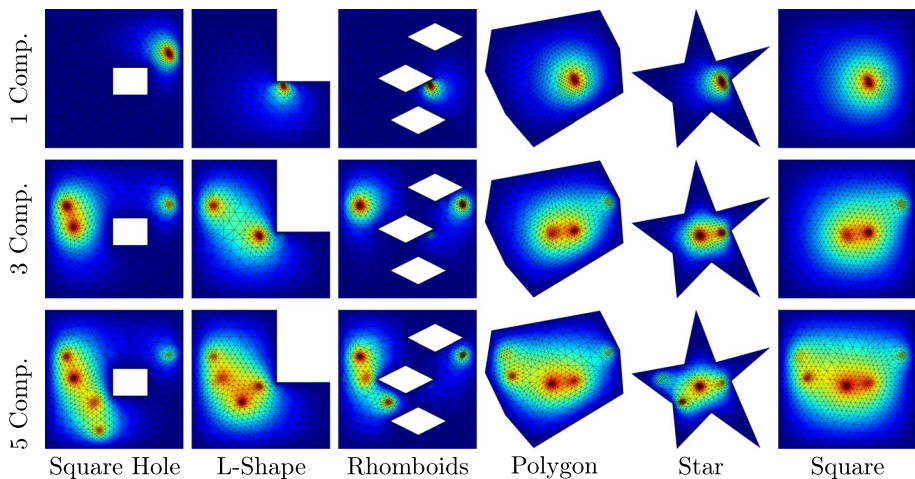


Fig. 27 *ASMR++* refinements for 6 different domains in $(0, 1)^2$ and Poisson's equation with a Gaussian mixture model load function with 1, 3 and 5 components. Even though the policy is only trained on L-shaped domains with 3 components, it easily generalizes to different domains and more components in the applied load function

Once trained, we evaluate the policy on spiral-shaped domains of increasing sizes, ensuring the volume of the initial mesh elements remains constant. To align the complexity of the load function with the enlarged mesh sizes, we adapt the number of Gaussian mixture model modes, using more modes for larger domains. Specifically, we use 16 components which are first placed on a uniform grid and then randomly perturbed, allowing for them to be positioned outside of the mesh. Figure 30 shows an exemplary ASMR++ refinement for a 5×5 domain. Figure 15 shows a final refined mesh on a 20×20 domain with 100 components, including detailed views of specific sections. Both figures use the same policy with a fixed element penalty.

Hyperparameters

General Hyperparameters

The following section lists important hyperparameters used for the training of all methods. Hyperparameters are kept consistent across all tasks and RL-AMR methods unless mentioned otherwise.

Neural Networks

We implement all neural networks in PyTorch (Paszke et al., 2019) and optimize them using ADAM (Kingma & Ba, 2014) with a learning rate of $3.0e-4$ and batch size of 32 unless mentioned otherwise. All MLPs use 2 hidden layers and a latent dimension of 64. Each MPN consists of 2 message passing steps, where each update function is represented as an MLP with *LeakyReLU* activation functions. We apply Layer Normalization (Ba et al., 2016) and Residual Connections (He et al., 2016) independently after each node and edge feature update, and use Edge Dropout (Rong et al., 2019) of 0.1 during training. The edge feature aggregations $\oplus$ are mean aggregations. We use separate architectures for the policy and the value function, i.e., we do not share weights between the policy and value function. The policy and value function heads are MLPs with *tanh* activation functions acting on the final latent node features of the MPN.

PPO

For PPO, we follow previous work (Andrychowicz et al., 2021) to select important hyperparameters and code-level optimizations. We train each PPO policy for a total of 400 iterations, except for the 3 d Poisson task, where we only use 200 iterations. In each iteration, the

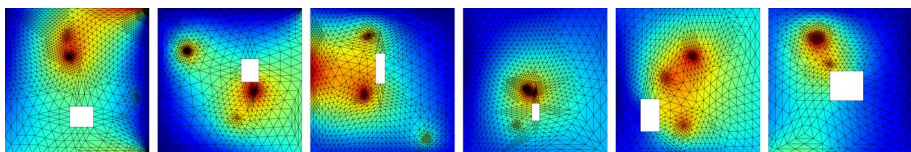


Fig. 28 Random augmented training PDEs and ASMR++ refinements for the domain size up-scaling experiments on Poisson's Equation

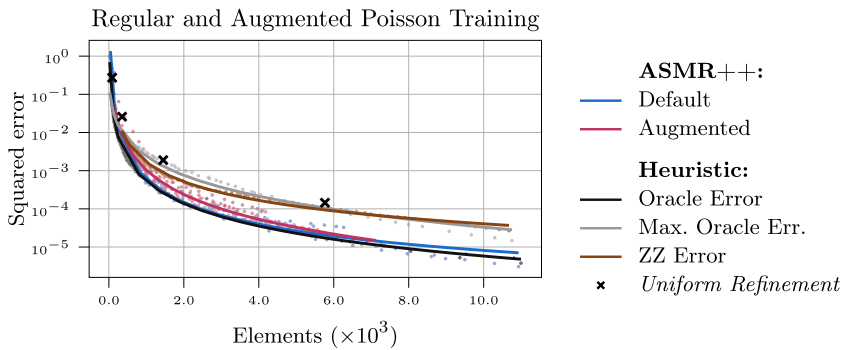


Fig. 29 Pareto plot of the normalized squared error for ASMR++ trained on regular and augmented environments evaluated on the regular evaluation environments. Augmenting the training environments facilitates inference-time up-scaling at the cost of slightly decreased refinement quality

algorithm samples 256 environment transitions and performs 5 epochs of optimization. The value function loss is multiplied with a factor of 0.5. We clip the gradient norm to 0.5 and choose clip ranges of 0.2 for the policy and value function. We normalize the observations with a running mean and standard deviation. Advantages are estimated via Generalized Advantage Estimate (Schulman et al., 2015) using $\lambda = 0.95$. We compute an agent's advantage by subtracting the agent-wise value estimates from the return in Eq. 7.

DQN

For DQN-based approaches, we draw 500 initial samples with a random policy for the replay buffer and then train for $24 * 400 = 9600$ steps, where each step consists of executing and storing an environment transition and then drawing a random mini-batch with replacement from the buffer for a single gradient update. We experimented with more training steps in preliminary experiments, finding that they do no significantly improve performance or stabilize training, but may lead to increased memory footprint and longer runtimes. We update the target networks using Polyak averaging at a rate of 0.99 per step. We select training actions using a Boltzmann distribution over the predicted Q-values per agent, where we linearly decrease the temperature of the distribution from 1 to 0.01 in the first half of training. This action selection strategy favors more correlated actions when compared to an epsilon greedy action sampling, which empirically stabilizes the training for our problem setting of iterative mesh refinement. Further, we follow previous work (Hessel et al., 2018) and combine a number of common improvements for DQNs, namely double Q-learning (Van Hasselt et al., 2016), dueling Q-networks (Wang et al., 2016) and prioritized experience replay (Schaul et al., 2016).

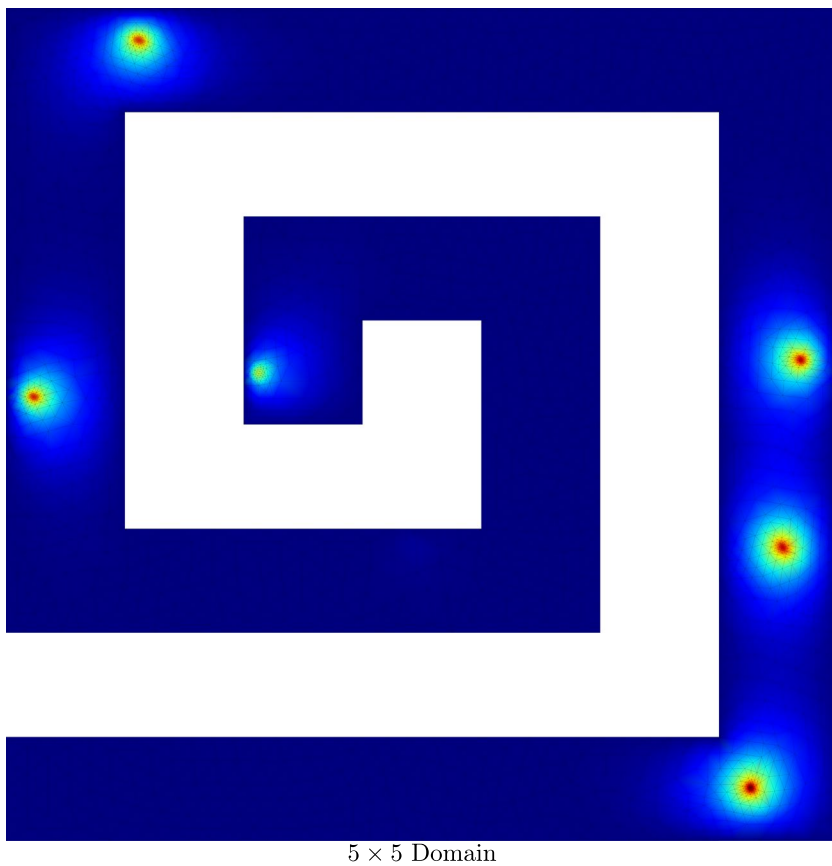


Fig. 30 Visualization of a final ASMR++ mesh on a 5×5 spiral domain with a randomly sampled load function. Our approach consistently provides high-quality refinements for both larger domains and more complex load functions during inference

Baseline-Specific Parameters

Single Agent

We use a maximum refinement depth of 10 refinements per element to avoid numerical instabilities during simulation, skipping actions that try to refine elements that have been refined too often. We consider environment sequences of up to $T = 400$ steps since the method marks only one element at a time, but train a separate policy for every value of T .

Sweep

The *Sweep* baseline moves a single agent to a random mesh element after each training step. The agent then decides if this element should be refined or not. We follow the proposed hyperparameters and have training rollouts of 200 steps (Foucart et al., 2023). Since *Sweep*

uses a local agent living on a single mesh element, we use an MLP instead of the MPN for both the policy and value function. Correspondingly, we adapt the input features, using our regular node features and the global resource budget proposed by the authors. We additionally add aggregated neighborhood information in the form of a mean solution and area of the element's neighbors and the average distance to them. The global budget is controlled via a maximum number of elements $N_{\max}$, allowing to train policies that produce refinements of different granularity. We increase the number of environment transitions per PPO step to 512, and the number of DQN steps to $96 * 400 = 38400$. This change is intended to compensate for decreased number of refined elements of each environment step while roughly equating for the computation time of the other methods.

VDGN-Like

We use a learning rate of $1.0\text{e-}5$ instead of $3.0\text{e-}4$ for the DQN variant of the *VDGN*-like baseline to stabilize its training.

ASMR

We use the reward of Eq. B.1, an infinite horizon RL setting with $\gamma = 0.99$ and omit the edge dropout in the observation graph for *ASMR*, and additionally use an agent mapping without the normalization factor $\frac{|\Omega^t|}{|\Omega^{t+1}|}$ in Eq. 4.

Mesh Resolution Parameters

All AMR methods in this work allow for some control over the desired refinement level of the final mesh. *ASMR++*, *ASMR* and *VDGN* use an element penalty α that trades off the cost of adding new elements with the benefit of a refinement. *Sweep* considers an element budget $N_{\max}$, attempting to minimize the error of the mesh while staying within this budget. *Single Agent* varies the number of rollout steps T , refining once for every step.

Both *ASMR++* and *VDGN* train on a range of α values and condition the policy on it, allowing for adaptive mesh resolutions during inference. Here, we evaluate 20 penalties that are log-uniformly sampled over the training penalty range to cover different final mesh resolutions. As the other methods do not directly support an adaptive penalty during inference, we train 10 separate policies with 10 repetitions each for different refinement parameters. The *Oracle*, *Maximum Oracle*, and *ZZ Error Heuristics* refine based on estimated errors per mesh element. Here, we evaluate 100 values for the error threshold θ , which specifies which elements to refine based on the ratio between their error and the maximum element error. Tables 2 and 3 list the minimum and maximum mesh resolution parameters for all tasks and RL-AMR methods and heuristics, respectively.

Table 2 Ranges for the different refinement hyperparameters for all task and RL-AMR methods

Parameter	ASMR++ α	ASMR α	VDGN-like α	Sweep $N_{\max}$	Single agent T
Laplace (Small)	[0.001, 0.1]	[0.01, 0.3]	[1e-5, 5e-2]	[50, 1000]	[5, 100]
Laplace	[0.001, 0.03]	[0.01, 0.5]	[1e-5, 1e-2]	[200, 3000]	[25, 400]
Poisson	[0.0001, 0.03]	[0.002, 0.1]	[2e-5, 5e-2]	[200, 3000]	[25, 400]
Stokes flow	[0.0005, 0.05]	[0.015, 0.3]	[3e-4, 2e-2]	[150, 2500]	[25, 400]
Linear elasticity	[0.00015, 0.03]	[0.01, 0.15]	[1e-5, 1e-2]	[500, 6000]	[25, 400]
Heat diffusion	[0.00005, 0.01]	[0.003, 0.3]	[1e-5, 1e-2]	[400, 5000]	[25, 400]
Neumann boundaries	[0.00005, 0.03]	[0.0075, 0.5]	[1e-5, 1e-1]	[200, 3000]	[25, 400]
Poisson 3d	[0.0001, 0.1]	[0.1, 20.0]	[1e-5, 1e-1]	[500, 5000]	[25, 400]

ASMR++, ASMR and VDG N apply a penalty α for each added element. Sweep uses an element budget $N_{\max}$. Single Agent varies the number of rollout steps T

Table 3 Ranges for the threshold θ of the error-based refinement strategy of the heuristic baselines

Parameter	Oracle error θ	Max. oracle err. θ	ZZ error θ
Laplace (Small)	[0.22, 1.0]	[0.16, 1.0]	[0.002, 1.0]
Laplace	[0.20, 1.0]	[0.12, 1.0]	[0.001, 1.0]
Poisson	[0.12, 1.0]	[0.15, 1.0]	[0.002, 1.0]
Stokes flow	[0.12, 1.0]	[0.04, 1.0]	[0.001, 1.0]
Linear elasticity	[0.06, 1.0]	[0.02, 1.0]	[0.001, 1.0]
Heat diffusion	[0.04, 1.0]	[0.02, 1.0]	[0.001, 1.0]
Neumann boundaries	[0.14, 1.0]	[0.24, 1.0]	[0.001, 1.0]
Poisson 3d	[0.02, 1.0]	[0.02, 1.0]	[0.001, 1.0]

All ranges are chosen to facilitate direct comparison to the RL-AMR methods in the quantitative evaluation, i.e., to produce meshes with a comparable number of elements

Baseline Visualizations

RL-AMR Visualizations

We visualize the final meshes and corresponding PDE solutions of all RL-AMR methods on the Stokes Flow task for 6 different refinement granularities on the same randomly selected system of equations. For ASMR++ and VDG N , which condition the policy on the element penalty parameter, we use the same trained model for the 6 mesh refinement granularities. Figure 31 visualizes the results.

We find that ASMR++ provides accurate refinements for different values of α using a single policy. The produced meshes are of higher quality than those created by ASMR, especially near the corners of the rhomboid obstacles in the domain. All other RL-AMR methods are unable to produce consistent meshes. Single Agent tends to over-refine uninteresting regions, likely due to its iterative refinement procedure. In contrast, Sweep often collapses to uniform or mostly uniform refinements, which may be a by-product of the difference in training on individual elements and sweeping over all elements at once during inference. VDG N produces high-quality refinements in some cases, but is inconsistent across seeds and in some cases fails to optimize for different element penalties with the same model, which

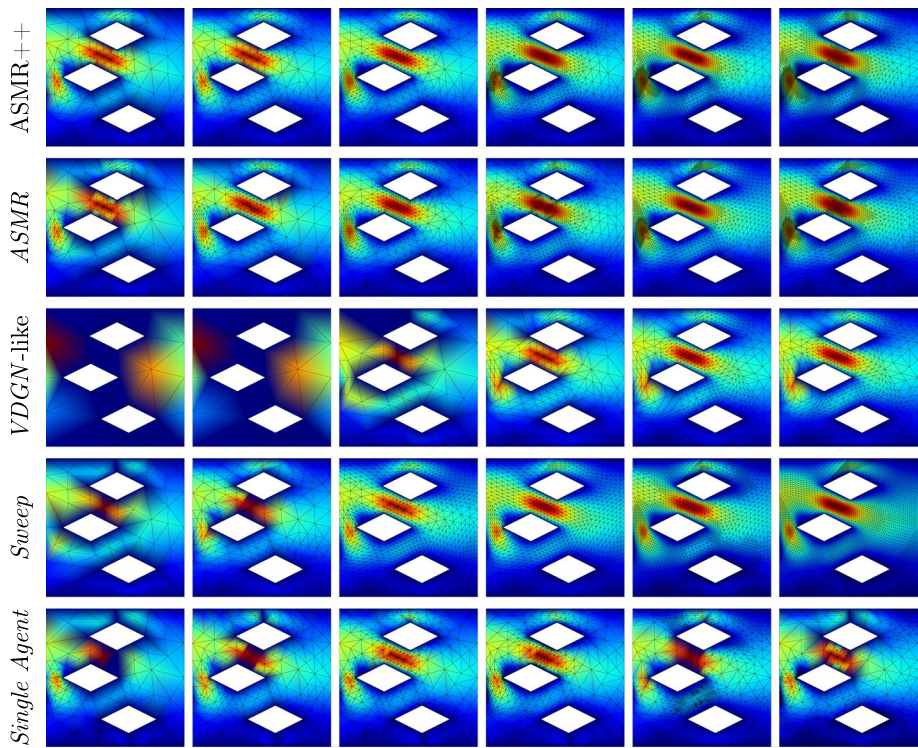


Fig. 31 Comparison of the different RL-AMR methods for the Stokes Flow task for the same PDE and different target mesh resolutions. The target mesh resolution increases from the left to the right for each row. ASMR++ and the *VDGn-like* baseline use the same model queried with different α values for each row. The other methods use a different policy for each produced mesh

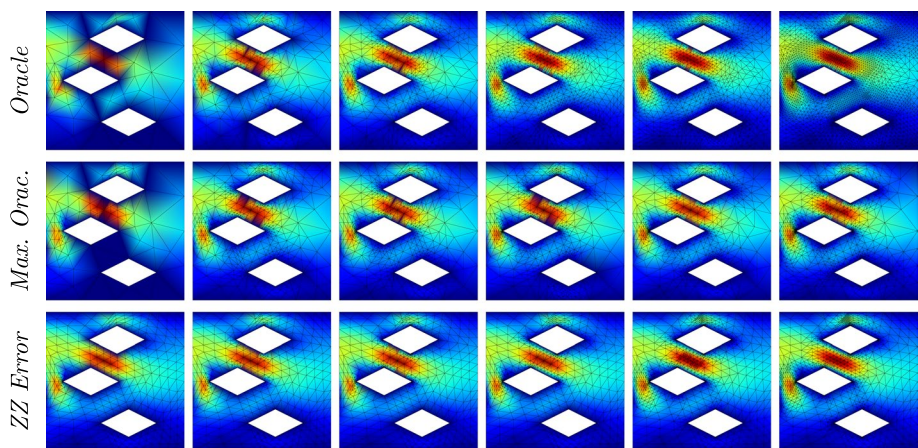


Fig. 32 Comparison of the different *Heuristic* methods for the same Stokes Flow task PDE and different error thresholds θ . The target mesh resolution increases from the left to the right for each row

could be because its value decomposition objective may scale poorly to large numbers of agents. These qualitative findings are consistent with the error measurements of Fig. 8.

Heuristic Visualizations

We visualize exemplary refinements for the error estimation-based threshold *Heuristics* in Fig. 32. All *Heuristics* greedily refine the elements with their respective largest error estimates, regardless of the resulting decrease in error. This behavior is fully local, which may cause issues for global dependencies (Strauss, 2007) and conforming refinements. The

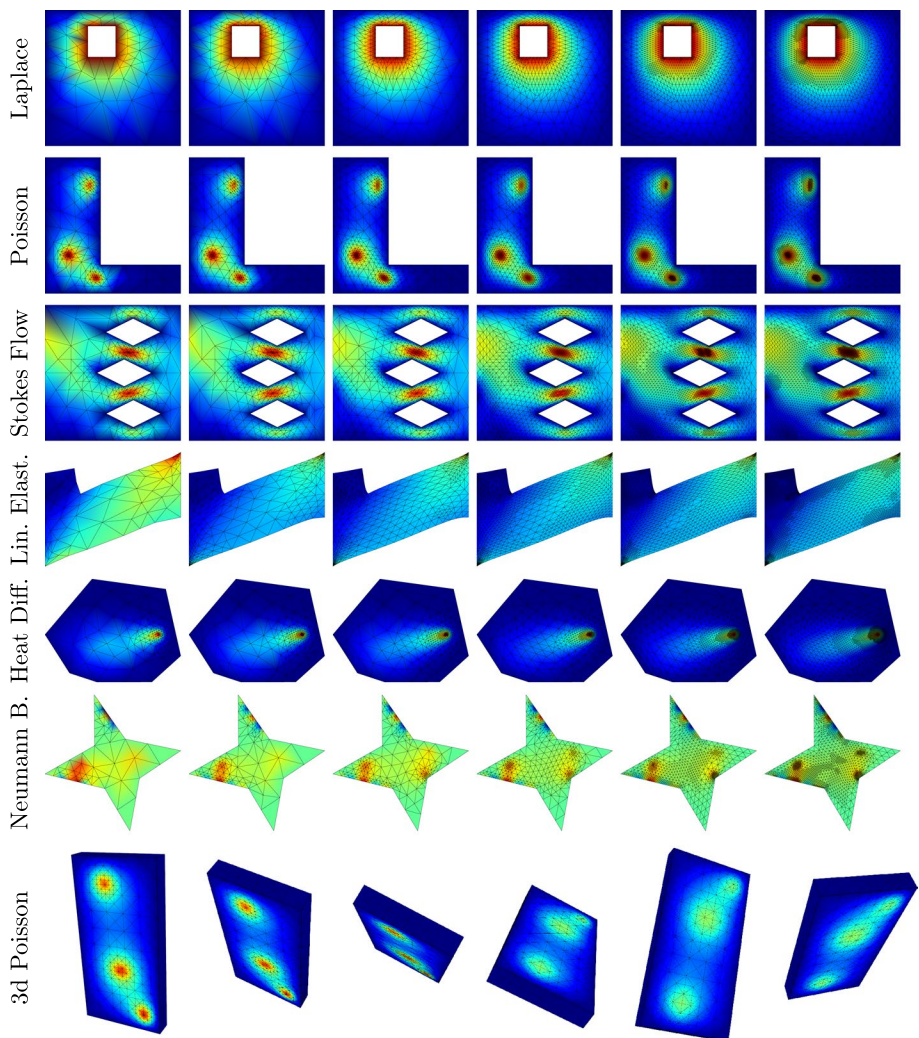


Fig. 33 ASMR++ meshes for the different tasks and varying element penalties α . Each row uses the same policy conditioned on a range of low (left) to high (right) α values for the 2-dimensional tasks. For the 3-dimensional Poisson task, we use the same PDE and α value, and show different camera angles

ZZ Error Heuristic uses a smoother error estimate than the oracle heuristics, which leads to potentially sub-optimal but generally more consistent refinements.

Additional ASMR++ Visualizations

We visualize meshes created by ASMR++ policies for all considered tasks. For the 2-dimension tasks, we follow Section F and use a fixed policy to create refinements with 6 different granularities for a random PDE for each task, only varying the element penalty α . For the 3-dimensional variant of Poisson's equation, we instead visualize the same random PDE and α value for 6 different camera angles. Figure 33 displays the final refined meshes, as well as the solutions of the PDE on the meshes. On all tasks, ASMR++ provides highly accurate refinements across target mesh refinement levels.

Acknowledgements NF was supported by the BMBF project Davis (Datengetriebene Vernetzung für die ingenieurtechnische Simulation). This work is also part of the DFG AI Research Unit 5339 regarding the combination of physics-based simulation with AI-based methodologies for the fast maturation of manufacturing processes. The financial support by German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) is gratefully acknowledged. The authors acknowledge support by the state of Baden-Württemberg through bwHPC, as well as the HoreKa supercomputer funded by the Ministry of Science, Research and the Arts Baden-Württemberg and by the German Federal Ministry of Education and Research.

Author Contributions The following contribution statement uses the CRediT system. N.F.: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing – Original Draft, Visualization. P.D.: Software, Validation, Writing – Review & Editing. T.W.: Conceptualization, Data Curation, Writing – Review & Editing. S.R.: Software, Visualization. L.K.: Writing - Review & Editing, Supervision. G.N.: Conceptualization, Methodology, Writing – Review & Editing, Supervision, Project administration, Funding acquisition.

Funding Open Access funding enabled and organized by Projekt DEAL. NF was supported by the BMBF project Davis (Datengetriebene Vernetzung für die ingenieurtechnische Simulation). This work is also part of the DFG AI Research Unit 5339 on the combination of physics-based simulation with AI-based methodologies for accelerating the maturation of manufacturing processes. Financial support by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) is gratefully acknowledged. Computational resources were provided by the state of Baden-Württemberg through bwHPC, and the HoreKa supercomputer funded by the Ministry of Science, Research and the Arts Baden-Württemberg and the German Federal Ministry of Education and Research.

Data Availability Our method uses Reinforcement Learning, which generates data on the fly. Thus, no explicit dataset exists. The Reinforcement Learning environments are made publicly available together with the rest of the code.

Code Availability We publish code under <http://www.github.com/NiklasFreymuth/ASMRplusplus>.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agarwal, R., Schwarzer, M., Castro, P. S., Courville, A. C., & Bellemare, M. (2021). Deep reinforcement learning at the edge of the statistical precipice. *Advances in neural information processing systems*, 34, 29304–29320.
- Alkin, B., Bleeker, M., Kurle, R., Kronlachner, T., Sonnleitner, R., Dorfer, M., & Brandstetter, J. (2025). Ab-upt: Scaling neural cfd surrogates for high-fidelity automotive aerodynamics simulations via anchored-branched universal physics transformers [arXiv:2502.09692](https://arxiv.org/abs/2502.09692) arXiv preprint.
- Alkin, B., Fürst, A., Schmid, S., Gruber, L., Holzleitner, M., & Brandstetter, J. (2024). Universal physics transformers: A framework for efficiently scaling neural operators. *Advances in Neural Information Processing Systems*, 37, 25152–25194.
- Anderson, R., Andrej, J., Barker, A., Bramwell, J., Camier, J.-S., Cervený, J., Dobrev, V., Dudouit, Y., Fisher, A., Kolev, T., et al. (2021). Mfem: A modular finite element methods library. *Computers & Mathematics with Applications*, 81, 42–74.
- Andrychowicz, M., Raichuk, A., Stańczyk, P., Orsini, M., Girgin, S., Mariner, R., Hussenot, L., Geist, M., Pietquin, O., Michalski, M., Gelly, S., & Bachem, O.: What matters for on-policy deep actor-critic methods? a large-scale study. In: International Conference on Learning Representations (2021). <https://openreview.net/forum?id=nIAxjsniDzg>
- Arnold, D. N., Mukherjee, A., & Pouly, L. (2000). Locally adapted tetrahedral meshes using bisection. *SIAM Journal on Scientific Computing*, 22(2), 431–448.
- Ba, J. L., Kiros, J. R., & Hinton, G. E. (2016). *Layer normalization*. *stat*, 1050, 21.
- Baker, T. J. (1997). Mesh adaptation strategies for problems in fluid dynamics. *Finite Elements in Analysis and Design*, 25(3–4), 243–273.
- Bangerth, W., & Rannacher, R. (2013). *Adaptive finite element methods for differential equations*. Birkhäuser.
- Bangerth, W., Burstedde, C., Heister, T., & Kronbichler, M. (2012). Algorithms and data structures for massively parallel generic adaptive finite element codes. *ACM Transactions on Mathematical Software (TOMS)*, 38(2), 1–28.
- Becker, R., Gantner, G., Innerberger, M., & Praetorius, D. (2023). Goal-oriented adaptive finite element methods with optimal computational complexity. *Numerische Mathematik*, 153(1), 111–140.
- Bentley, J. L. (1975). Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9), 509–517.
- Berger, M. J., & Colella, P. (1989). Local adaptive mesh refinement for shock hydrodynamics. *Journal of Computational Physics*, 82(1), 64–84.
- Binev, P., Dahmen, W., & DeVore, R. (2004). Adaptive finite element methods with convergence rates. *Numerische Mathematik*, 97, 219–268.
- Bohn, J., & Feischl, M. (2021). Recurrent neural networks as optimal mesh refinement strategies. *Computers & Mathematics with Applications*, 97, 61–76.
- Brandstetter, J., Worrall, D. E., & Welling, M. (2022). Message passing neural pde solvers. International Conference on Learning Representations
- Brenner, S. C., & Scott, L. R. (2008). *The mathematical theory of finite element methods* (Vol. Vol. 3). Springer.
- Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). Openai gym [arXiv:1606.01540](https://arxiv.org/abs/1606.01540) arXiv preprint.
- Bronstein, M. M., Bruna, J., Cohen, T., & Veličković, P. (2021). Geometric deep learning: Grids, groups, graphs, geodesics, and gauges [arXiv:2104.13478](https://arxiv.org/abs/2104.13478) arXiv preprint.
- Bryan, G. L., Norman, M. L., O’Shea, B. W., Abel, T., Wise, J. H., Turk, M. J., Reynolds, D. R., Collins, D. C., Wang, P., Skillman, S. W., et al. (2014). Enzo: An adaptive mesh refinement code for astrophysics. *The Astrophysical Journal Supplement Series*, 211(2), Article 19.
- Carstensen, C. (2004). An adaptive mesh-refining algorithm allowing for an h^1 stable l_2 projection onto courrant finite element spaces. *Constructive Approximation*, 20, 549–564.
- Cervený, J., Dobrev, V., & Kolev, T. (2019). Nonconforming mesh refinement for high-order finite elements. *SIAM Journal on Scientific Computing*, 41(4), 367–392.

- Cohen, A., Teng, E., Berges, V.-P., Dong, R.-P., Henry, H., Mattar, M., Zook, A., & Ganguly, S. (2021). On the use and misuse of absorbing states in multi-agent reinforcement learning [arXiv:2111.05992](#) arXiv preprint.
- Cunningham, A. J., Frank, A., Varnière, P., Mitran, S., & Jones, T. W. (2009). Simulating magnetohydrodynamical flow with constrained transport and adaptive mesh refinement: Algorithms and tests of the AstroBear code. *The Astrophysical Journal Supplement Series*, 182(2), 519.
- Dzanic, T., Mittal, K., Kim, D., Yang, J., Petrides, S., Keith, B., & Anderson, R. (2024). Dynamo: Multi-agent reinforcement learning for dynamic anticipatory mesh optimization with applications to hyperbolic conservation laws. *Journal of Computational Physics*. <https://doi.org/10.1016/j.jcp.2024.112924>
- Fidkowski, K. J., & Darmofal, D. L. (2011). Review of output-based error estimation and mesh adaptation in computational fluid dynamics. *ALAA journal*, 49(4), 673–694.
- Fortunato, M., Pfaff, T., Wirsnberger, P., Pritzel, A., & Battaglia, P.: Multiscale meshgraphnets. In: ICML 2022 2nd AI for Science Workshop (2022). <https://openreview.net/forum?id=G3TRlsmMhhf>
- Foucart, C., Charous, A., & Lermusiaux, P. F. (2023). Deep reinforcement learning for adaptive mesh refinement. *Journal of Computational Physics*, 491, Article 112381.
- Freytmuth, N., Dahlinger, P., Würth, T., Reisch, S., Kärger, L., & Neumann, G. (2023). Swarm reinforcement learning for adaptive mesh refinement. *Advances in Neural Information Processing Systems*, 36
- Gibert, G., Prabel, B., Gravouil, A., & Jacquemoud, C. (2019). A 3d automatic mesh refinement x-fem approach for fatigue crack propagation. *Finite Elements in Analysis and Design*, 157, 21–37.
- Gillette, A., Keith, B., & Petrides, S. (2022). Learning robust marking policies for adaptive mesh refinement [arXiv:2207.06339](#) arXiv preprint.
- Guillet, T., Pakmor, R., Springel, V., Chandrasekar, P., & Klingenberg, C. (2019). High-order magnetohydrodynamics for astrophysics with an adaptive mesh refinement discontinuous galerkin scheme. *Monthly Notices of the Royal Astronomical Society*, 485(3), 4209–4246.
- Guo, X., Li, W., Iorio, F.: Convolutional neural networks for steady flow approximation. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16, pp. 481–490. Association for Computing Machinery, New York, NY, USA (2016). <https://doi.org/10.1145/2939672.2939738>
- Gustafsson, T., & Mcbain, G. D. (2020). scikit-fem: A Python package for finite element assembly. *Journal of Open Source Software*, 5(52), Article 2369.
- Hayes, C. F., Rădulescu, R., Bargiacchi, E., Källström, J., Macfarlane, M., Reymond, M., Verstraeten, T., Zintgraf, L. M., Dazeley, R., Heintz, F., et al. (2022). A practical guide to multi-objective reinforcement learning and planning. *Autonomous Agents and Multi-Agent Systems*, 36, Article 26.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778)
- Hessel, M., Modayil, J., Van Hasselt, H., Schaul, T., Ostrovski, G., Dabney, W., Horgan, D., Piot, B., Azar, M., & Silver, D. (2018). Rainbow: Combining improvements in deep reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 32,
- Ho-Le, K. (1988). Finite element mesh generation methods: A review and classification. *Computer-aided design*, 20(1), 27–38.
- Huang, K., Krügener, M., Brown, A., Menhorn, F., Bungartz, H.-J., & Hartmann, D. (2021). Machine learning-based optimal mesh generation in computational fluid dynamics [arXiv:2102.12923](#) arXiv preprint.
- Huang, W., & Russell, R. D. (2010). *Adaptive Moving Mesh Methods* (Vol. Vol. 174). Springer.
- Huergo, D., Rubio, G., & Ferrer, E. (2024). A reinforcement learning strategy for p-adaptation in high order solvers. *Results in Engineering*, 21, Article 101693.
- Hüttenrauch, M., Adrian, V., & Neumann, G. (2019). Deep reinforcement learning for swarm systems. *Journal of Machine Learning Research*, 20(54), 1–31.
- Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International Conference on Machine Learning* (pp. 448–456) pmlr.
- John, V. (2016). *Finite element methods for incompressible flow problems* (Vol. Vol. 51). Springer.
- Jones, M. T., & Plassmann, P. E. (1997). Adaptive refinement of unstructured finite-element meshes. *Finite Elements in Analysis and Design*, 25(1–2), 41–60.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization [arXiv:1412.6980](#) arXiv preprint.
- Kita, E., & Kamiya, N. (2001). Error estimation and adaptive mesh refinement in boundary element method, an overview. *Engineering Analysis with Boundary Elements*, 25(7), 479–495.
- Larson, M.G., & Bengzon, F.: The Finite Element Method: Theory, Implementation, and Applications. Texts in Computational Science and Engineering, vol. 10. Springer, Berlin, Heidelberg (2013). <https://doi.org/10.1007/978-3-642-33287-6>
- Lee, H., Hwang, D., Kim, D., Kim, H., Tai, J.J., Subramanian, K., Wurman, P.R., Choo, J., Stone, P., & Seno, T.: Simba: Simplicity bias for scaling up parameters in deep reinforcement learning. In: The Thirteenth International Conference on Learning Representations (2024)

- Linkerhägner, J., Freymuth, N., Scheikl, P.M., Mathis-Ullrich, F., & Neumann, G.: Grounding graph network simulators using physical sensor observations. In The Eleventh International Conference on Learning Representations (ICLR) (2023). <https://openreview.net/forum?id=jsZsEd8VEY>
- Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning [arXiv:1312.5602](https://arxiv.org/abs/1312.5602) arXiv preprint.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533.
- Mukherjee, A. (1996). *An adaptive finite element code for elliptic boundary value problems in three dimensions with applications in numerical relativity*. The Pennsylvania State University.
- Nagarajan, A., & Soghrati, S. (2018). Conforming to interface structured adaptive mesh refinement: 3d algorithm and implementation. *Computational Mechanics*, 62, 1213–1238.
- Nauman, M., Ostaszewski, M., Jankowski, K., Miłoś, P., & Cygan, M. (2024). Bigger, regularized, optimistic: Scaling for compute and sample efficient continuous control. *Advances in neural information processing systems*, 37, 113038–113071.
- Ortiz, M., & Quigley, J., IV. (1991). Adaptive mesh refinement in strain localization problems. *Computer Methods in Applied Mechanics and Engineering*, 90(1–3), 781–804.
- Pan, J., Huang, J., Cheng, G., & Zeng, Y. (2023). Reinforcement learning for automatic quadrilateral mesh generation: A soft actor-critic approach. *Neural Networks*, 157, 288–304.
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., & Antiga, L. (2019). Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32.
- Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., & Battaglia, P. W. (2021). Learning mesh-based simulation with graph networks. International Conference on Learning Representations [arxiv:2010.03409](https://arxiv.org/abs/2010.03409).
- Plewa, T., Linde, T., & Weirs, V. G. (2005). *Adaptive mesh refinement-theory and applications*. Springer.
- Quarteroni, A., & Quarteroni, S. (2009). *Numerical models for differential problems* (Vol. Vol. 2). Springer.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational physics*, 378, 686–707.
- Reddy, J. N. (2019). *Introduction to the finite element method*. McGraw-Hill Education.
- Rivara, M. C. (1984). Algorithms for refining triangular grids suitable for adaptive and multigrid techniques. *International journal for numerical methods in Engineering*, 20(4), 745–756.
- Rong, Y., Huang, W., Xu, T., & Huang, J.: Dropedge: Towards deep graph convolutional networks on node classification. In International Conference on Learning Representations (2019)
- Roth, J., Schröder, M., & Wick, T. (2022). Neural network guided adjoint computations in dual weighted residual error estimation. *SN Applied Sciences*, 4(2), 62.
- Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., & Battaglia, P. (2020). Learning to simulate complex physics with graph networks. *Proceedings of the 37th International Conference on Machine Learning* (pp. 8459–8468). New York: PMLR.
- Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009). The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1), 61–80. <https://doi.org/10.1109/TNN.2008.2005605>
- Schaul, T., Quan, J., Antonoglou, I., & Silver, D.: Prioritized experience replay. In: Bengio, Y., LeCun, Y. (eds.) 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2–4, 2016, Conference Track Proceedings (2016)
- Schulman, J., Moritz, P., Levine, S., Jordan, M., & Abbeel, P. (2015). High-dimensional continuous control using generalized advantage estimation [arXiv:1506.02438](https://arxiv.org/abs/1506.02438) arXiv preprint.
- Schulman, J., Wolski, F., Dhariwal, P., Radford, A., & Klimov, O. (2017). Proximal policy optimization algorithms [arXiv:1707.06347](https://arxiv.org/abs/1707.06347) arXiv preprint.
- Służalec, T., Grzeszczuk, R., Rojas, S., Dzwiniel, W., & Paszyński, M. (2023). Quasi-optimal hp-finite element refinements towards singularities via deep neural network prediction. *Computers & Mathematics with Applications*, 142, 157–174.
- Šošić, A., KhudaBukhsh, W.R., Zoubir, A.M., & Koeppl, H.: Inverse reinforcement learning in swarm systems. In: Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems, pp. 1413–1421 (2017)
- Stein, E. (2007). *Adaptive finite elements in linear and nonlinear solid and structural mechanics* (Vol. Vol. 416). Springer.
- Stevenson, R. (2008). The completion of locally refined simplicial partitions created by bisection. *Mathematics of computation*, 77(261), 227–241.
- Strauss, W. A. (2007). *Partial differential equations: An introduction*. Wiley.

- Suárez, J. P., Abad, P., Plaza, A., & Padron, M. A. (2005). Computational aspects of the refinement of 3d tetrahedral meshes. *Journal of Computational Methods in Sciences and Engineering*, 5(4), 215–224.
- Sunehag, P., Lever, G., Gruslys, A., Czarnecki, W. M., Zambaldi, V., Jaderberg, M., Lanctot, M., Sonnerat, N., Leibo, J. Z., & Tuyls, K. (2017). Value-decomposition networks for cooperative multi-agent learning [arXiv:1706.05296](https://arxiv.org/abs/1706.05296) arXiv preprint.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT press.
- Um, K., Hu, X., & Thuerey, N. (2018). Liquid splash modeling with neural networks. *Computer Graphics Forum*, 37(8), 171–182. <https://doi.org/10.1111/cgf.13522>
- Van Hasselt, H., Guez, A., & Silver, D. (2016). Deep reinforcement learning with double q-learning. *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 30,
- Van Moffaert, K., & Nowé, A. (2014). Multi-objective reinforcement learning using sets of Pareto dominating policies. *The Journal of Machine Learning Research*, 15(1), 3483–3512.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., & Polosukhin, I.: Attention is all you need. *Advances in neural information processing systems* **30** (2017)
- Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2018). *Graph attention networks*. *stat*, 1050, 4.
- Wallwork, J. G. (2021). Mesh adaptation and adjoint methods for finite element coastal ocean modelling. Imperial College London. PhD thesis.
- Wallwork, J. G., Lu, J., Zhang, M., & Piggott, M. D. (2022). E2n: Error estimation networks for goal-oriented mesh adaptation [arXiv:2207.11233](https://arxiv.org/abs/2207.11233) arXiv preprint.
- Wang, Z., Schaul, T., Hessel, M., Hasselt, H., Lanctot, M., & Freitas, N. (2016). Dueling network architectures for deep reinforcement learning. *International Conference on Machine Learning* (pp. 1995–2003) PMLR.
- Wanner, G., & Hairer, E. (1996). *Solving ordinary differential equations II* (Vol. Vol. 375). Springer.
- Wu, T., Maruyama, T., Zhao, Q., Wetzstein, G., & Leskovec, J.: Learning controllable adaptive simulation for multi-resolution physics. In: The Eleventh International Conference on Learning Representations (2023). <https://openreview.net/forum?id=PbfgkZ2HdbE>
- Würth, T., Freymuth, N., Zimmerling, C., Neumann, G., & Kärgner, L. (2024). Physics-informed meshgraphs (pi-mgn): Neural finite element solvers for non-stationary and nonlinear simulations on arbitrary meshes. *Computer Methods in Applied Mechanics and Engineering*. <https://doi.org/10.1016/j.cma.2024.117102>
- Würth, T., Freymuth, N., Zimmerling, C., Neumann, G., & Kärgner, L. (2024). Physics-informed meshgraphs (pi-mgns): Neural finite element solvers for non-stationary and nonlinear simulations on arbitrary meshes [arXiv:2402.10681](https://arxiv.org/abs/2402.10681) arXiv preprint.
- Würth, T., Krauß, C., Zimmerling, C., & Kärgner, L. (2023). Physics-informed neural networks for data-free surrogate modelling and engineering optimization-an example from composite manufacturing. *Materials & Design*, 231, Article 112034.
- Yang, J., Dzanic, T., Petersen, B. K., Kudo, J., Mittal, K., Tomov, V., Camier, J.-S., Zhao, T., Zha, H., Kolev, T., Anderson, R., & Faissol, D. (2023). Reinforcement learning for adaptive mesh refinement. 26th International Conference on Artificial Intelligence and Statistics. AISTATS.
- Yang, J., Mittal, K., Dzanic, T., Petrides, S., Keith, B., Petersen, B., Faissol, D., & Anderson, R. (2023). Multi-agent reinforcement learning for adaptive mesh refinement. 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)
- Yano, M., & Darmofal, D. L. (2012). An optimization-based framework for anisotropic simplex mesh adaptation. *Journal of Computational Physics*, 231(22), 7626–7649.
- Zhang, Z., Wang, Y., Jimack, P.K., & Wang, H.: Meshingnet: A new mesh generation method based on deep learning. In *Computational Science—ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part III* 20, pp. 186–198 (2020). Springer
- Zienkiewicz, O. C., & Zhu, J. Z. (1992). The superconvergent patch recovery and a posteriori error estimates. Part I: The recovery technique. *International Journal for Numerical Methods in Engineering*, 33(7), 1331–1364.
- Zienkiewicz, O. C., Taylor, R. L., & Zhu, J. Z. (2005). *The finite element method: Its basis and fundamentals*. Elsevier.
- Zimmerling, C., Dörr, D., Henning, F., & Kärgner, L. (2019). A machine learning assisted approach for textile formability assessment and design improvement of composite components. *Compos Part A*, 124, Article Article 105459. <https://doi.org/10.1016/j.compositesa.2019.05.027>
- Zimmerling, C., Poppe, C., Stein, O., & Kärgner, L. (2022). Optimisation of manufacturing process parameters for variable component geometries using reinforcement learning. *Materials & Design*, 214, Article 110423. <https://doi.org/10.1016/j.matdes.2022.110423>

Zimmerling, C., Trippe, D., Fengler, B., Kärger, L.: An approach for rapid prediction of textile draping results for variable composite component geometries using deep neural networks. *AIP Conference Proceedings* **2113**(1) (2019) <https://doi.org/10.1063/1.5112512><https://pubs.aip.org/aip/acp/article-pdf/doi/10.1063/1.5112512.020007>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.