Commun. Comput. Phys.
doi: 10.4208/cicp.OA-2024-0260

Vol. **39**, No. 1, pp. 1-28
January 2026

First-Principle-Like Reinforcement Learning of Nonlinear Numerical Schemes for Conservation Laws

Hao-Chen Wang[1], Meilin Yu[2,*] and Heng Xiao[1]

[1]*Stuttgart Center for Simulation Science (SC SimTech), University of Stuttgart, Stuttgart, Germany.*
[2]*University of Maryland Baltimore County (UMBC), Baltimore, MD, USA.*

Received 24 October 2024; Accepted (in revised version) 9 July 2025

Abstract. In this study we present a universal nonlinear numerical scheme design method for nonlinear conservation laws, enabled by multi-agent reinforcement learning (MARL). Unlike contemporary approaches based on supervised learning or reinforcement learning, our method does not rely on reference data or empirical design. Instead, a first-principle-like approach using fundamental computational fluid dynamics (CFD) principles, including total variation diminishing (TVD) and k-exact reconstruction, is employed to design nonlinear numerical schemes. The third-order finite volume scheme is employed as the workhorse to test the performance of the MARL-based nonlinear numerical scheme design method. Numerical results demonstrate that the new MARL-based method can strike a balance between accuracy and numerical dissipation in nonlinear numerical scheme design, and outperforms the third-order MUSCL (Monotonic Upstream-centered Scheme for Conservation Laws) with the van Albada limiter for shock capturing. Furthermore, we demonstrate for the first time that a numerical scheme trained from one-dimensional (1D) Burgers' equation simulations can be directly used for numerical simulations of both 1D and 2D (two-dimensional constructions using the tensor product operation) Euler equations. The framework of the MARL-based numerical scheme design concepts can incorporate, in general, all types of numerical schemes as simulation machines.

AMS subject classifications: 65M08, 68T07, 76M12

Key words: Reinforcement learning, first-principle-like rewards, nonlinear numerical scheme, conservation law, shock capturing.

1 Introduction

Hyperbolic conservation laws, governed by nonlinear partial differential equations (PDEs), have extensive applications across numerous fields of science and engineer-

*Corresponding author. *Email addresses:* haochen.wang@itlr.uni-stuttgart.de (H.-C. Wang), mlyu@umbc.edu (M. Yu), heng.xiao@simtech.uni-stuttgart.de (H. Xiao)

ing, such as aero-hydrodynamics, astrophysics, plasma physics, advanced manufacture and transportation engineering [1]. One feature of nonlinear hyperbolic conservation laws is that their solutions admit singularities (e.g., shock waves and contact discontinuity), which can be developed in finite time from smooth initial data. This poses grand challenges on numerical simulations as nonlinear numerical schemes need to be developed to take both scheme stability and numerical resolution into consideration; see Godunov's pioneering work on numerical methods for shock capturing [2]. As a result, many nonlinear numerical scheme construction methods have been developed during the last half century, such as high-resolution schemes with TVD slope limiters [3], weighted essentially non-oscillatory (WENO) methods [4–6], total variation bounded (TVB) discontinuous Galerkin methods [7], hierarchical multi-dimensional limiting process (MLP) [8, 9], moving discontinuous Galerkin finite element method with interface condition enforcement (MDG-ICE) [10, 11] and localized artificial viscosity and diffusivity methods [12–14], just to name a few. However, many nonlinear numerical schemes developed so far have to introduce empirical components, such as the use of limiter functions in the MUSCL scheme and smoothness indicators in the WENO schemes.

1.1 Supervised learning of shock capturing schemes

To reduce the dependence on empirical designs in numerical methods, researchers have increasingly turned to machine learning techniques to develop data-driven models. These models differ from traditional approaches by replacing many empirical components with neural networks. For instance, Ray et al. [15, 16] developed a data-driven troubled-cell indicator by training an artificial neural network (ANN) and tested it on 1D grids and 2D unstructured grids. Beck et al. [17] developed a data-driven shock indicator by using image-based edge detection methods on 2D grids. Bezgin et al. [18] developed a data-driven nonlinear weight function for WENO3. Numerical results showed that these data-driven models can perform better than the empirical ones and do not need problem-dependent parameter tuning. However, one common issue shared by the aforementioned works is that special numerical treatments, such as certain auxiliary equations and their corresponding analytical solutions, needs to be used to encode desired numerical features into the machine learning model. The choice of specific equations is still largely based on the authors' experience and analytical solutions may not be available.

Researchers have also leveraged the tool of machine learning to design new flux limiters [19] and learn discretizations for PDEs directly [20]. Nguyen et al. [19] designed a framework to derive an optimal flux limiter for the coarse-grained Burgers' equation by learning from high-resolution data. Numerical results demonstrated that the trained flux limiter achieves better results than standard limiters, but only in Burgers' equation simulations. Therefore, the model's generalizability to different physics is questionable. Bar et al. [20] designed a data-driven discretization method to learn the optimal approximations to PDEs on a coarse grid directly from the solutions on a finer grid. Numerical results demonstrated that their proposed method outperforms the standard numerical