Numerische Mathematik (2019) 142:995–1047
https://doi.org/10.1007/s00211-019-01042-w

Numerische Mathematik

Provably positive high-order schemes for ideal magnetohydrodynamics: analysis on general meshes

Kailiang Wu[1] · **Chi-Wang Shu**[2]

Received: 17 August 2018 / Revised: 11 February 2019 / Published online: 3 May 2019

Abstract
This paper proposes and analyzes arbitrarily high-order discontinuous Galerkin (DG) and finite volume methods which provably preserve the positivity of density and pressure for the ideal magnetohydrodynamics (MHD) on general meshes. Unified auxiliary theories are built for rigorously analyzing the positivity-preserving (PP) property of numerical MHD schemes with a Harten–Lax–van Leer (HLL) type flux on polytopal meshes in any space dimension. The main challenges overcome here include establishing certain relation between the PP property and a discrete divergence of magnetic field on general meshes, and estimating proper wave speeds in the HLL flux to ensure the PP property. In the 1D case, we prove that the standard DG and finite volume methods with the proposed HLL flux are PP, under a condition accessible by a PP limiter. For the multidimensional conservative MHD system, the standard DG methods with a PP limiter are not PP in general, due to the effect of unavoidable divergence error in the magnetic field. We construct provably PP high-order DG and finite volume schemes by proper discretization of the symmetrizable MHD system, with two divergence-controlling techniques: the locally divergence-free elements and suitably discretized Godunov–Powell source term. The former technique leads to zero divergence within each cell, while the latter controls the divergence error across cell interfaces. Our analysis reveals in theory that a coupling of these two techniques is very important for positivity preservation, as they exactly contribute the discrete divergence terms which are absent in standard multidimensional DG schemes but crucial for ensuring the PP property. Several numerical tests further confirm the PP property and the effectiveness of the proposed PP schemes. Unlike the conservative MHD system, the exact smooth solutions of the symmetrizable MHD system are proved to retain the positivity even if the divergence-free condition is not satisfied. Our analysis and findings further the understanding, at both discrete and continuous levels, of the relation between the PP property and the divergence-free constraint.

Mathematics Subject Classification 65M60 · 65M08 · 35L65 · 76W05

Research is supported in part by ARO Grant W911NF-15-1-0226 and NSF Grant DMS-1719410.

Extended author information available on the last page of the article

1 Introduction

This paper is concerned with highly accurate and robust numerical methods for the ideal compressible magnetohydrodynamics (MHD), which play an important role in many fields including astrophysics, plasma physics and space physics. When viscous, resistive and relativistic effects can be neglected, the governing equations of ideal MHD, which combine the equations of gas dynamics with the Maxwell equations, have been widely used to model the dynamics of electrically conducting fluids in the presence of magnetic field. The ideal MHD system can be written as

$$\mathbf{U}_t + \nabla \cdot \mathbf{F}(\mathbf{U}) = \mathbf{0}, \tag{1}$$

with an additional divergence-free constraint on the magnetic field

$$\nabla \cdot \mathbf{B} = 0. \tag{2}$$

The conservative vector $\mathbf{U} = (\rho, \rho\mathbf{v}, \mathbf{B}, E)^\top$; in the d-dimensional case, the divergence operator $\nabla\cdot = \sum_{i=1}^{d} \frac{\partial}{\partial x_i}$, and the flux $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_d)$ with

$$\mathbf{F}_i(\mathbf{U}) = \Big(\rho v_i,\ \rho v_i \mathbf{v} - B_i \mathbf{B} + p_{\text{tot}} \mathbf{e}_i,\ v_i \mathbf{B} - B_i \mathbf{v},\ v_i (E + p_{\text{tot}}) - B_i (\mathbf{v} \cdot \mathbf{B})\Big)^\top.$$

Here ρ is the density, $\mathbf{v} = (v_1, v_2, v_3)$ is the fluid velocity, $\mathbf{B} = (B_1, B_2, B_3)$ denotes the magnetic field, $p_{\text{tot}} = p + \frac{|\mathbf{B}|^2}{2}$ is the total pressure consisting of the gas pressure p and the magnetic pressure, the vector $\mathbf{e}_i$ denotes the ith row of the unit matrix of size 3, $E = \rho e + \frac{1}{2}\left(\rho|\mathbf{v}|^2 + |\mathbf{B}|^2\right)$ is the total energy consisting of thermal, kinetic and magnetic energies, and e denotes the specific internal energy. The system (1) is closed with an equation of state (EOS). Although the ideal EOS, $p = (\gamma - 1)\rho e$, with a constant adiabatic index γ, is the most widely used choice, there are situations where it is more suitable to use other EOSs. A general EOS can be expressed as $p = p(\rho, e)$, which is assumed to satisfy the following condition (cf. [67]):

$$\text{if} \quad \rho > 0, \quad \text{then} \quad e > 0 \ \Leftrightarrow\ p(\rho, e) > 0. \tag{3}$$

This condition is reasonable and holds for the ideal EOS with $\gamma > 1$.

Although the satisfaction of the divergence-free condition (2) is not explicitly included in the system (1), the exact solution of (1) always preserves zero divergence in future time if the initial divergence is zero. However, most of the numerical MHD schemes for $d \geq 2$ lead to a nonzero divergence of numerical magnetic field due to truncation errors, even if the initial data satisfy (2). As it is widely known, large divergence error can lead to numerical instabilities or nonphysical features in the computed solutions, cf. [7,11,26,36,48]. In the past several decades, many numerical techniques were proposed to control the divergence error or enforce the divergence-free condition in the discrete sense, including but not limited to: the projection method [11], the hyperbolic divergence cleaning method [20], the locally divergence-free methods [36,63], the constrained transport method [26] and its variants (e.g., [1,3,7,14,27,29,37,45,47,61]), and

the eight-wave methods (e.g., [12,40,42,43]). The eight-wave method was first proposed by Powell [42,43], based on proper discretization of the Godunov form [30] of ideal MHD equations

$$\mathbf{U}_t + \nabla \cdot \mathbf{F}(\mathbf{U}) = -(\nabla \cdot \mathbf{B})\, \mathbf{S}(\mathbf{U}), \tag{4}$$

where $\mathbf{S}(\mathbf{U}) = (0,\ \mathbf{B},\ \mathbf{v},\ \mathbf{v}\cdot\mathbf{B})^\top$. In the literature, (4) is sometimes also called Powell's system. The right-hand side term of (4), termed as the Godunov–Powell source term in the following, is proportional to $\nabla\cdot\mathbf{B}$. This means, at the continuous level, the Godunov form (4) and conservative form (1) are equivalent under the condition (2). However, the Godunov–Powell source term modifies the character of the MHD equations, making the system (4) Galilean invariant (cf. [21]), symmetrizable [30] and useful for designing entropy stable schemes (see, e.g., [12,22,40]). These good properties do not hold anymore if the source term is dropped. As first demonstrated by Powell [43], the inclusion of the source term also helps advect the divergence away with the flow. This renders the eight-wave method stable to control the divergence error, although some drawbacks [48] may be caused due to the loss of conservativeness.

In physics, the density, pressure and internal energy are positive. An equivalent mathematical description is that, the conservative vector $\mathbf{U}$ should stay in the *set of physically admissible states* defined by

$$\mathcal{G} = \left\{ \mathbf{U} = (\rho, \mathbf{m}, \mathbf{B}, E)^\top : \rho > 0,\ \mathcal{E}(\mathbf{U}) := E - \frac{1}{2}\left(\frac{|\mathbf{m}|^2}{\rho} + |\mathbf{B}|^2 \right) > 0 \right\}, \tag{5}$$

where the condition (3) has been used, and $\mathcal{E}(\mathbf{U}) = \rho e$ denotes the internal energy. We are interested in *positivity-preserving* (PP) numerical schemes whose solutions always stay in $\mathcal{G}$. The motivation comes from that, once the negative density or negative pressure (internal energy) is obtained in the numerical simulations, the discrete problem becomes ill-posed due to the loss of hyperbolicity, causing the breakdown of the simulation codes. However, most of the existing MHD schemes are generally not PP, and thus may suffer from a large risk of failure when simulating MHD problems with low density, low internal energy, low plasma-beta and/or strong discontinuity. A few efforts were made to reduce such risk. Balsara and Spicer [6] tried to maintain positive pressure by switching the Riemann solvers for different wave situations. Janhunen [34] noticed the challenge of developing PP schemes for the conservative system (1), so he proposed a modified MHD system, which is similar to the Godunov form (4) but includes only the source term in the induction equation. Based on his modified system, Janhunen [34] presented a new 1D Riemann solver and numerically demonstrated the PP property. Bouchut et al. [9,10] derived several approximate multiwave Riemann solvers for the 1D ideal MHD, with sufficient conditions for the solvers to satisfy the PP property and discrete entropy inequalities. Waagan [50] noticed the importance of proper discretization on Janhunen's modified system, and developed a positive second-order scheme by the approximate Riemann solvers of [9,10] and a new linear reconstruction. The robustness of that scheme was further demonstrated in [51] by extensive tests and comparisons. Recent years have witnessed significant advances in developing bound-preserving high-order schemes for hyperbolic systems; see the pioneer works by Zhang and Shu [65,66,68], and more

recent works, e.g., [15,33,39,52,55,59,60,64]. Balsara [5] proposed a self-adjusting PP limiter to enforce the positivity of the reconstructed solutions in a finite volume method for (1). Cheng et al. [13] extended the PP limiter of [66,67] to enforce the positivity of DG solutions for (1). The PP limiters in [5,13] are based on a presumed proposition that the cell-averaged solutions computed by those schemes always belong to $\mathcal{G}$. Such a proposition has not yet been rigorously proved for those methods in [5,13], although it could be deduced for the 1D schemes in [13] under some assumptions. Using the presumed PP property of the Lax–Friedrichs (LF) scheme, Christlieb et al. [16,17] developed PP high-order finite difference methods for the ideal MHD by extending the parametrized flux limiters [46,58,59]. It was numerically demonstrated that all the above PP techniques could enhance the robustness of MHD codes, but few theoretical evidences were provided, especially in the multidimensional cases, to completely prove the PP property of fully discretized schemes. In fact, finite numerical tests could be insufficient to genuinely demonstrate that a scheme is always PP under all circumstances. It is highly significant to develop *provably* PP schemes and rigorous PP analysis for the ideal MHD.

Seeking provably PP schemes for the ideal MHD is quite difficult, largely due to the intrinsic complexity of the MHD equations as well as the lack of sufficient knowledge about the underlying relation between the PP property and the divergence-free condition (2). One can see from (5) that the difficulties mainly lie in maintaining the positivity of internal energy, whose computation nonlinearly involves all the conservative variables. In most numerical methods, the conservative quantities are themselves evolved according to their own conservation laws, which are seemingly unrelated to and numerically do not necessarily guarantee the positivity of the computed internal energy. In theory, it is indeed a challenge to make an a priori judgment on whether a scheme is always PP under all circumstances or not.

Recently, two progresses [53,54] were made to rigorously analyze, understand and design provably PP methods for the ideal MHD. The first rigorous PP analysis was carried out in [53] for conservative finite volume and DG schemes for (1). The analysis unveiled in theory that a discrete divergence-free (DDF) condition is crucial for designing the PP conservative schemes for (1). This finding is consistent with the relativistic MHD case [56]. It was also proved in [53] that if the proposed DDF condition is slightly violated, even the first-order multidimensional LF scheme for (1) is generally not PP, and using very small CFL number or many times larger numerical viscosity does not help to prevent this effect. The DDF condition relies on a combination of the information on adjacent cells, and thus is not ensured by a locally divergence-free approach. As a result, in the multidimensional cases, a usual PP limiter does not genuinely guarantee the PP property of the standard DG schemes for (1), even if the locally divergence-free DG element [36] is employed. Interestingly, on the other hand, at the PDE level the positivity preservation and the divergence-free condition (1) are also inextricably linked for the ideal MHD system. For the conservative system (1), Janhunen [34] pointed out that the exact solutions to 1D Riemann problems can have negative pressure if the initial data has a jump in the normal magnetic field (a nonzero divergence). Recently in [54], we first observed that the exact smooth solution of (1) may also fail to be PP if the divergence-free condition (2) is (slightly) violated. Fortunately, in the present paper we find that the smooth solutions of the

modified system (4) always retain the desired positivity even if the magnetic field is not divergence-free. *All these findings motivate us to seek the multidimensional PP schemes via proper discretization of the modified system* (4) *rather than the conservative system* (1). Although Janhunen's modified MHD system [34] may also preserve the positivity, some other physical considerations suggest that Godunov's form (4) is better than Janhunen's as demonstrated in [23]. Using the analysis techniques proposed in [53], we first successfully developed in [54] the multidimensional provably PP high-order DG methods for (4). Note that the study in [53,54] was restricted to the schemes with the *global LF flux* on *uniform Cartesian meshes*. It is desirable to construct provably PP high-order schemes with lower dissipative numerical fluxes and on more general/unstructured meshes.

The aim of this paper is to present the rigorous analysis and a general framework for constructing provably PP high-order DG and finite volume methods with the HLL-type flux for the ideal MHD on general meshes. As a nontrivial extension of [53,54] in which the PP analysis techniques only work for Cartesian meshes and global LF flux, this work improves the analysis techniques of [53] and gives deeper understanding of positivity preservation at both continuous and discrete levels. The new contributions and significant innovations of this work are outlined as follows:

1. We present unified auxiliary theories for PP analysis of schemes with the HLL-type flux on general meshes for the ideal MHD in any space dimension. These provide a novel way to analytically extract the underlying relation between the PP property and the discrete divergence of magnetic field on an arbitrary polytopal mesh. Explicit estimates of the wave speeds in the HLL flux are technically derived to guarantee the provably PP property.
2. For the 1D MHD system (1), we prove the PP property of the standard finite volume and DG methods with the proposed HLL flux, under a condition accessible by a simple PP limiter.
3. In the multidimensional cases, we construct provably PP high-order DG methods based on the proposed HLL flux, a PP limiter [13], and a proper discretization of the modified MHD system (4) with two divergence-controlling techniques: the locally divergence-free elements and a novel discretization of the Godunov–Powell source term in an upwind manner according to the associated local wave speeds in the HLL flux. The former technique leads to zero divergence within each cell, while the latter controls the divergence error across cell interfaces. Our analysis clearly reveals in theory that a coupling of these two techniques is very important for positivity preservation, as they exactly contribute the discrete divergence terms which are absent in standard multidimensional DG schemes but crucial for ensuring the PP property. We also generalize the DDF condition of [53] to general meshes and derive sufficient conditions for achieving PP conservative schemes in the multiple dimensions.
4. We prove that the strong solution to the initial-value problem of the modified MHD system (4) preserves the positivity of density and pressure even if the divergence-free condition (2) is not satisfied. This feature, not enjoyed by the conservative system (1) (see [54]), can serve as a justification for designing provably PP multidimensional schemes based on the modified system (4).

The efforts mentioned above are novel and highly nontrivial. A key difficulty is to analytically quantify the relation of the PP property to the discrete divergence on general meshes. Especially, in the analysis of the positivity of $\mathcal{E}(\mathbf{U})$, the discrete equations for the conservative variables are nonlinearly coupled, and the limiting values of the numerical solution at the interfaces of each cell are intrinsically connected by the discrete divergence. These make the PP analysis in the MHD case very complicated especially in the multidimensional cases, and some standard analysis techniques (cf. [66]) are inapplicable as demonstrated in [53]. We will skillfully address these challenges by a novel equivalent form of the set $\mathcal{G}$ and highly technical estimates. Note that a LF flux can be considered as a special HLL flux. Therefore, all the analyses in the this paper directly apply to the local and global LF fluxes. It is also worth mentioning that many multi-state or multi-wave HLL-type fluxes were developed or applied to the ideal MHD in the literature (e.g., [4,8,9,28,32,34,38,41]), but only a few of them (cf. [9,32,41]) were shown to be PP for some 1D schemes. Moreover, their PP property for higher order schemes, in the multidimensional cases, and its relation to the divergence-free condition in the discrete sense have not yet been rigorously proved.

The paper is organized as follows. After establishing the auxiliary theories for our PP analysis on general meshes in Sect. 2, we present the 1D and multidimensional provably PP methods in Sects. 3 and 4, respectively. We conduct numerical tests in Sect. 5 to verify the PP property and the effectiveness of the proposed PP techniques, before concluding the paper in Sect. 6. The positivity of strong solutions of the modified MHD system (4) is shown in "Appendix A".

2 Auxiliary theories

In this section, we present the auxiliary results for our PP analysis on general meshes.

2.1 Properties of admissible state set

The function $\mathcal{E}(\mathbf{U})$ in (5) is nonlinear with respect to $\mathbf{U}$, complicating the analysis of the PP property of a given scheme. The following equivalent form of $\mathcal{G}$ was proposed in [53].

Lemma 1 *The admissible state set $\mathcal{G}$ is equivalent to*

$$\mathcal{G}_* = \left\{ \mathbf{U} = (\rho, \mathbf{m}, \mathbf{B}, E)^\top : \rho > 0,\ \mathbf{U} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} > 0,\ \forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3 \right\}, \tag{6}$$

where

$$\mathbf{n}^* = \left(\frac{|\mathbf{v}^*|^2}{2},\ -\mathbf{v}^*,\ -\mathbf{B}^*,\ 1 \right)^\top.$$

The proof of Lemma 1 can be found in [53]. As we can see, the equivalent set $\mathcal{G}_*$ is defined *with two constraints linear with respect to* $\mathbf{U}$, which give it advantages over

the natural definition (5) in showing the PP property of numerical schemes. This novel equivalent form is a cornerstone of our PP analysis.

The convexity of admissible state set is useful in bound-preserving analysis, as it helps reduce the complexity of the analysis if the scheme can be rewritten into certain convex combinations; see e.g., [52,57,66,68]. The convexity holds for $\mathcal{G}_*$, cf. [53].

Lemma 2 *The set $\mathcal{G}_*$ is convex. Moreover, $\lambda\mathbf{U}_1 + (1-\lambda)\mathbf{U}_0 \in \mathcal{G}_*$ for any $\mathbf{U}_1 \in \mathcal{G}_*$, $\mathbf{U}_0 \in \overline{\mathcal{G}}_*$ and $\lambda \in (0, 1]$, where $\overline{\mathcal{G}}_*$ is the closure of $\mathcal{G}_*$.*

2.2 Technical estimates relative to flux

2.2.1 Main estimates

We summarize our main estimate result in this subsection with the proof of it given later.

For the sake of convenience, we introduce the following notations, which will be frequently used in this paper. For any vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_d) \in \mathbb{R}^d$, we define the inner products

$$\langle \boldsymbol{\xi}, \mathbf{v} \rangle := \sum_{k=1}^{d} \xi_k v_k, \qquad \langle \boldsymbol{\xi}, \mathbf{B} \rangle := \sum_{k=1}^{d} \xi_k B_k, \qquad \langle \boldsymbol{\xi}, \mathbf{F} \rangle := \sum_{k=1}^{d} \xi_k \mathbf{F}_k.$$

For any unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$, define

$$\mathscr{C}(\mathbf{U}; \boldsymbol{\xi}) := \frac{1}{\sqrt{2}} \left[\mathscr{C}_s^2 + \frac{|\mathbf{B}|^2}{\rho} + \sqrt{\left(\mathscr{C}_s^2 + \frac{|\mathbf{B}|^2}{\rho} \right)^2 - 4 \frac{\mathscr{C}_s^2 \langle \boldsymbol{\xi}, \mathbf{B} \rangle^2}{\rho}} \right]^{\frac{1}{2}},$$

where $\mathscr{C}_s := \frac{p}{\rho\sqrt{2e}}$. Note that, for the ideal EOS, $\mathscr{C}_s = \sqrt{\frac{(\gamma-1)p}{2\rho}}$.

Recall that a technical inequality constructed in [53, Lemma 2.6] has played a pivotal role in the PP analysis on Cartesian meshes in [53,54]. That inequality involves two states, which correspond to the numerical solutions at a couple of symmetric quadrature points on cell interfaces. The cells of a general mesh are generally non-symmetric, so that the results in [53] are inapplicable to the present analysis. To carry out PP analysis on a general mesh, we need to construct a (general) "multi-state" inequality, which is derived in the following theorem.

Theorem 1 *For $1 \le j \le N$, let $s_j > 0$ and the unit vector $\boldsymbol{\xi}^{(j)} \in \mathbb{R}^d$ satisfy*

$$\sum_{j=1}^{N} s_j \boldsymbol{\xi}^{(j)} = \mathbf{0}. \tag{7}$$

Given N admissible states $\mathbf{U}^{(j)}$, $1 \le j \le N$, we define

$$\widehat{\alpha}_j := \max\left\{\langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^{(j)}\rangle, \frac{1}{\sum_{i=1}^N s_i}\sum_{i=1}^N s_i \left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \frac{\sqrt{\rho^{(j)}}\mathbf{v}^{(j)} + \sqrt{\rho^{(i)}}\mathbf{v}^{(i)}}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}}\right\rangle\right\} + \mathscr{C}(\mathbf{U}^{(j)}; \boldsymbol{\xi}^{(j)}) + \frac{2}{\sum_{i=1}^N s_i}\sum_{i=1}^N s_i \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}}. \tag{8}$$

Then for any $\alpha_j \ge \widehat{\alpha}_j$, the state

$$\overline{\mathbf{U}} := \frac{1}{\sum_{j=1}^N s_j\alpha_j}\sum_{j=1}^N s_j\left(\alpha_j\mathbf{U}^{(j)} - \left\langle \boldsymbol{\xi}^{(j)}, \mathbf{F}(\mathbf{U}^{(j)})\right\rangle\right), \tag{9}$$

belongs to $\mathcal{G}_\rho := \{\mathbf{U} = (\rho, \mathbf{m}, \mathbf{B}, E)^\top : \rho > 0\}$, and satisfies

$$\overline{\mathbf{U}} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \ge -\frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\sum_{j=1}^N s_j\alpha_j}\sum_{j=1}^N s_j\langle \boldsymbol{\xi}^{(j)}, \mathbf{B}^{(j)}\rangle, \quad \forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3. \tag{10}$$

Furthermore, $\overline{\mathbf{U}} \in \overline{\mathcal{G}}_*$ if

$$\sum_{j=1}^N s_j\langle \boldsymbol{\xi}^{(j)}, \mathbf{B}^{(j)}\rangle = 0. \tag{11}$$

The proof of Theorem 1 and the construction of the inequality (10) are highly nontrivial and technical. For better legibility, we put the proof in Sect. 2.2.2. Here, we would like to briefly explain the result in Theorem 1, whose meaning will become more clear in the PP analysis in Sects. 3 and 4. Let us consider a cell of the computational mesh, and assume it is a non-self-intersecting d-polytope with N edges ($d = 2$) or faces ($d = 3$). The index j on the variables in Theorem 1 represents the jth edge or face of the polytope, and s_j and $\boldsymbol{\xi}^{(j)}$ respectively correspond to the $(d-1)$-dimensional Hausdorff measure and the unit outward normal vector of the jth edge or face. One can verify that the condition (7) holds naturally. In addition, $\mathbf{U}^{(j)}$ stands for the approximate values of $\mathbf{U}$ on the jth edge or face. The condition (11) is actually a DDF condition over the polytope.

Remark 1 In Theorem 1, $\sum_{j=1}^N s_j\alpha_j$ is always positive, because

$$\sum_{j=1}^N s_j\widehat{\alpha}_j > \frac{1}{\sum_{i=1}^N s_i}\sum_{j=1}^N s_j\sum_{i=1}^N s_i\left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \frac{\sqrt{\rho^{(j)}}\mathbf{v}^{(j)} + \sqrt{\rho^{(i)}}\mathbf{v}^{(i)}}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}}\right\rangle = 0.$$

Remark 2 Theorem 1, particularly the inequality (10), clearly establishes a connection between the PP property and the discrete divergence of magnetic field, i.e., $\sum_{j=1}^N s_j\langle \boldsymbol{\xi}^{(j)}, \mathbf{B}^{(j)}\rangle$. This will be a key point of our PP analysis. The right-hand side

term of (10) is very important. The construction of this term is highly technical. If it is dropped, the inequality (10) would become invalid. As we will see, this term provides a way to take into account the discrete divergence in the PP analysis.

The following results are immediate corollaries of Theorem 1, which are useful for deriving PP numerical fluxes.

For any unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$, and any pair of admissible states $\mathbf{U}$ and $\tilde{\mathbf{U}}$, we define

$$\alpha_r(\mathbf{U}, \tilde{\mathbf{U}}; \boldsymbol{\xi}) := \max\left\{\langle \boldsymbol{\xi}, \mathbf{v}\rangle, \frac{\sqrt{\rho}\langle \boldsymbol{\xi}, \mathbf{v}\rangle + \sqrt{\tilde{\rho}}\langle \boldsymbol{\xi}, \tilde{\mathbf{v}}\rangle}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}\right\} + \mathscr{C}(\mathbf{U}; \boldsymbol{\xi}) + \frac{|\mathbf{B}-\tilde{\mathbf{B}}|}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}, \tag{12}$$

$$\alpha_l(\mathbf{U}, \tilde{\mathbf{U}}; \boldsymbol{\xi}) := \min\left\{\langle \boldsymbol{\xi}, \mathbf{v}\rangle, \frac{\sqrt{\rho}\langle \boldsymbol{\xi}, \mathbf{v}\rangle + \sqrt{\tilde{\rho}}\langle \boldsymbol{\xi}, \tilde{\mathbf{v}}\rangle}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}\right\} - \mathscr{C}(\mathbf{U}; \boldsymbol{\xi}) - \frac{|\mathbf{B}-\tilde{\mathbf{B}}|}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}, \tag{13}$$

and

$$\alpha_\star(\mathbf{U}, \tilde{\mathbf{U}}; \boldsymbol{\xi}) := \max\left\{|\langle \boldsymbol{\xi}, \mathbf{v}\rangle|, \left|\frac{\sqrt{\rho}\langle \boldsymbol{\xi}, \mathbf{v}\rangle + \sqrt{\tilde{\rho}}\langle \boldsymbol{\xi}, \tilde{\mathbf{v}}\rangle}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}\right|\right\} + \mathscr{C}(\mathbf{U}; \boldsymbol{\xi}) + \frac{|\mathbf{B}-\tilde{\mathbf{B}}|}{\sqrt{\rho}+\sqrt{\tilde{\rho}}}. \tag{14}$$

Corollary 1 *For any $\mathbf{U}, \tilde{\mathbf{U}} \in \mathcal{G}$, any unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$, and*

$$\forall \alpha \geq \alpha_r(\mathbf{U}, \tilde{\mathbf{U}}; \boldsymbol{\xi}), \quad \forall \tilde{\alpha} \leq \alpha_l(\tilde{\mathbf{U}}, \mathbf{U}; \boldsymbol{\xi}),$$

the state

$$\overline{\mathbf{U}} := \frac{1}{\alpha - \tilde{\alpha}}\Big(\alpha\mathbf{U} - \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U})\rangle - \tilde{\alpha}\tilde{\mathbf{U}} + \langle \boldsymbol{\xi}, \mathbf{F}(\tilde{\mathbf{U}})\rangle\Big),$$

belongs to $\mathcal{G}_\rho$ and satisfies

$$\overline{\mathbf{U}} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} + \frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\alpha - \tilde{\alpha}}\left(\langle \boldsymbol{\xi}, \mathbf{B}\rangle - \langle \boldsymbol{\xi}, \tilde{\mathbf{B}}\rangle\right) \geq 0, \quad \forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3. \tag{15}$$

Furthermore, if $\langle \boldsymbol{\xi}, \mathbf{B}\rangle - \langle \boldsymbol{\xi}, \tilde{\mathbf{B}}\rangle = 0$, then $\overline{\mathbf{U}} \in \overline{\mathcal{G}}_*$.

Proof This directly follows from Theorem 1 with $N = 2$, by taking

$$s_1 = s_2 = 1, \ \ \boldsymbol{\xi}^{(1)} = -\boldsymbol{\xi}^{(2)} = \boldsymbol{\xi}, \ \ \mathbf{U}^{(1)} = \mathbf{U}, \ \ \mathbf{U}^{(2)} = \tilde{\mathbf{U}}, \ \ \alpha_1 = \alpha, \ \ \alpha_2 = -\tilde{\alpha}.$$

□

Corollary 2 *Let $\mathbf{U}, \tilde{\mathbf{U}} \in \mathcal{G}$, unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$. For $\forall \alpha \geq \alpha_\star(\mathbf{U}, \tilde{\mathbf{U}}; \boldsymbol{\xi})$, $\forall \tilde{\alpha} \geq \alpha_\star(\tilde{\mathbf{U}}, \mathbf{U}; \boldsymbol{\xi})$, the state*

$$\overline{\mathbf{U}} := \frac{1}{\alpha + \tilde{\alpha}}\Big(\alpha\mathbf{U} - \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U})\rangle + \tilde{\alpha}\tilde{\mathbf{U}} + \langle \boldsymbol{\xi}, \mathbf{F}(\tilde{\mathbf{U}})\rangle\Big),$$

belongs to $\mathcal{G}_\rho$ *and satisfies*

$$\overline{\mathbf{U}} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} + \frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\alpha + \tilde{\alpha}} \left(\langle \boldsymbol{\xi}, \mathbf{B} \rangle - \langle \boldsymbol{\xi}, \tilde{\mathbf{B}} \rangle \right) \geq 0, \quad \forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3. \tag{16}$$

Furthermore, if $\langle \boldsymbol{\xi}, \mathbf{B} \rangle - \langle \boldsymbol{\xi}, \tilde{\mathbf{B}} \rangle = 0$, *then* $\overline{\mathbf{U}} \in \overline{\mathcal{G}}_*$.

Proof This is a direct consequence of Corollary 1. □

Remark 3 The inequalities (10), (15) and (16) extend the inequality constructed in [53, Lemma 2.6]. Corollaries 1 and 2 are useful for estimating the wave speeds to ensure the PP property of the HLL flux and local Lax–Friedrichs flux, respectively; see Theorem 2.

2.2.2 Proof of Theorem 1

We first establish several technical lemmas as the stepping stones on the path to prove Theorem 1.

For any $\mathbf{U} \in \mathcal{G}$ and $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$, we define the nonzero vector $\boldsymbol{\theta} \in \mathbb{R}^7$ by

$$\boldsymbol{\theta}(\mathbf{U}, \mathbf{v}^*, \mathbf{B}^*) := \frac{1}{\sqrt{2}} \left(\mathbf{B} - \mathbf{B}^*, \ \sqrt{\rho}(\mathbf{v} - \mathbf{v}^*), \ \sqrt{2\rho e} \right)^\top.$$

As a novel point, introducing such a vector will bring much convenience in the following estimates and analyses. It is easy to verify that

$$\mathbf{U} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} = |\boldsymbol{\theta}|^2. \tag{17}$$

Lemma 3 *The set*

$$\mathcal{G}_\rho := \left\{ \mathbf{U} = (\rho, \mathbf{m}, \mathbf{B}, E)^\top : \rho > 0 \right\},$$

is a convex set. And for any $\mathbf{U} \in \mathcal{G}_\rho$, $\boldsymbol{\xi} \in \mathbb{R}^d$ *and* $\alpha > \langle \boldsymbol{\xi}, \mathbf{v} \rangle$, *it holds*

$$\alpha \mathbf{U} - \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}) \rangle \in \mathcal{G}_\rho.$$

Proof The result can be easily verified. □

Lemma 4 *For any* $\mathbf{U} \in \mathcal{G}$, *any* $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$ *and all* $i \in \{1, 2, 3\}$, *we have*

$$\mathbf{F}_i(\mathbf{U}) \cdot \mathbf{n}^* - B_i(\mathbf{v}^* \cdot \mathbf{B}^*) \leq v_i \sum_{k=4}^{7} \theta_k^2 + v_i^* \left(\frac{1}{2} |\mathbf{B}|^2 - \mathbf{B} \cdot \mathbf{B}^* \right) + \mathscr{C}_i |\boldsymbol{\theta}|^2, \tag{18}$$

where $\mathscr{C}_i := \mathscr{C}(\mathbf{U}; \mathbf{e}_i)$, *and the vector* $\mathbf{e}_i$ *is the* i*-th row of the unit matrix of size 3.*

Proof For any $i \in \{1, 2, 3\}$, we observe that

$$\mathbf{F}_i(\mathbf{U}) \cdot \mathbf{n}^* - B_i(\mathbf{v}^* \cdot \mathbf{B}^*) = v_i \sum_{k=4}^{7} \theta_k^2 + v_i^* \left(\frac{1}{2}|\mathbf{B}|^2 - \mathbf{B} \cdot \mathbf{B}^* \right) + \Phi_i, \tag{19}$$

where

$$\Phi_i(\mathbf{U}, \mathbf{v}^*, \mathbf{B}^*) := p(v_i - v_i^*) + \sum_{\substack{1 \le k \le 3 \\ k \ne i}} \Big(B_k(v_i - v_i^*) - B_i(v_k - v_k^*) \Big)(B_k - B_k^*).$$

Let us show that Φ_i is bounded by $\mathscr{C}_i|\boldsymbol{\theta}|^2$ from above. We further observe that Φ_i is a quadratic form in the variables θ_k, $1 \le k \le 7$, and moreover, the coefficients of the quadratic form do not depend on $\mathbf{v}^*$ and $\mathbf{B}^*$. Specifically, for the fixed i, we have

$$p(v_i - v_i^*) = 2\mathscr{C}_s \frac{\sqrt{\rho}}{\sqrt{2}} (v_i - v_i^*) \sqrt{\rho e} = 2\mathscr{C}_s \theta_{3+i} \theta_7,$$

$$\Big(B_k(v_i - v_i^*) - B_i(v_k - v_k^*) \Big)(B_k - B_k^*) = 2 \frac{B_k}{\sqrt{\rho}} \theta_{3+i} \theta_k - 2 \frac{B_i}{\sqrt{\rho}} \theta_{3+k} \theta_k, \quad \forall k \ne i.$$

Define $i_1 := i \bmod 3 + 1$ and $i_2 := (i + 1) \bmod 3 + 1$, and

$$\tilde{\boldsymbol{\theta}} := \big(\theta_{3+i},\ \theta_{3+i_1},\ \theta_{3+i_2},\ \theta_{i_1},\ \theta_{i_2},\ \theta_7 \big)^\top,$$

then

$$\Phi_i = 2\mathscr{C}_s \theta_{3+i} \theta_7 + 2 \sum_{k \in \{i_1, i_2\}} \left(\frac{B_k}{\sqrt{\rho}} \theta_{3+i} \theta_k - \frac{B_i}{\sqrt{\rho}} \theta_{3+k} \theta_k \right) = \tilde{\boldsymbol{\theta}}^\top \mathbf{A} \tilde{\boldsymbol{\theta}},$$

where

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & B_{i_1} \rho^{-\frac{1}{2}} & B_{i_2} \rho^{-\frac{1}{2}} & \mathscr{C}_s \\ 0 & 0 & 0 & -B_i \rho^{-\frac{1}{2}} & 0 & 0 \\ 0 & 0 & 0 & 0 & -B_i \rho^{-\frac{1}{2}} & 0 \\ B_{i_1} \rho^{-\frac{1}{2}} & -B_i \rho^{-\frac{1}{2}} & 0 & 0 & 0 & 0 \\ B_{i_2} \rho^{-\frac{1}{2}} & 0 & -B_i \rho^{-\frac{1}{2}} & 0 & 0 & 0 \\ \mathscr{C}_s & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The spectral radius of $\mathbf{A}$ is $\mathscr{C}_i$. Therefore,

$$|\Phi_i| \le |\tilde{\boldsymbol{\theta}}^\top \mathbf{A} \tilde{\boldsymbol{\theta}}| \le \mathscr{C}_i |\tilde{\boldsymbol{\theta}}|^2 = \mathscr{C}_i (|\boldsymbol{\theta}|^2 - \theta_i^2) \le \mathscr{C}_i |\boldsymbol{\theta}|^2,$$

which along with the identity (19) imply (18). □

For any unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$, we introduce a matrix $\mathbf{T}_{\boldsymbol{\xi}} := \operatorname{diag}\big\{1, \widehat{\mathbf{T}}_{\boldsymbol{\xi}}, \widehat{\mathbf{T}}_{\boldsymbol{\xi}}, 1\big\}$, with the rotational matrix $\widehat{\mathbf{T}}_{\boldsymbol{\xi}}$ defined as follows:

(i) In $d = 1$, $\boldsymbol{\xi} = \xi$ is a scalar of value 1 or -1, and $\widehat{\mathbf{T}}_{\boldsymbol{\xi}}$ is defined as $\mathrm{diag}\{\xi, 1, 1\}$.
(ii) In $d = 2$, let $(\cos\varphi, \sin\varphi)$ be the polar coordinate representation of $\boldsymbol{\xi}$, and

$$\widehat{\mathbf{T}}_{\boldsymbol{\xi}} := \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(iii) In $d = 3$, let $(\sin\phi\cos\varphi, \sin\phi\sin\varphi, \cos\phi)$ be the spherical coordinate representation of $\boldsymbol{\xi}$, and

$$\widehat{\mathbf{T}}_{\boldsymbol{\xi}} := \begin{pmatrix} \sin\phi\cos\varphi & \sin\phi\sin\varphi & \cos\phi \\ -\sin\varphi & \cos\varphi & 0 \\ -\cos\phi\cos\varphi & -\cos\phi\sin\varphi & \sin\phi \end{pmatrix}.$$

The rotational invariance property of the d-dimensional MHD system (1) implies

$$\langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U})\rangle = \mathbf{T}_{\boldsymbol{\xi}}^{-1}\mathbf{F}_1(\mathbf{T}_{\boldsymbol{\xi}}\mathbf{U}). \tag{20}$$

This helps us extend Lemma 4 to the following general case.

Lemma 5 *For any $\mathbf{U} \in \mathcal{G}$, any $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$ and any unit vector $\boldsymbol{\xi} \in \mathbb{R}^d$, it holds*

$$\langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U})\rangle\cdot\mathbf{n}^* - \langle \boldsymbol{\xi}, \mathbf{B}\rangle(\mathbf{v}^*\cdot\mathbf{B}^*) \le \langle \boldsymbol{\xi}, \mathbf{v}\rangle \sum_{k=4}^{7}\theta_k^2 + \langle \boldsymbol{\xi}, \mathbf{v}^*\rangle\Big(\frac{1}{2}|\mathbf{B}|^2 - \mathbf{B}\cdot\mathbf{B}^*\Big) + \mathscr{C}(\mathbf{U}; \boldsymbol{\xi})|\boldsymbol{\theta}|^2.$$

Proof Let $\hat{\mathbf{U}} := \mathbf{T}_{\boldsymbol{\xi}}\mathbf{U}$, $\hat{\mathbf{v}}^* := \widehat{\mathbf{T}}_{\boldsymbol{\xi}}\mathbf{v}^*$, $\hat{\mathbf{B}} := \widehat{\mathbf{T}}_{\boldsymbol{\xi}}\mathbf{B}^*$, $\hat{\boldsymbol{\theta}} := \boldsymbol{\theta}(\hat{\mathbf{U}}, \hat{\mathbf{v}}^*, \hat{\mathbf{B}}^*)$, and

$$\hat{\mathbf{n}}^* := \Big(\frac{|\hat{\mathbf{v}}^*|^2}{2}, \ -\hat{\mathbf{v}}^*, \ -\hat{\mathbf{B}}^*, \ 1\Big)^\top = \mathbf{T}_{\boldsymbol{\xi}}\mathbf{n}^*.$$

By the definition (5), one can easily verify $\hat{\mathbf{U}} \in \mathcal{G}$, which, together with the orthogonality of $\mathbf{T}_{\boldsymbol{\xi}}^{-1}$ and $\widehat{\mathbf{T}}_{\boldsymbol{\xi}}^{-1}$, imply

$$\begin{aligned}
&\langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U})\rangle \cdot \mathbf{n}^* - \langle \boldsymbol{\xi}, \mathbf{B}\rangle(\mathbf{v}^*\cdot\mathbf{B}^*) \\
&\overset{(20)}{=} \big(\mathbf{T}_{\boldsymbol{\xi}}^{-1}\mathbf{F}_1(\hat{\mathbf{U}})\big)\cdot\big(\mathbf{T}_{\boldsymbol{\xi}}^{-1}\hat{\mathbf{n}}^*\big) - \hat{B}_1(\widehat{\mathbf{T}}_{\boldsymbol{\xi}}^{-1}\hat{\mathbf{v}}^*)\cdot(\widehat{\mathbf{T}}_{\boldsymbol{\xi}}^{-1}\hat{\mathbf{B}}^*) \\
&= \mathbf{F}_1(\hat{\mathbf{U}})\cdot\hat{\mathbf{n}}^* - \hat{B}_1(\hat{\mathbf{v}}^*\cdot\hat{\mathbf{B}}^*) \\
&\overset{(18)}{\le} \hat{v}_1\sum_{k=4}^{7}\hat{\theta}_k^2 + \hat{v}_1^*\Big(\frac{1}{2}|\hat{\mathbf{B}}|^2 - \hat{\mathbf{B}}\cdot\hat{\mathbf{B}}^*\Big) + \mathscr{C}_1(\hat{\mathbf{U}})|\hat{\boldsymbol{\theta}}|^2 \\
&= \langle \boldsymbol{\xi}, \mathbf{v}\rangle\sum_{k=4}^{7}\theta_k^2 + \langle \boldsymbol{\xi}, \mathbf{v}^*\rangle\Big(\frac{1}{2}|\mathbf{B}|^2 - \mathbf{B}\cdot\mathbf{B}^*\Big) + \mathscr{C}(\mathbf{U}; \boldsymbol{\xi})|\boldsymbol{\theta}|^2.
\end{aligned}$$

The proof is completed. □

Lemma 6 *Assume that* $\mathbf{U} = (\rho, \rho\mathbf{v}, \mathbf{B}, E)^\top \in \mathcal{G}$, $\tilde{\mathbf{U}} = (\tilde{\rho}, \tilde{\rho}\tilde{\mathbf{v}}, \tilde{\mathbf{B}}, \tilde{E})^\top \in \mathcal{G}$. *For* $\forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$, $\forall \boldsymbol{\xi} \in \mathbb{R}^d$ *and* $\forall \delta \in \mathbb{R}$, *it holds*

$$\begin{aligned}
&\langle \boldsymbol{\xi}, \mathbf{v}^* \rangle \left[\left(\frac{|\mathbf{B}|^2}{2} - \mathbf{B} \cdot \mathbf{B}^* \right) - \left(\frac{|\tilde{\mathbf{B}}|^2}{2} - \tilde{\mathbf{B}} \cdot \mathbf{B}^* \right) \right] \\
&\quad \le \langle \boldsymbol{\xi}, \delta\mathbf{v} + (1-\delta)\tilde{\mathbf{v}} \rangle \sum_{k=1}^{3} \left(\theta_k^2 - \tilde{\theta}_k^2 \right) + |\boldsymbol{\xi}| f(\mathbf{U}, \tilde{\mathbf{U}}; \delta) \left(|\boldsymbol{\theta}|^2 + |\tilde{\boldsymbol{\theta}}|^2 \right),
\end{aligned} \tag{21}$$

where $\boldsymbol{\theta} := \boldsymbol{\theta}(\mathbf{U}, \mathbf{v}^*, \mathbf{B}^*)$ *and* $\tilde{\boldsymbol{\theta}} := \boldsymbol{\theta}(\tilde{\mathbf{U}}, \mathbf{v}^*, \mathbf{B}^*)$, *and* $f(\mathbf{U}, \tilde{\mathbf{U}}; \delta)$ *is defined by*

$$f(\mathbf{U}, \tilde{\mathbf{U}}; \delta) := \frac{|\tilde{\mathbf{B}} - \mathbf{B}|}{\sqrt{2}} \sqrt{\frac{\delta^2}{\rho} + \frac{(1-\delta)^2}{\tilde{\rho}}}. \tag{22}$$

Proof With the aid of the Cauchy–Schwarz inequality, we have

$$\begin{aligned}
&\langle \boldsymbol{\xi}, \mathbf{v}^* \rangle \left[\left(\frac{|\mathbf{B}|^2}{2} - \mathbf{B} \cdot \mathbf{B}^* \right) - \left(\frac{|\tilde{\mathbf{B}}|^2}{2} - \tilde{\mathbf{B}} \cdot \mathbf{B}^* \right) \right] - \langle \boldsymbol{\xi}, \delta\mathbf{v} + (1-\delta)\tilde{\mathbf{v}} \rangle \sum_{k=1}^{3} \left(\theta_k^2 - \tilde{\theta}_k^2 \right) \\
&= \left(\frac{\delta}{2} \langle \boldsymbol{\xi}, \mathbf{v} - \mathbf{v}^* \rangle + \frac{1-\delta}{2} \langle \boldsymbol{\xi}, \tilde{\mathbf{v}} - \mathbf{v}^* \rangle \right) (\tilde{\mathbf{B}} - \mathbf{B}) \cdot \left(\mathbf{B} + \tilde{\mathbf{B}} - 2\mathbf{B}^* \right) \\
&\le \frac{|\boldsymbol{\xi}|}{2} \left(\frac{|\delta|}{\sqrt{\rho}} \sqrt{\rho} |\mathbf{v} - \mathbf{v}^*| + \frac{|1-\delta|}{\sqrt{\tilde{\rho}}} \sqrt{\tilde{\rho}} |\tilde{\mathbf{v}} - \mathbf{v}^*| \right) |\tilde{\mathbf{B}} - \mathbf{B}| \left(|\mathbf{B} - \mathbf{B}^*| + |\tilde{\mathbf{B}} - \mathbf{B}^*| \right) \\
&\le \frac{|\boldsymbol{\xi}|}{2} \sqrt{\frac{\delta^2}{\rho} + \frac{(1-\delta)^2}{\tilde{\rho}}} \sqrt{\rho |\mathbf{v} - \mathbf{v}^*|^2 + \tilde{\rho} |\tilde{\mathbf{v}} - \mathbf{v}^*|^2} |\tilde{\mathbf{B}} - \mathbf{B}| \sqrt{2 \left(|\mathbf{B} - \mathbf{B}^*|^2 + |\tilde{\mathbf{B}} - \mathbf{B}^*|^2 \right)} \\
&= 2|\boldsymbol{\xi}| f(\mathbf{U}, \tilde{\mathbf{U}}; \delta) \sqrt{\sum_{k=4}^{6} \left(\theta_k^2 + \tilde{\theta}_k^2 \right)} \sqrt{\sum_{k=1}^{3} \left(\theta_k^2 + \tilde{\theta}_k^2 \right)} \\
&\le |\boldsymbol{\xi}| f(\mathbf{U}, \tilde{\mathbf{U}}; \delta) \sum_{k=1}^{6} \left(\theta_k^2 + \tilde{\theta}_k^2 \right) \le |\boldsymbol{\xi}| f(\mathbf{U}, \tilde{\mathbf{U}}; \delta) \left(|\boldsymbol{\theta}|^2 + |\tilde{\boldsymbol{\theta}}|^2 \right).
\end{aligned}$$

The proof is completed. □

We are now ready to prove Theorem 1.

Proof Note that $\alpha_j \ge \widehat{\alpha}_j > \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^{(j)} \rangle$. It follows from Lemma 3 that $\alpha_j \mathbf{U}^{(j)} - \langle \boldsymbol{\xi}^{(j)}, \mathbf{F}(\mathbf{U}^{(j)}) \rangle \in \mathcal{G}_\rho$, and furthermore $\overline{\mathbf{U}} \in \mathcal{G}_\rho$, by noting that $\sum_{j=1}^{N} s_j \alpha_j > 0$ (see Remark 1).

We then focus on proving the inequality (10), or equivalently,

$$\sum_{j=1}^{N} s_j \Pi^{(j)} \le \sum_{j=1}^{N} \alpha_j |\boldsymbol{\theta}^{(j)}|^2, \tag{23}$$

where $\boldsymbol{\theta}^{(j)} := \boldsymbol{\theta}(\mathbf{U}^{(j)}, \mathbf{v}^*, \mathbf{B}^*)$, and

$$\Pi^{(j)} := \langle \boldsymbol{\xi}^{(j)}, \mathbf{F}(\mathbf{U}^{(j)}) \rangle \cdot \mathbf{n}^* - \langle \boldsymbol{\xi}^{(j)}, \mathbf{B}^{(j)} \rangle (\mathbf{v}^* \cdot \mathbf{B}^*).$$

Using Lemma 5 gives

$$\begin{aligned}\sum_{j=1}^{N} s_j \Pi^{(j)} \le & \left\{ \sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^{(j)} \rangle \sum_{k=4}^{7} \left| \theta_k^{(j)} \right|^2 \right\} + \left\{ \sum_{j=1}^{N} s_j \mathscr{C}(\mathbf{U}^{(j)}; \boldsymbol{\xi}) |\boldsymbol{\theta}^{(j)}|^2 \right\} \\ & + \left\{ \sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \rangle \left(\frac{1}{2} |\mathbf{B}^{(j)}|^2 - \mathbf{B}^{(j)} \cdot \mathbf{B}^* \right) \right\} \\ =: & \, \Pi_1 + \Pi_2 + \Pi_3. \end{aligned} \tag{24}$$

Noting that, for any $1 \le i \le N$, the hypothesis (7) implies

$$\sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \rangle = \left\langle \sum_{j=1}^{N} s_j \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \right\rangle = 0.$$

Thus we can reformulate Π_3 as

$$\begin{aligned}\Pi_3 &= \sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \rangle \left(\frac{1}{2} |\mathbf{B}^{(j)}|^2 - \mathbf{B}^{(j)} \cdot \mathbf{B}^* \right) - \sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \rangle \left(\frac{1}{2} |\mathbf{B}^{(i)}|^2 - \mathbf{B}^{(i)} \cdot \mathbf{B}^* \right) \\ &= \sum_{j=1}^{N} s_j \langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^* \rangle \left[\left(\frac{1}{2} |\mathbf{B}^{(j)}|^2 - \mathbf{B}^{(j)} \cdot \mathbf{B}^* \right) - \left(\frac{1}{2} |\mathbf{B}^{(i)}|^2 - \mathbf{B}^{(i)} \cdot \mathbf{B}^* \right) \right] =: \sum_{j=1}^{N} s_j \Pi_3^{(ji)}, \end{aligned}$$

for any $1 \le i \le N$. For any $\delta \in \mathbb{R}$, it follows from Lemma 6 that

$$\begin{aligned}\Pi_3^{(ji)} \le & \langle \boldsymbol{\xi}^{(j)}, \delta \mathbf{v}^{(j)} + (1-\delta) \mathbf{v}^{(i)} \rangle \sum_{k=1}^{3} \left(|\theta_k^{(j)}|^2 - |\theta_k^{(i)}|^2 \right) \\ & + f(\mathbf{U}^{(j)}, \mathbf{U}^{(i)}; \delta) \left(|\boldsymbol{\theta}^{(j)}|^2 + |\boldsymbol{\theta}^{(i)}|^2 \right). \end{aligned} \tag{25}$$

In particular, we take the free variable δ as $\sqrt{\rho^{(j)}} / \left(\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}} \right)$, which gives the Roe-type weighted average. Let

$$\bar{\mathbf{v}}^{(ji)} := \frac{\sqrt{\rho^{(j)}} \mathbf{v}^{(j)} + \sqrt{\rho^{(i)}} \mathbf{v}^{(i)}}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}},$$

then the inequality (25) becomes

$$\Pi_3^{(ji)} \le \left\langle \boldsymbol{\xi}^{(j)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \sum_{k=1}^{3} \left(|\theta_k^{(j)}|^2 - |\theta_k^{(i)}|^2 \right) + \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} \left(|\boldsymbol{\theta}^{(j)}|^2 + |\boldsymbol{\theta}^{(i)}|^2 \right). \tag{26}$$

It follows that

$$\begin{aligned}\left(\sum_{i=1}^{N} s_i\right)\Pi_3 &= \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \Pi_3^{(ji)} \\ &\le \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \left\langle \boldsymbol{\xi}^{(j)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \sum_{k=1}^{3} \left(|\theta_k^{(j)}|^2 - |\theta_k^{(i)}|^2 \right) \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} \left(|\boldsymbol{\theta}^{(j)}|^2 + |\boldsymbol{\theta}^{(i)}|^2 \right). \end{aligned} \tag{27}$$

By $\bar{\mathbf{v}}^{(ji)} = \bar{\mathbf{v}}^{(ij)}$ and the technique of exchanging indexes i and j, we obtain

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \left\langle \boldsymbol{\xi}^{(j)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \sum_{k=1}^{3} |\theta_k^{(i)}|^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \left\langle \boldsymbol{\xi}^{(i)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \sum_{k=1}^{3} |\theta_k^{(j)}|^2, \\ &\sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} |\boldsymbol{\theta}^{(i)}|^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} |\boldsymbol{\theta}^{(j)}|^2. \end{aligned}$$

Therefore, the inequality (27) can be rewritten as

$$\begin{aligned}\left(\sum_{i=1}^{N} s_i\right)\Pi_3 &\le \sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \sum_{k=1}^{3} |\theta_k^{(j)}|^2 \\ &\quad + 2\sum_{i=1}^{N}\sum_{j=1}^{N} s_i s_j \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} |\boldsymbol{\theta}^{(j)}|^2, \end{aligned}$$

which further yields

$$\begin{aligned}\Pi_3 &\le \sum_{j=1}^{N} s_j \left(\frac{1}{\sum_{i=1}^{N} s_i} \sum_{i=1}^{N} s_i \left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \right) \sum_{k=1}^{3} |\theta_k^{(j)}|^2 \\ &\quad + \sum_{j=1}^{N} s_j \left(\frac{2}{\sum_{i=1}^{N} s_i} \sum_{i=1}^{N} s_i \frac{|\mathbf{B}^{(j)} - \mathbf{B}^{(i)}|}{\sqrt{\rho^{(j)}} + \sqrt{\rho^{(i)}}} \right) |\boldsymbol{\theta}^{(j)}|^2. \end{aligned} \tag{28}$$

Note that

$$\begin{aligned}&\Pi_1 + \sum_{j=1}^{N} s_j \left(\frac{1}{\sum_{i=1}^{N} s_i} \sum_{i=1}^{N} s_i \left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \right) \sum_{k=1}^{3} |\theta_k^{(j)}|^2 \\ &\le \sum_{j=1}^{N} s_j \max\left\{ \left\langle \boldsymbol{\xi}^{(j)}, \mathbf{v}^{(j)} \right\rangle, \frac{1}{\sum_{i=1}^{N} s_i} \sum_{i=1}^{N} s_i \left\langle \boldsymbol{\xi}^{(j)} - \boldsymbol{\xi}^{(i)}, \bar{\mathbf{v}}^{(ji)} \right\rangle \right\} \sum_{k=1}^{7} |\theta_k^{(j)}|^2, \end{aligned}$$

which along with (24) and (28) imply

$$\sum_{j=1}^{N} s_j \Pi^{(j)} \le \sum_{j=1}^{N} \widehat{\alpha}_j |\boldsymbol{\theta}^{(j)}|^2 \le \sum_{j=1}^{N} \alpha_j |\boldsymbol{\theta}^{(j)}|^2.$$

Hence the inequality (23) holds.

Under the condition (11), the inequality (10) becomes $\overline{\mathbf{U}} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \ge 0$, $\forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$, which together with $\overline{\mathbf{U}} \in \mathcal{G}_\rho$ imply $\overline{\mathbf{U}} \in \overline{\mathcal{G}}_*$. The proof is completed. □

2.3 Estimates relative to source term

We also need the following lemma, which was proposed in [54].

Lemma 7 *For any* $\mathbf{U} \in \mathcal{G}$ *and any* $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$, *we have*

$$\mathbf{S}(\mathbf{U}) \cdot \mathbf{n}^* = (\mathbf{v} - \mathbf{v}^*) \cdot (\mathbf{B} - \mathbf{B}^*) - \mathbf{v}^* \cdot \mathbf{B}^*, \tag{29}$$

$$|\sqrt{\rho}(\mathbf{v} - \mathbf{v}^*) \cdot (\mathbf{B} - \mathbf{B}^*)| < \mathbf{U} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}. \tag{30}$$

Furthermore, for any $b \in \mathbb{R}$, *it holds*

$$-b(\mathbf{S}(\mathbf{U}) \cdot \mathbf{n}^*) \ge b(\mathbf{v}^* \cdot \mathbf{B}^*) - \frac{|b|}{\sqrt{\rho}} \left(\mathbf{U} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \right). \tag{31}$$

2.4 Properties of the HLL flux

The Harten–Lax–van Leer (HLL) flux is derived from an approximate Riemann solver in the direction normal to each cell interface. Let $\boldsymbol{\xi} \in \mathbb{R}^d$ be the unit normal vector of the interface. Then the HLL flux at the interface is given by

$$\hat{\mathbf{F}}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) = \begin{cases} \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^-) \rangle, & 0 \le \sigma_l < \sigma_r, \\ \dfrac{\sigma_r \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^-) \rangle - \sigma_l \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^+) \rangle + \sigma_l \sigma_r (\mathbf{U}^+ - \mathbf{U}^-)}{\sigma_r - \sigma_l}, & \sigma_l < 0 < \sigma_r, \\ \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^+) \rangle, & \sigma_l < \sigma_r \le 0. \end{cases} \tag{32}$$

Here $\sigma_l(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi})$ and $\sigma_r(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi})$ are functions of $\mathbf{U}^-$, $\mathbf{U}^+$ and $\boldsymbol{\xi}$, denoting the estimates of the leftmost and rightmost wave speeds in the (rotated) Riemann problem in the direction of $\boldsymbol{\xi}$, where $\mathbf{U}^-$ and $\mathbf{U}^+$ are the left and right initial states respectively. We require $\sigma_r > \sigma_l$, and

$$\sigma_r(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) = -\sigma_l(\mathbf{U}^+, \mathbf{U}^-; -\boldsymbol{\xi}), \tag{33}$$

which ensures that the numerical flux (32) is conservative, that is,

$$\hat{\mathbf{F}}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) + \hat{\mathbf{F}}(\mathbf{U}^+, \mathbf{U}^-; -\boldsymbol{\xi}) = 0.$$

Let

$$\sigma^+ = \max\{\sigma_r, 0\}, \quad \sigma^- = \min\{\sigma_l, 0\},$$

then the flux (32) can be reformulated as

$$\hat{\mathbf{F}}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) = \frac{\sigma^+ \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^-)\rangle - \sigma^- \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^+)\rangle + \sigma^- \sigma^+ (\mathbf{U}^+ - \mathbf{U}^-)}{\sigma^+ - \sigma^-}. \tag{34}$$

Note that the LF flux can be considered as a special HLL flux with $\sigma_r = -\sigma_l = \sigma$, where σ is the maximum wave speed. Therefore, all the analysis in the present paper also applies to the local LF flux and global LF flux.

The following property is derived for the HLL flux (32) in the ideal MHD case.

Theorem 2 *Assume* $\mathbf{U}^-, \mathbf{U}^+ \in \mathcal{G}$. *If the parameters (approximate wave speeds) in the HLL flux* (32) *satisfy*

$$\sigma_r \geq \alpha_r(\mathbf{U}^+, \mathbf{U}^-; \boldsymbol{\xi}), \qquad \sigma_l \leq \alpha_l(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}), \tag{35}$$

then

$$\hat{\mathbf{F}}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) = \sigma^- \mathbf{H}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) + \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^-)\rangle - \sigma^- \mathbf{U}^-, \tag{36}$$

$$\hat{\mathbf{F}}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) = \sigma^+ \mathbf{H}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) + \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^+)\rangle - \sigma^+ \mathbf{U}^+, \tag{37}$$

and the intermediate state

$$\mathbf{H}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) := \frac{1}{\sigma^+ - \sigma^-}\Big(\sigma^+ \mathbf{U}^+ - \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^+)\rangle - \sigma^- \mathbf{U}^- + \langle \boldsymbol{\xi}, \mathbf{F}(\mathbf{U}^-)\rangle\Big) \tag{38}$$

belongs to $\mathcal{G}_\rho$ *and satisfies*

$$\mathbf{H} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} + \frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\sigma^+ - \sigma^-}\big(\langle \boldsymbol{\xi}, \mathbf{B}^+\rangle - \langle \boldsymbol{\xi}, \mathbf{B}^-\rangle\big) \geq 0, \quad \forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3. \tag{39}$$

Furthermore, if $\langle \boldsymbol{\xi}, \mathbf{B}^+\rangle = \langle \boldsymbol{\xi}, \mathbf{B}^-\rangle$, *then* $\mathbf{H} \in \overline{\mathcal{G}}_*$.

Proof The identities (36)–(37) can be verified by using (34). Under the condition (35), we have

$$\sigma^+ \geq \sigma_r \geq \alpha_r(\mathbf{U}^+, \mathbf{U}^-; \boldsymbol{\xi}), \quad \sigma^- \leq \sigma_l \leq \alpha_l(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}).$$

It follows from Corollary 1 that $\mathbf{H}(\mathbf{U}^-, \mathbf{U}^+; \boldsymbol{\xi}) \in \mathcal{G}_\rho$ and satisfies (39). □

Remark 4 It is observed from (39) that the admissibility of the intermediate state $\mathbf{H}$ is closely related to the jump in the normal magnetic field across the cell interface. If the jump is zero, then $\mathbf{H} \in \overline{\mathcal{G}}_*$; otherwise, $\mathbf{H}$ does not always belong to $\overline{\mathcal{G}}_*$ even if many times larger wave speeds are employed. However, in the multidimensional cases, a

standard finite volume or DG method cannot avoid jumps in normal magnetic field at cell interfaces although such jumps do not exist in the exact solution. This causes some challenges essentially different from 1D case. We will demonstrate that this issue can be overcome by coupling two divergence-controlling techniques: the locally divergence-free element and properly discretized Godunov–Powell source term. The former technique leads to zero divergence within each cell, while the latter controls the divergence error across cell interfaces.

Remark 5 The proposed condition (35) for the wave speeds σ_l and σ_r is crucial for the provably PP property of our schemes presented later. The condition (35) is acceptable, because α_l and α_r are respectively close to the minimum and maximum signal speeds of the system (4) in the direction of $\boldsymbol{\xi}$. Let σ_l^{std} and σ_r^{std} denote a standard choice of wave speeds in the HLL flux, for example, Davis [19] gave those speeds as

$$\sigma_l^{\text{std}} = \min\{\lambda_1(\mathbf{U}^-;\boldsymbol{\xi}), \lambda_1(\mathbf{U}^+;\boldsymbol{\xi})\}, \quad \sigma_r^{\text{std}} = \max\{\lambda_8(\mathbf{U}^-;\boldsymbol{\xi}), \lambda_8(\mathbf{U}^+;\boldsymbol{\xi})\}, \tag{40}$$

or Einfeldt et al. [25] suggested to use

$$\sigma_l^{\text{std}} = \min\{\lambda_1(\mathbf{U}^-;\boldsymbol{\xi}), \lambda_1(\mathbf{U}^{\text{Roe}};\boldsymbol{\xi})\}, \quad \sigma_r^{\text{std}} = \max\{\lambda_8(\mathbf{U}^+;\boldsymbol{\xi}), \lambda_8(\mathbf{U}^{\text{Roe}};\boldsymbol{\xi})\},$$

where $\lambda_1(\mathbf{U};\boldsymbol{\xi})$ amd $\lambda_8(\mathbf{U};\boldsymbol{\xi})$ are the minimum and maximum eigenvalues of the Jacobi matrix of the system (4) in the direction of $\boldsymbol{\xi}$, and $\lambda_i(\mathbf{U}^{\text{Roe}};\boldsymbol{\xi})$ is the estimate of eigenvalues based on the Roe matrix (cf. [43]). These choices may not necessarily give a PP flux in the MHD case and probably not satisfy (35). In practice, by considering the stability and the PP property, we suggest to use

$$\sigma_l = \min\{\alpha_l(\mathbf{U}^-, \mathbf{U}^+;\boldsymbol{\xi}), \sigma_l^{\text{std}}\}, \quad \sigma_r = \max\{\alpha_r(\mathbf{U}^+, \mathbf{U}^-;\boldsymbol{\xi}), \sigma_r^{\text{std}}\} \tag{41}$$

in the HLL flux, and use

$$\sigma_r = -\sigma_l = \max\left\{\alpha_\star(\mathbf{U}^-, \mathbf{U}^+;\boldsymbol{\xi}), \alpha_\star(\mathbf{U}^+, \mathbf{U}^-;\boldsymbol{\xi}), \sigma^{\text{std}}\right\},$$

in the local LF flux, where σ^{std} denotes a standard numerical viscosity parameter for the local LF flux.

3 Positivity-preserving schemes in one dimension

In this section, we propose provably PP finite volume and DG schemes with the proposed HLL flux for 1D MHD equations (1). Let x denote the spatial variable. The condition (2) and the fifth equation of (1) imply $B_1(x,t) \equiv$ constant (denoted by $\mathtt{B_{const}}$) for all x and $t \geq 0$.

Let $I_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $I = \cup_j I_j$ be a partition of the spatial domain. Denote $\Delta x_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. Let $\{t_0 = 0, t_{n+1} = t_n + \Delta t_n, n \geq 0\}$ be a partition of the time interval $[0, T]$, where the time step-size Δt_n is determined by some CFL condition. Let $\bar{\mathbf{U}}_j^n$ denote the numerical approximation to the cell average of the exact solution

$\mathbf{U}(x,t)$ over I_j at $t=t_n$. We would like to seek PP schemes with $\bar{\mathbf{U}}_j^n$ always preserved in the admissible state set $\mathcal{G}$.

3.1 First-order scheme

We consider the 1D first-order scheme

$$\bar{\mathbf{U}}_j^{n+1} = \bar{\mathbf{U}}_j^n - \frac{\Delta t_n}{\Delta x_j}\Big(\hat{\mathbf{F}}_1(\bar{\mathbf{U}}_j^n, \bar{\mathbf{U}}_{j+1}^n) - \hat{\mathbf{F}}_1(\bar{\mathbf{U}}_{j-1}^n, \bar{\mathbf{U}}_j^n)\Big), \tag{42}$$

where $\hat{\mathbf{F}}_1(\bar{\mathbf{U}}_j^n, \bar{\mathbf{U}}_{j+1}^n) := \hat{\mathbf{F}}(\bar{\mathbf{U}}_j^n, \bar{\mathbf{U}}_{j+1}^n; 1)$ is taken as the HLL flux in (34). It is worth noting that in the 1D case, since $B_1(x,t) \equiv$ constant, the Godunov–Powell source term does not exist.

Theorem 3 *Assume that $\bar{\mathbf{U}}_j^0 \in \mathcal{G}$ and $\bar{B}_{1,j}^0 = B_{const}$ for all j, and the wave speeds in the HLL flux satisfy (35). Then the state $\bar{\mathbf{U}}_j^n$, computed by the scheme (42) under the CFL condition*

$$\left(\sigma_{j-\frac{1}{2}}^{n,+} - \sigma_{j+\frac{1}{2}}^{n,-}\right)\frac{\Delta t_n}{\Delta x_j} < 1, \quad \forall j, \tag{43}$$

belongs to $\mathcal{G}$ and satisfies $\bar{B}_{1,j}^n = B_{const}$ for all j and $n \in \mathbb{N}$, where

$$\sigma_{j-\frac{1}{2}}^{n,+} := \sigma^+(\bar{\mathbf{U}}_{j-1}^n, \bar{\mathbf{U}}_j^n; 1), \qquad \sigma_{j+\frac{1}{2}}^{n,-} := \sigma^-(\bar{\mathbf{U}}_j^n, \bar{\mathbf{U}}_{j+1}^n; 1).$$

Proof Here the induction argument is used for the time level number n. It is obvious that the conclusion holds for $n=0$ under the hypothesis on the initial data. We now assume that $\bar{\mathbf{U}}_j^n \in \mathcal{G}$ with $\bar{B}_{1,j}^n = \mathrm{B}_{\mathrm{const}}$ for all j, and we check whether the conclusion holds for $n+1$. Let $\lambda := \Delta t_n/\Delta x_j$, and $\mathbf{H}_{j+\frac{1}{2}}^n := \mathbf{H}(\bar{\mathbf{U}}_j^n, \bar{\mathbf{U}}_{j+1}^n; 1)$; see (38) for the definition of $\mathbf{H}$. Under the induction hypothesis, we have that $\mathbf{H}_{j+\frac{1}{2}}^n \in \overline{\mathcal{G}}_*, \forall j$ according to Theorem 2, and the fifth component of $\mathbf{H}_{j+\frac{1}{2}}^n$ is $\mathrm{B}_{\mathrm{const}}$ for all j by noting that the fifth component of $\mathbf{F}_1$ is zero. Using the identities (36) and (37), one can rewrite the scheme (42) as

$$\begin{aligned}
\bar{\mathbf{U}}_j^{n+1} &= \bar{\mathbf{U}}_j^n - \lambda\Big[\Big(\sigma_{j+\frac{1}{2}}^{n,-}\mathbf{H}_{j+\frac{1}{2}}^n + \mathbf{F}_1(\bar{\mathbf{U}}_j^n) - \sigma_{j+\frac{1}{2}}^{n,-}\bar{\mathbf{U}}_j^n\Big) \\
&\quad - \Big(\sigma_{j-\frac{1}{2}}^{n,+}\mathbf{H}_{j-\frac{1}{2}}^n + \mathbf{F}_1(\bar{\mathbf{U}}_j^n) - \sigma_{j-\frac{1}{2}}^{n,+}\bar{\mathbf{U}}_j^n\Big)\Big] \\
&= \Big(1 + \lambda\Big(\sigma_{j+\frac{1}{2}}^{n,-} - \sigma_{j-\frac{1}{2}}^{n,+}\Big)\Big)\bar{\mathbf{U}}_j^n + \Big(-\lambda\sigma_{j+\frac{1}{2}}^{n,-}\Big)\mathbf{H}_{j+\frac{1}{2}}^n + \lambda\sigma_{j-\frac{1}{2}}^{n,+}\mathbf{H}_{j-\frac{1}{2}}^n.
\end{aligned} \tag{44}$$

Under the condition (43), $\bar{\mathbf{U}}_j^{n+1}$ is a convex combination of $\bar{\mathbf{U}}_j^n$, $\mathbf{H}_{j+\frac{1}{2}}^n$ and $\mathbf{H}_{j-\frac{1}{2}}^n$. Hence we have $\bar{\mathbf{U}}_j^{n+1} \in \mathcal{G}$ by Lemma 2. The fifth equation of (44) also implies

$$\bar{B}_{1,j}^{n+1} = \left(1+\lambda\left(\sigma_{j+\frac{1}{2}}^{n,-} - \sigma_{j-\frac{1}{2}}^{n,+}\right)\right)\mathtt{B}_{\mathtt{const}} - \lambda\sigma_{j+\frac{1}{2}}^{n,-}\mathtt{B}_{\mathtt{const}} + \lambda\sigma_{j-\frac{1}{2}}^{n,+}\mathtt{B}_{\mathtt{const}} = \mathtt{B}_{\mathtt{const}}.$$

Therefore, the conclusion holds for $n+1$. The proof is completed. □

3.2 High-order schemes

For convenience, we first focus on the forward Euler method for time discretization and will discuss the high-order time discretization later. We consider the high-order finite volume schemes as well as the scheme satisfied by the cell averages of a standard DG method for (1), which have the following form

$$\bar{\mathbf{U}}_j^{n+1} = \bar{\mathbf{U}}_j^n - \frac{\Delta t_n}{\Delta x_j}\left(\hat{\mathbf{F}}_1\left(\mathbf{U}_{j+\frac{1}{2}}^-, \mathbf{U}_{j+\frac{1}{2}}^+\right) - \hat{\mathbf{F}}_1\left(\mathbf{U}_{j-\frac{1}{2}}^-, \mathbf{U}_{j-\frac{1}{2}}^+\right)\right), \tag{45}$$

where $\hat{\mathbf{F}}_1(\mathbf{U}_{j+\frac{1}{2}}^-, \mathbf{U}_{j+\frac{1}{2}}^+) := \hat{\mathbf{F}}(\mathbf{U}_{j+\frac{1}{2}}^-, \mathbf{U}_{j+\frac{1}{2}}^+; 1)$ is taken as the HLL flux in (34). The quantities $\mathbf{U}_{j+\frac{1}{2}}^-$ and $\mathbf{U}_{j+\frac{1}{2}}^+$ are the high-order approximations of the point values $\mathbf{U}(x_{j+\frac{1}{2}}, t_n)$ within the cells I_j and I_{j+1}, respectively, computed by

$$\mathbf{U}_{j+\frac{1}{2}}^- = \mathbf{U}_j^n\left(x_{j+\frac{1}{2}} - 0\right), \quad \mathbf{U}_{j+\frac{1}{2}}^+ = \mathbf{U}_{j+1}^n\left(x_{j+\frac{1}{2}} + 0\right). \tag{46}$$

Here the function $\mathbf{U}_j^n(x)$ is a polynomial vector of degree k with the cell-averaged value of $\bar{\mathbf{U}}_j^n$. It approximates $\mathbf{U}(x, t_n)$ within the cell I_j, and is either reconstructed in the finite volume methods from $\{\bar{\mathbf{U}}_j^n\}$ or directly evolved in the DG methods. The evolution equations for the high-order "moments" of $\mathbf{U}_j^n(x)$ in the DG methods are omitted because here we are only concerned with the PP property of the schemes.

If the polynomial degree $k = 0$, i.e., $\mathbf{U}_j^n(x) = \bar{\mathbf{U}}_j^n$, $\forall x \in I_j$, then the scheme (45) reduces to the first-order scheme (42), which has been proved to be PP under the CFL condition (43).

When the polynomial degree $k \geq 1$, the solution $\bar{\mathbf{U}}_j^{n+1}$ of the high-order scheme (45) does not always belong to $\mathcal{G}$ even if $\bar{\mathbf{U}}_j^n \in \mathcal{G}$ for all j. In the following theorem, we give a satisfiable condition for achieving the provably PP property of the scheme (45) when $k \geq 1$.

Let $\{\widehat{x}_j^{(\mu)}\}_{\mu=1}^{\mathtt{L}}$ be the $\mathtt{L}$-point Gauss–Lobatto quadrature nodes in the interval I_j, and the associated weights denoted by $\{\widehat{\omega}_\mu\}_{\mu=1}^{\mathtt{L}}$ with $\sum_{\mu=1}^{\mathtt{L}} \widehat{\omega}_\mu = 1$. We require $2\mathtt{L} - 3 \geq k$ such that the algebraic precision of corresponding quadrature is at least k, for example, one can particularly take $\mathtt{L} = \lceil \frac{k+3}{2} \rceil$.

Theorem 4 *Let the wave speeds in the HLL flux satisfy (35). If the polynomial vectors* $\{\mathbf{U}_j^n(x)\}$ *satisfy*

$$B^{\pm}_{1,j+\frac{1}{2}} = \mathtt{B_{const}}, \quad \forall j, \tag{47}$$

$$\mathbf{U}_j^n(\widehat{x}_j^{(\mu)}) \in \mathcal{G}, \quad \forall \mu \in \{1, 2, \ldots, \mathtt{L}\}, \ \forall j, \tag{48}$$

then the high-order scheme (45) is PP under the CFL condition

$$\frac{\Delta t_n}{\Delta x_j} \max\left\{\alpha_j^\star + \sigma^{n,+}_{j-\frac{1}{2}}, \alpha_j^\star - \sigma^{n,-}_{j+\frac{1}{2}}\right\} \le \widehat{\omega}_1, \quad \forall j, \tag{49}$$

where $\sigma^{n,\pm}_{j+\frac{1}{2}} := \sigma^{\pm}(\mathbf{U}^-_{j+\frac{1}{2}}, \mathbf{U}^+_{j+\frac{1}{2}}; 1)$, *and*

$$\alpha_j^\star := \max\left\{\alpha_\star\left(\mathbf{U}^+_{j-\frac{1}{2}}, \mathbf{U}^-_{j+\frac{1}{2}}; 1\right), \alpha_\star\left(\mathbf{U}^-_{j+\frac{1}{2}}, \mathbf{U}^+_{j-\frac{1}{2}}; 1\right)\right\}.$$

Proof Using (36)–(37), we can reformulate the numerical fluxes in (45) as

$$\hat{\mathbf{F}}_1\left(\mathbf{U}^-_{j+\frac{1}{2}}, \mathbf{U}^+_{j+\frac{1}{2}}\right) = \sigma^{n,-}_{j+\frac{1}{2}}\mathbf{H}_{j+\frac{1}{2}} + \mathbf{F}_1\left(\mathbf{U}^-_{j+\frac{1}{2}}\right) - \sigma^{n,-}_{j+\frac{1}{2}}\mathbf{U}^-_{j+\frac{1}{2}}, \tag{50}$$

$$\hat{\mathbf{F}}_1\left(\mathbf{U}^-_{j-\frac{1}{2}}, \mathbf{U}^+_{j-\frac{1}{2}}\right) = \sigma^{n,+}_{j-\frac{1}{2}}\mathbf{H}_{j-\frac{1}{2}} + \mathbf{F}_1\left(\mathbf{U}^+_{j-\frac{1}{2}}\right) - \sigma^{n,+}_{j-\frac{1}{2}}\mathbf{U}^+_{j-\frac{1}{2}}, \tag{51}$$

where $\mathbf{H}_{j+\frac{1}{2}} = \mathbf{H}(\mathbf{U}^-_{j+\frac{1}{2}}, \mathbf{U}^+_{j+\frac{1}{2}}; 1)$. Under the conditions (47)–(48), we have $\mathbf{H}_{j+\frac{1}{2}} \in \overline{\mathcal{G}}_*$ for all j by using Theorem 2. The exactness of the $\mathtt{L}$-point Gauss–Lobatto quadrature rule for the polynomials of degree k implies

$$\bar{\mathbf{U}}_j^n = \frac{1}{\Delta x_j}\int_{I_j}\mathbf{U}_j^n(x)dx = \sum_{\mu=1}^{\mathtt{L}}\widehat{\omega}_\mu\mathbf{U}_j^n\left(\widehat{x}_j^{(\mu)}\right).$$

Noting $\widehat{\omega}_1 = \widehat{\omega}_{\mathtt{L}}$ and $\widehat{x}_j^{1,\mathtt{L}} = x_{j\mp\frac{1}{2}}$ and using (50)–(51), we can rewrite the scheme (45) into the following convex combination form

$$\begin{aligned}\bar{\mathbf{U}}_j^{n+1} = &\sum_{\mu=2}^{\mathtt{L}-1}\widehat{\omega}_\mu\mathbf{U}_j^n\left(\widehat{x}_j^{(\mu)}\right) + \left(2\widehat{\omega}_1 + \lambda\sigma^{n,-}_{j+\frac{1}{2}} - \lambda\sigma^{n,+}_{j-\frac{1}{2}}\right)\boldsymbol{\Xi}\\ &+ \left(-\lambda\sigma^{n,-}_{j+\frac{1}{2}}\right)\mathbf{H}_{j+\frac{1}{2}} + \lambda\sigma^{n,+}_{j-\frac{1}{2}}\mathbf{H}_{j-\frac{1}{2}},\end{aligned} \tag{52}$$

where $\lambda := \Delta t_n / \Delta x_j$, and

$$\Xi := \frac{\left(\lambda^{-1}\widehat{\omega}_1 + \sigma^{n,-}_{j+\frac{1}{2}}\right)\mathbf{U}^-_{j+\frac{1}{2}} - \mathbf{F}_1\left(\mathbf{U}^-_{j+\frac{1}{2}}\right) + \left(\lambda^{-1}\widehat{\omega}_1 - \sigma^{n,+}_{j-\frac{1}{2}}\right)\mathbf{U}^+_{j-\frac{1}{2}} + \mathbf{F}_1\left(\mathbf{U}^+_{j-\frac{1}{2}}\right)}{\lambda^{-1}\widehat{\omega}_1 + \sigma^{n,-}_{j+\frac{1}{2}} + \lambda^{-1}\widehat{\omega}_1 - \sigma^{n,+}_{j-\frac{1}{2}}}.$$

The condition (49) implies

$$\lambda^{-1}\widehat{\omega}_1 + \sigma^{n,-}_{j+\frac{1}{2}} \geq \alpha^\star_j \geq \alpha_\star\left(\mathbf{U}^-_{j+\frac{1}{2}}, \mathbf{U}^+_{j-\frac{1}{2}}; 1\right)$$
$$\lambda^{-1}\widehat{\omega}_1 - \sigma^{n,+}_{j-\frac{1}{2}} \geq \alpha^\star_j \geq \alpha_\star\left(\mathbf{U}^+_{j-\frac{1}{2}}, \mathbf{U}^-_{j+\frac{1}{2}}; 1\right),$$

which together with the condition (47) yield $\Xi \in \overline{\mathcal{G}}_*$ by Corollary 2. We therefore conclude $\bar{\mathbf{U}}^{n+1}_j \in \mathcal{G}$ from (52) according to the convexity of $\mathcal{G}_*$ and Lemma 1. □

Remark 6 In practice, it is easy to ensure the condition (47), since the exact solution $B_1(x,t) \equiv \mathtt{B_{const}}$ and the flux for B_1 in the x-direction is zero. The condition (48) can be enforced by a simple scaling limiter, which was designed in [13] by extending the techniques in [65–67]. For readers' convenience, the PP limiter is briefly reviewed in "Appendix B".

The above PP analysis is focused on first-order time discretization. In fact, it is also valid for the high-order explicit time discretization using strong stability-preserving (SSP) methods (cf. [31]), because $\mathcal{G}$ is convex and an SSP method is a convex combination of the forward Euler method.

4 Positivity-preserving schemes in multiple dimensions

In this section, we develop provably PP methods for the multidimensional ideal MHD. We remark that the design of multidimensional PP schemes have challenges essentially different from the 1D case, due to the divergence-free condition (2). For the sake of clarity, we shall restrict ourselves to the 2D case ($d = 2$), keeping in mind that our PP methods and analyses are extendable to the 3D case. We will use $\mathbf{x} \in \mathbb{R}^d$ to denote the spatial coordinate vector.

Assume that the 2D spatial domain is partitioned into a mesh $\mathcal{T}_h$, which can be unstructured and consists of polygonal cells. An illustration of two special meshes is given in Fig. 1. Let $K \in \mathcal{T}_h$ be a polygonal cell with edges $\mathscr{E}^j_K$, $j = 1, \ldots, N_K$, and K_j be the adjacent cell which shares the edge $\mathscr{E}^j_K$ with K. We denote by $\boldsymbol{\xi}^{(j)}_K = \left(\xi^{(j)}_{1,K}, \ldots, \xi^{(j)}_{d,K}\right)$ the unit normal vector of $\mathscr{E}^j_K$ pointing from K to K_j. The notations $|K|$ and $|\mathscr{E}^j_K|$ are used to denote the area of K and the length of $\mathscr{E}^j_K$, respectively. The time interval is also divided into the mesh $\{t_0 = 0, t_{n+1} = t_n + \Delta t_n, n \geq 0\}$ with the time step-size Δt_n determined by some CFL condition. Throughout this section, the lower-case letter k will be used to denote the DG polynomial degree, while the capital letter K always represents a cell.

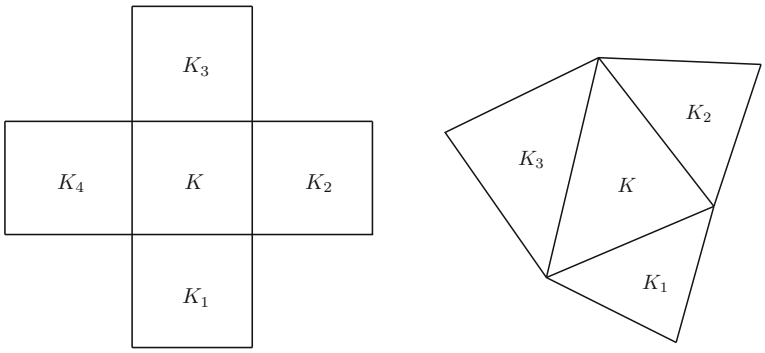

Fig. 1 Illustration of a rectangular mesh (left) and a triangular mesh (right)

4.1 First-order schemes

We consider the following first-order scheme for the Godunov form (4) of the ideal MHD equations

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n - \frac{\Delta t_n}{|K|} \sum_{j=1}^{N_K} \left|\mathscr{E}_K^j\right| \hat{\mathbf{F}}\left(\bar{\mathbf{U}}_K^n, \bar{\mathbf{U}}_{K_j}^n; \boldsymbol{\xi}_K^{(j)}\right) - \Delta t_n \left(\mathrm{div}_K \bar{\mathbf{B}}^n\right) \mathbf{S}(\bar{\mathbf{U}}_K^n), \tag{53}$$

where $\bar{\mathbf{U}}_K^n$ is the numerical approximation to the cell average of $\mathbf{U}(\mathbf{x}, t_n)$ over the cell K, and the numerical flux $\hat{\mathbf{F}}$ is taken as the HLL flux in (34). The last term at the right-hand side of (53) is suitably discretized from the Godunov–Powell source, with $\mathrm{div}_K \bar{\mathbf{B}}^n$ defined by

$$\mathrm{div}_K \bar{\mathbf{B}}^n := \frac{1}{|K|} \sum_{j=1}^{N_K} \left|\mathscr{E}_K^j\right| \left\langle \boldsymbol{\xi}_K^{(j)}, \frac{\sigma_{K,j}^{n,+} \bar{\mathbf{B}}_K^n - \sigma_{K,j}^{n,-} \bar{\mathbf{B}}_{K_j}^n}{\sigma_{K,j}^{n,+} - \sigma_{K,j}^{n,-}} \right\rangle, \tag{54}$$

where $\sigma_{K,j}^{n,\pm} := \sigma^{\pm}(\bar{\mathbf{U}}_K^n, \bar{\mathbf{U}}_{K_j}^n; \boldsymbol{\xi}_K^{(j)})$. The quantity $\mathrm{div}_K \bar{\mathbf{B}}^n$ can be considered as a discrete divergence of magnetic field, because it is a first-order accurate approximation to the left-hand side of

$$\frac{1}{|K|} \sum_{j=1}^{N_K} \int_{\mathscr{E}_K^j} \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}(\mathbf{x}, t_n) \right\rangle ds = \frac{1}{|K|} \int_K \nabla \cdot \mathbf{B} d\mathbf{x} = 0.$$

In the special case of using the LF type fluxes, $\sigma_{K,j}^{n,+} = -\sigma_{K,j}^{n,-}$, then the discrete divergence becomes

$$\mathrm{div}_K \bar{\mathbf{B}}^n = \frac{1}{|K|} \sum_{j=1}^{N_K} \left|\mathscr{E}_K^j\right| \left\langle \boldsymbol{\xi}_K^{(j)}, \frac{\bar{\mathbf{B}}_K^n + \bar{\mathbf{B}}_{K_j}^n}{2} \right\rangle,$$

which is consistent with the one introduced in [53,54] on the Cartesian meshes.

The PP property of the scheme (53) is shown as follows.

Theorem 5 *Let the wave speeds in the HLL flux satisfy* (35). *If* $\bar{\mathbf{U}}_K^n \in \mathcal{G}, \forall K \in \mathcal{T}_h$, *then the solution* $\bar{\mathbf{U}}_K^{n+1}$ *of* (53) *belongs to* $\mathcal{G}$ *for all* $K \in \mathcal{T}_h$ *under the CFL-type condition*

$$\Delta t_n \left(\frac{1}{|K|} \sum_{j=1}^{N_K} \left| \mathscr{E}_K^j \right| \left(-\sigma_{K,j}^{n,-} \right) + \frac{\left| \mathrm{div}_K \bar{\mathbf{B}}^n \right|}{\sqrt{\bar{\rho}_K^n}} \right) < 1, \quad \forall K \in \mathcal{T}_h. \tag{55}$$

Proof Let $\mathbf{H}_{K,j}^n := \mathbf{H}(\bar{\mathbf{U}}_K^n, \bar{\mathbf{U}}_{K_j}^n; \boldsymbol{\xi}_K^{(j)})$. Then the identity (36) implies

$$\hat{\mathbf{F}}\left(\bar{\mathbf{U}}_K^n, \bar{\mathbf{U}}_{K_j}^n; \boldsymbol{\xi}_K^{(j)}\right) = \sigma_{K,j}^{n,-} \mathbf{H}_{K,j}^n + \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{F}(\bar{\mathbf{U}}_K^n) \right\rangle - \sigma_{K,j}^{n,-} \bar{\mathbf{U}}_K^n. \tag{56}$$

Using (56) and the identity

$$\sum_{j=1}^{N_k} \left| \mathscr{E}_K^j \right| \boldsymbol{\xi}_K^{(j)} = \mathbf{0}, \tag{57}$$

one can rewrite the scheme (53) as

$$\bar{\mathbf{U}}_K^{n+1} = \frac{\Delta t_n}{|K|} \sum_{j=1}^{N_K} \left| \mathscr{E}_K^j \right| \left(-\sigma_{K,j}^{n,-} \right) \mathbf{H}_{K,j}^n + (1 - \lambda_K) \bar{\mathbf{U}}_K^n - \Delta t_n \left(\mathrm{div}_K \bar{\mathbf{B}}^n \right) \mathbf{S}(\bar{\mathbf{U}}_K^n), \tag{58}$$

where $\lambda_K := \frac{\Delta t_n}{|K|} \sum_{j=1}^{N_K} \left| \mathscr{E}_K^j \right| \left(-\sigma_{K,j}^{n,-} \right) \in [0, 1)$. Thanks to Theorem 2, we have $\mathbf{H}_{K,j}^n \in \mathcal{G}_\rho$ and for any $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$,

$$\mathbf{H}_{K,j}^n \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \geq - \frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\sigma_{K,j}^{n,+} - \sigma_{K,j}^{n,-}} \left\langle \boldsymbol{\xi}_K^{(j)}, \bar{\mathbf{B}}_{K_j}^n - \bar{\mathbf{B}}_K^n \right\rangle. \tag{59}$$

Since $\mathbf{H}_{K,j}^n \in \mathcal{G}_\rho$ and the first component of $\mathbf{S}(\bar{\mathbf{U}}_K^n)$ is zero, we have $\bar{\rho}_K^{n+1} \geq (1 - \lambda_K) \bar{\rho}_K^n > 0$. For any $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$, using (29) we derive from (58) that

$$\bar{\mathbf{U}}_K^{n+1} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} = \Pi_1 + \Pi_2,$$

where

$$\Pi_1 := \frac{\Delta t_n}{|K|} \sum_{j=1}^{N_K} \left| \mathscr{E}_K^j \right| \left(-\sigma_{K,j}^{n,-} \right) \left(\mathbf{H}_{K,j}^n \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \right) + \Delta t_n \left(\mathrm{div}_K \bar{\mathbf{B}}^n \right) \left(\mathbf{v}^* \cdot \mathbf{B}^* \right),$$

$$\Pi_2 := (1 - \lambda_K) \left(\bar{\mathbf{U}}_K^n \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \right) - \Delta t_n \left(\mathrm{div}_K \bar{\mathbf{B}}^n \right) (\bar{\mathbf{v}}_K^n - \mathbf{v}^*) \cdot (\bar{\mathbf{B}}_K^n - \mathbf{B}^*).$$

Let us estimate the lower bounds of Π_1 and Π_2 respectively. Using (59) and (57) gives

$$\begin{aligned}
\Pi_1 &\overset{(59)}{\geq} \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\sigma_{K,j}^{n,-}\frac{\langle \boldsymbol{\xi}_K^{(j)}, \bar{\mathbf{B}}_{K_j}^n - \bar{\mathbf{B}}_K^n\rangle}{\sigma_{K,j}^{n,+}-\sigma_{K,j}^{n,-}}(\mathbf{v}^*\cdot\mathbf{B}^*) + \Delta t_n\big(\mathrm{div}_K\bar{\mathbf{B}}^n\big)(\mathbf{v}^*\cdot\mathbf{B}^*)\\
&\overset{(54)}{=} \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\left(\sigma_{K,j}^{n,-}\frac{\langle \boldsymbol{\xi}_K^{(j)}, \bar{\mathbf{B}}_{K_j}^n - \bar{\mathbf{B}}_K^n\rangle}{\sigma_{K,j}^{n,+}-\sigma_{K,j}^{n,-}} + \left\langle \boldsymbol{\xi}_K^{(j)}, \frac{\sigma_{K,j}^{n,+}\bar{\mathbf{B}}_K^n - \sigma_{K,j}^{n,-}\bar{\mathbf{B}}_{K_j}^n}{\sigma_{K,j}^{n,+}-\sigma_{K,j}^{n,-}}\right\rangle\right)(\mathbf{v}^*\cdot\mathbf{B}^*)\\
&= \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\left\langle \boldsymbol{\xi}_K^{(j)}, \bar{\mathbf{B}}_K^n\right\rangle(\mathbf{v}^*\cdot\mathbf{B}^*) \overset{(57)}{=} 0.
\end{aligned}$$

It follows from (30) that

$$\begin{aligned}
\Pi_2 &\geq (1-\lambda_K)\left(\bar{\mathbf{U}}_K^n\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}\right) - \Delta t_n\frac{\big|\mathrm{div}_K\bar{\mathbf{B}}^n\big|}{\sqrt{\bar{\rho}_K^n}}\Big|\sqrt{\bar{\rho}_K^n}(\bar{\mathbf{v}}_K^n - \mathbf{v}^*)\cdot(\bar{\mathbf{B}}_K^n - \mathbf{B}^*)\Big|\\
&\overset{(30)}{\geq} \left(1-\lambda_K - \Delta t_n\frac{\big|\mathrm{div}_K\bar{\mathbf{B}}^n\big|}{\sqrt{\bar{\rho}_K^n}}\right)\left(\bar{\mathbf{U}}_K^n\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}\right) > 0.
\end{aligned}$$

Therefore, $\bar{\mathbf{U}}_K^{n+1}\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} > 0$, $\forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$.

Hence $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$ by Lemma 1. □

It is worth emphasizing that the suitably discretized Godunov–Powell source term is crucial for guaranteeing the PP property of the scheme (53). While the scheme (53) without this term reduces to the 2D HLL scheme for the conservative MHD system (1), specifically,

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n - \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\hat{\mathbf{F}}\big(\bar{\mathbf{U}}_K^n, \bar{\mathbf{U}}_{K_j}^n; \boldsymbol{\xi}_K^{(j)}\big). \tag{60}$$

For the LF flux, the analysis in [53] on Cartesian meshes showed that the scheme (60) is generally not PP, unless a discrete divergence-free (DDF) condition is satisfied. We find that, on a general mesh $\mathcal{T}_h$, the corresponding DDF condition is

$$\mathrm{div}_K\bar{\mathbf{B}}^n = 0, \quad \forall K \in \mathcal{T}_h. \tag{61}$$

As a direct consequence of Theorem 5, we immediately have the following corollary.

Corollary 3 *Let the wave speeds in the HLL flux satisfy (35). If $\bar{\mathbf{U}}_K^n \in \mathcal{G}$, $\forall K \in \mathcal{T}_h$, and satisfy the DDF condition (61), then under the CFL condition*

$$\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\left(-\sigma_{K,j}^{n,-}\right) < 1, \quad \forall K \in \mathcal{T}_h,$$

the solution $\bar{\mathbf{U}}_K^{n+1}$ of (60) belongs to $\mathcal{G}$ for all $K \in \mathcal{T}_h$.

If $\mathcal{T}_h$ is a Cartesian mesh and the numerical flux $\hat{\mathbf{F}}$ is taken as the global LF flux, then the scheme (60) preserves the DDF condition (61) provided that the DDF condition is satisfied by the initial data [53]. It was also shown in [53] that even slightly violating the DDF condition can cause the failure of the scheme (60) to preserve the positivity of pressure. Unfortunately, on general meshes the scheme (60) does not necessarily preserve the DDF condition (61), and it is generally not PP.

4.2 High-order schemes

We are now in the position to discuss provably PP high-order schemes for the multidimensional ideal MHD. We mainly focus on the PP high-order DG methods, keeping in mind that the analysis and framework also apply to high-order finite volume schemes.

4.2.1 Locally divergence-free schemes

We first propose locally divergence free schemes for the modified ideal MHD equations (4), as they are the base schemes of our PP high-order schemes presented later. To achieve high-order spatial accuracy, we approximate the exact solution $\mathbf{U}(\mathbf{x}, t_n)$ with a discontinuous piecewise polynomial function $\mathbf{U}_h^n(\mathbf{x})$, which is sought in the locally divergence-free space [36]

$$\mathbb{V}_h^k = \left\{ \mathbf{u} = (u_1, \ldots, u_8)^\top \;\middle|\; u_\ell\big|_K \in \mathbb{P}^k(K), \forall \ell, \; \sum_{i=1}^{d} \frac{\partial u_{4+i}}{\partial x_i}\bigg|_K = 0, \; \forall K \in \mathcal{T}_h \right\},$$

where $\mathbb{P}^k(K)$ denotes the space of polynomials in K of degree at most k.

We consider the $\mathbb{P}^k$-based locally divergence-free DG method for the Godunov form (4) of the ideal MHD equations. Specifically, the DG solution $\mathbf{U}_h^n \in \mathbb{V}_h^k$ is evolved forward by

$$\begin{aligned} \int_K \mathbf{u} \cdot \frac{\mathbf{U}_h^{n+1} - \mathbf{U}_h^n}{\Delta t_n} dxdy &= \int_K \nabla \mathbf{u} \cdot \mathbf{F}(\mathbf{U}_h^n) d\mathbf{x} \\ &- \sum_{j=1}^{N_K} \int_{\mathscr{E}_K^j} \mathbf{u}^{\mathrm{int}(K)} \cdot \left\{ \hat{\mathbf{F}}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}, \mathbf{U}_h^{n,\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\right) \right. \\ &\left. - \left[\eta_K(\mathbf{x}) \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_h^{n,\mathrm{ext}(K)} - \mathbf{B}_h^{n,\mathrm{int}(K)} \right\rangle \mathbf{S}\big(\mathbf{U}_h^{n,\mathrm{int}(K)}\big) \right] \right\} ds, \quad \forall \mathbf{u} \in \mathbb{V}_h^k, \end{aligned} \tag{62}$$

where the numerical flux $\hat{\mathbf{F}}$ is taken as the HLL flux in (34), and the factor

$$\eta_K(\mathbf{x}) := \frac{\sigma^-\big(\mathbf{U}_h^{n,\mathrm{int}(K)}, \mathbf{U}_h^{n,\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\big)}{\sigma^+\big(\mathbf{U}_h^{n,\mathrm{int}(K)}, \mathbf{U}_h^{n,\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\big) - \sigma^-\big(\mathbf{U}_h^{n,\mathrm{int}(K)}, \mathbf{U}_h^{n,\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\big)}, \quad \forall \mathbf{x} \in \mathscr{E}_K^j.$$

Here the superscripts "$\mathrm{int}(K)$" and "$\mathrm{ext}(K)$" indicate that the associated limits at the interface $\mathscr{E}_K^j$ are taken from the interior and exterior of K, respectively. The term

inside the square bracket in (62) is suitably discretized from the Godunov–Powell source term. The factor η_K is carefully devised in an upwind manner according to the local wave speeds in the HLL flux. This is very important, and is motivated from our following theoretical analysis for achieving the provably PP property, as we will see the proof of Theorem 6. If the LF flux is employed, i.e., $\sigma^- = -\sigma^+$, then $\eta_K(\mathbf{x}) \equiv -\frac{1}{2}$, and the discretized Godunov–Powell source term reduces to the one used in [54].

In the practical computations, the boundary and element integrals at the right-hand side of (62) are discretized by certain quadratures of sufficiently high order accuracy (specifically, the algebraic degree of accuracy should be at least $2k$). For example, we can employ the Gauss quadrature with $Q = k + 1$ points for the boundary integral:

$$\begin{aligned}
&\int_{\mathscr{E}_K^j} \mathbf{u}^{\mathrm{int}(K)} \cdot \Big[\hat{\mathbf{F}}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}, \mathbf{U}_h^{n,\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\right) \\
&\quad - \eta_K(\mathbf{x}) \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_h^{n,\mathrm{ext}(K)} - \mathbf{B}_h^{n,\mathrm{int}(K)} \right\rangle \mathbf{S}\big(\mathbf{U}_h^{n,\mathrm{int}(K)}\big)\Big] ds \\
&\approx |\mathscr{E}_K^j| \sum_{q=1}^{Q} \omega_q \mathbf{u}^{\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right) \cdot \Big[\hat{\mathbf{F}}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right), \mathbf{U}_h^{n,\mathrm{ext}(K)}\left(\mathbf{x}_K^{(jq)}\right); \boldsymbol{\xi}_K^{(j)}\right) \\
&\quad - \eta_K\left(\mathbf{x}_K^{(jq)}\right) \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_h^{n,\mathrm{ext}(K)}\left(\mathbf{x}_K^{(jq)}\right) - \mathbf{B}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right) \right\rangle \mathbf{S}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right)\right) \Big],
\end{aligned}$$

where $\{\mathbf{x}_K^{(jq)}\}_{1\le q\le Q}$ are the quadrature points on the interface $\mathscr{E}_K^j$, and $\{\omega_q\}_{1\le q\le Q}$ are the associated weights.

Let

$$\mathbf{U}_h^n\big|_K =: \mathbf{U}_K^n(\mathbf{x}),$$

and its cell average over K be $\bar{\mathbf{U}}_K^n$. Then we can derive from (62) the evolution equations for the cell averages $\{\bar{\mathbf{U}}_K^n\}$ as follows

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n + \Delta t_n \mathbf{L}_K(\mathbf{U}_h^n), \tag{63}$$

where

$$\begin{aligned}
\mathbf{L}_K(\mathbf{U}_h^n) := -\frac{1}{|K|} \sum_{j=1}^{N_K} \sum_{q=1}^{Q} |\mathscr{E}_K^j| \omega_q \Big[&\hat{\mathbf{F}}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right), \mathbf{U}_h^{n,\mathrm{ext}(K)}\left(\mathbf{x}_K^{(jq)}\right); \boldsymbol{\xi}_K^{(j)}\right) \\
&- \eta_K(\mathbf{x}_K^{(jq)}) \left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_h^{n,\mathrm{ext}(K)}\left(\mathbf{x}_K^{(jq)}\right) - \mathbf{B}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right) \right\rangle \mathbf{S}\left(\mathbf{U}_h^{n,\mathrm{int}(K)}\left(\mathbf{x}_K^{(jq)}\right)\right) \Big].
\end{aligned}$$

The discrete equations (63) can also be derived from a finite volume method for (4), if the approximate function $\mathbf{U}_h^n$ in (63) is reconstructed from the cell averages $\{\bar{\mathbf{U}}_K^n\}$ by a locally divergence-free approach (cf. [62,69]) such that $\mathbf{U}_h^n \in \mathbb{V}_h^k$.

When $k = 0$, the above DG and finite volume schemes reduce to the first-order scheme (53), whose PP property has been proved in Theorem 5. When $k \ge 1$, the above high-order DG and finite volume schemes are not PP in general. However, we

find that these locally divergence-free schemes can be rendered provably PP by a simple limiting procedure, as demonstrated in the following.

4.2.2 Positivity-preserving schemes

We first assume that there exists a special 2D quadrature on each cell $K \in \mathcal{T}_h$ satisfying:

- The quadrature rule is with positive weights and exact for integrals of polynomials of degree up to k on the cell K.
- The set of the quadrature points, denoted by $\mathbb{S}_K$, must include all the Gauss quadrature points $\mathbf{x}_K^{(jq)}$, $j = 1, \ldots, N_K$, $q = 1, \ldots, Q$, on the cell interface.

In other words, we would like to have a special quadrature such that

$$\frac{1}{|K|}\int_K u(\mathbf{x})d\mathbf{x} = \sum_{j=1}^{N_K}\sum_{q=1}^{Q} \varpi_{jq} u\left(\mathbf{x}_K^{(jq)}\right) + \sum_{q=1}^{\widetilde{Q}} \widetilde{\varpi}_q u\left(\widetilde{\mathbf{x}}_K^{(q)}\right), \quad \forall u \in \mathbb{P}^k(K), \tag{64}$$

where $\{\widetilde{\mathbf{x}}_K^{(q)}\}$ are the other (possible) quadrature nodes in K, and the quadrature weights $\varpi_{jq}, \widetilde{\varpi}_q$ are positive and satisfy $\sum_{j=1}^{N_K}\sum_{q=1}^{Q} \varpi_{jq} + \sum_{q=1}^{\widetilde{Q}} \widetilde{\varpi}_q = 1$. For rectangular cells, such a quadrature was constructed in [65,66] by tensor products of Gauss quadrature and Gauss–Lobatto quadrature. For triangular cells, it can be constructed by a Dubinar transform from rectangles to triangles [68]. For more general polygonal cells, one can always decompose the polygons into non-overlapping triangles, and then build the above quadrature rule by gathering those on the small triangles; see, for example, [24,49]. An illustration of the special quadrature on rectangle and triangle for $k = 2$ is shown in Fig. 2, where the (red) solid points are $\{\mathbf{x}_K^{(jq)}\}$ and the (blue) hollow circles denote $\{\widetilde{\mathbf{x}}_K^{(q)}\}$. We remark that such a special quadrature is not employed for computing any integral, but only used in the PP limiter and theoretical analysis as it decomposes the cell average into a convex combination of the desired point values.

Based on the high-order locally divergence-free schemes in Sect. 4.2.1 and the above special quadrature, we construct the provably PP high-order DG and finite volume schemes as follows. The rigorous proof of the PP property is very technical and will be given later.

Step 0 Initialization. Set $t = 0$ and $n = 0$. Using the initial data computes $\{\bar{\mathbf{U}}_K^0\}$ and $\{\mathbf{U}_K^0(\mathbf{x})\}$. $\bar{\mathbf{U}}_K^0 \in \mathcal{G}$ can be ensured by the convexity of $\mathcal{G}$, and $\mathbf{U}_h^0 \in \mathbb{V}_h^k$ is guaranteed if a local L^2-projection of the initial data onto $\mathbb{V}_h^k$ is used.

Step 1 Given admissible cell averages $\left\{\bar{\mathbf{U}}_K^n\right\}$ and $\mathbf{U}_h^n \in \mathbb{V}_h^k$, perform the PP limiting procedure. Use the PP limiter in [13] to modify the polynomials $\left\{\mathbf{U}_K^n(\mathbf{x})\right\}$, such that the modified polynomials $\left\{\widetilde{\mathbf{U}}_K^n(\mathbf{x})\right\}$ satisfy

$$\widetilde{\mathbf{U}}_K^n(\mathbf{x}) \in \mathcal{G}, \quad \forall \mathbf{x} \in \mathbb{S}_K := \left\{\widetilde{\mathbf{x}}_K^{(q)}\right\}_{1\le q\le \widetilde{Q}} \bigcup \left\{\mathbf{x}_K^{(jq)}\right\}_{1\le j\le N_K, 1\le q\le Q}. \tag{65}$$

For readers' convenience, the PP limiter is briefly reviewed in "Appendix B". Let $\widetilde{\mathbf{U}}_h^n(\mathbf{x})$ denote the discontinuous piecewise polynomial function defined by $\widetilde{\mathbf{U}}_K^n(\mathbf{x})$. We

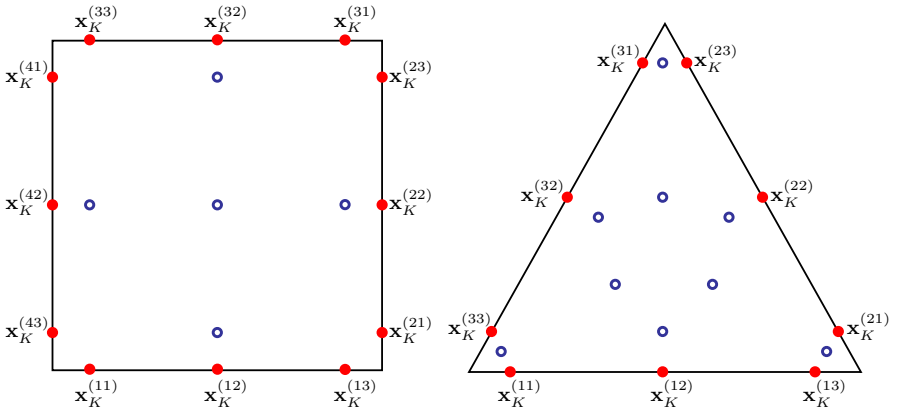

Fig. 2 Illustration of the quadrature (64) on a rectangular cell (left) and a triangular cell (right) for $k = 2$. The (red) solid points are $\{\mathbf{x}_K^{(jq)}\}$ and the (blue) hollow circles denote $\{\widetilde{\mathbf{x}}_K^{(q)}\}$; all of them constitute the point set $\mathbb{S}_K$ (color figure online)

have $\widetilde{\mathbf{U}}_h^n \in \mathbb{V}_h^k$, because the PP limiter only involves element and component wise convex combination of $\mathbf{U}_K^n(\mathbf{x})$ and its cell average.

Step 2 Update the cell averages by the scheme

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n + \Delta t_n \mathbf{L}_K(\widetilde{\mathbf{U}}_h^n), \tag{66}$$

As will be shown in Theorem 6, the PP limiting procedure in Step 1 can ensure the computed $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$, *which meets the condition of performing PP limiting procedure in the next time-forward step.*

Step 3 Build the discontinuous piecewise polynomial function $\mathbf{U}_h^{n+1}$. For our $\mathbb{P}^k$-based DG method ($k \geq 1$), evolve the high-order "moments" of the polynomials $\{\mathbf{U}_K^{n+1}(\mathbf{x})\}$ by (62) with $\mathbf{U}_h^n$ replaced by $\widetilde{\mathbf{U}}_h^n$. For a high-order finite volume scheme, reconstruct the approximate solution polynomials $\{\mathbf{U}_K^{n+1}(\mathbf{x})\}$ from the cell averages $\{\bar{\mathbf{U}}_K^{n+1}\}$ by a locally divergence-free approach such that $\mathbf{U}_h^{n+1} \in \mathbb{V}_h^k$. The details are omitted here, as these does not affect the PP property of the proposed schemes.

Step 4 Set $t_{n+1} = t_n + \Delta t_n$. If $t_{n+1} < T$, assign $n \leftarrow n+1$ and go to Step 1, where $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$ has been ensured in Step 2; otherwise, output numerical results.

We now prove the PP property of the above schemes, i.e., show that the cell average $\bar{\mathbf{U}}_K^{n+1}$ computed by (66) always belongs to $\mathcal{G}$ under the condition (65). It is worth emphasizing that the locally divergence-free spatial discretization and the suitably discretized Godunov–Powell source term in (62) are crucial for achieving the provably PP scheme, as will be seen from the proof of Theorem 6.

To shorten the notations, we define

$$\mathbf{U}_{jq}^{\mathrm{int}(K)} := \widetilde{\mathbf{U}}_h^{n,\mathrm{int}(K)}(\mathbf{x}_K^{(jq)}), \qquad \mathbf{U}_{jq}^{\mathrm{ext}(K)} := \widetilde{\mathbf{U}}_h^{n,\mathrm{ext}(K)}(\mathbf{x}_K^{(jq)}),$$

where the dependence on n is omitted. Let

$$\sigma_{jq}^{K,\pm} := \sigma^{\pm}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}, \mathbf{U}_{jq}^{\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\big).$$

For $\forall K \in \mathcal{T}_h$, we define

$$\begin{aligned}\widehat{\alpha}_{jq}^{\mathrm{int}(K)} &:= \mathscr{C}(\mathbf{U}_{jq}^{\mathrm{int}(K)}; \boldsymbol{\xi}_K^{(j)}) + \frac{2}{|\partial K|}\sum_{i=1}^{N_K}|\mathscr{E}_K^i|\frac{|\mathbf{B}_{jq}^{\mathrm{int}(K)} - \mathbf{B}_{iq}^{\mathrm{int}(K)}|}{\sqrt{\rho_{jq}^{\mathrm{int}(K)}} + \sqrt{\rho_{iq}^{\mathrm{int}(K)}}} \\ &+ \max\left\{\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{v}_{jq}^{\mathrm{int}(K)}\rangle, \frac{1}{|\partial K|}\sum_{i=1}^{N_K}|\mathscr{E}_K^i|\left\langle \boldsymbol{\xi}_K^{(j)} - \boldsymbol{\xi}_K^{(i)}, \frac{\sqrt{\rho_{jq}^{\mathrm{int}(K)}}\mathbf{v}_{jq}^{\mathrm{int}(K)} + \sqrt{\rho_{iq}^{\mathrm{int}(K)}}\mathbf{v}_{iq}^{\mathrm{int}(K)}}{\sqrt{\rho_{jq}^{\mathrm{int}(K)}} + \sqrt{\rho_{iq}^{\mathrm{int}(K)}}}\right\rangle\right\},\end{aligned}$$

with $|\partial K| := \sum_{i=1}^{N_K}|\mathscr{E}_K^i|$ denoting the circumference of the cell K.

Theorem 6 *Let the wave speeds in the HLL flux satisfy* (35)*. If the polynomial vectors* $\{\widetilde{\mathbf{U}}_K^n(\mathbf{x})\}$ *are locally divergence-free and satisfy the condition* (65)*, then the scheme* (66) *preserves* $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$ *under the CFL-type condition*

$$\Delta t_n \frac{|\mathscr{E}_K^j|}{|K|}\alpha_{jq}^K < \frac{\varpi_{jq}}{\omega_q}, \qquad \forall K \in \mathcal{T}_h,\ 1 \le j \le N_K,\ 1 \le q \le Q, \tag{67}$$

with

$$\alpha_{jq}^K := \widehat{\alpha}_{jq}^{\mathrm{int}(K)} - \sigma_{jq}^{K,-} - \eta_K\big(\mathbf{x}_K^{(jq)}\big)\Big(\rho_{jq}^{\mathrm{int}(K)}\Big)^{-\frac{1}{2}}\left|\big\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_{jq}^{\mathrm{int}(K)} - \mathbf{B}_{jq}^{\mathrm{ext}(K)}\big\rangle\right|. \tag{68}$$

Note that $\sigma_{jq}^{K,-} \le 0$ and $-1 \le \eta_K\big(\mathbf{x}_K^{(jq)}\big) \le 0$. The last term in (68) is relatively small compared to the maximum signal speed, and thus does not cause strict restriction on the time step-size; see the detailed justification and numerical evidence in [54].

We now present the proof of Theorem 6.

Proof Recalling the identity (36) and Theorem 2, one has

$$\begin{aligned}\hat{\mathbf{F}}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}, \mathbf{U}_{jq}^{\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\big) &= \sigma_{jq}^{K,-}\mathbf{H}_{jq}^K + \Big\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{F}\Big(\mathbf{U}_{jq}^{\mathrm{int}(K)}\Big)\Big\rangle - \sigma_{jq}^{K,-}\mathbf{U}_{jq}^{\mathrm{int}(K)} \\ &= \big(\widehat{\alpha}_{jq}^{\mathrm{int}(K)} - \sigma_{jq}^{K,-}\big)\mathbf{U}_{jq}^{\mathrm{int}(K)} - \Big(\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\mathbf{U}_{jq}^{\mathrm{int}(K)} - \big\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{F}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}\big)\big\rangle\Big) + \sigma_{jq}^{K,-}\mathbf{H}_{jq}^K,\end{aligned}$$

where $\mathbf{H}_{jq}^K := \mathbf{H}(\mathbf{U}_{jq}^{\mathrm{int}(K)}, \mathbf{U}_{jq}^{\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}) \in \mathcal{G}_\rho$ and for $\forall \mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$,

$$\mathbf{H}_{jq}^K \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \ge -\frac{\mathbf{v}^* \cdot \mathbf{B}^*}{\sigma_{jq}^{K,+} - \sigma_{jq}^{K,-}}\Big\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_{jq}^{\mathrm{ext}(K)} - \mathbf{B}_{jq}^{\mathrm{int}(K)}\Big\rangle. \tag{69}$$

Plugging the above formula of $\hat{\mathbf{F}}$ into (66), we can rewrite the scheme (66) as

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n + \boldsymbol{\Xi}_1 + \boldsymbol{\Xi}_2 + \boldsymbol{\Xi}_3 + \boldsymbol{\Xi}_4, \tag{70}$$

with

$$\begin{aligned}
\boldsymbol{\Xi}_1 &:= \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(\sigma_{jq}^{K,-}-\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\mathbf{U}_{jq}^{\mathrm{int}(K)}\\
\boldsymbol{\Xi}_2 &:= \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\mathbf{U}_{jq}^{\mathrm{int}(K)}-\langle\boldsymbol{\xi}_K^{(j)},\mathbf{F}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}\big)\rangle\right),\\
\boldsymbol{\Xi}_3 &:= \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(-\sigma_{jq}^{K,-}\right)\mathbf{H}_{jq}^{K},\\
\boldsymbol{\Xi}_4 &:= \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\eta_K\big(\mathbf{x}_K^{(jq)}\big)\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{ext}(K)}-\mathbf{B}_{jq}^{\mathrm{int}(K)}\right\rangle\mathbf{S}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}\big).
\end{aligned}$$

For $1\le q\le Q$, let

$$\overline{\mathbf{U}}_q^{\mathrm{int}(K)} := \frac{1}{\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\widehat{\alpha}_{jq}^{\mathrm{int}(K)}}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\left(\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\mathbf{U}_{jq}^{\mathrm{int}(K)}-\langle\boldsymbol{\xi}_K^{(j)},\mathbf{F}\big(\mathbf{U}_{jq}^{\mathrm{int}(K)}\big)\rangle\right),$$

then $\boldsymbol{\Xi}_2$ can be reformulated as

$$\boldsymbol{\Xi}_2=\frac{\Delta t_n}{|K|}\sum_{q=1}^{Q}\omega_q\left(\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\overline{\mathbf{U}}_q^{\mathrm{int}(K)}. \tag{71}$$

Thanks to Theorem 1 and Eq. (57), we have, for all $1\le q\le Q$, $\overline{\mathbf{U}}_q^{\mathrm{int}(K)}\in\mathcal{G}_\rho$ and

$$\overline{\mathbf{U}}_q^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\ge-\frac{\mathbf{v}^*\cdot\mathbf{B}^*}{\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\widehat{\alpha}_{jq}^{\mathrm{int}(K)}}\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}\rangle,\quad\forall\mathbf{v}^*,\mathbf{B}^*\in\mathbb{R}^3.$$

Note $\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\widehat{\alpha}_{jq}^{\mathrm{int}(K)}>0$ as indicated in Remark 1. Therefore, $\boldsymbol{\Xi}_2\in\mathcal{G}_\rho$, and

$$\begin{aligned}
\Pi_2 &:= \frac{\Delta t_n}{|K|}\sum_{q=1}^{Q}\omega_q\left(\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\left(\overline{\mathbf{U}}_q^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)\\
&\ge-\frac{\Delta t_n}{|K|}\big(\mathbf{v}^*\cdot\mathbf{B}^*\big)\sum_{q=1}^{Q}\omega_q\sum_{j=1}^{N_K}|\mathscr{E}_K^j|\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}\rangle.
\end{aligned}\tag{72}$$

It follows that

$$\begin{aligned}\Pi_2 &\geq -\frac{\Delta t_n}{|K|}\left(\mathbf{v}^*\cdot\mathbf{B}^*\right)\sum_{j=1}^{N_K}\int_{\mathscr{E}_K^j}\langle\boldsymbol{\xi}_K^{(j)},\widetilde{\mathbf{B}}_K^n\rangle ds\\ &= -\frac{\Delta t_n}{|K|}\left(\mathbf{v}^*\cdot\mathbf{B}^*\right)\int_K\left(\nabla\cdot\widetilde{\mathbf{B}}_K^n\right)d\mathbf{x}=0,\end{aligned}\tag{73}$$

where we have sequentially used the exactness of the Q-point quadrature rule on each interface for polynomials of degree up to k, Green's theorem and the locally divergence-free property of the polynomial vector $\widetilde{\mathbf{B}}_K^n(\mathbf{x})$.

Now, we first show $\bar{\rho}_K^{n+1}>0$. Recalling that the first component of $\mathbf{S}(\mathbf{U})$ is zero, we know that the first component of $\boldsymbol{\Xi}_4$ is zero. Since $\boldsymbol{\Xi}_2\in\mathcal{G}_\rho$ and $\mathbf{H}_{jq}^K\in\mathcal{G}_\rho$, $1\leq j\leq N_K$, $1\leq q\leq Q$, we deduce from (70) that

$$\begin{aligned}\bar{\rho}_K^{n+1} &> \bar{\rho}_K^n+\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(\sigma_{jq}^{K,-}-\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\rho_{jq}^{\mathrm{int}(K)}\\ &=\sum_{q=1}^{\widetilde{Q}}\widetilde{\varpi}_q\widetilde{\rho}_K^n\left(\widetilde{\mathbf{x}}_K^{(q)}\right)+\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\varpi_{jq}\rho_{jq}^{\mathrm{int}(K)}\\ &\quad+\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(\sigma_{jq}^{K,-}-\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\rho_{jq}^{\mathrm{int}(K)}\\ &\geq\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\omega_q\rho_{jq}^{\mathrm{int}(K)}\left(\frac{\varpi_{jq}}{\omega_q}-\frac{\Delta t_n}{|K|}|\mathscr{E}_K^j|\left(\widehat{\alpha}_{jq}^{\mathrm{int}(K)}-\sigma_{jq}^{K,-}\right)\right)\geq 0,\end{aligned}$$

where we have used in the above equality the exactness of the quadrature rule (64) for polynomials of degree up to k, and in the last inequality the condition (67).

We then prove for any $\mathbf{v}^*,\mathbf{B}^*\in\mathbb{R}^3$ that $\bar{\mathbf{U}}_K^{n+1}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}>0$. It follows from (70) that

$$\bar{\mathbf{U}}_K^{n+1}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}=\Pi_0+\Pi_1+\Pi_2+\Pi_3+\Pi_4,\tag{74}$$

where $\Pi_2\geq 0$ is defined in (72), $\Pi_4:=\boldsymbol{\Xi}_4\cdot\mathbf{n}^*$, and

$$\Pi_0:=\bar{\mathbf{U}}_K^n\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2},\tag{75}$$

$$\Pi_1:=\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(\sigma_{jq}^{K,-}-\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right),\tag{76}$$

$$\Pi_3:=\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\left(-\sigma_{jq}^{K,-}\right)\left(\mathbf{H}_{jq}^K\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right).\tag{77}$$

We now estimate the lower bounds of Π_0, Π_3 and Π_4 respectively. Based on the exactness of the quadrature rule (64) for polynomials of degree up to k, we can decompose the cell average as

$$\bar{\mathbf{U}}_K^n = \frac{1}{|K|}\int_K \widetilde{\mathbf{U}}_h^n(\mathbf{x})d\mathbf{x} = \sum_{q=1}^{\widetilde{Q}} \widetilde{\varpi}_q \widetilde{\mathbf{U}}_h^n\left(\widetilde{\mathbf{x}}_K^{(q)}\right) + \sum_{j=1}^{N_K}\sum_{q=1}^{Q} \varpi_{jq}\mathbf{U}_{jq}^{\mathrm{int}(K)}.$$

It follows that

$$\begin{aligned}\Pi_0 &= \sum_{q=1}^{\widetilde{Q}} \widetilde{\varpi}_q \left(\widetilde{\mathbf{U}}_h^n\left(\widetilde{\mathbf{x}}_K^{(q)}\right)\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}\right) + \sum_{j=1}^{N_K}\sum_{q=1}^{Q}\varpi_{jq}\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}\right)\\ &\geq \sum_{j=1}^{N_K}\sum_{q=1}^{Q}\varpi_{jq}\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2}\right),\end{aligned} \tag{78}$$

where the inequality follows from Lemma 1 and $\widetilde{\mathbf{U}}_h^n(\widetilde{\mathbf{x}}_K^{(q)}) \in \mathcal{G}$ according to (65). Noting $\sigma_{jq}^{K,-} \leq 0$ and using (69) give a lower bound of Π_3 as

$$\begin{aligned}\Pi_3 &\geq \frac{\Delta t_n}{|K|}(\mathbf{v}^*\cdot\mathbf{B}^*)\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\frac{\sigma_{jq}^{K,-}}{\sigma_{jq}^{K,+}-\sigma_{jq}^{K,-}}\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{ext}(K)}-\mathbf{B}_{jq}^{\mathrm{int}(K)}\right\rangle\\ &= \frac{\Delta t_n}{|K|}(\mathbf{v}^*\cdot\mathbf{B}^*)\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\eta_K\left(\mathbf{x}_K^{(jq)}\right)\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{ext}(K)}-\mathbf{B}_{jq}^{\mathrm{int}(K)}\right\rangle.\end{aligned} \tag{79}$$

A lower bound of Π_4 can be derived by using the inequality (31) as

$$\begin{aligned}\Pi_4 \geq \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\Bigg[&\eta_K\left(\mathbf{x}_K^{(jq)}\right)\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}-\mathbf{B}_{jq}^{\mathrm{ext}(K)}\right\rangle(\mathbf{v}^*\cdot\mathbf{B}^*)\\ &-\left(\rho_{jq}^{\mathrm{int}(K)}\right)^{-\frac{1}{2}}\left|\eta_K\left(\mathbf{x}_K^{(jq)}\right)\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}-\mathbf{B}_{jq}^{\mathrm{ext}(K)}\right\rangle\right|\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)\Bigg],\end{aligned}$$

which, along with (79) and $\eta_K(\mathbf{x}_K^{(jq)}) \leq 0$, further imply

$$\begin{aligned}\Pi_3+\Pi_4 \geq \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\Bigg[&|\mathscr{E}_K^j|\omega_q\eta_K\left(\mathbf{x}_K^{(jq)}\right)\left(\rho_{jq}^{\mathrm{int}(K)}\right)^{-\frac{1}{2}}\\ &\times\left|\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}-\mathbf{B}_{jq}^{\mathrm{ext}(K)}\right\rangle\right|\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)\Bigg].\end{aligned} \tag{80}$$

Combining the lower bounds in (73), (78), (80), with (74), we obtain

$$\begin{aligned}\bar{\mathbf{U}}_K^{n+1}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}&\geq\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\varpi_{jq}\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)\\&+\frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\Bigg[\left(\sigma_{jq}^{K,-}-\widehat{\alpha}_{jq}^{\mathrm{int}(K)}\right)+\eta_K\left(\mathbf{x}_K^{(jq)}\right)\\&\times\frac{\left|\left\langle\boldsymbol{\xi}_K^{(j)},\mathbf{B}_{jq}^{\mathrm{int}(K)}-\mathbf{B}_{jq}^{\mathrm{ext}(K)}\right\rangle\right|}{\sqrt{\rho_{jq}^{\mathrm{int}(K)}}}\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)\Bigg]\\&=\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\left(\varpi_{jq}-\frac{\Delta t_n}{|K|}|\mathscr{E}_K^j|\omega_q\alpha_{jq}^K\right)\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}\right)>0,\end{aligned}$$

where the CFL condition (67) and $\mathbf{U}_{jq}^{\mathrm{int}(K)}\in\mathcal{G}=\mathcal{G}_*$ have been used in the last inequality. Therefore, we have

$$\bar{\mathbf{U}}_K^{n+1}\cdot\mathbf{n}^*+\frac{|\mathbf{B}^*|^2}{2}>0,\quad\forall\mathbf{v}^*,\mathbf{B}^*\in\mathbb{R}^3,$$

which, along with $\bar{\rho}_K^{n+1}>0$, imply $\bar{\mathbf{U}}_K^{n+1}\in\mathcal{G}$ by Lemma 1.

The proof is completed. □

Let us further understand the above PP DG schemes and the result in Theorem 6 on two special meshes.

Example 1 Assume that the mesh is rectangular with cells $\{[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}]\times[y_{\ell-\frac{1}{2}},y_{\ell+\frac{1}{2}}]\}$ and spatial step-sizes $\Delta x_i:=x_{i+\frac{1}{2}}-x_{i-\frac{1}{2}}$ and $\Delta y_\ell:=y_{\ell+\frac{1}{2}}-y_{\ell-\frac{1}{2}}$ in x- and y-directions respectively, where (x,y) denotes the 2D spatial coordinate variables. Let $\mathbb{S}_i^x=\{x_i^{(q)}\}_{q=1}^Q$ and $\mathbb{S}_\ell^y=\{y_\ell^{(q)}\}_{q=1}^Q$ denote the Q-point Gauss quadrature points in the intervals $[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}]$ and $[y_{\ell-\frac{1}{2}},y_{\ell+\frac{1}{2}}]$ respectively. For the cell $K=[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}]\times[y_{\ell-\frac{1}{2}},y_{\ell+\frac{1}{2}}]$, the point set $\mathbb{S}_K$ in (65) is given by (cf. [65,66])

$$\mathbb{S}_K=\left(\widehat{\mathbb{S}}_i^x\otimes\mathbb{S}_\ell^y\right)\cup\left(\mathbb{S}_i^x\otimes\widehat{\mathbb{S}}_\ell^y\right),\tag{81}$$

where $\widehat{\mathbb{S}}_i^x=\{\widehat{x}_i^{(\mu)}\}_{\mu=1}^{\mathtt{L}}$ and $\widehat{\mathbb{S}}_\ell^y=\{\widehat{y}_\ell^{(\mu)}\}_{\mu=1}^{\mathtt{L}}$ denote the $\mathtt{L}$-point Gauss–Lobatto quadrature points in the intervals $[x_{i-\frac{1}{2}},x_{i+\frac{1}{2}}]$ and $[y_{\ell-\frac{1}{2}},y_{\ell+\frac{1}{2}}]$ respectively, where $\mathtt{L}\geq\frac{k+3}{2}$ such that the associated quadrature has algebraic accuracy of at least degree k. See Fig. 2 for an illustration of $\mathbb{S}_K$ for $k=2$. With $\mathbb{S}_K$ in (81), a special quadrature (cf. [65,66]) satisfying (64) can be constructed:

$$\begin{aligned}\frac{1}{|K|}\int_K u(\mathbf{x})d\mathbf{x} = &\frac{\Delta x_i\widehat{\omega}_1}{\Delta x_i+\Delta y_\ell}\sum_{q=1}^{Q}\omega_q\left(u\big(x_i^{(q)},y_{\ell-\frac{1}{2}}\big)+u\big(x_i^{(q)},y_{\ell+\frac{1}{2}}\big)\right)\\ &+\frac{\Delta y_\ell\widehat{\omega}_1}{\Delta x_i+\Delta y_\ell}\sum_{q=1}^{Q}\omega_q\left(u\big(x_{i-\frac{1}{2}},y_\ell^{(q)}\big)+u\big(x_{i+\frac{1}{2}},y_\ell^{(q)}\big)\right)\\ &+\frac{\Delta x_i}{\Delta x_i+\Delta y_\ell}\sum_{\mu=2}^{\mathtt{L}-1}\sum_{q=1}^{Q}\widehat{\omega}_\mu\omega_q u\big(x_i^{(q)},\widehat{y}_\ell^{(\mu)}\big)\\ &+\frac{\Delta y_\ell}{\Delta x_i+\Delta y_\ell}\sum_{\mu=2}^{\mathtt{L}-1}\sum_{q=1}^{Q}\widehat{\omega}_\mu\omega_q u\big(\widehat{x}_i^{(\mu)},y_\ell^{(q)}\big),\quad \forall u\in\mathbb{P}^k(K),\end{aligned}\tag{82}$$

where $\{\widehat{w}_\mu\}_{\mu=1}^{\mathtt{L}}$ are the weights of the $\mathtt{L}$-point Gauss–Lobatto quadrature. If labeling the bottom, right, top and left adjacent cells of K as K_1, K_2, K_3 and K_4, respectively, as illustrated in Fig. 1, then (82) implies

$$\varpi_{jq}=\frac{\Delta x_i\widehat{\omega}_1\omega_q}{\Delta x_i+\Delta y_\ell},\quad j=1,3;\qquad \varpi_{jq}=\frac{\Delta y_\ell\widehat{\omega}_1\omega_q}{\Delta x_i+\Delta y_\ell},\quad j=2,4.$$

Then according to Theorem 6, the CFL condition (67) for our PP DG schemes on rectangular meshes is

$$\Delta t_n\left(\frac{1}{\Delta x_i}+\frac{1}{\Delta y_\ell}\right)\alpha_{jq}^K<\widehat{\omega}_1=\frac{1}{\mathtt{L}(\mathtt{L}-1)},\quad \forall K\in\mathcal{T}_h,\ 1\le j\le 4,\ 1\le q\le Q.$$

Example 2 Assume that the mesh is triangular. A special quadrature satisfying (64) was introduced in [68], with the point set $\mathbb{S}_K$, denoted by local barycentric coordinates, as

$$\begin{aligned}&\left\{\left(\frac{1}{2}+\zeta_q,\left(\frac{1}{2}+\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right),\left(\frac{1}{2}-\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right)\right),\right.\\ &\quad\left(\left(\frac{1}{2}-\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right),\frac{1}{2}+\zeta_q,\left(\frac{1}{2}+\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right)\right),\\ &\quad\left.\left(\left(\frac{1}{2}+\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right),\left(\frac{1}{2}-\widehat{\zeta}_\mu\right)\left(\frac{1}{2}-\zeta_q\right),\frac{1}{2}+\zeta_q\right),1\le q\le Q,1\le\mu\le\mathtt{L}\right\},\end{aligned}$$

where $\{\zeta_q\}_{q=1}^{Q}$ and $\{\widehat{\zeta}_\mu\}_{\mu=1}^{\mathtt{L}}$ are the Gauss quadrature points and the Gauss–Lobatto quadrature points on $\left[-\frac{1}{2},\frac{1}{2}\right]$ respectively, and $\mathtt{L}\ge\frac{k+3}{2}$. For this quadrature, (64) becomes (cf. [68])

$$\frac{1}{|K|}\int_K u(\mathbf{x})d\mathbf{x}=\frac{2}{3}\widehat{\omega}_1\sum_{j=1}^{3}\sum_{q=1}^{Q}\omega_q u(\mathbf{x}_K^{(jq)})+\sum_{q=1}^{\widetilde{Q}}\widetilde{\varpi}_q u(\widetilde{\mathbf{x}}_K^{(q)}),\quad \forall u\in\mathbb{P}^k(K),\tag{83}$$

where $\widetilde{Q} = 3(\mathtt{L}-2)Q$. The specific expressions of the weights $\widetilde{\varpi}_q$ at quadrature points in the interior of K are omitted here. Eq. (83) implies

$$\varpi_{jq} = \frac{2}{3}\widehat{\omega}_1\omega_q, \quad 1 \le j \le 3.$$

Then, according to Theorem 6, the CFL condition (67) for our PP DG schemes on triangular meshes is

$$\Delta t_n \frac{|\mathscr{E}_K^j|}{|K|}\alpha_{jq}^K < \frac{2}{3}\widehat{\omega}_1 = \frac{2}{3\mathtt{L}(\mathtt{L}-1)}, \quad \forall K \in \mathcal{T}_h,\ 1 \le j \le 3,\ 1 \le q \le Q.$$

4.3 Why do we need the Godunov–Powell source term?

There are two features in our PP schemes: the locally divergence-free spatial discretization and the properly discretized Godunov–Powell source term. The former leads to zero divergence within each cell, while the latter controls the divergence error across the cell interfaces. The proof of Theorem 6 shows that, thanks to these two features, the PP property is obtained without requiring the DDF condition, which is needed for the PP property of the conservative schemes without the discretized Godunov–Powell source, see the following theorem.

The scheme (66) without the discretized Godunov–Powell source term becomes

$$\bar{\mathbf{U}}_K^{n+1} = \bar{\mathbf{U}}_K^n - \frac{\Delta t_n}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}|\mathscr{E}_K^j|\omega_q\hat{\mathbf{F}}\left(\mathbf{U}_{jq}^{\mathrm{int}(K)}, \mathbf{U}_{jq}^{\mathrm{ext}(K)}; \boldsymbol{\xi}_K^{(j)}\right), \tag{84}$$

which is a conservative finite volume scheme or the scheme satisfied by the cell averages of a DG method for the conservative MHD system (1). As mentioned before, even the first-order version ($k = 0$) of the scheme (84) is generally not PP unless a DDF condition is satisfied by the numerical magnetic field. The DDF condition can also be generalized to high-order schemes ($k \ge 1$), as shown in Theorem 7.

Theorem 7 *Let the wave speeds in the HLL flux satisfy* (35)*. If the polynomial vectors* $\{\widetilde{\mathbf{U}}_K^n(\mathbf{x})\}$ *satisfy the condition* (65)*, then under the CFL-type condition*

$$\Delta t_n \frac{|\mathscr{E}_K^j|}{|K|}\left(\widehat{\alpha}_{jq}^{\mathrm{int}(K)} - \sigma_{jq}^{K,-}\right) < \frac{\varpi_{jq}}{\omega_q}, \qquad \forall K \in \mathcal{T}_h,\ 1 \le j \le N_K,\ 1 \le q \le Q, \tag{85}$$

the solution $\bar{\mathbf{U}}_K^{n+1}$ *of the scheme* (84) *satisfies that* $\bar{\rho}_K^{n+1} > 0$ *and*

$$\mathcal{E}(\bar{\mathbf{U}}_K^{n+1}) > -\Delta t_n\left(\bar{\rho}_K^{n+1}\right)^{-1}(\bar{\mathbf{m}}_K^{n+1}\cdot\bar{\mathbf{B}}_K^{n+1})\left(\mathrm{div}_K\widetilde{\mathbf{B}}_h^n\right), \tag{86}$$

where $\mathrm{div}_K \widetilde{\mathbf{B}}_h^n$ *is the discrete divergence defined by*

$$\mathrm{div}_K \widetilde{\mathbf{B}}_h^n := \frac{1}{|K|} \sum_{j=1}^{N_K} \sum_{q=1}^{Q} |\mathscr{E}_K^j| \omega_q \left\langle \boldsymbol{\xi}_K^{(j)}, \frac{\sigma_{jq}^{K,+} \mathbf{B}_{jq}^{\mathrm{int}(K)} - \sigma_{jq}^{K,-} \mathbf{B}_{jq}^{\mathrm{ext}(K)}}{\sigma_{jq}^{K,+} - \sigma_{jq}^{K,-}} \right\rangle. \tag{87}$$

Furthermore, if the magnetic field $\widetilde{\mathbf{B}}_h^n(\mathbf{x})$ *satisfies the DDF condition*

$$\mathrm{div}_K \widetilde{\mathbf{B}}_h^n = 0, \tag{88}$$

then $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$.

Proof Since the first component of $\mathbf{S}(\mathbf{U})$ is zero, the discrete equations for ρ in the two schemes (66) and (84) are the same. Hence $\bar{\rho}_K^{n+1} > 0$ directly follows from the proof of Theorem 7.

Similar to the proof of Theorem 7, it can be derived for any $\mathbf{v}^*, \mathbf{B}^* \in \mathbb{R}^3$ that

$$\bar{\mathbf{U}}_K^{n+1} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} = \Pi_0 + \Pi_1 + \Pi_2 + \Pi_3, \tag{89}$$

where Π_2 is defined (72), and Π_0, Π_1 and Π_3 are defined in (75)–(77), respectively. Combining the estimates (72), (78) and (79), gives

$$\begin{aligned}
\bar{\mathbf{U}}_K^{n+1} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} &\geq -\Delta t_n (\mathbf{v}^* \cdot \mathbf{B}^*) \left(\mathrm{div}_K \widetilde{\mathbf{B}}_h^n\right) \\
&+ \sum_{j=1}^{N_K} \sum_{q=1}^{Q} \left(\varpi_{jq} - \frac{\Delta t_n}{|K|} |\mathscr{E}_K^j| \omega_q \left(\widehat{\alpha}_{jq}^{\mathrm{int}(K)} - \sigma_{jq}^{K,-} \right) \right) \left(\mathbf{U}_{jq}^{\mathrm{int}(K)} \cdot \mathbf{n}^* + \frac{|\mathbf{B}^*|^2}{2} \right) \\
&> -\Delta t_n (\mathbf{v}^* \cdot \mathbf{B}^*) \left(\mathrm{div}_K \widetilde{\mathbf{B}}_h^n\right).
\end{aligned}$$

Taking $\mathbf{v}^* = \bar{\mathbf{m}}_K^{n+1} / \bar{\rho}_K^{n+1}$ and $\mathbf{B}^* = \bar{\mathbf{B}}_K^{n+1}$ gives (86).

Under the DDF condition (88), the estimate (86) becomes $\mathcal{E}(\bar{\mathbf{U}}_K^{n+1}) > 0$, which along with $\bar{\rho}_K^{n+1} > 0$ imply $\bar{\mathbf{U}}_K^{n+1} \in \mathcal{G}$. □

In practice, it is not easy to meet the DDF condition (88), because it depends on the limiting values of the magnetic field calculated from the adjacent cells of K. The locally divergence-free property cannot ensure the DDF condition (88). If $\mathbf{B}_h^n(\mathbf{x})$ is globally divergence-free, i.e., locally divergence-free in each cell with normal magnetic component continuous across the cell interfaces, then by Green's theorem, the DDF condition $\mathrm{div}_K \mathbf{B}_h^n = 0$ is naturally satisfied and the Godunov–Powell source vanishes. There exist a few numerical techniques to maintain globally divergence-free property in the literature (e.g., [27,37,61]). However, unfortunately, the usual PP limiting technique (cf. [13,66]) with local scaling may destroy the globally divergence-free property of $\mathbf{B}_h^n(\mathbf{x})$. It is nontrivial and still open to design a limiting procedure which can enforce the conditions (65) and (88) at the same time.

Let us split the discrete divergence into two parts:

$$\begin{aligned}\mathrm{div}_K \widetilde{\mathbf{B}}_h^n = &\frac{1}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\left|\mathscr{E}_K^j\right|\omega_q\left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_{jq}^{\mathrm{int}(K)}\right\rangle \\ &+\frac{1}{|K|}\sum_{j=1}^{N_K}\sum_{q=1}^{Q}\left|\mathscr{E}_K^j\right|\omega_q\eta_K(\mathbf{x}_K^{(jq)})\left\langle \boldsymbol{\xi}_K^{(j)}, \mathbf{B}_{jq}^{\mathrm{int}(K)} - \mathbf{B}_{jq}^{\mathrm{ext}(K)}\right\rangle.\end{aligned}$$

The first part becomes zero if the locally divergence-free discretization is used, while the second part, which involves the divergence error across the cell interfaces, can be handled by including our properly discretized Godunov–Powell source term. As we have seen in the above analysis, a coupling of these two divergence-controlling techniques is very important in our PP DG methods, because they exactly contribute the discrete divergence terms which are absent in a standard multidimensional DG scheme (84) but crucial for ensuring the PP property. In other words, the suitably discretized Godunov–Powell source term helps eliminate the effect of divergence error on the positivity preservation. This is similar to the continuous case that the inclusion of Godunov–Powell source makes the modified MHD system (4) able to preserve the positivity even if the magnetic field is not divergence-free. It is also worth mentioning that, once the discretized Godunov–Powell source term is dropped, even the $\mathbb{P}^0$-based DG scheme (which is locally divergence-free) is not PP in general, and using arbitrary times larger wave speeds and/or any given small CFL number does not help to guarantee the PP property [53], unless the DDF condition is rigorously satisfied.

Remark 7 It is worth noting that in the 1D case, the divergence-free condition (2) and the fifth equation of (1) imply $B_1(x,t) \equiv$ constant for all x and $t \geq 0$. The proposed 1D schemes exactly preserve the 1D globally divergence-free property, and the Godunov–Powell source term does not exist in the proposed 1D PP schemes.

Remark 8 In the above discussions, we restrict ourselves to the first-order forward Euler time discretization. One can also use SSP high-order time discretizations (cf. [31]) to solve the ODE system $\frac{d}{dt}\mathbf{U}_h = \mathbf{L}(\mathbf{U}_h)$. For instance, the explicit third-order SSP Runge–Kutta method reads

$$\begin{aligned}\mathbf{U}_h^* &= \widetilde{\mathbf{U}}_h^n + \Delta t_n \mathbf{L}(\widetilde{\mathbf{U}}_h^n), \\ \mathbf{U}_h^{**} &= \frac{3}{4}\widetilde{\mathbf{U}}_h^n + \frac{1}{4}\Big(\widetilde{\mathbf{U}}_h^* + \Delta t_n \mathbf{L}(\widetilde{\mathbf{U}}_h^*)\Big), \\ \mathbf{U}_h^{n+1} &= \frac{1}{3}\widetilde{\mathbf{U}}_h^n + \frac{2}{3}\Big(\widetilde{\mathbf{U}}_h^{**} + \Delta t_n \mathbf{L}(\widetilde{\mathbf{U}}_h^{**})\Big),\end{aligned} \tag{90}$$

where the numerical solutions with "$\sim$" at above denote the PP limited solutions. Since an SSP method is a convex combination of the forward Euler method, our PP analysis of the proposed schemes remains valid according to the convexity of $\mathcal{G}$.

5 Numerical tests

In this section, we present some numerical results of the proposed PP DG schemes for several extreme MHD problems involving low density, low pressure, low plasma-beta $\beta := 2p/|\mathbf{B}|^2$, and/or strong discontinuity, to verify the provenly PP property and to demonstrate the effectiveness of our HLL flux and the proposed discretization of the Godunov–Powell source term. The tests below are conducted on uniform 1D meshes or 2D rectangular meshes, while the implementation of our PP schemes on unstructured triangular meshes is ongoing and will be reported in a separate paper. Without loss of generality, we focus on the proposed PP third-order ($\mathbb{P}^2$) DG methods with the SSP Runge–Kutta time discretization (90). Although our analysis has suggested a CFL condition for guaranteeing the provably PP property, we observe that our PP DG methods still work robustly and maintain the desired positivity with suitably larger time step-size in the tested cases. Unless otherwise stated, the following computations are restricted to the ideal EOS $p = (\gamma - 1)\rho e$ with $\gamma = 1.4$, and the CFL number is set as 0.15. The HLL flux is always used with the local wave speeds given by (41).

5.1 Smooth problems

A 1D and a 2D smooth problems are respectively solved on the uniform meshes of M^d cells to test the accuracy of the PP third-order DG methods. The 1D problem is similar to the one simulated in [66] for testing the PP DG scheme for the Euler equations, and has the exact solution

$$(\rho, \mathbf{v}, p, \mathbf{B})(x,t) = (1 + 0.99\sin(x-t),\ 1,\ ,0,\ ,0,\ 1,\ 0.1,\ 0,\ 0), \quad x \in [0, 2\pi],\ t \geq 0,$$

which describes a sine wave propagating with low density. The 2D problem is the vortex problem with the same setup as in [16] and has a extremely low pressure (about 5.3×10^{-12}) in the vortex center; the adiabatic index $\gamma = \frac{5}{3}$; the computational domain is $[-10, 10]^2$ with periodic boundary condition. Figure 3 displays the numerical errors obtained by the third-order DG method with the PP limiter at different grid resolutions. The results show that the expected convergence order is achieved.

Next, we simulate several MHD problems involving discontinuities. Before using the PP limiter, the WENO limiter [44] is also implemented with the aid of the local characteristic decomposition, to enhance the numerical stability of high-order DG schemes in resolving the strong discontinuities and their interactions. The 2D WENO limiter is combined with the locally divergence-free reconstruction approach in [69]. The WENO limiter is only employed in the "trouble" cells adaptively detected by the indicator of [35].

5.2 Riemann problems

Two 1D Riemann problems are solved. The first is a 1D vacuum shock tube problem (cf. [16]) with the initial data

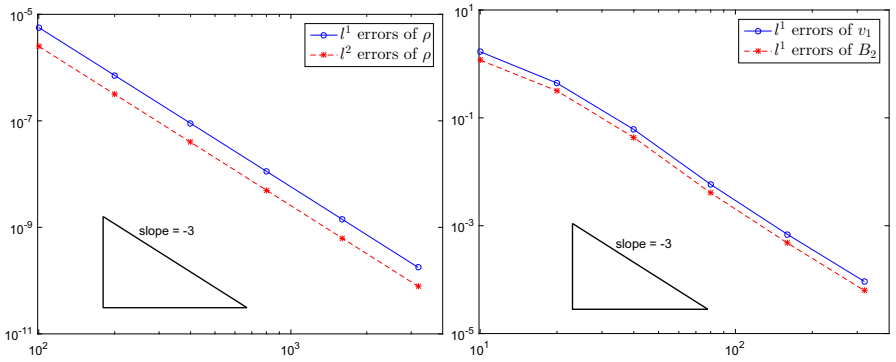

Fig. 3 Numerical errors obtained by the PP third-order DG method at different grid resolutions with M^d cells. Left: the 1D smooth problem at $t = 0.1$; right: the 2D smooth problem at $t = 0.05$. The horizontal axis denotes the value of M

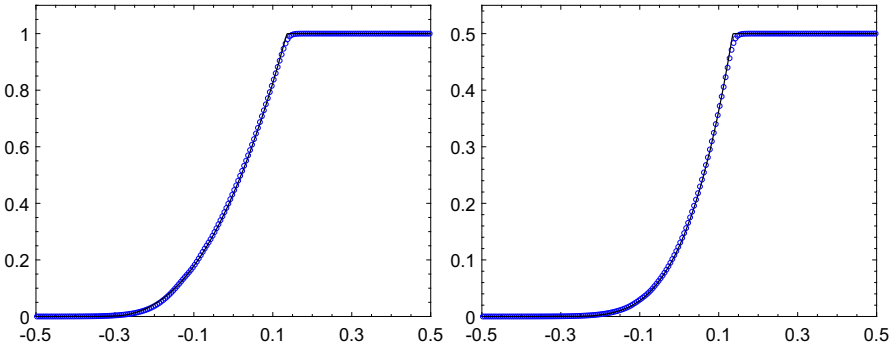

Fig. 4 The density (left) and pressure (right) obtained by the PP third-order DG method on the meshes of 200 cells (symbols "○") and 5000 cells (solid lines), respectively.

$$(\rho, \mathbf{v}, p, \mathbf{B})(x, 0) = \begin{cases} (10^{-12},\ 0,\ 0,\ 0,\ 10^{-12},\ 0,\ 0,\ 0), & x < 0, \\ (1,\ 0,\ 0,\ 0,\ 0.5,\ 0,\ 1,\ 0), & x > 0. \end{cases}$$

It is used to demonstrate that our 1D PP DG scheme can handle extremely low density and pressure. The adiabatic index $\gamma = \frac{5}{3}$, and the computational domain is set as $[-0.5, 0.5]$. Figure 4 shows the density and pressure of the numerical solution on the mesh of 200 cells as well as those of a highly resolved solution with 5000 cells at time $t = 0.1$. One can observe that the solutions of low resolution and high resolution are in good agreement. We confirm that the low pressure and the low density are both correctly captured by comparing with the results in [16]. The PP third-order DG method works very robustly during the simulation. It is noticed that, if the PP limiter is not used to enforce the condition (48), the method breaks down within a few time steps due to unphysical solution.

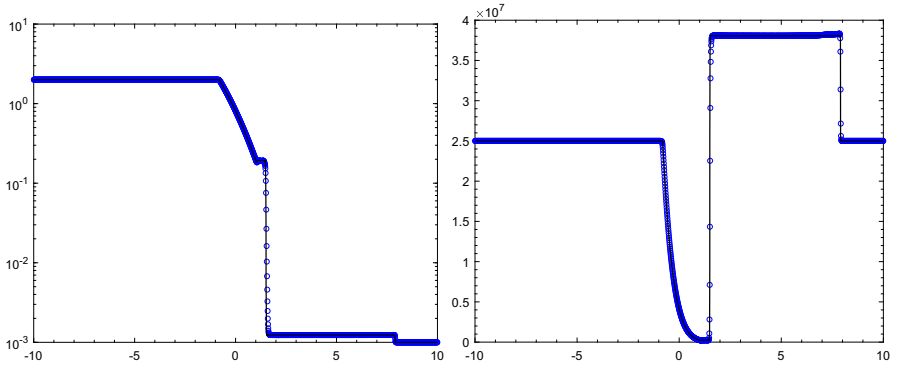

Fig. 5 Numerical results at $t = 0.00003$ obtained by the PP third-order DG method with 2000 cells (symbols "○") and 10,000 cells (solid lines). Left: log plot of density; right: magnetic pressure

The second Riemann problem is a variant of the Leblanc problem (cf. [66]) of gas dynamics by adding a strong magnetic field. The initial condition is

$$(\rho, \mathbf{v}, p, \mathbf{B})(x, 0) = \begin{cases} (2,\ 0,\ 0,\ 0,\ 10^9,\ 0,\ 5000,\ 5000), & x < 0, \\ (0.001,\ 0,\ 0,\ 0,\ 1,\ 0,\ 5000,\ 5000), & x > 0. \end{cases}$$

The initial pressure has a very large jump, and the plasma-beta at the right state is extremely low ($\beta = 4 \times 10^{-8}$), making the successful simulation of this problem a challenge. The computational domain is taken as $[-10, 10]$. To fully resolve the wave structure, a fine mesh is required for this test [66]. Figure 5 displays the numerical results at $t = 0.00003$ obtained by the PP third-order DG method using 2000 cells and 10,000 cells, respectively. We observe that the strong discontinuities are well captured, and the low resolution and high resolution are in good agreement. Figure 6 gives a comparison of the numerical solutions resolved by using the proposed HLL flux and the global LF flux of [53], respectively. As expected, the PP DG method with the HLL flux exhibits better resolution. In this extreme test, it is also necessary to enforce the condition (48) by the PP limiting procedure, otherwise negative pressure will appear in the cell averages of the DG solution.

5.3 Blast problem

This test was first introduced by Balsara and Spicer [7], and has become a benchmark for testing 2D MHD codes. If the low gas pressure, strong magnetic field or low plasma-beta is involved, then simulating such MHD blast problems can be very challenging. Therefore, it is often used to check the robustness of MHD schemes; see e.g., [13,16].

The simulation is implemented in $[-0.5, 0.5]^2$ with outflow boundary conditions. Our setup is the same as in [7,13]. Initially, the domain is filled with plasma at rest with unit density. The explosion zone ($r < 0.1$) has a pressure of 1000, while the ambient medium ($r > 0.1$) has a pressure of 0.1, where $r = \sqrt{x^2 + y^2}$. The magnetic field is initialized in the x-direction as $100/\sqrt{4\pi}$. For this setup, the ambient medium has a

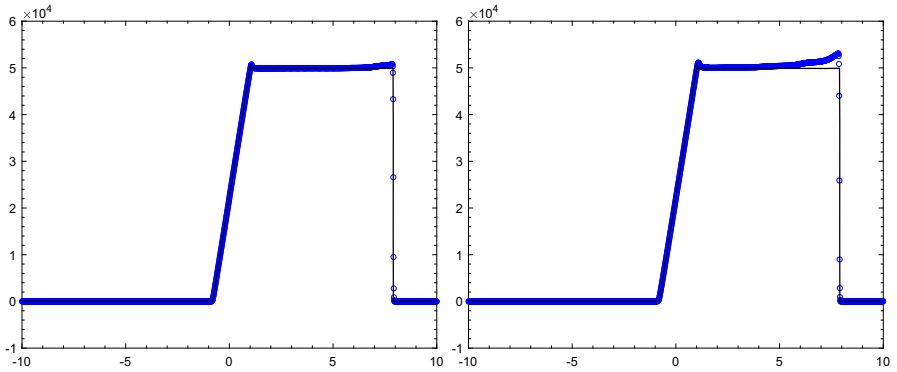

Fig. 6 Same as Fig. 5 except for the velocity v_1 obtained by using the proposed HLL flux (left) and the global LF flux (right)

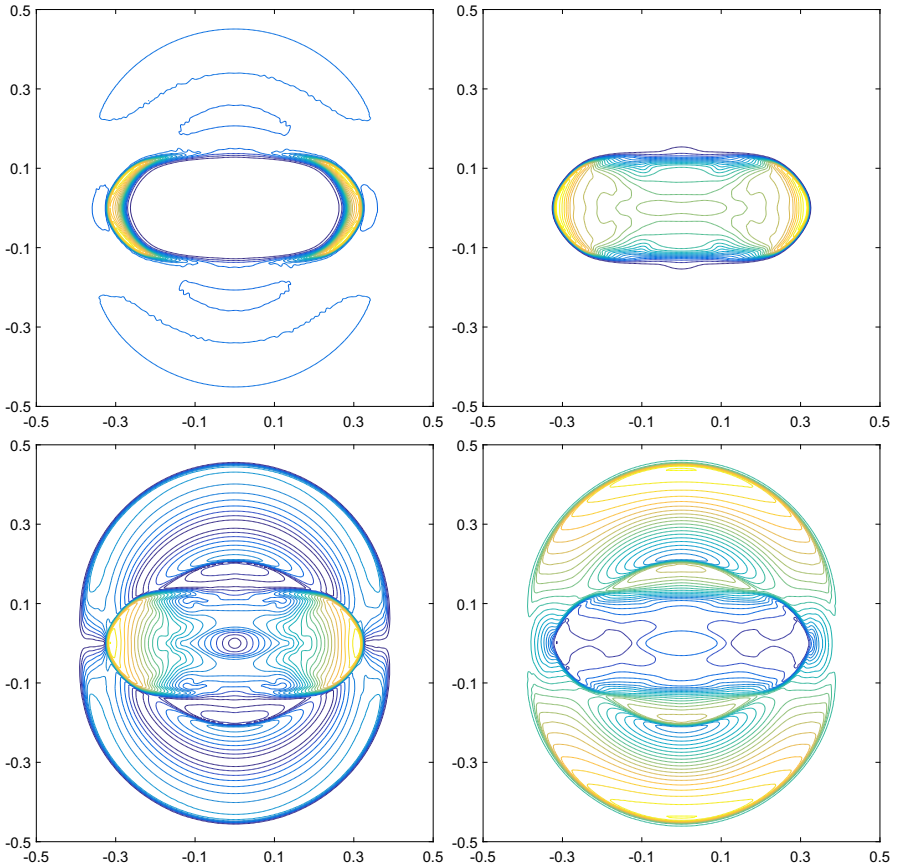

Fig. 7 The contour plots of density (top left), pressure (top right), velocity $|\mathbf{v}|$ (bottom left) and magnetic pressure (bottom right) at time $t = 0.01$ for the blast problem

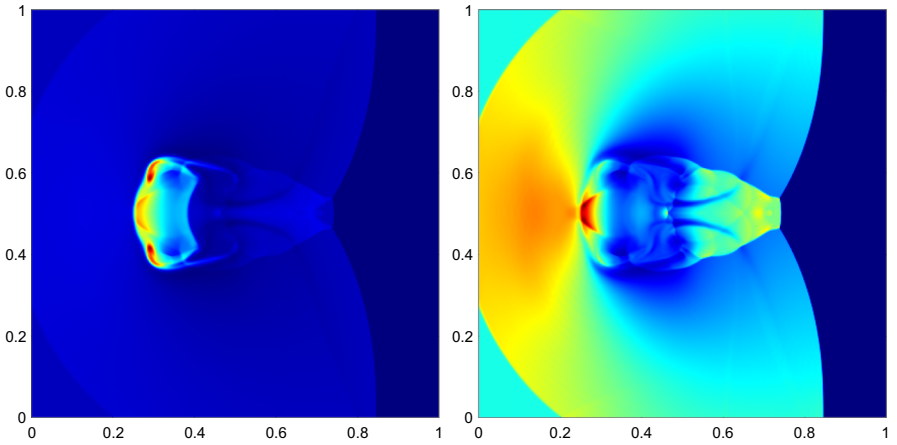

Fig. 8 The schlieren images of density (left) and pressure (right) at time $t = 0.06$ for the shock cloud interaction problem

small plasma-beta (about 2.51×10^{-4}). Our numerical results at $t = 0.01$, obtained by the PP third-order DG method with 320×320 cells, are displayed in Fig. 7. Our results agree well with those in [7,16,37], and the density profile is well captured with much less oscillations than those shown in [7,16]. The velocity profile clearly shows higher resolution than that in [54] obtained by the same DG method but with the global LF flux. We also notice that, if the PP limiter is turned off, the condition (65) will be violated since $t \approx 2.24 \times 10^{-4}$, and the method will fail due to negative numerical pressure.

5.4 Shock cloud interaction

This problem [18] describes the disruption of a high density cloud by a strong shock wave, and has been widely simulated in the literature (e.g., [2,48]). We employ the same setup as in [2,48]. The simulation is implemented in the domain $\Omega = [0, 1]^2$ with the right boundary specified as supersonic inflow condition and the others as outflow conditions. The adiabatic index $\gamma = \frac{5}{3}$, and the initial conditions are given by the two states

$$(\rho, \mathbf{v}, p, \mathbf{B}) = \begin{cases} (3.86859, 0, 0, 0, 167.345, 0, 2.1826182, -2.1826182), & x < 0.6, \\ (1, -11.2536, 0, 0, 1, 0, 0.56418958, 0.56418958), & x > 0.6, \end{cases}$$

separated by a discontinuity parallel to the y-axis at $x = 0.6$. To the right of the discontinuity there is a circular cloud of radius 0.15, centered at $x = 0.8$ and $y = 0.5$. The cloud has the same states as the surrounding fluid except for a higher density of 10.

We simulate this problem by using our PP third-order DG method with 400×400 cells. The numerical results at time $t = 0.06$ are shown in Fig. 8. It is seen that

the complex flow structures and interactions are captured with high resolution, and the results agree well with those in the literature, e.g., [2,48]. In this test, it is also necessary to employ the PP limiter to enforce the condition (65). We also observe that, if the discretized Godunov–Powell source term is dropped from our PP DG method, negative pressure will appear in the cell average of the DG solutions and the code breaks down at $t \approx 0.014$, because the resulting scheme (namely the locally divergence-free DG method with the proposed HLL flux and the PP and WENO limiters) is not PP in general. This further confirms the importance of the discretized Godunov–Powell source term.

5.5 Astrophysical jets

The last test is to simulate jet flow, which is relevant in astrophysics. In a high Mach number jet with strong magnetic field, the internal energy is very small compared to the huge kinetic and magnetic energy, thus negative pressure is very likely to be produced in the numerical simulations. Moreover, there may exist strong shock wave, shear flow and interface instabilities in high-speed jet flows. Successfully simulating such jet flows is indeed a challenge, cf. [5,55,57,66].

We consider the Mach 800 MHD jets proposed in [53,54] and extended from the gas dynamical jet of Balsara [5] by adding a magnetic field. Initially, the domain $[-0.5, 0.5] \times [0, 1.5]$ is full of the static ambient medium with $(\rho, p) = (0.1\gamma, 1)$. The adiabatic index $\gamma = 1.4$. A Mach 800 dense jet is injected in the y-direction through the inlet part ($|x| < 0.05$) on the bottom boundary ($y = 0$). The fixed inflow condition with $(\rho, p, v_1, v_2, v_3) = (\gamma, 1, 0, 800, 0)$ is specified on the nozzle $\{y = 0, |x| < 0.05\}$, while the other boundary conditions are outflow. A magnetic field $(0, B_a, 0)$ is initialized along the y-direction. As B_a is set larger, this test becomes more challenging. We set computational domain as $[0, 0.5] \times [0, 1.5]$ with the reflecting boundary condition specified at $x = 0$, and divided it into 200×600 cells. We here show our numerical results in two strongly magnetized cases: (i) $B_a = \sqrt{2000}$, and the corresponding plasma-beta $\beta_a = 10^{-3}$; (ii) $B_a = \sqrt{20{,}000}$, and the corresponding plasma-beta $\beta_a = 10^{-4}$. The schlieren images of the numerical solutions for these two cases are respectively displayed in Figs. 9 and 10 within the domain $[-0.5, 0.5] \times [0, 1.5]$. Those plots clearly show the time evolution of the jets. It is seen that the flow structures in different magnetized cases are very different. The present method well captures the Mach shock wave at the jet head and other discontinuities with high resolution. The results agree with those in [54] computed by the PP DG method with a global LF flux. In these extreme tests, our PP method exhibits good robustness without using any artificial treatment. We also perform the tests with varied Mach numbers, and the method also works very robustly. For example, the numerical result for a Mach 2000 jet with $B_a = \sqrt{20{,}000}$ is displayed in Fig. 11. Interestingly, the flow structures are similar to those in Fig. 9 of the Mach 800 jet with a weaker magnetic field $B_a = \sqrt{2000}$. This is probably due to the huge kinetic energy, which becomes dominant and weakens the effect of magnetic field. The dynamics of the Mach 2000 jet evolve much faster than the Mach 800 jet, as expected. A higher Mach (Mach 10,000)

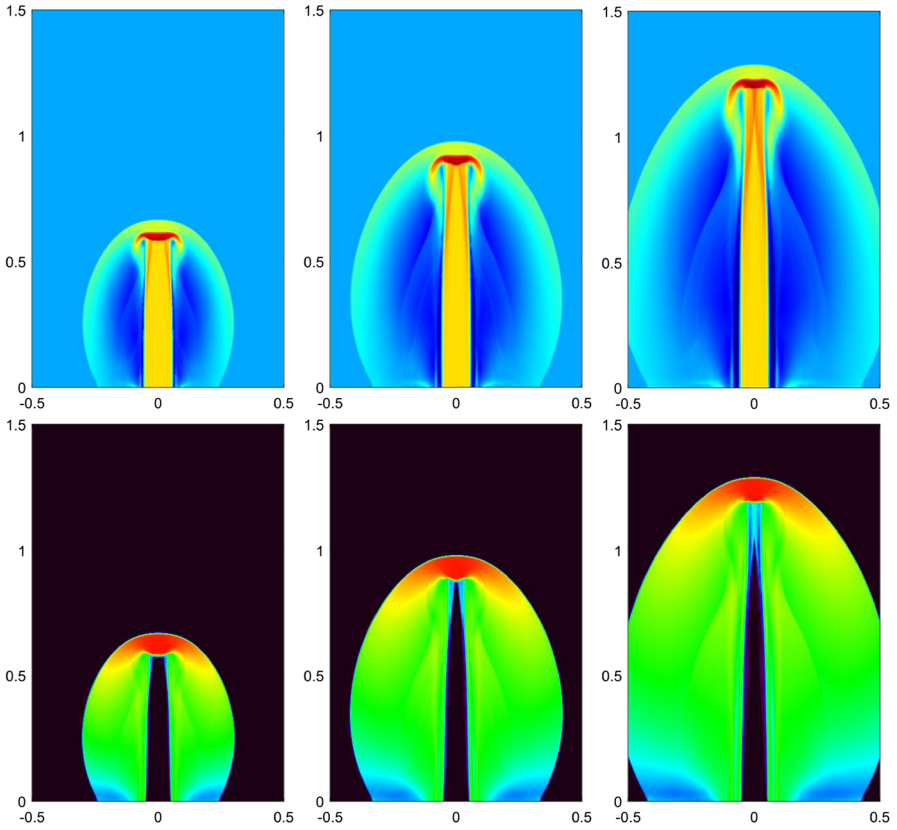

Fig. 9 The schlieren images of density logarithm (top) and gas pressure logarithm (bottom) for the Mach 800 jet problem with $B_a = \sqrt{2000}$. From left to right: $t = 0.001$, 0.0015 and 0.002

jet with $B_a = \sqrt{20{,}000}$ is further simulated and shown in Fig. 12. We see that this jet shape is thinner.

In the above simulations, it is necessary to employ the PP limiting procedure to meet the condition (65), which is not satisfied automatically. To confirm the importance of the suitably discretized Godunov–Powell source term in our PP schemes, we have also performed the above tests by dropping this term and keeping the PP and WENO limiters turned on. The resulting scheme is actually the locally divergence-free, conservative, third-order DG method with PP and WENO limiters. We find that this scheme with either the proposed HLL flux or the global LF flux, which is generally not PP in theory, cannot run the above jet tests. The failure results from negative numerical pressure produced in the cell averages of the DG solution. We observe that, without the discretized Godunov–Powell source term, the code also fails on a refined mesh, and also for more strongly magnetized cases. This, again, demonstrates that the suitably discretized Godunov–Powell source term is really crucial for guaranteeing the PP property.

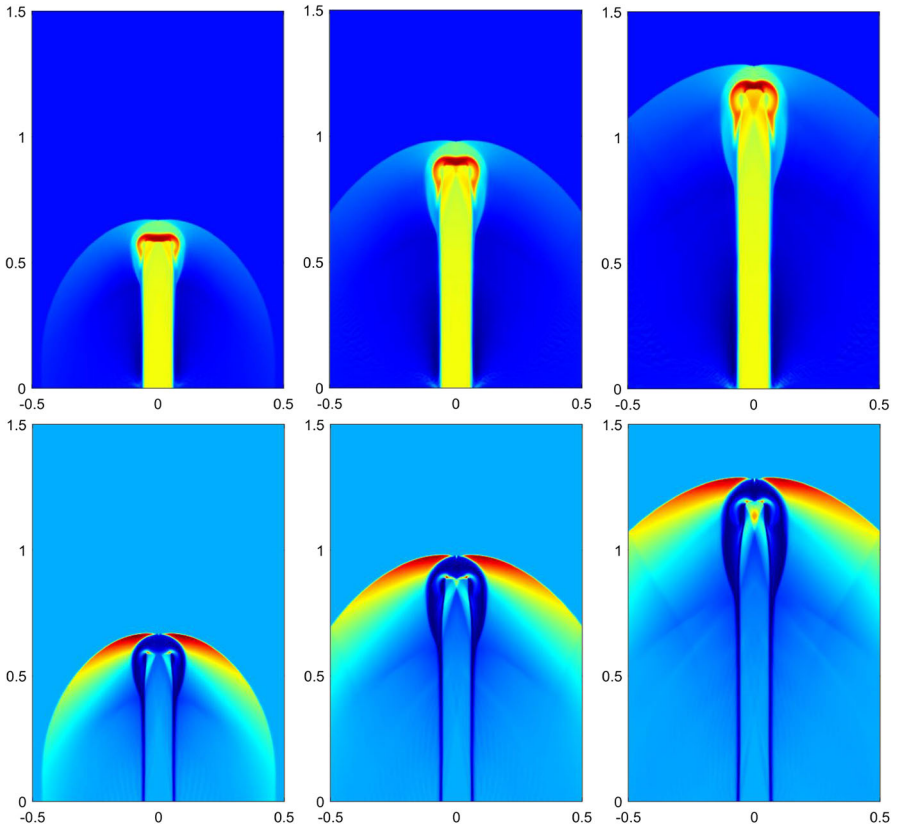

Fig. 10 The schlieren images of density logarithm (top) and magnetic pressure (bottom) for the Mach 800 jet problem with $B_a = \sqrt{20{,}000}$. From left to right: $t = 0.001$, 0.0015 and 0.002

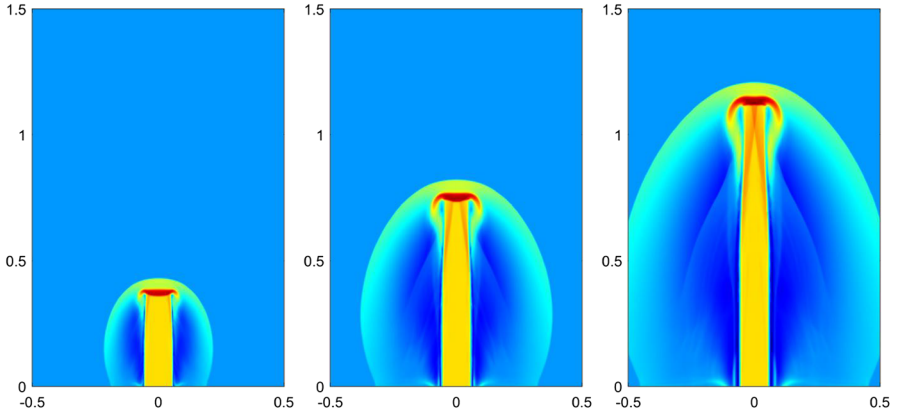

Fig. 11 The schlieren images of density logarithm for the Mach 2000 jet problem with $B_a = \sqrt{20{,}000}$. From left to right: $t = 0.00025$, 0.0005 and 0.00075

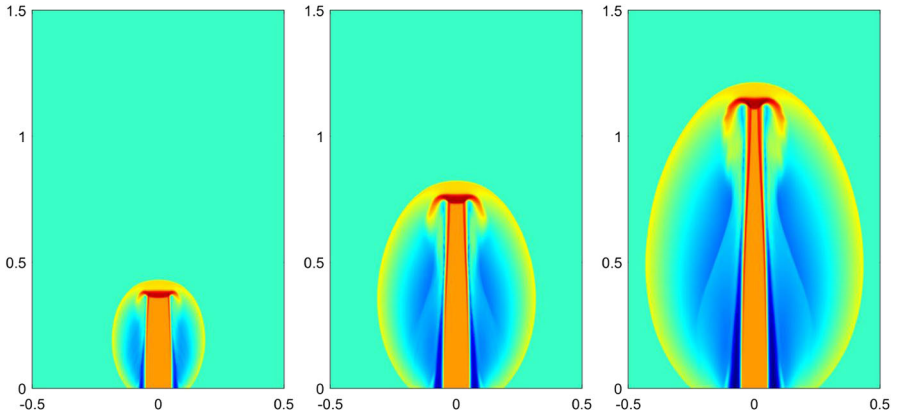

Fig. 12 The schlieren images of density logarithm for the Mach 10,000 jet problem with $B_a = \sqrt{20,000}$. From left to right: $t = 0.00005$, 0.0001 and 0.00015

6 Conclusions

In this paper, we proposed and analyzed provably PP high-order DG and finite volume schemes for the ideal MHD on general meshes. The unified auxiliary theories were built for rigorous PP analysis of numerical schemes with HLL-type flux on an arbitrary polytopal mesh. A close relation was established between the PP property and the discrete divergence of magnetic field on general meshes. We also derived explicit estimates of the wave speeds in the HLL flux to ensure the provably PP property. In the 1D case, we proved that the standard finite volume and DG methods with the proposed HLL flux are PP, under a condition accessible by a PP limiter. In the multidimensional cases, we constructed provably PP high-order DG schemes based on suitable discretization of the modified MHD system (4). In addition to the proper wave speeds in the numerical flux and a standard PP limiter, we demonstrated that a coupling of two divergence-controlling techniques is also crucial for achieving the provably PP property. The two techniques are the locally divergence-free DG element and a properly discretized Godunov–Powell source term, which control the divergence error within each cell and across the cell interfaces, respectively. Our analysis clearly revealed that these two techniques exactly contribute the discrete divergence terms which are absent in a standard multidimensional DG schemes but very important for ensuring the PP property. We also proved in "Appendix A" the positivity of the strong solution of the modified MHD system (1). Such a feature, not enjoyed by the conservative system (1) (see [54]), can serve as a justification for designing provably PP multidimensional schemes based on the modified system (4). The analysis and findings in this paper provide a clear understanding, at both discrete and continuous levels, of the relation between the PP property and the divergence-free constraint. The proposed framework and analysis techniques as well as the provenly PP schemes can also be useful for investigating or designing other PP schemes for the ideal MHD.

Several numerical tests were conducted on 1D mesh and 2D rectangular mesh, to confirm the provenly PP property and to demonstrate the effectiveness of the proposed

PP techniques. The implementation of our PP DG schemes on unstructured triangular meshes is ongoing and will be reported separately in the future.

A Positivity of strong solutions of the modified MHD system

In [54], we analytically demonstrated that the exact smooth solution of the conservative MHD system (1) may fail to be PP if the divergence-free condition (2) is violated. Here we would like to show that the strong solutions of the modified MHD system (4) always retain the positivity of density and pressure even if the divergence-free condition (2) is not satisfied. It is reasonable to hope that such a claim may also hold for the weak entropy solutions of (4).

Consider the initial-value problem of the system (4), for $\mathbf{x} \in \mathbb{R}^d$ and $t > 0$, with initial data

$$(\rho, \mathbf{v}, p, \mathbf{B})(\mathbf{x}, 0) = (\rho_0, \mathbf{v}_0, p_0, \mathbf{B}_0)(\mathbf{x}), \tag{91}$$

and the ideal EOS $p = (\gamma - 1)\rho e$, where $\gamma > 1$. Using the method of characteristics, one can show the following result.

Proposition 1 *Assume that the initial data* (91) *are in* $C^1(\mathbb{R}^d)$ *with* $\rho_0(\mathbf{x}) > 0$ *and* $p_0(\mathbf{x}) > 0$, $\forall \mathbf{x} \in \mathbb{R}^d$. *If the initial-value problem of* (4) *with* (91) *has a* C^1 *solution* $(\rho, \mathbf{v}, p, \mathbf{B})(\mathbf{x}, t)$ *for* $\mathbf{x} \in \mathbb{R}^d$ *and* $0 \le t < T$, *then the solution satisfies* $\rho(\mathbf{x}, t) > 0$ *and* $p(\mathbf{x}, t) > 0$ *for all* $\mathbf{x} \in \mathbb{R}^d$ *and* $0 \le t < T$.

Proof Let $\frac{D}{Dt} := \partial_t + \mathbf{v}(\mathbf{x}, t)\nabla\cdot$ be the directional derivative along the direction

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}, t). \tag{92}$$

For any $(\bar{\mathbf{x}}, \bar{t}) \in \mathbb{R}^d \times \mathbb{R}_+$, let $\mathbf{x} = \mathbf{x}(t; \bar{\mathbf{x}}, \bar{t})$ be the integral curve of (92) through the point $(\bar{\mathbf{x}}, \bar{t})$. Denote $\mathbf{x}_0(\bar{\mathbf{x}}, \bar{t}) := \mathbf{x}(0; \bar{\mathbf{x}}, \bar{t})$, then, at $t = 0$, the curve passes through the point $(\mathbf{x}_0(\bar{\mathbf{x}}, \bar{t}), 0)$. Recall that, for smooth solutions, the first equation of the system (4) can be reformulated as

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \mathbf{v}. \tag{93}$$

Integrating Eq. (93) along the curve $\mathbf{x} = \mathbf{x}(t; \bar{\mathbf{x}}, \bar{t})$ gives

$$\rho(\bar{\mathbf{x}}, \bar{t}) = \rho_0(\mathbf{x}_0(\bar{\mathbf{x}}, \bar{t})) \exp\left(- \int_0^{\bar{t}} \nabla \cdot \mathbf{v}(\mathbf{x}(t; \bar{\mathbf{x}}, \bar{t}), t) dt \right) > 0.$$

For smooth solutions, we derive from the modified system (4) the pressure equation

$$\frac{Dp}{Dt} = -\gamma p \nabla \cdot \mathbf{v}, \tag{94}$$

which implies $p(\bar{\mathbf{x}}, \bar{t}) = p_0(\mathbf{x}_0(\bar{\mathbf{x}}, \bar{t})) \exp\left(-\gamma \int_0^{\bar{t}} \nabla \cdot \mathbf{v}(\mathbf{x}(t; \bar{\mathbf{x}}, \bar{t}), t) dt\right) > 0.$ □

Remark 9 By similar arguments one can show that the above proposition also holds for the modified MHD equations introduced by Janhunen [34], because the corresponding equations for density and pressure are exactly also (93) and (94), respectively. This may explain why it is also possible to develop PP schemes based on proper discretization of Janhunen's MHD system, cf. [10,34,50,51].

Recall that the pressure equation associated with the conservative system (1) is

$$\frac{Dp}{Dt} = -\gamma p \nabla \cdot \mathbf{v} - (\gamma - 1)(\mathbf{v} \cdot \mathbf{B}) \nabla \cdot \mathbf{B},$$

which, in comparison with (94), has an additional term proportional to $\nabla \cdot \mathbf{B}$. As shown in [54], due to this term, negative pressure can appear in the exact smooth solution of the conservative MHD system (1) if $\nabla \cdot \mathbf{B} \neq 0$.

B Review of the positivity-preserving limiter

We employ a simple PP limiter to enforce the condition (48) or (65) for our 1D or 2D PP schemes. The limiter was originally proposed by Zhang and Shu [65–67] for scalar conservation laws and the compressible Euler equations. It was extended to the ideal MHD case in [13]. For readers' convenience, we here briefly review this limiter. It is worth noting that the PP limiter works only when the cell averages of the numerical solutions always stay in $\mathcal{G}$. This is rigorously proved for our PP high-order schemes, but does not always hold for the standard multidimensional DG schemes without the suitably discretized Godunov–Powell source term.

We perform the PP limiter separately for each cell. Let K denote a cell, and $\mathbb{S}_K$ be the quadrature points involved in the condition (48) or (65) in K. Let $\mathbf{U}_K^n(\mathbf{x})$ be the approximate polynomial solution within K, and $\bar{\mathbf{U}}_K^n$ be the cell average which is always preserved in $\mathcal{G}$ by our PP schemes. If $\mathbf{U}_K^n(\mathbf{x}) \notin \mathcal{G}$ for some $\mathbf{x} \in \mathbb{S}_K$, then we seek the modified polynomial $\widetilde{\mathbf{U}}_K^n(\mathbf{x})$ with the same cell average such that $\widetilde{\mathbf{U}}_K^n(\mathbf{x}) \in \mathcal{G}$ for all $\mathbf{x} \in \mathbb{S}_K$. To avoid the effect of the rounding error, we introduce two sufficiently small positive numbers, ϵ_1 and ϵ_2, as the desired lower bounds for density and internal energy, respectively, such that $\bar{\mathbf{U}}_K^n \in \mathcal{G}_\epsilon = \{\mathbf{U} : \rho \geq \epsilon_1, \mathcal{E}(\mathbf{U}) \geq \epsilon_2\}$; e.g., take $\epsilon_1 = \min\{10^{-13}, \bar{\rho}_K^n\}$ and $\epsilon_2 = \min\{10^{-13}, \mathcal{E}(\bar{\mathbf{U}}_K^n)\}$.

The PP limiting procedure consists of two steps. First, modify the density to enforce the positivity by

$$\widehat{\rho}_K(\mathbf{x}) = \theta_1(\rho_K^n(\mathbf{x}) - \bar{\rho}_K^n) + \bar{\rho}_K^n, \quad \theta_1 = \min\left\{1, \frac{\bar{\rho}_K^n - \epsilon_1}{\bar{\rho}_K^n - \min_{\mathbf{x}\in\mathbb{S}_K} \rho_K^n(\mathbf{x})}\right\}.$$

Then modify $\widehat{\mathbf{U}}_K(\mathbf{x}) := (\widehat{\rho}_K(\mathbf{x}), \mathbf{m}_K^n(\mathbf{x}), \mathbf{B}_K^n(\mathbf{x}), E_K^n(\mathbf{x}))^\top$ to enforce the positivity of internal energy by

$$\widetilde{\mathbf{U}}_K^n(\mathbf{x}) = \theta_2(\widehat{\mathbf{U}}_K(\mathbf{x}) - \bar{\mathbf{U}}_K^n) + \bar{\mathbf{U}}_K^n, \quad \theta_2 = \min\left\{1, \frac{\mathcal{E}(\bar{\mathbf{U}}_K^n) - \epsilon_2}{\mathcal{E}(\bar{\mathbf{U}}_K^n) - \min_{\mathbf{x}\in\mathbb{S}_K} \mathcal{E}\big(\widehat{\mathbf{U}}_K(\mathbf{x})\big)}\right\}.$$

It is easy to verify that $\widetilde{\mathbf{U}}_K^n(\mathbf{x})$ belongs to $\mathcal{G}_\epsilon$ for all $\mathbf{x} \in \mathbb{S}_K$ and has the cell average $\bar{\mathbf{U}}_K^n$. Such a limiter can also maintain the approximation accuracy; see [64–66].

References

1. Artebrant, R., Torrilhon, M.: Increasing the accuracy in locally divergence-preserving finite volume schemes for MHD. J. Comput. Phys. **227**(6), 3405–3427 (2008)
2. Balbás, J., Tadmor, E.: Nonoscillatory central schemes for one- and two-dimensional magnetohydrodynamics equations. II: high-order semidiscrete schemes. SIAM J. Sci. Comput. **28**(2), 533–560 (2006)
3. Balsara, D.S.: Second-order-accurate schemes for magnetohydrodynamics with divergence-free reconstruction. Astrophys. J. Suppl. Ser. **151**, 149–184 (2004)
4. Balsara, D.S.: Multidimensional HLLE Riemann solver: application to Euler and magnetohydrodynamic flows. J. Comput. Phys. **229**(6), 1970–1993 (2010)
5. Balsara, D.S.: Self-adjusting, positivity preserving high order schemes for hydrodynamics and magnetohydrodynamics. J. Comput. Phys. **231**, 7504–7517 (2012)
6. Balsara, D.S., Spicer, D.: Maintaining pressure positivity in magnetohydrodynamic simulations. J. Comput. Phys. **148**, 133–148 (1999a)
7. Balsara, D.S., Spicer, D.: A staggered mesh algorithm using high order Godunov fluxes to ensure solenoidal magnetic fields in magnetohydrodynamic simulations. J. Comput. Phys. **149**, 270–292 (1999b)
8. Balsara, D.S., Dumbser, M., Abgrall, R.: Multidimensional HLLC Riemann solver for unstructured meshes—with application to Euler and MHD flows. J. Comput. Phys. **261**, 172–208 (2014)
9. Bouchut, F., Klingenberg, C., Waagan, K.: A multiwave approximate Riemann solver for ideal MHD based on relaxation. I: theoretical framework. Numer. Math. **108**(1), 7–42 (2007)
10. Bouchut, F., Klingenberg, C., Waagan, K.: A multiwave approximate Riemann solver for ideal MHD based on relaxation II: numerical implementation with 3 and 5 waves. Numer. Math. **115**(4), 647–679 (2010)
11. Brackbill, J.U., Barnes, D.C.: The effect of nonzero $\nabla \cdot \mathbf{B}$ on the numerical solution of the magnetodydrodynamic equations. J. Comput. Phys. **35**(3), 426–430 (1980)
12. Chandrashekar, P., Klingenberg, C.: Entropy stable finite volume scheme for ideal compressible MHD on 2-D Cartesian meshes. SIAM J. Numer. Anal. **54**(2), 1313–1340 (2016)
13. Cheng, Y., Li, F., Qiu, J., Xu, L.: Positivity-preserving DG and central DG methods for ideal MHD equations. J. Comput. Phys. **238**, 255–280 (2013)
14. Christlieb, A.J., Rossmanith, J.A., Tang, Q.: Finite difference weighted essentially non-oscillatory schemes with constrained transport for ideal magnetohydrodynamics. J. Comput. Phys. **268**, 302–325 (2014)
15. Christlieb, A.J., Liu, Y., Tang, Q., Xu, Z.: High order parametrized maximum-principle-preserving and positivity-preserving WENO schemes on unstructured meshes. J. Comput. Phys. **281**, 334–351 (2015a)
16. Christlieb, A.J., Liu, Y., Tang, Q., Xu, Z.: Positivity-preserving finite difference weighted ENO schemes with constrained transport for ideal magnetohydrodynamic equations. SIAM J. Sci. Comput. **37**(4), A1825–A1845 (2015b)
17. Christlieb, A.J., Feng, X., Seal, D.C., Tang, Q.: A high-order positivity-preserving single-stage single-step method for the ideal magnetohydrodynamic equations. J. Comput. Phys. **316**, 218–242 (2016)
18. Dai, W., Woodward, P.R.: A simple finite difference scheme for multidimensional magnetohydrodynamical equations. J. Comput. Phys. **142**(2), 331–369 (1998)
19. Davis, S.F.: Simplified second-order Godunov-type methods. SIAM J. Sci. Stat. Comput. **9**(3), 445–473 (1988)
20. Dedner, A., Kemm, F., Kröner, D., Munz, C.D., Schnitzer, T., Wesenberg, M.: Hyperbolic divergence cleaning for the MHD equations. J. Comput. Phys. **175**(2), 645–673 (2002)
21. Dellar, P.J.: A note on magnetic monopoles and the one-dimensional MHD Riemann problem. J. Comput. Phys. **172**(1), 392–398 (2001)

22. Derigs, D., Gassner, G.J., Walch, S., Winters, A.R.: Entropy stable finite volume approximations for ideal magnetohydrodynamics. Jahresbericht der Deutschen Mathematiker-Vereinigung **120**, 153–219 (2018a)
23. Derigs, D., Winters, A.R., Gassner, G.J., Walch, S., Bohm, M.: Ideal GLM-MHD: about the entropy consistent nine-wave magnetic field divergence diminishing ideal magnetohydrodynamics equations. J. Comput. Phys. **364**, 420–467 (2018b)
24. Du, J., Shu, C.W.: Positivity-preserving high-order schemes for conservation laws on arbitrarily distributed point clouds with a simple WENO limiter. Int. J. Numer. Anal. Model. **15**, 1–25 (2018)
25. Einfeldt, B., Munz, C.D., Roe, P.L., Sjögreen, B.: On Godunov-type methods near low densities. J. Comput. Phys. **92**(2), 273–295 (1991)
26. Evans, C.R., Hawley, J.F.: Simulation of magnetohydrodynamic flows: a constrained transport method. Astrophys. J. **332**, 659–677 (1988)
27. Fu, P., Li, F., Xu, Y.: Globally divergence-free discontinuous Galerkin methods for ideal magnetohydrodynamic equations. J. Sci. Comput. **77**, 1621–1659 (2018)
28. Fuchs, F.G., McMurry, A.D., Mishra, S., Risebro, N.H., Waagan, K.: Approximate Riemann solvers and robust high-order finite volume schemes for multi-dimensional ideal MHD equations. Commun. Comput. Phys. **9**(2), 324–362 (2011)
29. Gardiner, T.A., Stone, J.M.: An unsplit Godunov method for ideal MHD via constrained transport. J. Comput. Phys. **205**(2), 509–539 (2005)
30. Godunov, S.K.: Symmetric form of the equations of magnetohydrodynamics. Numer. Methods Mech. Contin. Medium **1**, 26–34 (1972)
31. Gottlieb, S., Ketcheson, D.I., Shu, C.W.: High order strong stability preserving time discretizations. J. Sci. Comput. **38**(3), 251–289 (2009)
32. Gurski, K.: An HLLC-type approximate Riemann solver for ideal magnetohydrodynamics. SIAM J. Sci. Comput. **25**(6), 2165–2187 (2004)
33. Hu, X.Y., Adams, N.A., Shu, C.W.: Positivity-preserving method for high-order conservative schemes solving compressible Euler equations. J. Comput. Phys. **242**, 169–180 (2013)
34. Janhunen, P.: A positive conservative method for magnetohydrodynamics based on HLL and Roe methods. J. Comput. Phys. **160**(2), 649–661 (2000)
35. Krivodonova, L., Xin, J., Remacle, J.F., Chevaugeon, N., Flaherty, J.E.: Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws. Appl. Numer. Math. **48**(3–4), 323–338 (2004)
36. Li, F., Shu, C.W.: Locally divergence-free discontinuous Galerkin methods for MHD equations. J. Sci. Comput. **22**(1–3), 413–442 (2005)
37. Li, F., Xu, L., Yakovlev, S.: Central discontinuous Galerkin methods for ideal MHD equations with the exactly divergence-free magnetic field. J. Comput. Phys. **230**(12), 4828–4847 (2011)
38. Li, S.: An HLLC Riemann solver for magneto-hydrodynamics. J. Comput. Phys. **203**(1), 344–357 (2005)
39. Liang, C., Xu, Z.: Parametrized maximum principle preserving flux limiters for high order schemes solving multi-dimensional scalar hyperbolic conservation laws. J. Sci. Comput. **58**(1), 41–60 (2014)
40. Liu, Y., Shu, C.W., Zhang, M.: Entropy stable high order discontinuous Galerkin methods for ideal compressible MHD on structured meshes. J. Comput. Phys. **354**, 163–178 (2018)
41. Miyoshi, T., Kusano, K.: A multi-state HLL approximate Riemann solver for ideal magnetohydrodynamics. J. Comput. Phys. **208**(1), 315–344 (2005)
42. Powell, K.G.: An approximate Riemann solver for magnetohydrodynamics (that works in more than one dimension). Tech. Rep. ICASE Report No. 94-24, NASA Langley, VA (1994)
43. Powell, K.G., Roe, P., Myong, R., Gombosi, T.: An upwind scheme for magnetohydrodynamics. In: 12th Computational Fluid Dynamics Conference, p. 1704 (1995)
44. Qiu, J., Shu, C.W.: Runge–Kutta discontinuous Galerkin method using WENO limiters. SIAM J. Sci. Comput. **26**(3), 907–929 (2005)
45. Ryu, D., Miniati, F., Jones, T., Frank, A.: A divergence-free upwind code for multidimensional magnetohydrodynamic flows. Astrophys. J. **509**(1), 244–255 (1998)
46. Seal, D.C., Tang, Q., Xu, Z., Christlieb, A.J.: An explicit high-order single-stage single-step positivity-preserving finite difference WENO method for the compressible Euler equations. J. Sci. Comput. **68**(1), 171–190 (2016)
47. Torrilhon, M.: Locally divergence-preserving upwind finite volume schemes for magnetohydrodynamic equations. SIAM J. Sci. Comput. **26**(4), 1166–1191 (2005)

48. Tóth, G.: The $\nabla \cdot \mathbf{B} = 0$ constraint in shock-capturing magnetohydrodynamics codes. J. Comput. Phys. **161**(2), 605–652 (2000)
49. Vilar, F., Shu, C.W., Maire, P.H.: Positivity-preserving cell-centered lagrangian schemes for multi-material compressible flows: from first-order to high-orders. Part II: the two-dimensional case. J. Comput. Phys. **312**, 416–442 (2016)
50. Waagan, K.: A positive MUSCL-Hancock scheme for ideal magnetohydrodynamics. J. Comput. Phys. **228**(23), 8609–8626 (2009)
51. Waagan, K., Federrath, C., Klingenberg, C.: A robust numerical scheme for highly compressible magnetohydrodynamics: nonlinear stability, implementation and tests. J. Comput. Phys. **230**(9), 3331–3351 (2011)
52. Wu, K.: Design of provably physical-constraint-preserving methods for general relativistic hydrodynamics. Phys. Rev. D **95**(10), 103001 (2017)
53. Wu, K.: Positivity-preserving analysis of numerical schemes for ideal magnetohydrodynamics. SIAM J. Numer. Anal. **56**(4), 2124–2147 (2018)
54. Wu, K., Shu, C.W.: A provably positive discontinuous Galerkin method for multidimensional ideal magnetohydrodynamics. SIAM J. Sci. Comput. **40**(5), B1302–B1329 (2018)
55. Wu, K., Tang, H.: High-order accurate physical-constraints-preserving finite difference WENO schemes for special relativistic hydrodynamics. J. Comput. Phys. **298**, 539–564 (2015)
56. Wu, K., Tang, H.: Admissible states and physical-constraints-preserving schemes for relativistic magnetohydrodynamic equations. Math. Models Methods Appl. Sci. **27**(10), 1871–1928 (2017a)
57. Wu, K., Tang, H.: Physical-constraint-preserving central discontinuous Galerkin methods for special relativistic hydrodynamics with a general equation of state. Astrophys. J. Suppl. Ser. **228**(1), 3 (2017b)
58. Xiong, T., Qiu, J.M., Xu, Z.: Parametrized positivity preserving flux limiters for the high order finite difference WENO scheme solving compressible Euler equations. J. Sci. Comput. **67**(3), 1066–1088 (2016)
59. Xu, Z.: Parametrized maximum principle preserving flux limiters for high order schemes solving hyperbolic conservation laws: one-dimensional scalar problem. Math. Comput. **83**(289), 2213–2238 (2014)
60. Xu, Z., Zhang, X.: Bound-preserving high order schemes. In: Abgrall, R., Shu, C.-W. (eds.) Handbook of Numerical Methods for Hyperbolic Problems: Applied and Modern Issues, vol. 18. Elsevier, North-Holland, Amsterdam (2017)
61. Xu, Z.L., Liu, Y.J.: New central and central discontinuous Galerkin schemes on overlapping cells of unstructured grids for solving ideal magnetohydrodynamic equations with globally divergence-free magnetic field. J. Comput. Phys. **327**, 203–224 (2016)
62. Xu, Z.L., Balsara, D.S., Du, H.: Divergence-free WENO reconstruction-based finite volume scheme for solving ideal MHD equations on triangular meshes. Commun. Comput. Phys. **19**(4), 841–880 (2016)
63. Yakovlev, S., Xu, L., Li, F.: Locally divergence-free central discontinuous Galerkin methods for ideal MHD equations. J. Comput. Sci. **4**(1–2), 80–91 (2013)
64. Zhang, X.: On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations. J. Comput. Phys. **328**, 301–343 (2017)
65. Zhang, X., Shu, C.W.: On maximum-principle-satisfying high order schemes for scalar conservation laws. J. Comput. Phys. **229**(9), 3091–3120 (2010a)
66. Zhang, X., Shu, C.W.: On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. J. Comput. Phys. **229**(23), 8918–8934 (2010b)
67. Zhang, X., Shu, C.W.: Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. J. Comput. Phys. **230**(4), 1238–1248 (2011)
68. Zhang, X., Xia, Y., Shu, C.W.: Maximum-principle-satisfying and positivity-preserving high order discontinuous galerkin schemes for conservation laws on triangular meshes. J. Sci. Comput. **50**(1), 29–62 (2012)
69. Zhao, J., Tang, H.: Runge–Kutta discontinuous Galerkin methods for the special relativistic magnetohydrodynamics. J. Comput. Phys. **343**, 33–72 (2017)

Affiliations

Kailiang Wu[1] · **Chi-Wang Shu**[2]

✉ Kailiang Wu
wu.3423@osu.edu

Chi-Wang Shu
shu@dam.brown.edu

1 Department of Mathematics, The Ohio State University, Columbus, OH 43210, USA

2 Division of Applied Mathematics, Brown University, Providence, RI 02912, USA