

# Structure-preserving Local Discontinuous Galerkin method for nonlinear cross-diffusion systems

Sergio Gómez\*

Ansgar Jüngel†

Ilaria Perugia‡

2024-07-21

## Abstract

We present and analyze a structure-preserving method for the approximation of solutions to nonlinear cross-diffusion systems, which combines a Local Discontinuous Galerkin spatial discretization with the backward Euler time stepping scheme. The proposed method makes use of the underlying entropy structure of the system, expressing the main unknown in terms of the entropy variable by means of a nonlinear transformation. Such a transformation allows for imposing the physical positivity or boundedness constraints on the approximate solution in a strong sense. Moreover, nonlinearities do not appear explicitly within differential operators or interface terms in the scheme, which significantly improves its efficiency and ease its implementation. We prove the existence of discrete solutions and their asymptotic convergence to a weak solution to the continuous problem. Numerical results for some one- and two-dimensional problems illustrate the accuracy and entropy stability of the proposed method.

**Keywords.** Structure-preserving method, entropy stability, nonlinear cross-diffusion systems, Local Discontinuous Galerkin method.

**Mathematics Subject Classification.** 65M60, 65M12, 35K51, 35K55, 35Q92.

## 1 Introduction

We consider the following nonlinear reaction-diffusion system on a space-time cylinder  $Q_T = \Omega \times (0, T]$ , where  $\Omega \subset \mathbb{R}^d$  ( $d \in \{1, 2, 3\}$ ) is a bounded, polytopic domain with Lipschitz boundary  $\partial\Omega$ , and  $T > 0$ :

$$\partial_t \boldsymbol{\rho} - \nabla \circ (A(\boldsymbol{\rho}) \nabla \boldsymbol{\rho}) = \mathbf{f}(\boldsymbol{\rho}) \quad \text{in } Q_T, \quad (1.1a)$$

$$(A(\boldsymbol{\rho}) \nabla \boldsymbol{\rho}) \underline{n}_\Omega = \mathbf{0} \quad \text{on } \partial\Omega \times (0, T), \quad (1.1b)$$

$$\boldsymbol{\rho} = \boldsymbol{\rho}_0 \quad \text{on } \Omega \times \{0\}. \quad (1.1c)$$

Here, the unknown  $\boldsymbol{\rho} := (\rho_1, \dots, \rho_N) : Q_T \rightarrow \mathbb{R}^N$  for some number of species  $N \in \mathbb{N}$ ,  $A : \mathbb{R}^N \rightarrow \mathbb{R}^{N \times N}$  is the diffusion matrix,  $\mathbf{f} : \mathbb{R}^N \rightarrow \mathbb{R}^N$  describes the nonlinear interaction between the  $N$  species, and  $\boldsymbol{\rho}_0 \in L^\infty(\Omega)^N$  is a given initial datum. We denote by  $\nabla(\cdot)$  the  $\mathbb{R}^{N \times d}$  matrix, whose rows contain the componentwise spatial gradients, by  $\nabla \circ (\cdot)$  the row-wise spatial divergence operator, and by  $\underline{n}_\Omega$  the  $d$ -dimensional vector of the spatial components of the unit normal vector at  $\partial\Omega \times (0, T)$  pointing outside  $\Omega \times (0, T)$ .

The main challenges in the numerical approximation of the solution to nonlinear cross-diffusion systems are twofold: *i*) the diffusion matrix  $A(\cdot)$  may not be symmetric nor positive definite, and *ii*) a maximum principle may not be available. These issues prevent the use of standard techniques for the analysis of such systems, and make it difficult to guarantee that even weak solutions to the continuous problem respect the positivity or boundedness constraints of the physical unknowns. The boundedness-by-entropy framework in [14], which we describe below, circumvents these issues by exploiting the underlying entropy structure of the system.

We focus on discontinuous Galerkin (DG) methods, which are characterized by the use of discrete broken spaces without any enforced conformity. Among many other advantages, DG methods offer great versatility

---

\*Department of Mathematics and Applications, University of Milano-Bicocca, 20125 Milan, Italy ([sergio.gomezmacias@unimib.it](mailto:sergio.gomezmacias@unimib.it))

†Institute of Analysis and Scientific Computing, Vienna University of Technology, 1040 Vienna, Austria ([juengel@tuwien.ac.at](mailto:juengel@tuwien.ac.at))

‡Faculty of Mathematics, University of Vienna, 1090 Vienna, Austria ([ilaria.perugia@univie.ac.at](mailto:ilaria.perugia@univie.ac.at))

for the treatment of nonlinearities. In particular, the Local Discontinuous Galerkin (LDG) method, originally introduced in [11] for nonlinear convection-diffusion systems, does not require nonlinearities to appear within differential operators or interface terms, leading to nonlinear operators that can be evaluated naturally in parallel. Such a property is the result of appropriately rewriting the original problem in terms of auxiliary variables, and making use of  $L^2$ -orthogonal projections in the discrete space of the nonlinear terms (see, e.g., [11, 13]).

In order to obtain physically consistent discrete solutions, it is of utmost importance to design numerical methods that are not only accurate and efficient, but also reproduce, at the discrete level, the geometric and physical properties of the phenomenon being modeled. Such numerical methods are called *structure preserving*. One of the most difficult properties to reproduce at the discrete level is the physically expected positivity or boundedness of the continuous solution in finite element discretizations, especially for high-order approximations. Although this is a well-known issue (see, e.g., the recent review in [4] on finite element methods (FEM) respecting the discrete maximum principle for convection-diffusion equations), only in last years has major progress been made in the literature. We briefly mention some recent works on this subject that does not rely on slope limiters or postprocessing techniques. In [3], the authors proposed a nodally bound-preserving FEM, whose discrete solution belongs to the convex set of piecewise polynomials satisfying the physical bound constraints on the mesh nodes. While this suffices to ensure strong (pointwise) positivity of the discrete solution for linear approximations, it does not provide any control on the values of the discrete solution away from the mesh nodes for higher-order approximations. Motivated by the underlying entropy structure of the concerned PDEs, nonlinear transformations in terms of the entropy variable have been used to enforce positivity on the approximate solution of interior-penalty DG [6, 12], conforming FEM [7], and hybrid high-order (HHO) [19] discretizations.

In this work, we propose an LDG method for the numerical approximation of the nonlinear cross-diffusion system (1.1), which is based on the framework of [14], and possesses the following desirable properties:

- it allows for arbitrary degrees of approximation in space;
- it preserves the boundedness of the physical unknowns without requiring any postprocessing or slope limiter;
- nonlinearities do not appear explicitly within differential operators or interface terms, which endows the method with a natural parallelizable structure and high efficiency;
- it respects a discrete version of the entropy stability estimate of the continuous problem.

Although numerical methods for nonlinear cross-diffusion systems with some of these properties can be found in the literature, to the best of our knowledge, the proposed method is the first one satisfying all of them. For instance, finite volume methods for cross-diffusion systems have been proposed in [1, 9, 18], but at most second-order convergence rates in space are numerically obtained, whereas the entropy stable high-order DG method introduced in [24] guarantees only weak positivity on Cartesian meshes by means of scaling limiters.

**The boundedness-by-entropy framework.** Henceforth, we make the following assumptions:

(H1)  $A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$  and  $\mathbf{f} \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^N)$ , for a bounded domain  $\mathcal{D} \subset (0, \infty)^N$ .

(H2) There exists a convex function  $s \in \mathcal{C}^2(\mathcal{D}; (0, \infty)) \cap \mathcal{C}^0(\overline{\mathcal{D}}; (0, \infty))$ , with  $s' : \mathcal{D} \rightarrow \mathbb{R}^N$  invertible and inverse  $\mathbf{u} := (s')^{-1} \in \mathcal{C}^1(\mathbb{R}^N; \mathcal{D})$ , such that the following three conditions are satisfied:

(H2a) There exists a constant  $\gamma > 0$  such that

$$\mathbf{z} \cdot (s''(\boldsymbol{\rho})A(\boldsymbol{\rho})\mathbf{z}) \geq \gamma |\mathbf{z}|^2 \quad \forall \mathbf{z} \in \mathbb{R}^N, \boldsymbol{\rho} \in \mathcal{D}.$$

(H2b) There exists a constant  $C_f \geq 0$  such that

$$\mathbf{f}(\boldsymbol{\rho}) \cdot s'(\boldsymbol{\rho}) \leq C_f \quad \forall \boldsymbol{\rho} \in \mathcal{D}.$$

(H2c) The initial datum  $\boldsymbol{\rho}_0$  is integrable and satisfies  $\boldsymbol{\rho}_0(\mathbf{x}) \in \mathcal{D}$  for a.e.  $\mathbf{x} \in \Omega$ .

Observe that Assumption (H2c) implies that

$$\int_{\Omega} s(\boldsymbol{\rho}_0) d\mathbf{x} < \infty,$$

since  $\Omega$  is bounded,  $s$  is bounded on  $\overline{\mathcal{D}}$ , and  $\boldsymbol{\rho}_0$  is bounded.

Let us comment on these assumptions. The boundedness of the domain  $\mathcal{D}$  guarantees the boundedness of the solutions; see below. The function  $s$  can be interpreted as the entropy or free energy of the underlying physical problem. Assumption (H2a) requires that the product  $s''(\boldsymbol{\rho})A(\boldsymbol{\rho})$  is positive definite. This condition can be generalized to include degenerate or singular continuous problems [14, Ch. 4]. We present some examples of cross-diffusion systems that satisfy Assumption (H2a) in Appendix A. Assumption (H2b) can be generalized to  $\mathbf{f}(\boldsymbol{\rho}) \cdot s'(\boldsymbol{\rho}) \leq C_f(1+s(\boldsymbol{\rho}))$  for all  $\boldsymbol{\rho} \in \mathcal{D}$ , which yields an additional constant depending on time in inequality (1.3) below. The condition that the initial datum lies pointwise in  $\mathcal{D}$  means that vacuum is initially excluded. This condition can be generalized to include vacuum, i.e.  $\boldsymbol{\rho}_0 \in \overline{\mathcal{D}}$ ; see Remark 2.9.

The main idea of the boundedness-by-entropy framework in [14] consists in introducing the entropy variable  $\mathbf{w} := s'(\boldsymbol{\rho})$  and then use the invertibility of  $s'(\cdot)$  in Assumption (H2) to write the original unknown as  $\boldsymbol{\rho} = (s')^{-1}(\mathbf{w}) = \mathbf{u}(\mathbf{w})$ . In this way, the boundedness of  $\mathcal{D}$  in Assumption (H1) implies the pointwise boundedness of  $\mathbf{u}(\mathbf{w})$ , without requiring a maximum principle. Due to the regularity of the entropy density function  $s(\cdot)$  in Assumption (H2), the following chain rule is valid:

$$\nabla \mathbf{w} = \nabla (s'(\boldsymbol{\rho})) = s''(\boldsymbol{\rho}) \nabla \boldsymbol{\rho}. \quad (1.2)$$

Taking  $\mathbf{w}$  as the test function of the weak formulation of (1.1), the following entropy stability estimate follows from Assumptions (H2a)–(H2c) and the chain rule in (1.2):

$$\int_{\Omega} s(\boldsymbol{\rho}(\mathbf{x}, \tau)) d\mathbf{x} + \gamma \int_0^{\tau} \|\nabla \boldsymbol{\rho}\|_{[L^2(\Omega)]^d}^2 dt \leq \int_{\Omega} s(\boldsymbol{\rho}_0) d\mathbf{x} + C_f \tau |\Omega| \quad \text{for all } 0 < \tau \leq T. \quad (1.3)$$

**Outline of the paper.** In Section 2, we first rewrite the nonlinear cross-diffusion system in (1.1) in terms of some suitably chosen auxiliary variables. In Section 2.1, we present an LDG semidiscrete-in-space formulation of the rewritten system and prove its entropy stability. In Section 2.2, such a semidiscrete LDG formulation is combined with the backward Euler time discretization and a regularizing term to get a fully discrete scheme. Section 3 is devoted to the proof of the existence of discrete solutions. In Section 4.1, we introduce the assumptions on the regularizing term and the discrete spaces that are used to prove the convergence to semidiscrete-in-time solutions in Section 4.2, and to weak solutions to the continuous problem in Section 4.3. The validity of such assumptions for different cases is discussed in Section 4.4. Some numerical experiments in one and two dimensions are presented in Section 5 to assess the accuracy and entropy stability of the scheme. We finish with some concluding remarks in Section 6.

## 2 Definition of the method

We use the following notation for functions with  $N$  scalar-valued components and with  $N$   $d$ -vector-valued components, respectively:

$$\boldsymbol{\mu} = (\mu_1, \dots, \mu_N), \quad \underline{\boldsymbol{\mu}} = (\underline{\mu}_1, \dots, \underline{\mu}_N).$$

For the discretization in space, we introduce a DG approximation of problem (1.1), where nonlinearities do not appear within differential operators or interface terms, and a discrete version of the chain rule in (1.2) is satisfied. To this aim, we introduce the auxiliary variables  $\mathbf{w}$ ,  $\boldsymbol{\zeta}$ ,  $\boldsymbol{\sigma}$ , and  $\mathbf{q}$  defined by

$$\boldsymbol{\rho} := \mathbf{u}(\mathbf{w}), \quad (2.1a)$$

$$\boldsymbol{\zeta} := -\nabla \mathbf{w}, \quad (2.1b)$$

$$A(\boldsymbol{\rho})^T s''(\boldsymbol{\rho}) \boldsymbol{\sigma} := -A(\boldsymbol{\rho})^T s''(\boldsymbol{\rho}) \nabla \boldsymbol{\rho} = A(\boldsymbol{\rho})^T \boldsymbol{\zeta}, \quad (2.1c)$$

$$\mathbf{q} := A(\boldsymbol{\rho}) \boldsymbol{\sigma}, \quad (2.1d)$$

and rewrite problem (1.1) as

$$\begin{cases} \partial_t \boldsymbol{\rho} + \nabla \circ \mathbf{q} = \mathbf{f}(\boldsymbol{\rho}) & \text{in } Q_T, \\ \mathbf{q} \circ \mathbf{n}_{\Omega} = \mathbf{0} & \text{on } \partial\Omega \times (0, T), \\ \boldsymbol{\rho} = \boldsymbol{\rho}_0 & \text{on } \Omega \times \{0\}, \end{cases}$$

where  $\circ$  denotes the product

$$\boldsymbol{\mu} \circ \mathbf{n} := (\mu_1 \cdot \mathbf{n}, \dots, \mu_N \cdot \mathbf{n}),$$

for all vectors of  $N$   $d$ -vector-valued components  $\underline{\mu} = (\mu_1, \dots, \mu_N)$  and  $\underline{n} \in \mathbb{R}^d$ .

As  $A(\underline{\rho})^T s''(\underline{\rho})$  is positive definite by assumption (H2a), on the continuous level, definition (2.1c) is equivalent to  $\underline{\sigma} = -\nabla \underline{\rho}$ . Moreover, from (2.1a) and (2.1b), we have that  $\underline{\zeta} = -\nabla (s'(\underline{\rho}))$ . Therefore, definition (2.1c) is a reformulation of the chain rule (1.2) in terms of the auxiliary variables, which will guarantee that a discrete version of (1.2) suitable for the analysis of the method is satisfied.

## 2.1 Semi-discretization in space

Let  $\{\mathcal{T}_h\}_{h>0}$  be a family of conforming simplicial meshes of the spatial domain  $\Omega$  with maximum element diameter (mesh size)  $h$ . If  $d = 2, 3$ , we assume that the family  $\{\mathcal{T}_h\}_{h>0}$  satisfies the shape-regularity condition, i.e., there exists a constant  $\Upsilon > 0$  independent of  $h$  such that, for all  $K \in \mathcal{T}_h$ ,

$$\Upsilon h_K \leq \varrho_K, \quad (2.2)$$

where  $h_K$  denotes the diameter of  $K$  and  $\varrho_K$  is the radius of the incircle of  $K$ . We denote the set of all the mesh facets in  $\mathcal{T}_h$  by  $\mathcal{F}_h = \mathcal{F}_h^I \cup \mathcal{F}_h^N$ , where  $\mathcal{F}_h^I$  and  $\mathcal{F}_h^N$  are the sets of internal and (Neumann) boundary facets, respectively. We define the following piecewise polynomial spaces

$$\mathcal{S}_p(\mathcal{T}_h) := \prod_{K \in \mathcal{T}_h} \mathbb{P}^p(K), \quad \mathcal{M}_p(\mathcal{T}_h) := \prod_{K \in \mathcal{T}_h} \mathbb{P}^p(K)^d,$$

where  $\mathbb{P}^p(K)$  denotes the space of scalar-valued polynomials of degree at most  $p$  on the spatial domain  $K$ . We further denote by  $(\partial K)^\circ$  the union of the facets of  $K$  that belong to  $\mathcal{F}_h^I$  and define the piecewise constant function  $\mathbf{h} \in L^\infty(\mathcal{F}_h^I)$  as

$$\mathbf{h}(\mathbf{x}) := \eta^{-1} \min\{h_{K_1}, h_{K_2}\} \quad \text{if } \mathbf{x} \in F, \text{ and } F \in \mathcal{F}_h^I \text{ is shared by } K_1, K_2 \in \mathcal{T}_h, \quad (2.3)$$

for some constant  $\eta > 0$  independent of the mesh size.

For any element  $K \in \mathcal{T}_h$ , let  $\underline{n}_K$  be the unit normal  $d$ -dimensional vector to  $\partial K$  pointing outside  $K$ . For any piecewise smooth, scalar-valued function  $\mu$  and any  $\alpha_F \in [0, 1]$ , we define jumps and weighted mean-values, respectively, on each facet  $F \in \mathcal{F}_h^I$  by

$$[[\mu]]_{\mathbf{N}} := \mu|_{K_1} \underline{n}_{K_1} + \mu|_{K_2} \underline{n}_{K_2}, \quad \{\{\mu\}\}_{\alpha_F} := (1 - \alpha_F) \mu|_{K_1} + \alpha_F \mu|_{K_2} \quad \text{if } F = \partial K_1 \cap \partial K_2 \text{ for some } K_1, K_2 \in \mathcal{T}_h.$$

For piecewise smooth functions  $\underline{\mu}$  with  $N$  scalar-valued components,  $[[\underline{\mu}]]_{\mathbf{N}}$  and  $\{\{\underline{\mu}\}\}_{\alpha_F}$  are defined componentwise. Similarly, for piecewise smooth functions  $\underline{\mu}$  with  $N$   $d$ -vector-valued components,  $\{\{\underline{\mu}\}\}_{\alpha_F}$  is defined componentwise.

We propose the following structure-preserving LDG-like semidiscrete formulation: for any fixed  $t \in (0, T]$ , find  $(\mathbf{w}_h(\cdot, t), \underline{\zeta}_h(\cdot, t), \underline{\sigma}_h(\cdot, t), \underline{\mathbf{q}}_h(\cdot, t)) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N$  such that, on each element  $K \in \mathcal{T}_h$ ,

$$\int_K \underline{\zeta}_h \cdot \underline{\psi}_h \, d\mathbf{x} = - \int_{\partial K} \widehat{\mathbf{w}}_h \cdot (\underline{\psi}_h \circ \underline{n}_K) \, dS + \int_K \mathbf{w}_h \cdot (\nabla \circ \underline{\psi}_h) \, d\mathbf{x}, \quad (2.4a)$$

$$\int_K A(\mathbf{u}(\mathbf{w}_h))^T s''(\mathbf{u}(\mathbf{w}_h)) \underline{\sigma}_h \cdot \underline{\varphi}_h \, d\mathbf{x} = \int_K A(\mathbf{u}(\mathbf{w}_h))^T \underline{\zeta}_h \cdot \underline{\varphi}_h \, d\mathbf{x}, \quad (2.4b)$$

$$\int_K \underline{\mathbf{q}}_h \cdot \underline{\theta}_h \, d\mathbf{x} = \int_K A(\mathbf{u}(\mathbf{w}_h)) \underline{\sigma}_h \cdot \underline{\theta}_h \, d\mathbf{x}, \quad (2.4c)$$

$$\int_K \partial_t(\mathbf{u}(\mathbf{w}_h)) \cdot \underline{\lambda}_h \, d\mathbf{x} + \int_{\partial K} (\widehat{\underline{\mathbf{q}}}_h \circ \underline{n}_K) \cdot \underline{\lambda}_h \, dS - \int_K \underline{\mathbf{q}}_h \cdot \nabla \underline{\lambda}_h \, d\mathbf{x} = \int_K \mathbf{f}(\mathbf{u}(\mathbf{w}_h)) \cdot \underline{\lambda}_h \, d\mathbf{x}, \quad (2.4d)$$

for all test functions  $(\underline{\lambda}_h, \underline{\psi}_h, \underline{\varphi}_h, \underline{\theta}_h) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N$ , with  $\mathbf{w}_h(\cdot, 0) \in \mathcal{S}_p(\mathcal{T}_h)^N$  an approximation of  $s'(\underline{\rho}_0)$ . Here, the numerical fluxes  $\widehat{\mathbf{w}}_h$ ,  $\widehat{\underline{\mathbf{q}}}_h$  are approximations of the traces of  $\mathbf{w}_h$  and  $\underline{\mathbf{q}}_h$ , respectively, on the skeleton of  $\mathcal{T}_h$ . They are defined on each facet  $F \in \mathcal{F}_h$  as

$$\widehat{\mathbf{w}}_h := \begin{cases} \{\{\mathbf{w}_h\}\}_{\alpha_F} & \text{if } F \in \mathcal{F}_h^I \text{ and } F = \partial K_1 \cap \partial K_2 \text{ for some } K_1, K_2 \in \mathcal{T}_h, \\ \mathbf{w}_h & \text{if } F \in \mathcal{F}_h^N, \end{cases} \quad (2.5a)$$

$$\widehat{\underline{\mathbf{q}}}_h := \begin{cases} \{\{\underline{\mathbf{q}}_h\}\}_{1-\alpha_F} + \eta_F [[\underline{\mathbf{w}}_h]]_{\mathbf{N}} & \text{if } F \in \mathcal{F}_h^I \text{ and } F = \partial K_1 \cap \partial K_2 \text{ for some } K_1, K_2 \in \mathcal{T}_h, \\ \mathbf{0} & \text{if } F \in \mathcal{F}_h^N, \end{cases} \quad (2.5b)$$

where the weighted-mean parameter  $\alpha_F \in [0, 1]$  and the stabilization function  $\eta_F$  are defined on each facet  $F \in \mathcal{F}_h^I$ . We define  $\eta_F$  as

$$\eta_F = \mathbf{h}^{-1} \|A\|_{L^\infty(\mathcal{D})^{N \times N}}. \quad (2.6)$$

Taking the  $L^\infty$  norm of  $A$  in (2.6) may introduce additional diffusion. However, it avoids a nonlinear dependence of the stability term on  $\mathbf{u}(\mathbf{w}_h)$ .

**Remark 2.1** (Computation of  $\underline{\sigma}_h$ ). *The definition of  $\underline{\sigma}_h$  in (2.4b) is local. More precisely, given  $\mathbf{w}_h$ , the construction of  $\underline{\sigma}_h$  requires only the solution of completely independent (naturally parallelizable) linear (in  $\underline{\sigma}_h$ ) problems on each element  $K \in \mathcal{T}_h$ . In each of these local problems, the components of  $\underline{\sigma}_h$  for the  $N$  species are coupled. This is a consequence of the presence of the matrices  $A(\mathbf{u}(\mathbf{w}_h))^T s''(\mathbf{u}(\mathbf{w}_h))$  and  $A(\mathbf{u}(\mathbf{w}_h))^T$  on the left- and right-hand side integrals of (2.4b), respectively. ■*

Let  $M$ ,  $B$ , and  $S$  be the matrix representations, respectively, of the bilinear forms

$$\mathbf{m}_h(\underline{\zeta}_h, \underline{\psi}_h) := \int_{\Omega} \underline{\zeta}_h \cdot \underline{\psi}_h \, d\mathbf{x} \quad \forall \underline{\zeta}_h, \underline{\psi}_h \in \mathcal{M}_p(\mathcal{T}_h), \quad (2.7a)$$

$$\begin{aligned} b_h(w_h, \underline{\psi}_h) &:= - \int_{\mathcal{F}_h^I} \{\{w_h\}\}_{\alpha_F} \llbracket \underline{\psi}_h \rrbracket_{\mathbf{N}} \, dS - \int_{\mathcal{F}_h^N} w_h \underline{\psi}_h \cdot \mathbf{n}_{\Omega} \, dS \\ &\quad + \sum_{K \in \mathcal{T}_h} \int_K w_h \nabla \cdot \underline{\psi}_h \, d\mathbf{x} \quad \forall w_h \in \mathcal{S}_p(\mathcal{T}_h), \underline{\psi}_h \in \mathcal{M}_p(\mathcal{T}_h), \end{aligned} \quad (2.7b)$$

$$s_h(w_h, \lambda_h) := \int_{\mathcal{F}_h^I} \eta_F \llbracket w_h \rrbracket_{\mathbf{N}} \cdot \llbracket \lambda_h \rrbracket_{\mathbf{N}} \, dS \quad \forall w_h, \lambda_h \in \mathcal{S}_p(\mathcal{T}_h), \quad (2.7c)$$

and let  $\mathbf{u}_h$ ,  $\mathbf{n}_h$ ,  $\tilde{\mathbf{a}}_h$ ,  $\mathbf{a}_h$ , and  $\mathbf{f}_h$  be the operators associated with the nonlinear functionals

$$\mathbf{u}_h(\mathbf{w}_h, \phi_h) := \int_{\Omega} \mathbf{u}(\mathbf{w}_h) \cdot \phi_h \, d\mathbf{x} \quad \forall \mathbf{w}_h, \phi_h \in \mathcal{S}_p(\mathcal{T}_h)^N \quad (2.8a)$$

$$\mathbf{n}_h(\mathbf{w}_h; \underline{\sigma}_h, \underline{\varphi}_h) := \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h))^T s''(\mathbf{u}(\mathbf{w}_h)) \underline{\sigma}_h \cdot \underline{\varphi}_h \, d\mathbf{x} \quad \forall (\mathbf{w}_h, \underline{\sigma}_h, \underline{\varphi}_h) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N, \quad (2.8b)$$

$$\tilde{\mathbf{a}}_h(\mathbf{w}_h; \underline{\zeta}_h, \underline{\varphi}_h) := \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h))^T \underline{\zeta}_h \cdot \underline{\varphi}_h \, d\mathbf{x} \quad \forall (\mathbf{w}_h, \underline{\zeta}_h, \underline{\varphi}_h) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N, \quad (2.8c)$$

$$\mathbf{a}_h(\mathbf{w}_h; \underline{\sigma}_h, \underline{\theta}_h) := \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h)) \underline{\sigma}_h \cdot \underline{\theta}_h \, d\mathbf{x} \quad \forall (\mathbf{w}_h, \underline{\sigma}_h, \underline{\theta}_h) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N, \quad (2.8d)$$

$$\mathbf{f}_h(\mathbf{w}_h, \lambda_h) := \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}_h)) \cdot \lambda_h \, d\mathbf{x} \quad \forall \mathbf{w}_h, \lambda_h \in \mathcal{S}_p(\mathcal{T}_h)^N. \quad (2.8e)$$

After summing (2.4a)–(2.4d) over all the elements  $K \in \mathcal{T}_h$ , by the average-jump identity

$$\{\{\lambda_h\}\}_{\alpha_F} \llbracket \underline{\psi}_h \rrbracket_{\mathbf{N}} + \{\{\underline{\psi}_h\}\}_{1-\alpha_F} \cdot \llbracket \lambda_h \rrbracket_{\mathbf{N}} = \llbracket \lambda_h \underline{\psi}_h \rrbracket_{\mathbf{N}},$$

we get

$$\sum_{i=1}^N \mathbf{m}_h(\underline{\zeta}_{h,i}, \underline{\psi}_{h,i}) = \sum_{i=1}^N b_h(w_{h,i}, \underline{\psi}_{h,i}), \quad (2.9a)$$

$$\mathbf{n}_h(\mathbf{w}_h; \underline{\sigma}_h, \underline{\varphi}_h) = \tilde{\mathbf{a}}_h(\mathbf{w}_h; \underline{\zeta}_h, \underline{\varphi}_h), \quad (2.9b)$$

$$\sum_{i=1}^N \mathbf{m}_h(q_{h,i}, \underline{\theta}_{h,i}) = \mathbf{a}_h(\mathbf{w}_h; \underline{\sigma}_h, \underline{\theta}_h), \quad (2.9c)$$

$$\frac{d}{dt} \mathbf{u}_h(\mathbf{w}_h, \lambda_h) + \sum_{i=1}^N b_h(\lambda_{h,i}, q_{h,i}) + \sum_{i=1}^N s_h(w_{h,i}, \lambda_{h,i}) = \mathbf{f}_h(\mathbf{w}_h, \lambda_h). \quad (2.9d)$$

The ordinary differential equation (ODE) system (2.9a)–(2.9d) can be written in operator form as

$$(I_N \otimes M) \mathbf{Z}_h = (I_N \otimes B) \mathbf{W}_h, \quad (2.10a)$$

$$\mathcal{N}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h) = \tilde{\mathcal{A}}_h(\mathbf{W}_h; \mathbf{Z}_h), \quad (2.10b)$$

$$(I_N \otimes M) \mathbf{Q}_h = \mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h), \quad (2.10c)$$

$$\frac{d}{dt} \mathbf{U}_h(\mathbf{W}_h) + (I_N \otimes B^T) \mathbf{Q}_h + (I_N \otimes S) \mathbf{W}_h = \mathcal{F}_h(\mathbf{W}_h), \quad (2.10d)$$

where  $I_N$  denotes the identity matrix of size  $N$ ,  $\otimes$  the Kronecker product and  $\mathbf{W}_h, \mathbf{Z}_h, \boldsymbol{\Sigma}_h, \mathbf{Q}_h$  are the vector representations of  $\mathbf{w}_h, \underline{\zeta}_h, \boldsymbol{\sigma}_h, \mathbf{q}_h$ , respectively.

Since the nonlinear operators  $\mathcal{A}_h, \tilde{\mathcal{A}}_h$ , and  $\mathcal{N}_h$  are linear with respect to their second argument, equations (2.10b) and (2.10c) can be rewritten as

$$\widehat{\mathcal{N}}_h(\mathbf{W}_h) \boldsymbol{\Sigma}_h = \widehat{\mathcal{A}}_h(\mathbf{W}_h)^T \mathbf{Z}_h, \quad (I_N \otimes M) \mathbf{Q}_h = \widehat{\mathcal{A}}_h(\mathbf{W}_h) \boldsymbol{\Sigma}_h,$$

for some block-diagonal matrices  $\widehat{\mathcal{N}}_h(\mathbf{W}_h)$  and  $\widehat{\mathcal{A}}_h(\mathbf{W}_h)$ . Moreover, due to Assumption (H2a), the matrix  $\widehat{\mathcal{N}}_h(\mathbf{W}_h)$  is positive definite.

Eliminating  $\mathbf{Z}_h$  and  $\mathbf{Q}_h$ , we can write the ODE system (2.10) in the compact form

$$\mathcal{N}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h) = \tilde{\mathcal{A}}_h(\mathbf{W}_h; (I_N \otimes M^{-1} B) \mathbf{W}_h), \quad (2.11a)$$

$$\frac{d}{dt} \mathbf{U}_h(\mathbf{W}_h) + (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h) + (I_N \otimes S) \mathbf{W}_h = \mathcal{F}_h(\mathbf{W}_h). \quad (2.11b)$$

In the following Lemma 2.2, we prove some properties of the bilinear forms and nonlinear functionals defined in (2.7) and (2.8), respectively. From here on, we denote by  $\nabla_h(\cdot)$  the elementwise  $\nabla(\cdot)$  operator.

**Lemma 2.2.** *The bilinear forms defined in (2.7) and the nonlinear functionals defined in (2.8) satisfy the following continuity bounds:*

$$\sum_{i=1}^N m_h(\zeta_{h,i}, \psi_{h,i}) \leq \|\underline{\zeta}_h\|_{[L^2(\Omega)^{d_1}]^N} \|\underline{\psi}_h\|_{[L^2(\Omega)^{d_1}]^N}, \quad (2.12a)$$

$$\sum_{i=1}^N b_h(w_{h,i}, \psi_{h,i}) \lesssim \left( \|\nabla_h \mathbf{w}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 + \|\mathbf{h}^{-\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}^2 \right)^{\frac{1}{2}} \|\underline{\psi}_h\|_{[L^2(\Omega)^{d_1}]^N}, \quad (2.12b)$$

$$\sum_{i=1}^N s_h(w_{h,i}, \lambda_{h,i}) \lesssim \|\eta_F^{\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N} \|\eta_F^{\frac{1}{2}} \llbracket \boldsymbol{\lambda}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}, \quad (2.12c)$$

$$\mathbf{u}_h(\mathbf{w}_h, \boldsymbol{\phi}_h) \lesssim \|\boldsymbol{\phi}_h\|_{L^2(\Omega)^N}, \quad (2.12d)$$

$$\tilde{\mathbf{a}}_h(\mathbf{w}_h; \underline{\zeta}_h, \boldsymbol{\varphi}_h) \lesssim \|\underline{\zeta}_h\|_{[L^2(\Omega)^{d_1}]^N} \|\boldsymbol{\varphi}_h\|_{[L^2(\Omega)^{d_1}]^N}, \quad (2.12e)$$

$$\mathbf{a}_h(\mathbf{w}_h; \boldsymbol{\sigma}_h, \boldsymbol{\theta}_h) \lesssim \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)^{d_1}]^N} \|\boldsymbol{\theta}_h\|_{[L^2(\Omega)^{d_1}]^N}, \quad (2.12f)$$

$$\mathbf{f}_h(\mathbf{w}_h, \boldsymbol{\lambda}_h) \lesssim \|\boldsymbol{\lambda}_h\|_{L^2(\Omega)^N} \quad (2.12g)$$

for all functions in the corresponding discrete spaces, with hidden constants independent of the mesh size  $h$ . Moreover, the nonlinear functional  $\mathbf{n}_h(\cdot; \cdot, \cdot)$  satisfies the following coercivity property: for all  $\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$ ,

$$\mathbf{n}_h(\mathbf{w}_h; \boldsymbol{\sigma}_h, \boldsymbol{\sigma}_h) \geq \gamma \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 \quad \forall \boldsymbol{\sigma}_h \in \mathcal{M}_p(\mathcal{T}_h)^N, \quad (2.13)$$

where  $\gamma$  is the constant in Assumption (H2a).

*Proof.* The coercivity property (2.13) follows from Assumption (H2a). For (2.12b), the definition of the numerical flux  $\widehat{\mathbf{w}}_h$  and integration by parts give

$$b_h(w_{h,i}, \psi_{h,i}) = - \int_{\Omega} \nabla_h w_{h,i} \cdot \underline{\psi}_{h,i} d\mathbf{x} + \int_{\mathcal{F}_h^{\mathcal{I}}} \llbracket w_{h,i} \rrbracket_{\mathbf{N}} \cdot \{\!\! \{ \psi_{h,i} \}\!\! \}_{1-\alpha_F}.$$

We estimate the volume term on the right-hand side with the Cauchy–Schwarz inequality. For the interface term, on each  $F \in \mathcal{F}_h^{\mathcal{I}}$ , we use the weighted Cauchy–Schwarz inequality with weights  $\eta_F^{1/2}$  and  $\eta_F^{-1/2}$  and

the inverse trace inequality for  $\psi_{h,i}$ , taking into account that, due to the definition of  $\eta_F$  in (2.6),  $\eta_F^{-1/2} \lesssim \min\{h_{K_1}^{1/2}, h_{K_2}^{1/2}\}$ , where  $K_1$  and  $K_2$  are the two elements sharing  $F$ . Estimate (2.12b) readily follows. The remaining bounds in (2.12) follow from Assumptions (H1), the boundedness of  $\mathbf{u}$  (see (H2)), and the Cauchy–Schwarz inequality.  $\square$

We prove that, given  $\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$ , equations (2.9a) and (2.9b) define  $\boldsymbol{\sigma}_h \in \mathcal{M}_p(\mathcal{T}_h)^N$  in a unique way. In vector representation, this entails that, given  $\mathbf{W}_h$ , equation (2.11a) defines  $\boldsymbol{\Sigma}_h = \boldsymbol{\Sigma}_h(\mathbf{W}_h)$  in a unique way.

**Proposition 2.3.** *Given  $\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$ , equations (2.9a) and (2.9b) define  $\boldsymbol{\sigma}_h \in \mathcal{M}_p(\mathcal{T}_h)^N$  in a unique way. Moreover,  $\boldsymbol{\sigma}_h$  satisfies*

$$\|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 \lesssim \|\nabla_h \mathbf{w}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 + \|\mathbf{h}^{-\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}^2, \quad (2.14)$$

with hidden constant independent of the mesh size  $h$ .

*Proof.* (i) Given  $\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$ , there exists a unique  $\boldsymbol{\zeta}_h \in \mathcal{M}_p(\mathcal{T}_h)^N$  solution to (2.9a). Moreover,  $\boldsymbol{\zeta}_h$  satisfies

$$\|\boldsymbol{\zeta}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 \lesssim \|\nabla_h \mathbf{w}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 + \|\mathbf{h}^{-\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}^2.$$

This follows from the Lax–Milgram lemma, which is applicable owing to (2.12a) and (2.12b).

(ii) Given  $\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$  and  $\boldsymbol{\zeta}_h \in \mathcal{M}_p(\mathcal{T}_h)^N$  from step (i), there exists a unique  $\boldsymbol{\sigma}_h = \boldsymbol{\sigma}_h(\mathbf{w}_h) \in \mathcal{M}_p(\mathcal{T}_h)^N$  solution to (2.9b) that satisfies (2.14). This follows again from the Lax–Milgram lemma, which is applicable owing to (2.13), (2.12e), and (2.12b).  $\square$

We prove the following space-discrete entropy inequality, which is a discrete version of inequality (1.3).

**Proposition 2.4.** *Any solution  $(\mathbf{w}_h, \boldsymbol{\sigma}_h)$  to the semidiscrete formulation (2.11) satisfies the following entropy inequality for all  $\tau \in (0, T]$ :*

$$\int_{\Omega} s(\mathbf{u}(\mathbf{w}_h(\mathbf{x}, \tau))) d\mathbf{x} + \gamma \int_0^{\tau} \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)^{d_1}]^N}^2 dt + \int_0^{\tau} \|\eta_F^{\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}^2 dt \leq \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h(\mathbf{x}, 0))) d\mathbf{x} + \tau C_f |\Omega|.$$

*Proof.* Let  $\tau \in (0, T]$ . Multiplying (2.11b) by  $\mathbf{W}_h$  we get

$$\left\langle \frac{d}{dt} \mathcal{U}_h(\mathbf{W}_h), \mathbf{W}_h \right\rangle + \langle (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h), \mathbf{W}_h \rangle + \langle (I_N \otimes S) \mathbf{W}_h, \mathbf{W}_h \rangle = \langle \mathcal{F}_h(\mathbf{W}_h), \mathbf{W}_h \rangle. \quad (2.15)$$

We treat each term in identity (2.15) separately.

Since  $\mathbf{u} = (s')^{-1}$ , we can write  $\mathbf{w}_h$  as  $s'(\mathbf{u}(\mathbf{w}_h))$ . This, together with the chain rule, gives

$$\begin{aligned} \left\langle \frac{d}{dt} \mathcal{U}_h(\mathbf{W}_h), \mathbf{W}_h \right\rangle &= \int_{\Omega} \partial_t(\mathbf{u}(\mathbf{w}_h)) \cdot \mathbf{w}_h d\mathbf{x} = \int_{\Omega} \partial_t(\mathbf{u}(\mathbf{w}_h)) \cdot s'(\mathbf{u}(\mathbf{w}_h)) d\mathbf{x} \\ &= \int_{\Omega} \partial_t(s(\mathbf{u}(\mathbf{w}_h))) d\mathbf{x}. \end{aligned} \quad (2.16)$$

By using standard algebraic manipulations, equation (2.4b), and Assumption (H2a), we obtain

$$\begin{aligned} \langle (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h), \mathbf{W}_h \rangle &= \langle \mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h), (I_N \otimes M^{-1} B) \mathbf{W}_h \rangle \\ &= \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h)) \boldsymbol{\sigma}_h \cdot \boldsymbol{\zeta}_h d\mathbf{x} = \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h))^T \boldsymbol{\zeta}_h \cdot \boldsymbol{\sigma}_h d\mathbf{x} \\ &= \int_{\Omega} A(\mathbf{u}(\mathbf{w}_h))^T s''(\mathbf{u}(\mathbf{w}_h)) \boldsymbol{\sigma}_h \cdot \boldsymbol{\sigma}_h d\mathbf{x} \geq \gamma \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)^{d_1}]^N}^2. \end{aligned} \quad (2.17)$$

By the definition of the bilinear form  $s_h(\cdot, \cdot)$  in (2.7c), we have

$$\langle (I_N \otimes S) \mathbf{W}_h, \mathbf{W}_h \rangle = \|\eta_F^{\frac{1}{2}} \llbracket \mathbf{w}_h \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})^{d_1}]^N}^2. \quad (2.18)$$

Finally, the following upper bound follows from Assumption (H2b):

$$\langle \mathcal{F}_h(\mathbf{W}_h), \mathbf{W}_h \rangle = \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}_h)) \cdot \mathbf{w}_h d\mathbf{x} = \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}_h)) \cdot s'(\mathbf{u}(\mathbf{w}_h)) d\mathbf{x} \leq C_f |\Omega|. \quad (2.19)$$

Integrating in time (2.15) from 0 to  $\tau$ , and using bounds (2.16), (2.17), (2.18), and (2.19), we obtain the desired result.  $\square$

**Remark 2.5.** The definition of  $\mathbf{u}$  in Assumption (H2) guarantees that, in the semidiscrete formulation (2.4), the argument  $\mathbf{u}(\mathbf{w}_h)$  in the nonlinear terms  $A(\cdot)$ ,  $s''(\cdot)$ , and  $\mathbf{f}(\cdot)$  takes values in  $\mathcal{D}$ . Such a property is essential in the existence and convergence results in Theorems 3.2 and 4.1, and could not be guaranteed if a discrete approximation  $\boldsymbol{\rho}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$  of  $\boldsymbol{\rho} = \mathbf{u}(\mathbf{w})$  were used instead. ■

**Remark 2.6** (Constant diffusion tensor  $A$ ). If  $A$  is a constant diffusion tensor, the semidiscrete formulation (2.4) reduces to

$$\begin{aligned} \widetilde{\mathcal{N}}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h) &= (I_N \otimes B)\mathbf{W}_h, \\ \frac{d}{dt}\mathbf{u}_h(\mathbf{W}_h) + (A \otimes B^T M^{-1})\boldsymbol{\Sigma}_h + (I_N \otimes S)\mathbf{W}_h &= \mathcal{F}_h(\mathbf{W}_h), \end{aligned}$$

where  $\widetilde{\mathcal{N}}_h(\cdot, \cdot)$  is the operator associated with the nonlinear functional

$$\tilde{\mathbf{n}}_h(\mathbf{w}_h; \boldsymbol{\sigma}_h, \boldsymbol{\varphi}_h) := \int_{\Omega} s''(\mathbf{u}(\mathbf{w}_h))\boldsymbol{\sigma}_h \cdot \boldsymbol{\varphi}_h dx \quad \forall (\mathbf{w}_h, \boldsymbol{\sigma}_h, \boldsymbol{\varphi}_h) \in \mathcal{S}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N \times \mathcal{M}_p(\mathcal{T}_h)^N.$$

Moreover, if the entropy density is given by

$$s(\boldsymbol{\rho}) = \sum_{i=1}^N s_i(\rho_i),$$

matrix  $s''(\boldsymbol{\rho})$  is diagonal. In such a case, the  $N$  components of  $\boldsymbol{\sigma}_h$  are no longer coupled. ■

**Remark 2.7** (Differential operators). Rewriting model (1.1) in terms of the auxiliary variables  $\mathbf{w}_h$ ,  $\boldsymbol{\zeta}_h$ ,  $\boldsymbol{\sigma}_h$ , and  $\mathbf{q}_h$  allows us to localize the influence of the nonlinear terms in the semidiscrete formulation (2.4). More precisely, nonlinearities do not appear in interface terms, but only on local volume integrals. Consequently, the only non-block-diagonal operators in the method that have to be computed are the scalar matrices  $B$  and  $S$ , which are the standard LDG gradient and stability matrices, respectively. The resulting method is more efficient, compared to interior-penalty discretizations with nonlinearities under the differential operators (and thus in the interface terms); cf. [6, 12]. ■

**Remark 2.8** (Discrete positivity and boundedness). Obtaining a discrete approximation  $\boldsymbol{\rho}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$  that respects the positivity (or boundedness) of the physical unknown  $\boldsymbol{\rho}$  in a strong sense (i.e., pointwise) is a very difficult task. In fact, for high-order approximations, even if  $\boldsymbol{\rho}_h$  is enforced to satisfy such bounds on the nodes (weak positivity), the physical constraints might still be violated; cf. [3]. Our method provides an approximate solution  $\tilde{\boldsymbol{\rho}}_h = \mathbf{u}(\mathbf{w}_h) \notin \mathcal{S}_p(\mathcal{T}_h)^N$  that satisfies the physical constraints for any degree of approximation. ■

## 2.2 Fully discrete scheme

We discretize the ODE system (2.11) in time by the backward Euler method on a partition of the time interval  $(0, T)$  into  $N_t$  subintervals  $\{(t_{n-1}, t_n)\}_{n=1}^{N_t}$ , with  $t_0 = 0$ ,  $t_{N_t} = T$  and time steps  $\tau_n := t_n - t_{n-1} > 0$ . Moreover, we add a regularizing term with multiplicative parameter  $\varepsilon > 0$ , which is defined in terms of a symmetric,  $h$ -uniformly positive definite matrix  $C$  only depending on the space discretization. Such a regularizing term is essential in the existence and convergence results in Theorems 3.2 and 4.1. The fully discrete, regularized method reads as follows:

- define  $\mathbf{R}_h^0$  as the vector representation of the  $L^2(\Omega)$ -orthogonal projection of  $\boldsymbol{\rho}_0$  in  $\mathcal{S}_p(\mathcal{T}_h)^N$  denoted by  $\Pi_p^0 \boldsymbol{\rho}_0$ , and compute  $(\mathbf{W}_h^{\varepsilon,1}, \boldsymbol{\Sigma}_h^{\varepsilon,1})$  by solving

$$\mathcal{N}_h(\mathbf{W}_h^{\varepsilon,1}; \boldsymbol{\Sigma}_h^{\varepsilon,1}) = \widetilde{\mathcal{A}}_h(\mathbf{W}_h^{\varepsilon,1}; (I_N \otimes M^{-1}B)\mathbf{W}_h^{\varepsilon,1}), \quad (2.20a)$$

$$\begin{aligned} \varepsilon \tau_1 (I_N \otimes C)\mathbf{W}_h^{\varepsilon,1} + \left( \mathbf{u}_h(\mathbf{W}_h^{\varepsilon,1}) - \mathbf{R}_h^0 \right) \\ + \tau_1 (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{W}_h^{\varepsilon,1}; \boldsymbol{\Sigma}_h^{\varepsilon,1}) + \tau_1 (I_N \otimes S)\mathbf{W}_h^{\varepsilon,1} = \tau_1 \mathcal{F}_h(\mathbf{W}_h^{\varepsilon,1}); \end{aligned} \quad (2.20b)$$

- for  $n = 1, \dots, N_t - 1$ , compute  $(\mathbf{W}_h^{\varepsilon,n+1}, \boldsymbol{\Sigma}_h^{\varepsilon,n+1})$  by solving

$$\mathcal{N}_h(\mathbf{W}_h^{\varepsilon,n+1}; \boldsymbol{\Sigma}_h^{\varepsilon,n+1}) = \widetilde{\mathcal{A}}_h(\mathbf{W}_h^{\varepsilon,n+1}; (I_N \otimes M^{-1}B)\mathbf{W}_h^{\varepsilon,n+1}), \quad (2.21a)$$



$$\begin{aligned} & \varepsilon \tau_{n+1} (I_N \otimes C) \mathbf{W}_h^{\varepsilon, n+1} + \left( \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n+1}) - \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n}) \right) \\ & + \tau_{n+1} (I_N \otimes B^T M^{-1}) \mathcal{A}_h \left( \mathbf{W}_h^{\varepsilon, n+1}; \Sigma_h^{\varepsilon, n+1} \right) + \tau_{n+1} (I_N \otimes S) \mathbf{W}_h^{\varepsilon, n+1} = \tau_{n+1} \mathcal{F}_h(\mathbf{W}_h^{\varepsilon, n+1}). \end{aligned} \quad (2.21b)$$

The symmetric, positive definite matrix  $C$  defines a scalar product and a norm in  $\mathcal{S}_p(\mathcal{T}_h)$ : given  $\mathbf{w}_h$  and  $\mathbf{v}_h$  in  $\mathcal{S}_p(\mathcal{T}_h)^N$  with vector representations  $\mathbf{W}_h$  and  $\mathbf{V}_h$ , respectively, we set

$$\sum_{i=1}^N c_h(w_{h,i}, v_{h,i}) := c_h(\mathbf{w}_h, \mathbf{v}_h) := \langle (I_N \otimes C) \mathbf{W}_h, \mathbf{V}_h \rangle \quad \text{and} \quad \|\mathbf{w}_h\|_C^2 := \langle (I_N \otimes C) \mathbf{W}_h, \mathbf{W}_h \rangle. \quad (2.22)$$

**Remark 2.9** (Discrete initial condition). *The use of  $\mathbf{R}_h^0$  in the first step of the fully discrete scheme (2.20)–(2.21) has two motivations:*

- it allows for an initial datum  $\rho_0(\mathbf{x}) \in \overline{\mathcal{D}}$  for a.e.  $\mathbf{x} \in \Omega$ , whereas  $w_0 = s'(\rho_0)$  may be not well defined if  $\rho_0$  takes values on  $\partial\mathcal{D}$ ;
- it leads to an  $h$ -independent bound in the discrete entropy inequality in Theorem 3.1 below. ■

**Remark 2.10** (Reduced nonlinear system). *Setting  $\widehat{\mathcal{E}}_h(\mathbf{W}_h) := \widehat{\mathcal{A}}_h(\mathbf{W}_h) \widehat{\mathcal{N}}_h(\mathbf{W}_h)^{-1} \widehat{\mathcal{A}}_h(\mathbf{W}_h)^T$ , the fully discrete scheme (2.20)–(2.21) can be written in terms of the  $\mathbf{W}_h^{\varepsilon, n+1}$ -unknown only:*

- define  $\mathbf{R}_h^0$  as the vector representation of  $\Pi_p^0 \rho_0$ , and compute  $\mathbf{W}_h^{\varepsilon, 1}$  by solving

$$\begin{aligned} & \varepsilon \tau_1 (I_N \otimes C) \mathbf{W}_h^{\varepsilon, 1} + \left( \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, 1}) - \mathbf{R}_h^0 \right) \\ & + \tau_1 \left[ (I_N \otimes B^T M^{-1}) \widehat{\mathcal{E}}_h(\mathbf{W}_h^{\varepsilon, 1}) (I_N \otimes M^{-1} B) + (I_N \otimes S) \right] \mathbf{W}_h^{\varepsilon, 1} = \tau_1 \mathcal{F}_h(\mathbf{W}_h^{\varepsilon, 1}); \end{aligned}$$

- for  $n = 1, \dots, N_t - 1$ , compute  $\mathbf{W}_h^{\varepsilon, n+1}$  by solving

$$\begin{aligned} & \varepsilon \tau_{n+1} (I_N \otimes C) \mathbf{W}_h^{\varepsilon, n+1} + \left( \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n+1}) - \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n}) \right) \\ & + \tau_{n+1} \left[ (I_N \otimes B^T M^{-1}) \widehat{\mathcal{E}}_h(\mathbf{W}_h^{\varepsilon, n+1}) (I_N \otimes M^{-1} B) + (I_N \otimes S) \right] \mathbf{W}_h^{\varepsilon, n+1} = \tau_{n+1} \mathcal{F}_h(\mathbf{W}_h^{\varepsilon, n+1}). \end{aligned}$$

Due to the structure of  $\widehat{\mathcal{N}}_h(\mathbf{W}_h)$  and  $\widehat{\mathcal{A}}_h(\mathbf{W}_h)$ , the matrix  $\widehat{\mathcal{E}}_h(\mathbf{W}_h)$  is block diagonal. ■

**Remark 2.11** (Higher-order time discretizations). *The convexity of the entropy in Assumption (H2) allows for a proof of a discrete entropy inequality for the first-order backward Euler scheme via the elementary inequality  $s'(\mathbf{y}) \cdot (\mathbf{y} - \mathbf{x}) \geq s(\mathbf{y}) - s(\mathbf{x})$  for all  $\mathbf{x}, \mathbf{y} \in \mathcal{D}$ . Surprisingly, the use of higher-order time discretizations and their analysis is more delicate and mainly open. We refer to [16, 17] for structure-preserving second-order BDF (Backward Differentiation Formula) schemes for certain classes of cross-diffusion systems. ■*

### 3 Discrete entropy stability and existence of discrete solutions

In this section, we prove the entropy stability and existence of solutions to the fully discrete, regularized problem (2.20)–(2.21).

**Theorem 3.1** (Discrete entropy inequalities). *Any solution  $\{\mathbf{W}_h^{\varepsilon, n}\}_{n=1}^{N_t}$  to problem (2.20)–(2.21) satisfies*

$$\begin{aligned} \varepsilon \tau_1 \|\mathbf{w}_h^{\varepsilon, 1}\|_C^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, 1})) d\mathbf{x} + \gamma \tau_1 \|\boldsymbol{\sigma}_h^{\varepsilon, 1}\|_{[L^2(\Omega)^d]^N}^2 + \tau_1 \|\eta_F^{\frac{1}{2}}[\mathbf{w}_h^{\varepsilon, 1}]\|_{[L^2(\mathcal{F}_h^x)^d]^N}^2 \\ \leq \int_{\Omega} s(\rho_0) d\mathbf{x} + C_f \tau_1 |\Omega|, \end{aligned} \quad (3.1)$$

$$\begin{aligned} \varepsilon \tau_{n+1} \|\mathbf{w}_h^{\varepsilon, n+1}\|_C^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, n+1})) d\mathbf{x} + \gamma \tau_{n+1} \|\boldsymbol{\sigma}_h^{\varepsilon, n+1}\|_{[L^2(\Omega)^d]^N}^2 + \tau_{n+1} \|\eta_F^{\frac{1}{2}}[\mathbf{w}_h^{\varepsilon, n+1}]\|_{[L^2(\mathcal{F}_h^x)^d]^N}^2 \\ \leq \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, n})) d\mathbf{x} + C_f \tau_{n+1} |\Omega|, \end{aligned} \quad (3.2)$$

and

$$\begin{aligned} \varepsilon \sum_{n=0}^{N_t-1} \tau_{n+1} \|\mathbf{w}_h^{\varepsilon, n+1}\|_C^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, N_t})) \mathrm{d}\mathbf{x} + \gamma \sum_{n=0}^{N_t-1} \tau_{n+1} \|\boldsymbol{\sigma}_h^{\varepsilon, n+1}\|_{[L^2(\Omega)]^{d_1 N}}^2 \\ + \sum_{n=0}^{N_t-1} \tau_{n+1} \|\eta_{\tilde{F}}^{\frac{1}{2}} [\mathbf{w}_h^{\varepsilon, n+1}]_{\mathbb{N}}\|_{[L^2(\mathcal{F}_h^{\tilde{I}})]^{d_1 N}}^2 \leq \int_{\Omega} s(\boldsymbol{\rho}_0) \mathrm{d}\mathbf{x} + C_f T |\Omega|. \end{aligned} \quad (3.3)$$

*Proof.* We multiply (2.20b) by  $\mathbf{W}_h^{\varepsilon, 1}$ . For the first two terms, using the  $L^2(\Omega)$ -orthogonality of  $\Pi_p^0$ , the fact that  $\mathbf{u} = (s')^{-1}$ , and the convexity of  $s$ , we find that

$$\begin{aligned} \varepsilon \tau_1 \left\langle (I_N \otimes C) \mathbf{W}_h^{\varepsilon, 1}, \mathbf{W}_h^{\varepsilon, 1} \right\rangle + \left\langle \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, 1}) - \mathbf{R}_h^0, \mathbf{W}_h^{\varepsilon, 1} \right\rangle &= \varepsilon \tau_1 \|\mathbf{w}_h^{\varepsilon, 1}\|_C^2 + \int_{\Omega} \left( \mathbf{u}(\mathbf{w}_h^{\varepsilon, 1}) - \Pi_p^0 \boldsymbol{\rho}_0 \right) \cdot \mathbf{w}_h^{\varepsilon, 1} \mathrm{d}\mathbf{x} \\ &= \varepsilon \tau_1 \|\mathbf{w}_h^{\varepsilon, 1}\|_C^2 + \int_{\Omega} \left( \mathbf{u}(\mathbf{w}_h^{\varepsilon, 1}) - \boldsymbol{\rho}_0 \right) \cdot \mathbf{w}_h^{\varepsilon, 1} \mathrm{d}\mathbf{x} \\ &= \varepsilon \tau_1 \|\mathbf{w}_h^{\varepsilon, 1}\|_C^2 + \int_{\Omega} \left( \mathbf{u}(\mathbf{w}_h^{\varepsilon, 1}) - \boldsymbol{\rho}_0 \right) \cdot s'(\mathbf{u}(\mathbf{w}_h^{\varepsilon, 1})) \mathrm{d}\mathbf{x} \\ &\geq \varepsilon \tau_1 \|\mathbf{w}_h^{\varepsilon, 1}\|_C^2 + \int_{\Omega} \left( s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, 1})) - s(\boldsymbol{\rho}_0) \right) \mathrm{d}\mathbf{x}. \end{aligned}$$

For the remaining terms, proceeding exactly as in the proof of Proposition 2.4, we obtain

$$\begin{aligned} \tau_1 \left\langle (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{W}_h^{\varepsilon, 1}; \boldsymbol{\Sigma}_h^{\varepsilon, 1}), \mathbf{W}_h^{\varepsilon, 1} \right\rangle &\geq \tau_1 \gamma \|\boldsymbol{\sigma}_h^{\varepsilon, 1}\|_{[L^2(\Omega)]^{d_1 N}}^2, \\ \tau_1 \left\langle (I_N \otimes S) \mathbf{W}_h^{\varepsilon, 1}, \mathbf{W}_h^{\varepsilon, 1} \right\rangle &= \tau_1 \|\eta_{\tilde{F}}^{\frac{1}{2}} [\mathbf{w}_h^{\varepsilon, 1}]_{\mathbb{N}}\|_{[L^2(\mathcal{F}_h^{\tilde{I}})]^{d_1 N}}^2, \\ \tau_1 \left\langle \mathcal{F}_h(\mathbf{W}_h^{\varepsilon, 1}), \mathbf{W}_h^{\varepsilon, 1} \right\rangle &\leq \tau_1 C_f |\Omega|. \end{aligned}$$

All the above estimates immediately give (3.1). In order to prove (3.2), we proceed as above. We write explicitly the estimate of the first two terms for completeness:

$$\begin{aligned} \varepsilon \tau_{n+1} \left\langle (I_N \otimes C) \mathbf{W}_h^{\varepsilon, n+1}, \mathbf{W}_h^{\varepsilon, n+1} \right\rangle + \left\langle \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n+1}) - \mathcal{U}_h(\mathbf{W}_h^{\varepsilon, n}), \mathbf{W}_h^{\varepsilon, 1} \right\rangle \\ = \varepsilon \tau_{n+1} \|\mathbf{w}_h^{\varepsilon, n+1}\|_C^2 + \int_{\Omega} \left( \mathbf{u}(\mathbf{w}_h^{\varepsilon, n+1}) - \mathbf{u}(\mathbf{w}_h^{\varepsilon, n}) \right) \cdot \mathbf{w}_h^{\varepsilon, n+1} \mathrm{d}\mathbf{x} \\ = \varepsilon \tau_{n+1} \|\mathbf{w}_h^{\varepsilon, n+1}\|_C^2 + \int_{\Omega} \left( \mathbf{u}(\mathbf{w}_h^{\varepsilon, n+1}) - \mathbf{u}(\mathbf{w}_h^{\varepsilon, n}) \right) \cdot s'(\mathbf{u}(\mathbf{w}_h^{\varepsilon, n+1})) \mathrm{d}\mathbf{x} \\ \geq \varepsilon \tau_{n+1} \|\mathbf{w}_h^{\varepsilon, n+1}\|_C^2 + \int_{\Omega} \left( s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, n+1})) - s(\mathbf{u}(\mathbf{w}_h^{\varepsilon, n})) \right) \mathrm{d}\mathbf{x}. \end{aligned}$$

Finally, to obtain (3.3), we multiply (2.20b) and (2.21b) by  $\mathbf{W}_h^{\varepsilon, 1}$  and  $\mathbf{W}_h^{\varepsilon, n+1}$ , respectively, sum over all indices  $n = 0, \dots, N_t - 1$ , and use the same arguments as above.  $\square$

**Theorem 3.2.** *For  $n = 0, \dots, N_t - 1$ , there exists a solution  $\mathbf{W}_h^{\varepsilon, n+1}$  to problem (2.20) ( $n = 0$ ) or to problem (2.21) ( $n \geq 1$ ).*

*Proof.* We begin with  $n = 0$ . Consider the linearized problem: given  $\mathbf{V}_h \in \mathbb{R}^{\dim(\mathcal{S}_p(\mathcal{T}_h)^N)}$ , find  $\mathbf{W}_h^{\varepsilon} \in \mathbb{R}^{\dim(\mathcal{S}_p(\mathcal{T}_h)^N)}$  such that

$$\varepsilon \tau_1 (I_N \otimes C) \mathbf{W}_h^{\varepsilon} = -\mathcal{U}_h(\mathbf{V}_h) + \mathbf{R}_h^0 - \tau_1 (I_N \otimes B^T M^{-1}) \mathcal{A}_h(\mathbf{V}_h; \boldsymbol{\Sigma}_h(\mathbf{V}_h)) - \tau_1 (I_N \otimes S) \mathbf{V}_h + \tau_1 \mathcal{F}_h(\mathbf{V}_h),$$

where  $\boldsymbol{\Sigma}_h(\mathbf{V}_h)$  is the unique solution to

$$\mathcal{N}_h(\mathbf{V}_h; \boldsymbol{\Sigma}_h(\mathbf{V}_h)) = \tilde{\mathcal{A}}_h(\mathbf{V}_h; (I_N \otimes M^{-1} B) \mathbf{V}_h);$$

see the text above Proposition 2.3. As  $C$  is positive definite,  $\mathbf{W}_h^{\varepsilon}$  is uniquely defined. This defines a function

$$\Phi : \mathcal{S}_p(\mathcal{T}_h)^N \rightarrow \mathcal{S}_p(\mathcal{T}_h)^N, \quad \mathbf{v}_h \mapsto \mathbf{w}_h^{\varepsilon},$$

where  $\mathbf{v}_h \in \mathcal{S}_p(\mathcal{T}_h)^N$  and  $\mathbf{w}_h^\varepsilon \in \mathcal{S}_p(\mathcal{T}_h)^N$  are the functions whose coefficient vectors are  $\mathbf{V}_h$  and  $\mathbf{W}_h^\varepsilon$ , respectively. Due to the continuity of  $A$ ,  $\mathbf{f}$ , and  $\mathbf{u}$ , and to estimate (2.14),  $\Phi$  is continuous.

We apply the Schaefer fixed-point theorem [23, Thm. 4.3.2] to prove that  $\Phi$  has a fixed point, which implies the existence of solutions to (2.20). In order to do so, it only remains to prove that the following set is bounded:

$$\{\mathbf{w}_h \in \mathcal{S}_p(\mathcal{T}_h)^N : \mathbf{w}_h = \delta\Phi(\mathbf{w}_h), \delta \in [0, 1]\}.$$

Let  $\mathbf{w}_h \neq 0$  be in this set, and let  $\mathbf{W}_h$  be its coefficient vector. Then,  $\mathbf{w}_h = \delta\Phi(\mathbf{w}_h)$  for some  $\delta \in (0, 1]$ , namely  $\mathbf{W}_h$  satisfies

$$\begin{aligned} & \frac{\varepsilon\tau_1}{\delta}(I_N \otimes C)\mathbf{W}_h + (\mathbf{u}_h(\mathbf{W}_h) - \mathbf{R}_h^0) \\ & + \tau_1(I_N \otimes B^T M^{-1})\mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h(\mathbf{W}_h)) + \tau_1(I_N \otimes S)\mathbf{W}_h = \tau_1\mathcal{F}_h(\mathbf{W}_h). \end{aligned}$$

We multiply the previous equation by  $\mathbf{W}_h$ . It follows as in the proof of Theorem 3.1 that

$$\begin{aligned} \frac{\varepsilon\tau_1}{\delta} \langle (I_N \otimes C)\mathbf{W}_h, \mathbf{W}_h \rangle + \langle \mathbf{u}_h(\mathbf{W}_h) - \mathbf{R}_h^0, \mathbf{W}_h \rangle &\geq \frac{\varepsilon\tau_1}{\delta} \|\mathbf{w}_h\|_C^2 + \int_{\Omega} (s(\mathbf{u}(\mathbf{w}_h)) - s(\rho_0)) \, d\mathbf{x}, \\ \tau_1 \langle (I_N \otimes B^T M^{-1})\mathcal{A}_h(\mathbf{W}_h; \boldsymbol{\Sigma}_h(\mathbf{W}_h)), \mathbf{W}_h \rangle &\geq \tau_1\gamma \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)]^{d_1 N}}^2, \\ \tau_1 \langle (I_N \otimes S)\mathbf{W}_h, \mathbf{W}_h \rangle &= \tau_1 \|\eta_F^{\frac{1}{2}}[\mathbf{w}_h]_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})]^{d_1 N}}^2, \\ \tau_1 \langle \mathcal{F}_h(\mathbf{W}_h), \mathbf{W}_h \rangle &\leq \tau_1 C_f |\Omega|, \end{aligned}$$

from which we obtain

$$\frac{\varepsilon\tau_1}{\delta} \|\mathbf{w}_h\|_C^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h)) \, d\mathbf{x} + \tau_1\gamma \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)]^{d_1 N}}^2 + \tau_1 \|\eta_F^{\frac{1}{2}}[\mathbf{w}_h]_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})]^{d_1 N}}^2 \leq \int_{\Omega} s(\rho_0) \, d\mathbf{x} + \tau_1 C_f |\Omega|.$$

Due to Assumption (H2c),  $\|\mathbf{w}_h\|_C$  is uniformly bounded with respect to  $\delta$ . Therefore, the Schaefer fixed-point theorem implies the existence of a fixed point of  $\Phi$  ( $\delta = 1$ ) and therefore the existence of a solution  $\mathbf{W}_h^{\varepsilon,1}$  to problem (2.20). In particular, for the function  $\mathbf{w}_h^{\varepsilon,1}$  corresponding to the coefficient vector  $\mathbf{W}_h^{\varepsilon,1}$ , we have

$$\int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,1})) \, d\mathbf{x} \leq \int_{\Omega} s(\rho_0) \, d\mathbf{x} + \tau_1 C_f |\Omega|.$$

For  $n \geq 1$ , we proceed by induction. Assuming the existence of  $\mathbf{W}_h^{\varepsilon,n}$  and the boundedness of  $\int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,n})) \, d\mathbf{x}$ , we apply the same arguments as above to the linearized problem

$$\begin{aligned} \varepsilon\tau_{n+1}(I_N \otimes C)\mathbf{W}_h^\varepsilon &= -\mathbf{u}_h(\mathbf{V}_h) + \mathbf{u}_h(\mathbf{W}_h^{\varepsilon,n}) - \tau_{n+1}(I_N \otimes B^T M^{-1})\mathcal{A}_h(\mathbf{V}_h; \boldsymbol{\Sigma}_h(\mathbf{V}_h)) \\ &\quad - \tau_{n+1}(I_N \otimes S)\mathbf{V}_h + \tau_{n+1}\mathcal{F}_h(\mathbf{V}_h), \end{aligned}$$

to deduce that

$$\begin{aligned} \frac{\varepsilon\tau_{n+1}}{\delta} \|\mathbf{w}_h\|_C^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h)) \, d\mathbf{x} + \tau_{n+1}\gamma \|\boldsymbol{\sigma}_h\|_{[L^2(\Omega)]^{d_1 N}}^2 + \tau_{n+1} \|\eta_F^{\frac{1}{2}}[\mathbf{w}_h]_{\mathbf{N}}\|_{[L^2(\mathcal{F}_h^{\mathcal{I}})]^{d_1 N}}^2 \\ \leq \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,n})) \, d\mathbf{x} + \tau_{n+1} C_f |\Omega|. \end{aligned}$$

The boundedness of  $\int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,n})) \, d\mathbf{x}$  entails the uniform boundedness of  $\|\mathbf{w}_h\|_C$ , and the existence of a solution  $\mathbf{W}_h^{\varepsilon,n+1}$  to problem (2.21) is derived as above. Moreover,

$$\int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,n+1})) \, d\mathbf{x} \leq \int_{\Omega} s(\mathbf{u}(\mathbf{w}_h^{\varepsilon,n})) \, d\mathbf{x} + \tau_{n+1} C_f |\Omega|,$$

which completes the proof.  $\square$

**Remark 3.3** (Regularizing term). *The regularizing term with multiplicative parameter  $\varepsilon > 0$  in the fully discrete scheme (2.21) is a discrete version of the one introduced for the semidiscrete-in-time formulation in [14, §3]. Such a term is introduced so as to enforce a numerical control on the  $L^\infty(\Omega)$ -norm of the entropy variable  $\mathbf{w}_h$ ; see Section 4.1 for more details.  $\blacksquare$*

## 4 Convergence of the fully discrete scheme

We fix  $\varepsilon > 0$  and a partition  $\mathcal{I}_\tau$  of the time interval  $(0, T)$  defined as in Section 2.2, where the index  $\tau$  denotes the maximum element length. Consider a sequence of spatial meshes indexed by  $m \in \mathbb{N}$ ,  $\{\mathcal{T}_{h_m}\}_m$ , where  $h_m$  is the maximum element diameter of  $\mathcal{T}_{h_m}$ . We assume that  $\{h_m\}_m$  is a decreasing sequence with  $h_m \leq 1$  for all  $m \in \mathbb{N}$  and  $\lim_{m \rightarrow \infty} h_m = 0$ . We introduce the Local Discontinuous Galerkin gradient operator  $\nabla_{\text{DG}} : \mathcal{S}_p(\mathcal{T}_{h_m}) \rightarrow \mathcal{M}_p(\mathcal{T}_{h_m})$ , which is defined by

$$\int_{\Omega} \nabla_{\text{DG}} \lambda_m \cdot \boldsymbol{\theta}_m \, d\mathbf{x} = \int_{\Omega} (\nabla_h \lambda_m - \mathcal{L}(\lambda_m)) \cdot \boldsymbol{\theta}_m \, d\mathbf{x} \quad \forall \boldsymbol{\theta}_m \in \mathcal{M}_p(\mathcal{T}_{h_m}), \quad (4.1)$$

with the jump lifting operator  $\mathcal{L} : \mathcal{S}_p(\mathcal{T}_{h_m}) \rightarrow \mathcal{M}_p(\mathcal{T}_{h_m})$  given by

$$\int_{\Omega} \mathcal{L}(\lambda_m) \cdot \boldsymbol{\theta}_m \, d\mathbf{x} = \int_{\mathcal{F}_{h_m}^{\text{I}}} \llbracket \lambda_m \rrbracket_{\mathbb{N}} \cdot \{\!\! \{ \boldsymbol{\theta}_m \} \!\!\}_{1-\alpha_F} \, dS \quad \forall \boldsymbol{\theta}_m \in \mathcal{M}_p(\mathcal{T}_{h_m}).$$

### 4.1 Assumptions for $h$ -convergence

In the following Section 4.2, we prove the convergence of fully discrete solutions to semidiscrete-in-time functions, as  $m \rightarrow \infty$ . To this aim, we make the following abstract assumption, whose validity is discussed in Section 4.4 below.

**Assumption 4.1.** *We set  $\ell = 1$  if  $d = 1$  and, if  $d = 2, 3$ ,*

$$\ell = \begin{cases} 1 & \text{if } s''A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N}), \\ 2 & \text{otherwise.} \end{cases}$$

*We assume that  $\mathcal{T}_h$  and  $p$  are such that*

$$\mathcal{S}_\ell^{\text{cont}}(\mathcal{T}_h) := \mathcal{S}_\ell(\mathcal{T}_h) \cap H^1(\Omega) \subset \mathcal{S}_p(\mathcal{T}_h), \quad (4.2)$$

*and that there exists a DG norm  $\|\cdot\|_{\text{DG}}$  in  $\mathcal{S}_p(\mathcal{T}_{h_m})^N$ , which satisfies the following conditions:*

*i) There exists a positive constant  $C_{\text{DG}}$  independent of  $h_m$  such that*

$$\sum_{i=1}^N c_{h_m}(w_{m,i}, w_{m,i}) \geq C_{\text{DG}} \|\mathbf{w}_m\|_{\text{DG}}^2 \quad \forall \mathbf{w}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N.$$

*ii) If  $d = 1$  or  $\ell = 2$ , the following discrete Sobolev embedding is valid: there exists a positive constant  $C_S$  independent of  $h_m$  such that*

$$\|\mathbf{w}_m\|_{L^\infty(\Omega)^N} \leq C_S \|\mathbf{w}_m\|_{\text{DG}} \quad \forall \mathbf{w}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N. \quad (4.3)$$

*iii) For any sequence  $\{\mathbf{w}_m\}_m$  with  $\mathbf{w}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N$  that is uniformly bounded in the DG norm, there exist a subsequence still denoted by  $\{\mathbf{w}_m\}_m$  and a function  $\mathbf{w} \in H^\ell(\Omega)^N$  such that, as  $m \rightarrow \infty$ ,*

$$\begin{aligned} \nabla_{\text{DG}} \mathbf{w}_m &\rightharpoonup \nabla \mathbf{w} && \text{weakly in } [L^2(\Omega)^d]^N, \\ \mathbf{w}_m &\rightarrow \mathbf{w} && \text{strongly in } L^q(\Omega)^N, \end{aligned}$$

*with  $1 \leq q < 6$ , if  $d = 3$ , or  $1 \leq q < \infty$ , if  $d = 1, 2$ . Moreover, for any  $\boldsymbol{\lambda} \in H^\ell(\Omega)^N$  there exists a sequence  $\{\boldsymbol{\lambda}_m\}_m$  with  $\boldsymbol{\lambda}_m \in \mathcal{S}_\ell^{\text{cont}}(\mathcal{T}_{h_m})^N$  such that, as  $m \rightarrow \infty$ , it converges strongly in  $H^1(\Omega)^N$  to  $\boldsymbol{\lambda}$  and*

$$c_{h_m}(\mathbf{w}_m, \boldsymbol{\lambda}_m) \rightarrow \int_{\Omega} \left( \sum_{|\alpha|=\ell} D^\alpha \mathbf{w} \cdot D^\alpha \boldsymbol{\lambda} + \mathbf{w} \cdot \boldsymbol{\lambda} \right) d\mathbf{x}. \quad (4.4)$$

## 4.2 $h$ -convergence

For  $n = 0, \dots, N_t$ , we denote by  $\mathbf{w}_m^{\varepsilon, n}$  a solution to the fully discrete scheme (2.20)–(2.21) on the spatial mesh  $\mathcal{T}_{h_m}$  at the discrete time  $t_n$  of the fixed temporal mesh  $\mathcal{I}_\tau$ .

**Theorem 4.1** ( $h$ -convergence). *Fix  $\varepsilon > 0$  and a temporal mesh  $\mathcal{I}_\tau$ . Let Assumption 4.1 be satisfied. Then:*

I) *Setting  $\boldsymbol{\rho}_m^{\varepsilon, 0} := \Pi_p^0 \boldsymbol{\rho}_0$ , for  $m \rightarrow \infty$ , we have*

$$\int_{\Omega} (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_m^{\varepsilon, 0}) \cdot \mathbf{w} \, d\mathbf{x} \rightarrow 0 \quad \forall \mathbf{w} \in H^1(\Omega)^N. \quad (4.5)$$

Moreover, for any  $n = 1, \dots, N_t$ , there exists  $\mathbf{w}^{\varepsilon, n} \in H^\ell(\Omega)^N$  with  $\mathbf{u}(\mathbf{w}^{\varepsilon, n}) \in H^1(\Omega)^N$  and a subsequence of  $\{\mathcal{T}_{h_m}\}_m$  still denoted by  $\{\mathcal{T}_{h_m}\}_m$  such that, as  $m \rightarrow \infty$ ,

$$\boldsymbol{\rho}_m^{\varepsilon, n} := \mathbf{u}(\mathbf{w}_m^{\varepsilon, n}) \rightarrow \boldsymbol{\rho}^{\varepsilon, n} := \mathbf{u}(\mathbf{w}^{\varepsilon, n}) \quad \text{strongly in } L^r(\Omega)^N \text{ for all } r \in [1, \infty).$$

II) *Set, for convenience,  $\boldsymbol{\rho}^{\varepsilon, 0} := \boldsymbol{\rho}_0$ . For  $n = 0, \dots, N_t - 1$ ,  $\mathbf{w}^{\varepsilon, n+1}$  solves*

$$\begin{aligned} \varepsilon \tau_{n+1} \int_{\Omega} \left( \sum_{|\alpha|=\ell} D^\alpha \mathbf{w}^{\varepsilon, n+1} \cdot D^\alpha \boldsymbol{\lambda} + \mathbf{w}^{\varepsilon, n+1} \cdot \boldsymbol{\lambda} \right) d\mathbf{x} + \int_{\Omega} (\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}) - \boldsymbol{\rho}^{\varepsilon, n}) \cdot \boldsymbol{\lambda} \, d\mathbf{x} \\ + \tau_{n+1} \int_{\Omega} A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \nabla \mathbf{u}(\mathbf{w}^{\varepsilon, n+1}) \cdot \nabla \boldsymbol{\lambda} \, d\mathbf{x} = \tau_{n+1} \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \cdot \boldsymbol{\lambda} \, d\mathbf{x} \quad \forall \boldsymbol{\lambda} \in H^\ell(\Omega)^N. \end{aligned} \quad (4.6)$$

III) *For  $n = 0, \dots, N_t - 1$ ,  $\mathbf{w}^{\varepsilon, n+1}$  satisfies*

$$\begin{aligned} \varepsilon \tau_{n+1} \|\mathbf{w}^{\varepsilon, n+1}\|_{H^\ell(\Omega)^N}^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \, d\mathbf{x} + \gamma \tau_{n+1} \|\nabla \mathbf{u}(\mathbf{w}^{\varepsilon, n+1})\|_{[L^2(\Omega)^d]^N}^2 \\ \leq \int_{\Omega} s(\boldsymbol{\rho}^{\varepsilon, n}) \, d\mathbf{x} + C_f \tau_{n+1} |\Omega|, \end{aligned}$$

and

$$\begin{aligned} \varepsilon \sum_{n=0}^{N_t-1} \tau_{n+1} \|\mathbf{w}^{\varepsilon, n+1}\|_{H^\ell(\Omega)^N}^2 + \int_{\Omega} s(\mathbf{u}(\mathbf{w}^{\varepsilon, N_t})) \, d\mathbf{x} + \gamma \sum_{n=0}^{N_t-1} \tau_{n+1} \|\nabla \mathbf{u}(\mathbf{w}^{\varepsilon, n+1})\|_{[L^2(\Omega)^d]^N}^2 \\ \leq \int_{\Omega} s(\boldsymbol{\rho}_0) \, d\mathbf{x} + C_f T |\Omega|. \end{aligned}$$

*Proof. Part I)* The limit in (4.5) follows from the estimate

$$\left| \int_{\Omega} (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_m^{\varepsilon, 0}) \cdot \mathbf{w} \, d\mathbf{x} \right| = \left| \int_{\Omega} \boldsymbol{\rho}_0 \cdot (\mathbf{w} - \Pi_p^0 \mathbf{w}) \, d\mathbf{x} \right| \leq C h_m \|\boldsymbol{\rho}_0\|_{L^2(\Omega)^N} \|\mathbf{w}\|_{H^1(\Omega)^N},$$

where  $C > 0$  is independent of  $h_m$ . Since the right-hand side of (3.3) is uniformly bounded, estimate (3.2), together with Assumption 4.1, *i*), implies that  $\{\mathbf{w}_m^{\varepsilon, n}\}_m$  is bounded in the DG norm, uniformly with respect to  $h_m$ . Then, by Assumption 4.1, *iii*), there exist a function  $\mathbf{w}^{\varepsilon, n} \in H^\ell(\Omega)^N$  and a subsequence of  $\{\mathbf{w}_m^{\varepsilon, n}\}_m$ , still denoted by  $\{\mathbf{w}_m^{\varepsilon, n}\}_m$ , such that, as  $m \rightarrow \infty$ ,

$$\mathbf{w}_m^{\varepsilon, n} \rightarrow \mathbf{w}^{\varepsilon, n} \quad \text{strongly in } L^q(\Omega)^N,$$

with  $1 \leq q < 6$ , if  $d = 3$ , or  $1 \leq q < \infty$ , if  $d = 1, 2$ . Up to extraction of another subsequence, we can also assume that  $\mathbf{w}_m^{\varepsilon, n}$  converges to  $\mathbf{w}^{\varepsilon, n}$  almost everywhere in  $\Omega$ . As  $\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}) \in L^\infty(\Omega)^N$ , the dominated convergence theorem implies that  $\boldsymbol{\rho}_m^{\varepsilon, n} := \mathbf{u}(\mathbf{w}_m^{\varepsilon, n})$  converges strongly to  $\mathbf{u}(\mathbf{w}^{\varepsilon, n}) =: \boldsymbol{\rho}^{\varepsilon, n}$  in  $L^r(\Omega)^N$  for all  $r \in [1, \infty)$ . This proves the first part of the theorem.

*Part II)* Now we prove that the limit  $\mathbf{w}^{\varepsilon, n}$  solves problem (4.6) for  $n = 0, \dots, N_t - 1$ . We write (2.20)–(2.21) as a variational problem:

$$\begin{aligned} \int_{\Omega} A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T s''(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \boldsymbol{\sigma}_m^{\varepsilon, n+1} \cdot \boldsymbol{\varphi}_m \, d\mathbf{x} \\ = - \int_{\Omega} A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T \nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1} \cdot \boldsymbol{\varphi}_m \, d\mathbf{x} \quad \forall \boldsymbol{\varphi}_m \in \mathcal{M}_p(\mathcal{T}_{h_m})^N, \end{aligned} \quad (4.7a)$$

$$\begin{aligned}
& \varepsilon \tau_{n+1} \mathbf{c}_{h_m}(\mathbf{w}_m^{\varepsilon, n+1}, \boldsymbol{\lambda}_m) + \int_{\Omega} \mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}) \cdot \boldsymbol{\lambda}_m \, d\mathbf{x} \\
& - \tau_{n+1} \int_{\Omega} A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \boldsymbol{\sigma}_m^{\varepsilon, n+1} \cdot \nabla_{\text{DG}} \boldsymbol{\lambda}_m \, d\mathbf{x} + \tau_{n+1} \int_{\mathcal{F}_{h_m}^I} \eta_F \llbracket \mathbf{w}_m^{\varepsilon, n+1} \rrbracket_{\mathbf{N}} \cdot \llbracket \boldsymbol{\lambda}_m \rrbracket_{\mathbf{N}} \, dS \\
& = \int_{\Omega} \boldsymbol{\rho}_m^{\varepsilon, n} \cdot \boldsymbol{\lambda}_m \, d\mathbf{x} + \tau_n \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \cdot \boldsymbol{\lambda}_m \, d\mathbf{x} \quad \forall \boldsymbol{\lambda}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N.
\end{aligned} \tag{4.7b}$$

Here,  $\mathbf{c}_{h_m}(\cdot, \cdot)$  is the bilinear form in (2.22) and  $\nabla_{\text{DG}}(\cdot)$  is the LDG gradient defined in (4.1).

We infer from the discrete entropy inequalities in Theorem 3.1 that  $\{\boldsymbol{\sigma}_m^{\varepsilon, n+1}\}_m$  is bounded in the  $L^2(\Omega)$ -norm. This implies that there exists  $\boldsymbol{\sigma}^{\varepsilon, n+1} \in [L^2(\Omega)^d]^N$  such that, up to extracting a subsequence,

$$\boldsymbol{\sigma}_m^{\varepsilon, n+1} \rightharpoonup \boldsymbol{\sigma}^{\varepsilon, n+1} \quad \text{weakly in } [L^2(\Omega)^d]^N.$$

Moreover, as  $\{\mathbf{w}_m^{\varepsilon, n+1}\}_m$  is bounded in the DG norm, by Assumption 4.1, *iii*), there exists  $\mathbf{w}^{\varepsilon, n+1} \in H^1(\Omega)^N$  such that, up to a subsequence,

$$\begin{aligned}
\nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1} & \rightharpoonup \nabla \mathbf{w}^{\varepsilon, n+1} & \text{weakly in } [L^2(\Omega)^d]^N, \\
\mathbf{w}_m^{\varepsilon, n+1} & \rightarrow \mathbf{w}^{\varepsilon, n+1} & \text{strongly in } L^q(\Omega)^N,
\end{aligned}$$

with  $1 \leq q < 6$  if  $d = 3$ , and  $1 \leq q < \infty$  if  $d = 1, 2$ . From Part *I*), we have that  $\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}) \rightarrow \mathbf{u}(\mathbf{w}^{\varepsilon, n+1})$  strongly in  $L^r(\Omega)^N$  for any  $r \in [1, \infty)$  and therefore almost everywhere. Due to the continuity of  $A$ , we also have that  $A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \rightarrow A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}))$  almost everywhere. Furthermore, as  $A$  is continuous in  $\overline{\mathcal{D}}$  (see (H1)) and  $\mathbf{u} : \mathbb{R}^N \rightarrow \mathcal{D}$ , the sequence  $\{A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))\}_m$  is uniformly bounded. Therefore,

$$A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \rightarrow A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \quad \text{strongly in } [L^r(\Omega)^d]^{N \times N} \text{ for all } r \in [1, \infty).$$

Similarly, we deduce that

$$\mathbf{f}(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \rightarrow \mathbf{f}(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \quad \text{strongly in } [L^r(\Omega)^d]^N \text{ for all } r \in [1, \infty). \tag{4.8}$$

The boundedness of  $\{A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T \nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1}\}_m$  in  $[L^2(\Omega)^d]^N$  implies that there exists  $\Phi \in [L^2(\Omega)^d]^N$  such that, up to extracting a subsequence,

$$A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T \nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1} \rightharpoonup \Phi \quad \text{weakly in } [L^2(\Omega)^d]^N.$$

As  $A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T \nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1}$  is the product of a term that converges strongly in  $[L^r(\Omega)^d]^{N \times N}$  for any  $r \in [1, \infty)$  and a term that converges weakly in  $[L^2(\Omega)^d]^N$ , it converges weakly in  $[L^s(\Omega)^d]^N$  for any  $s < 2$  ( $\frac{1}{r} + \frac{1}{2} = \frac{1}{s}$ ) to the product of the two limits. Therefore, for the uniqueness of the weak limit,  $\Phi$  must be equal to the product of the two limits. This proves that

$$A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T \nabla_{\text{DG}} \mathbf{w}_m^{\varepsilon, n+1} \rightharpoonup A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}))^T \nabla \mathbf{w}^{\varepsilon, n+1} \quad \text{weakly in } [L^2(\Omega)^d]^N. \tag{4.9}$$

Similarly, we have the convergence

$$A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \boldsymbol{\sigma}_m^{\varepsilon, n+1} \rightharpoonup A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \boldsymbol{\sigma}^{\varepsilon, n+1} \quad \text{weakly in } [L^2(\Omega)^d]^N. \tag{4.10}$$

Moreover, if  $d = 1$  or  $\ell = 2$ , Assumption 4.1, *ii*), implies that  $\mathbf{w}_m^{\varepsilon, n+1}(\mathbf{x}) \in \mathcal{K}$  a.e. in  $\Omega$ , for some compact  $\mathcal{K} \subset \mathbb{R}^N$  and all  $m \geq 0$ . Therefore,  $\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})(\mathbf{x}) \in \tilde{\mathcal{K}}$  a.e. in  $\mathbb{R}^N$ , for some compact  $\tilde{\mathcal{K}} \subset \mathcal{D}$  and all  $m \geq 0$ . Since  $A^T s''$  is continuous in  $\tilde{\mathcal{K}}$ , proceeding again as for (4.9), it follows that

$$A(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1}))^T s''(\mathbf{u}(\mathbf{w}_m^{\varepsilon, n+1})) \boldsymbol{\sigma}_m^{\varepsilon, n+1} \rightharpoonup A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}))^T s''(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \boldsymbol{\sigma}^{\varepsilon, n+1} \quad \text{weakly in } [L^2(\Omega)^d]^N. \tag{4.11}$$

When  $s'' A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$ , the weak convergence in (4.11) follows from the boundedness of  $\mathcal{D}$  without requiring Assumption 4.1, *ii*), to be satisfied.

In order to pass to the limit in both sides of equation (4.7a), we observe that, for every  $\boldsymbol{\varphi} \in [L^2(\Omega)^d]^N$ , there exists a sequence  $\{\boldsymbol{\varphi}_m\}_m \subset \mathcal{M}_p(\mathcal{T}_{h_m})^N$  that converges to  $\boldsymbol{\varphi}$  strongly in  $[L^2(\Omega)^d]^N$ . We test (4.7a) with  $\boldsymbol{\varphi}_m$ . Then, the weak convergence (4.11) and the strong convergence of  $\boldsymbol{\varphi}_m$  imply that

$$\int_{\Omega} A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}))^T s''(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \boldsymbol{\sigma}^{\varepsilon, n+1} \cdot \boldsymbol{\varphi} \, d\mathbf{x} = - \int_{\Omega} A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}))^T \nabla \mathbf{w}^{\varepsilon, n+1} \cdot \boldsymbol{\varphi} \, d\mathbf{x} \quad \forall \boldsymbol{\varphi} \in [L^2(\Omega)^d]^N.$$

This, together with the chain rule  $\nabla \mathbf{w}^{\varepsilon, n+1} = \nabla s'(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) = s''(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \nabla \mathbf{u}(\mathbf{w}^{\varepsilon, n+1})$  and Assumption (H2a) leads to

$$\underline{\boldsymbol{\sigma}}^{\varepsilon, n+1} = -\nabla \mathbf{u}(\mathbf{w}^{\varepsilon, n+1}). \quad (4.12)$$

Next, we consider equation (4.7b). For any  $\boldsymbol{\lambda} \in H^\ell(\Omega)^N$ , let  $\{\boldsymbol{\lambda}_m\}_m \subset \mathcal{S}_\ell^{\text{cont}}(\mathcal{T}_{h_m})^N$  be a sequence as in Assumption 4.1, *iii*). Due to the assumption  $\mathcal{S}_\ell^{\text{cont}}(\mathcal{T}_{h_m}) \subset \mathcal{S}_p(\mathcal{T}_{h_m})$ , we can test (4.7b) with  $\boldsymbol{\lambda}_m$ . Taking into account that  $\boldsymbol{\lambda}_m$  has zero jumps across interelement boundaries, the last term on the left-hand side in (4.7b), which involves  $\llbracket \boldsymbol{\lambda}_m \rrbracket_{\mathbb{N}}$ , is equal to zero. We deduce from Assumption 4.1 (in particular, (4.4)), part *I*) of the present theorem, and the limits (4.8) and (4.10), that the weak convergence to the appropriate limits of the remaining terms that involve trial functions. Together with the strong convergence of the terms containing test functions, we find that

$$\begin{aligned} \tau_{n+1} \int_{\Omega} \left( \sum_{|\alpha|=\ell} D^\alpha \mathbf{w}^{\varepsilon, n+1} \cdot D^\alpha \boldsymbol{\lambda} + \mathbf{w}^{\varepsilon, n+1} \cdot \boldsymbol{\lambda} \right) dx + \int_{\Omega} (\mathbf{u}(\mathbf{w}^{\varepsilon, n+1}) \cdot \boldsymbol{\lambda} - \tau_{n+1} A(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \underline{\boldsymbol{\sigma}}^{\varepsilon, n+1} \cdot \nabla \boldsymbol{\lambda}) dx \\ = \int_{\Omega} \boldsymbol{\rho}^{\varepsilon, n} \cdot \boldsymbol{\lambda} dx + \tau_{n+1} \int_{\Omega} \mathbf{f}(\mathbf{u}(\mathbf{w}^{\varepsilon, n+1})) \cdot \boldsymbol{\lambda} dx \quad \forall \boldsymbol{\lambda} \in [H^\ell(\Omega)]^N. \end{aligned}$$

The combination of this with identity (4.12) implies that solves (4.6) for  $n = 0, \dots, N_t - 1$ ,  $\mathbf{w}^{\varepsilon, n+1}$ . This completes the proof of second part of the theorem.

*Part III*) This part follows from (4.6) in Part *II*), by proceeding as in Theorem 3.1.  $\square$

**Remark 4.2** (Unbounded domains  $\mathcal{D}$ ). *In the proof of the existence of discrete solutions and of the convergence to a solution to the  $\varepsilon$ -perturbed continuous problem (4.6), we repeatedly use the boundedness of  $\mathcal{D}$  and the continuity of  $A$  and  $\mathbf{f}$  on  $\overline{\mathcal{D}}$  (see Assumption (H2a)). Such a restriction can be lifted by using the argument employed to prove limit (4.11). More precisely, the presence of the regularizing term in the fully discrete scheme (2.21) and Assumption *ii*) guarantee that  $\mathbf{w}_h^{\varepsilon, n}(\mathbf{x}) \in \mathcal{K}$  a.e. in  $\Omega$ , for some compact  $\mathcal{K} \subset \mathbb{R}^N$ , which implies that  $\mathbf{u}(\mathbf{w}_h^{\varepsilon, n})(\mathbf{x}) \in \tilde{\mathcal{K}}$  a.e. in  $\mathbb{R}^N$ , for some compact  $\tilde{\mathcal{K}} \subset \mathcal{D}$  and all  $h > 0$ . Therefore, at each occurrence, the assumption of the boundedness of  $\mathcal{D}$  can be replaced by the boundedness of  $\tilde{\mathcal{K}}$  and the fact that the compact  $\tilde{\mathcal{K}}$  is independent of  $h$ .  $\blacksquare$*

### 4.3 Convergence to a weak solution to the continuous problem

Let  $\mathcal{I}_\tau$  be a temporal mesh and  $\{\boldsymbol{\rho}^{\varepsilon, n}\}_{n=0}^{N_t}$  be the corresponding semidiscrete-in-time solution from Theorem 4.1. For simplicity, we assume  $\mathcal{I}_\tau$  to be uniform. We define  $\boldsymbol{\rho}^{(\varepsilon, \tau)} \in L^2(0, T; H^1(\Omega)^N)$  as the piecewise linear reconstruction in time of  $\{\boldsymbol{\rho}^{\varepsilon, n}\}_{n=0}^{N_t}$  defined by

$$\boldsymbol{\rho}^{(\varepsilon, \tau)}(\cdot, t) := \boldsymbol{\rho}^{\varepsilon, n+1}(\cdot) - ((n+1)\tau - t)(\boldsymbol{\rho}^{\varepsilon, n+1}(\cdot) - \boldsymbol{\rho}^{\varepsilon, n}(\cdot))/\tau \quad \text{for } n\tau \leq t \leq (n+1)\tau, \quad 0 \leq n \leq N_t - 1. \quad (4.13)$$

We also define the shift  $\mathbf{s}_\tau \boldsymbol{\rho}^{(\varepsilon, \tau)}(\cdot, t) = \boldsymbol{\rho}^{\varepsilon, n}(\cdot)$  for  $n\tau \leq t \leq (n+1)\tau$ ,  $0 \leq n \leq N_t - 1$ .

We say that  $\boldsymbol{\rho}$  is a *weak solution to the continuous problem (1.1)* if it satisfies

- $\boldsymbol{\rho} \in L^2(0, T; H^1(\Omega)^N) \cap H^1(0, T; [H^1(\Omega)^N]')$ ;
- $\boldsymbol{\rho}(\mathbf{x}, t) \in \overline{\mathcal{D}}$  a.e. in  $\Omega \times (0, T]$  (in particular,  $\boldsymbol{\rho} \in L^\infty(0, T; [L^\infty(\Omega)]^N)$ );
- $\boldsymbol{\rho}(\cdot, 0) = \boldsymbol{\rho}_0(\cdot)$  in the sense of  $[H^1(\Omega)^N]'$ ;
- $\int_0^T \langle \partial_t \boldsymbol{\rho}, \boldsymbol{\lambda} \rangle dt + \int_0^T \int_{\Omega} A(\boldsymbol{\rho}) \nabla \boldsymbol{\rho} : \nabla \boldsymbol{\lambda} dx dt = \int_0^T \int_{\Omega} \mathbf{f}(\boldsymbol{\rho}) \cdot \boldsymbol{\lambda} dx dt \quad \forall \boldsymbol{\lambda} \in L^2(0, T; H^1(\Omega)^N)$ ,

where  $\langle \cdot, \cdot \rangle$  denotes the duality between  $[H^1(\Omega)^N]'$  and  $H^1(\Omega)^N$ .

**Theorem 4.3.** *Let Assumption 4.1 be satisfied, and let  $\boldsymbol{\rho}^{(\varepsilon, \tau)}$  be the piecewise linear reconstruction of the semidiscrete-in-time solution from Theorem 4.1. Then, there exists a continuous weak solution  $\boldsymbol{\rho}$  to problem (1.1) such that, up to a subsequence that is not relabeled, for  $(\varepsilon, \tau) \rightarrow (0, 0)$ , we have*

$$\begin{aligned} \boldsymbol{\rho}^{(\varepsilon, \tau)} &\rightarrow \boldsymbol{\rho} && \text{strongly in } L^r(0, T; L^r(\Omega)^N) \text{ for any } r < \infty \text{ and a.e. in } \Omega \times (0, T], \\ \nabla \boldsymbol{\rho}^{(\varepsilon, \tau)} &\rightharpoonup \nabla \boldsymbol{\rho} && \text{weakly in } L^2(0, T; [L^2(\Omega)^d]^N), \\ \tau^{-1}(\boldsymbol{\rho}^{(\varepsilon, \tau)} - \mathbf{s}_\tau \boldsymbol{\rho}^{(\varepsilon, \tau)}) &\rightharpoonup \partial_t \boldsymbol{\rho} && \text{weakly in } L^2(0, T; [H^\ell(\Omega)^N]'), \end{aligned}$$

where the integer  $\ell$  is as in Assumption 4.1.

*Proof.* The proof follows closely the arguments in steps 2 and 3 of [14, §3].  $\square$

#### 4.4 Validity of Assumption 4.1

The proof of Theorem 4.1 strongly relies on the validity of Assumption 4.1. Due to our mesh assumptions, inclusion (4.2) is satisfied whenever  $p \geq \ell$ . Before discussing the existence of a bilinear form  $c_h(\cdot, \cdot)$  and a DG norm  $\|\cdot\|_{\text{DG}}$  with the properties *i)–iii)* in Assumption 4.1, we prove the following estimate, which is an extension of [8, Thm. 4.4] to the cases  $q = 1$ ,  $p = 1, 2$ .

**Lemma 4.4** (Broken trace estimate). *For  $r \in \{1, 2\}$ , we have the following estimate:*

$$\|v\|_{L^1(\partial\Omega)} \lesssim \|v\|_{L^r(\Omega)} + \|\nabla_h v\|_{L^r(\Omega)^d} + \|\mathbf{h}^{\frac{1-r}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^r(\mathcal{F}_h^{\mathcal{I}})^d} \quad \forall v \in \mathcal{S}_p(\mathcal{T}_h),$$

where the hidden constant is independent of  $h$  and  $v$ , but it depends on  $\Omega$ .

*Proof.* Let  $v \in \mathcal{S}_p(\mathcal{T}_h)$  and  $Q_h : \mathcal{S}_p(\mathcal{T}_h) \rightarrow \mathcal{S}_1^{\text{cont}}(\mathcal{T}_h)$  be the reconstruction operator defined in [8, §3]. By the triangle inequality,

$$\|v\|_{L^1(\partial\Omega)} \leq \|Q_h v\|_{L^1(\partial\Omega)} + \|v - Q_h v\|_{L^1(\partial\Omega)}. \quad (4.14)$$

The trace theorem in  $W^{1,1}(\Omega)$  gives

$$\|Q_h v\|_{L^1(\partial\Omega)} \lesssim \|Q_h v\|_{L^1(\Omega)} + \|\nabla Q_h v\|_{L^1(\Omega)^d}. \quad (4.15)$$

Thus, it follows from (4.14) and (4.15), by applying the triangle inequality and [8, Thm. 3.1], that

$$\begin{aligned} \|v\|_{L^1(\partial\Omega)} &\lesssim \|v\|_{L^1(\Omega)} + \|v - Q_h v\|_{L^1(\Omega)} + \|\nabla Q_h v\|_{L^1(\Omega)^d} + \|v - Q_h v\|_{L^1(\partial\Omega)} \\ &\lesssim \|v\|_{L^1(\Omega)} + \|\nabla_h v\|_{L^1(\Omega)^d} + \|\llbracket v \rrbracket_{\mathbf{N}}\|_{L^1(\mathcal{F}_h^{\mathcal{I}})^d}, \end{aligned} \quad (4.16)$$

which completes the proof for  $r = 1$ .

We now consider the case  $r = 2$ . We infer from the Cauchy–Schwarz inequality that

$$\|v\|_{L^1(\Omega)} = \sum_{K \in \mathcal{T}_h} \int_K |v| d\mathbf{x} \leq \left( \sum_{K \in \mathcal{T}_h} |K| \right)^{\frac{1}{2}} \|v\|_{L^2(\Omega)} = |\Omega|^{\frac{1}{2}} \|v\|_{L^2(\Omega)}, \quad (4.17)$$

and similarly,

$$\|\nabla_h v\|_{L^1(\Omega)^d} \leq |\Omega|^{\frac{1}{2}} \|\nabla_h v\|_{L^2(\Omega)^d}. \quad (4.18)$$

Moreover, by the definition of  $\mathbf{h}$  in (2.3) and, for  $d = 2, 3$ , the shape-regularity assumption, we have<sup>1</sup>

$$\int_{\mathcal{F}_h^{\mathcal{I}}} \mathbf{h} dS \lesssim \sum_{K \in \mathcal{T}_h} |K|,$$

from which we deduce that

$$\begin{aligned} \|\llbracket v \rrbracket_{\mathbf{N}}\|_{L^1(\mathcal{F}_h^{\mathcal{I}})^d} &= \int_{\mathcal{F}_h^{\mathcal{I}}} \mathbf{h}^{\frac{1}{2}} \mathbf{h}^{-\frac{1}{2}} |\llbracket v \rrbracket_{\mathbf{N}}| dS \leq \left( \int_{\mathcal{F}_h^{\mathcal{I}}} \mathbf{h} dS \right)^{\frac{1}{2}} \|\mathbf{h}^{-\frac{1}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \left( \sum_{K \in \mathcal{T}_h} |K| \right)^{\frac{1}{2}} \|\mathbf{h}^{-\frac{1}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \\ &\lesssim \|\mathbf{h}^{-\frac{1}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d}. \end{aligned} \quad (4.19)$$

Combining the broken trace estimate for  $r = 1$  in (4.16) with bounds (4.17), (4.18), and (4.19), the desired result for  $r = 2$  follows, completing the proof.  $\square$

We now discuss the validity of Assumption 4.1. For this, we distinguish three cases.

<sup>1</sup>For  $d = 1$ ,  $\int_{\mathcal{F}_h^{\mathcal{I}}} \mathbf{h} dS = \sum_{i=1}^{M-1} \mathbf{h}(x_i) \lesssim \sum_{K \in \mathcal{T}_h} |K|$ , where  $\{x_i\}_{i=0}^M$  are the meshpoints.

For  $d = 2, 3$ ,  $|K| = (\text{sum of facet } (d-1)\text{-measures}) \times \varrho_K/d$ , with  $\varrho_K$  being the inradius of  $K$ . From the shape-regularity assumption (2.2), we deduce that  $|K| \geq \Upsilon h_K h_F/d$  for any facet  $F$  of  $K$ , and obtain  $\int_{\mathcal{F}_h^{\mathcal{I}}} \mathbf{h} dS \lesssim \sum_{F \subset \mathcal{F}_h^{\mathcal{I}}} h|F| \lesssim \sum_{K \in \mathcal{T}_h} |K|$ .



- **Case  $d = 1$  ( $\ell = 1$ ).** Choose  $c_{h_m}(\cdot, \cdot)$  and the  $H^1$ -type norm  $\|\cdot\|_{\text{DG}}$  as

$$c_{h_m}(w_m, v_m) := \int_{\Omega} w_m v_m d\mathbf{x} + \int_{\Omega} \nabla_{\text{DG}} w_m \cdot \nabla_{\text{DG}} v_m d\mathbf{x} + \int_{\mathcal{F}_{h_m}^{\mathcal{I}}} h^{-1} \llbracket w_m \rrbracket_{\mathbf{N}} \cdot \llbracket v_m \rrbracket_{\mathbf{N}} dS \quad \forall w_m, v_m \in \mathcal{S}_p(\mathcal{T}_{h_m}), \quad (4.20)$$

$$\|\mathbf{w}_m\|_{\text{DG}}^2 := \|\mathbf{w}_m\|_{L^2(\Omega)^N}^2 + \|\nabla_h \mathbf{w}_m\|_{[L^2(\Omega)^d]^N}^2 + \|\mathbf{h}^{-\frac{1}{2}} \llbracket \mathbf{w}_m \rrbracket_{\mathbf{N}}\|_{[L^2(\mathcal{F}_{h_m}^{\mathcal{I}})^d]^N}^2 \quad \forall \mathbf{w}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N. \quad (4.21)$$

With this choice, property *i*) follows from the coercivity of the LDG discretization of the Laplace operator (see, e.g., [20, Prop. 3.1]), property *iii*) follows from [8, Thm. 5.2 and Lemma 8], and property *ii*) follows from the following proposition.

**Proposition 4.5** (Discrete Sobolev embedding in 1D). *Let  $(a, b)$  be an interval in  $\mathbb{R}$ , and let  $\mathcal{S}_p(\mathcal{T}_h)$  be defined on the partition  $\mathcal{T}_h$  given by  $a =: x_0 < x_1 < \dots < x_M =: b$ . If the DG norm is chosen as in (4.21) with  $N = 1$ , then, for all  $v \in \mathcal{S}_p(\mathcal{T}_h)$ ,*

$$\|v\|_{L^\infty(a,b)} \lesssim \|v\|_{\text{DG}},$$

where the hidden constant is independent of  $h$  and  $v$ .

*Proof.* Let  $K_i := (x_{i-1}, x_i)$  and  $h_i := x_i - x_{i-1}$ , for  $i = 1, \dots, M$ . For any  $v \in \mathcal{S}_p(\mathcal{T}_h)$  and  $j = 1, \dots, M$ , by the Fundamental Theorem of Calculus and the Hölder inequality, we have for all  $x \in K_j$ ,

$$\begin{aligned} v(x) &= v(a) + \sum_{i=1}^{j-1} \left( \int_{K_i} v'(x) dx + v(x_i^+) - v(x_i^-) \right) + \int_{x_{j-1}}^x v'(x) dx \\ &\leq |v(a)| + \sum_{i=1}^M \|v'\|_{L^1(K_i)} + \sum_{i=1}^{M-1} |v(x_i^+) - v(x_i^-)| \\ &\leq |v(a)| + \sum_{i=1}^M h_i^{\frac{1}{2}} \|v'\|_{L^2(K_i)} + \sum_{i=1}^{M-1} h^{\frac{1}{2}}(x_i) h^{-\frac{1}{2}}(x_i) |v(x_i^+) - v(x_i^-)| \\ &\leq |v(a)| + \left( \sum_{i=1}^M h_i \right)^{\frac{1}{2}} \|v'\|_{L^2(a,b)} + \left( \sum_{i=1}^{M-1} h(x_i) \right)^{\frac{1}{2}} \|\mathbf{h}^{-\frac{1}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})} \\ &\lesssim |v(a)| + \|v'\|_{L^2(a,b)} + \|\mathbf{h}^{-\frac{1}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})} \lesssim |v(a)| + \|v\|_{\text{DG}}. \end{aligned}$$

Lemma 4.4 with  $r = 2$  implies that  $|v(a)| \lesssim \|v\|_{\text{DG}}$ , and the proof is complete.  $\square$

- **Case  $d = 2, 3$  and  $s''A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$  ( $\ell = 1$ ).** In this case, the enforcement of the  $L^\infty(\Omega)$ -boundedness on the discrete entropy variable  $\mathbf{w}_m^{\varepsilon, n}$ , which is a consequence of property *ii*), is no longer necessary, as the weak convergence in (4.11) follows from the boundedness of  $\mathcal{D}$  and the continuity of  $s''A$  on  $\overline{\mathcal{D}}$ . Moreover, for the bilinear form  $c_{h_m}(\cdot, \cdot)$  and the norm  $\|\cdot\|_{\text{DG}}$  defined in (4.20) and (4.21), respectively, properties *i*) and *iii*) follow from the same results as in the case  $d = 1$ . We present some examples for cross-diffusion systems satisfying  $s''A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$  in Appendix A.

- **Case  $d = 2, 3$  and  $\ell = 2$ .** We define the discrete LDG Hessian operator  $\mathcal{H}_{\text{DG}} : \mathcal{S}_p(\mathcal{T}_{h_m}) \rightarrow [L^2(\Omega)]^{d \times d}$  as

$$\int_{\Omega} \mathcal{H}_{\text{DG}} \lambda_m : \Theta_m d\mathbf{x} = \int_{\Omega} (D_h^2 \lambda_m - \mathcal{R}(\lambda_m) + \mathcal{B}(\lambda_m)) : \Theta_m d\mathbf{x} \quad \forall \Theta_m \in \prod_{K \in \mathcal{T}_{h_m}} \mathbb{P}^p(K)^{d \times d},$$

where  $D_h^2$  denotes the elementwise Hessian operator, and the lifting operators  $\mathcal{R} : \mathcal{S}_p(\mathcal{T}_{h_m}) \rightarrow [L^2(\Omega)]^{d \times d}$  and  $\mathcal{B} : \mathcal{S}_p(\mathcal{T}_{h_m}) \rightarrow [L^2(\Omega)]^{d \times d}$  are defined by

$$\begin{aligned} \int_{\Omega} \mathcal{R}(\lambda_m) : \Theta_m d\mathbf{x} &= \sum_{K \in \mathcal{T}_{h_m}} \int_{(\partial K)^\circ} \{\Theta_m\}_{\mathbf{n}_K} \cdot \nabla \lambda_m dS & \forall \Theta_m \in \prod_{K \in \mathcal{T}_{h_m}} \mathbb{P}^p(K)^{d \times d}, \\ \int_{\Omega} \mathcal{B}(\lambda_m) : \Theta_m d\mathbf{x} &= \int_{\mathcal{F}_{h_m}^{\mathcal{I}}} \{\nabla_h \cdot \Theta_m\} \cdot \llbracket \lambda_m \rrbracket_{\mathbf{N}} dS & \forall \Theta_m \in \prod_{K \in \mathcal{T}_{h_m}} \mathbb{P}^p(K)^{d \times d}. \end{aligned}$$

For piecewise smooth functions  $w$ ,  $\mathbf{w}$ , and  $\boldsymbol{w}$  with  $d$ ,  $N$ , and  $d \times N$  components, respectively, we define the (vector-valued) total jump on each facet  $F = \partial K_1 \cap \partial K_2 \in \mathcal{F}_h^\mathcal{I}$ , for some  $K_1, K_2 \in \mathcal{T}_h$ , with a prescribed unit normal vector, say, pointing from  $K_1$  to  $K_2$ , as

$$[[w]] := w|_{K_1} - w|_{K_2}, \quad [[\mathbf{w}]] := \mathbf{w}|_{K_1} - \mathbf{w}|_{K_2}, \quad [[\boldsymbol{w}]] := \boldsymbol{w}|_{K_1} - \boldsymbol{w}|_{K_2}.$$

Finally, we choose  $c_{h_m}(\cdot, \cdot)$  and the  $H^2$ -type norm  $\|\cdot\|_{\text{DG}}$  as

$$\begin{aligned} c_{h_m}(w_m, v_m) &:= \int_{\Omega} w_m v_m d\mathbf{x} + \int_{\Omega} \nabla_{\text{DG}} w_m \cdot \nabla_{\text{DG}} v_m d\mathbf{x} + \int_{\Omega} \mathcal{H}_{\text{DG}} w_m : \mathcal{H}_{\text{DG}} v_m d\mathbf{x} \\ &\quad + \int_{\mathcal{F}_{h_m}^\mathcal{I}} h^{-1} [[\nabla_h w_m]] \cdot [[\nabla_h v_m]] dS + \int_{\mathcal{F}_{h_m}^\mathcal{I}} h^{-3} [[w_m]]_{\mathbf{N}} \cdot [[v_m]]_{\mathbf{N}} dS \quad \forall w_m, v_m \in \mathcal{S}_p(\mathcal{T}_{h_m}), \end{aligned} \quad (4.22)$$

$$\begin{aligned} \|\mathbf{w}_m\|_{\text{DG}}^2 &:= \|\mathbf{w}_m\|_{L^2(\Omega)^N}^2 + \|\nabla_h \mathbf{w}_m\|_{[L^2(\Omega)^{d \times N}]^N}^2 + \|D_h^2 \mathbf{w}_m\|_{[L^2(\Omega)^{d \times d \times N}]^N}^2 \\ &\quad + \|h^{-\frac{1}{2}} [[\nabla_h \mathbf{w}_m]]\|_{[L^2(\mathcal{F}_{h_m}^\mathcal{I})^{d \times N}]^N}^2 + \|h^{-\frac{3}{2}} [[\mathbf{w}_m]]_{\mathbf{N}}\|_{[L^2(\mathcal{F}_{h_m}^\mathcal{I})^{d \times N}]^N}^2 \quad \forall \mathbf{w}_m \in \mathcal{S}_p(\mathcal{T}_{h_m})^N. \end{aligned} \quad (4.23)$$

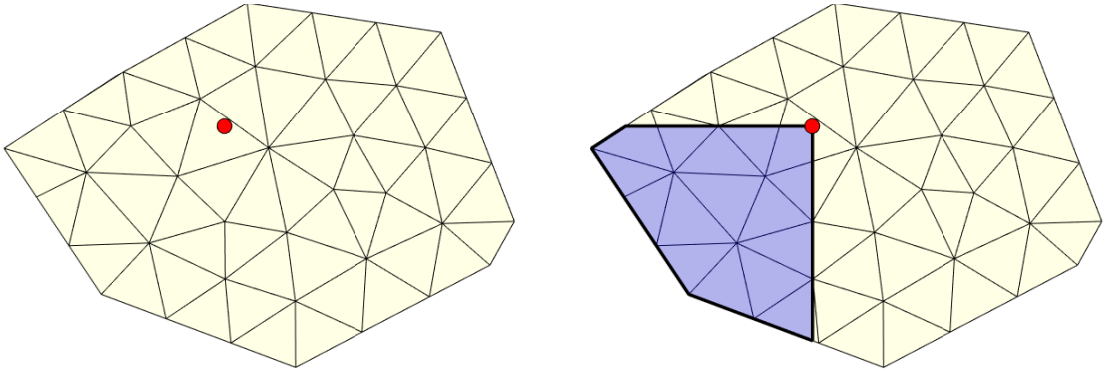
Then, property *i*) follows from [5, Lemma 2.6]. The discrete compactness argument in Assumption 4.1, *iii*), can be proven similarly as in [5, Lemma 2.2] (see also [5, Appendix C]), whereas (4.4) follows from [5, Lemmas 2.4 and 2.5] and from the second estimate in Step 2 of the proof of [5, Lemma 2.5]. For  $d = 2$ , Property *ii*) is proven in the following proposition.<sup>2</sup>

**Proposition 4.6** (Discrete Sobolev embedding in 2D). *Let  $\Omega \subset \mathbb{R}^2$  be an open, bounded polytopic domain, and let the DG-norm be defined as in (4.23) with  $N = 1$ . Then, for all  $v \in \mathcal{S}_p(\mathcal{T}_h)$ ,*

$$\|v\|_{L^\infty(\Omega)} \lesssim \|v\|_{\text{DG}},$$

where the hidden constant is independent of  $h$  and  $v$ .

*Proof.* Let  $v \in \mathcal{S}_p(\mathcal{T}_h)$  and  $(\bar{x}, \bar{y})$  be an interior point of some element  $K \in \mathcal{T}_h$ . If  $\Omega$  is convex, we define an auxiliary domain  $\tilde{\Omega} := [(-\infty, \bar{x}) \times (-\infty, \bar{y})] \cap \Omega$ , and an auxiliary mesh  $\tilde{\mathcal{T}}_h$  given by the “intersection” of  $\mathcal{T}_h$  and  $\tilde{\Omega}$ . We illustrate these definitions in Figure 1. If  $\Omega$  is not convex, let  $(\bar{x}, y_{\partial\Omega})$  be the intersection of the half-line  $(-\infty, \bar{x})$  with  $\partial\Omega$  having the largest  $y$ -coordinate, and  $(x_{\partial\Omega}, \bar{y})$  be the intersection of the half-line  $(-\infty, \bar{y})$  with  $\partial\Omega$  having the largest  $x$ -coordinate. We let  $\Gamma_x$  and  $\Gamma_y$  be the segments with endpoints  $(\bar{x}, \bar{y})$  and  $(x_{\partial\Omega}, \bar{y})$ , respectively. Then, we define  $\tilde{\Omega}$  as the connected subregion of  $\Omega$  delimited by  $\Gamma_x$ ,  $\Gamma_y$  on the side where the angle between  $\Gamma_x$  and  $\Gamma_y$  equals  $\pi/2$ .



**Figure 1:** Example of a two dimensional domain  $\Omega$  (in yellow). **Left panel:** Triangular mesh  $\mathcal{T}_h$  of  $\Omega$  and an interior point  $(\bar{x}, \bar{y})$  (depicted with a red dot) of some element  $K \in \mathcal{T}_h$ . **Right panel:** Auxiliary domain  $\tilde{\Omega}$  (in blue) and auxiliary mesh  $\tilde{\mathcal{T}}_h$ .

Integration by parts with respect to  $x$  gives

$$\sum_{\tilde{K} \in \tilde{\mathcal{T}}_h} \int_{\tilde{K}} \partial_{xy} v d\mathbf{x} = \int_{\tilde{\mathcal{F}}_h^\mathcal{I}} [[\partial_{y,h} v]]_{n_x} dS + \int_{\tilde{\mathcal{F}}_h^{\mathcal{N}}} \partial_{y,h} v n_{\tilde{\Omega}}^x dS,$$

<sup>2</sup>For  $d = 3$ , one could develop a similar proof based on the Fundamental Theorem of Calculus, provided that terms with third-order derivatives are added to the regularization form and to the DG norm. However, comparing with standard Sobolev embeddings, one expects the discrete Sobolev embedding to be valid also in 3D with definitions (4.22)–(4.23). This issue remains open.

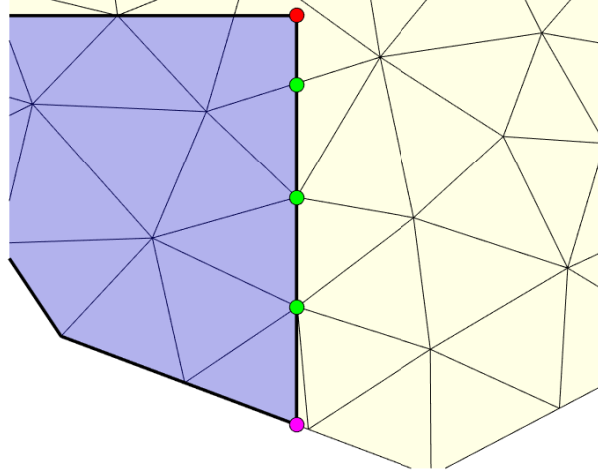
where  $[[\cdot]]_{n_x}$  denotes the first component of the normal jump  $[[\cdot]]_{\mathbf{N}}$ ,  $\partial_{y,h}$  the elementwise partial  $y$ -derivative, and  $n_{\Omega}^x$  the first component of the unit normal vector pointing outside  $\Omega$ . The boundary of  $\tilde{\Omega}$  can be split into three parts as  $\partial\tilde{\Omega} = (\partial\Omega \cap \partial\tilde{\Omega}) \cup \partial\tilde{\Omega}^x \cup \partial\tilde{\Omega}^y$ , where  $\tilde{\Omega}^x$  and  $\tilde{\Omega}^y$  are the parts of  $\partial\tilde{\Omega}$  along the lines  $x = \bar{x}$  and  $y = \bar{y}$ , respectively. Observe that

$$n_{\Omega}^x = \begin{cases} 0 & \text{on } \partial\tilde{\Omega}^y, \\ 1 & \text{on } \partial\tilde{\Omega}^x, \\ n_{\Omega}^x & \text{on } \partial\Omega \cap \partial\tilde{\Omega}, \end{cases}$$

whence,

$$\sum_{\tilde{K} \in \tilde{\mathcal{T}}_h} \int_{\tilde{K}} \partial_{xy} v \, d\mathbf{x} = \int_{\tilde{\mathcal{F}}_h^x} [[\partial_{y,h} v]]_{n_x} dS + \int_{\tilde{\mathcal{F}}_h^{\mathcal{N}} \cap \partial\Omega} \partial_{y,h} v n_{\Omega}^x dS + \int_{\tilde{\mathcal{F}}_h^{\mathcal{N}} \cap \partial\tilde{\Omega}^x} \partial_{y,h} v dS.$$

We now focus on the last term of the previous identity. Let  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$ , with  $\ell \in \mathbb{N}$ , be the set containing all internal vertices of  $\mathcal{T}_h$  that lie on  $\partial\tilde{\Omega}^x$ , as well as all intersections between  $\partial\tilde{\Omega}^x$  and those edges in  $\mathcal{F}_h^x$  that do not lie along  $\partial\tilde{\Omega}^x$ . We assume that the points in  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$  are ordered with decreasing  $y$ -coordinate. Furthermore, we denote by  $(\bar{x}, y^{\partial})$  the intersection between  $\partial\tilde{\Omega}^x$  and  $\partial\Omega$ . In Figure 2, we illustrate the notation used for the vertices of  $\tilde{\mathcal{T}}_h$  lying on  $\partial\tilde{\Omega}^x$ .<sup>3</sup>



**Figure 2:** Example of vertices of  $\tilde{\mathcal{T}}_h$  lying along  $\partial\tilde{\Omega}^x$ . The red dot has the coordinates  $(\bar{x}, \bar{y})$ ; the green dots have the coordinates  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$  for some  $\ell \in \mathbb{N}$ ; the purple dot belongs to  $\partial\Omega$  and has the coordinates  $(\bar{x}, y^{\partial})$ .

By the Fundamental Theorem of Calculus in one dimension, we have

$$\int_{\tilde{\mathcal{F}}_h^{\mathcal{N}} \cap \partial\tilde{\Omega}^x} \partial_{y,h} v dS = v(\bar{x}, \bar{y}) - \sum_{j=1}^{\ell} [[v(\bar{x}, y_j)]] - v(\bar{x}, y^{\partial}),$$

where  $[[v(\bar{x}, y_j)]] := \lim_{\varepsilon \rightarrow 0} |v(\bar{x}, y_j + \varepsilon) - v(\bar{x}, y_j - \varepsilon)|$ . Therefore,

$$\begin{aligned} v(\bar{x}, \bar{y}) &= \sum_{\tilde{K} \in \tilde{\mathcal{T}}_h} \int_{\tilde{K}} \partial_{xy} v \, d\mathbf{x} - \int_{\tilde{\mathcal{F}}_h^x} [[\partial_{y,h} v]]_{n_x} dS - \int_{\tilde{\mathcal{F}}_h^{\mathcal{N}} \cap \partial\Omega} \partial_{y,h} v n_{\Omega}^x dS + \sum_{j=1}^{\ell} [[v(\bar{x}, y_j)]] + v(\bar{x}, y^{\partial}) \\ &= T_1 + T_2 + T_3 + T_4 + T_5. \end{aligned}$$

We estimate the terms  $T_i$ ,  $i = 1, \dots, 5$ , separately.

<sup>3</sup>The boundary  $\partial\tilde{\Omega}^x$  crosses a vertex of  $\mathcal{T}_h$  (green dot in the middle) and an internal edge of  $\mathcal{T}_h$  (between the two green dots from the bottom up). This is not an issue, as the domain  $\tilde{\Omega}$  sees  $\partial\tilde{\Omega}^x$  only from the interior.

**Bound for  $T_1$ :** Proceeding as in (4.17), we find that

$$T_1 \leq |\Omega|^{\frac{1}{2}} \left( \sum_{K \in \mathcal{T}_h} \|\partial_{xy} v\|_{L^2(K)}^2 \right)^{\frac{1}{2}} \lesssim \|v\|_{\text{DG}}.$$

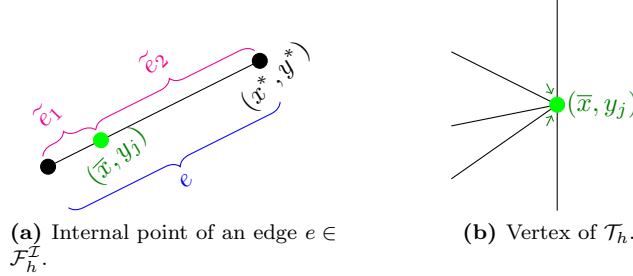
**Bound for  $T_2$ :** Since  $|n_x| \leq \sqrt{n_x^2 + n_y^2} \leq 1$ , proceeding as in (4.19) gives

$$T_2 \leq \int_{\tilde{\mathcal{F}}_h^{\mathcal{I}}} |[\![\partial_{y,h} v]\!]_{\mathbf{n}}| \, dS \leq \int_{\mathcal{F}_h^{\mathcal{I}}} |[\![\partial_{y,h} v]\!]| \, dS \lesssim \|\mathbf{h}^{-\frac{1}{2}} [\![\partial_{y,h} v]\!]\|_{L^2(\mathcal{F}_h^{\mathcal{I}})} \lesssim \|\mathbf{h}^{-\frac{1}{2}} [\![\nabla_h v]\!]\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \|v\|_{\text{DG}}.$$

**Bound for  $T_3$ :** The broken trace estimate in Lemma 4.4 for  $r = 2$  implies that

$$T_3 \leq \|\partial_{y,h} v\|_{L^1(\partial\Omega)} \lesssim \|\partial_{y,h} v\|_{L^2(K)} + \|\nabla_h(\partial_{y,h} v)\|_{L^2(K)^d} + \|\mathbf{h}^{-\frac{1}{2}} [\![\partial_{y,h} v]\!]\|_{L^2(\mathcal{F}_h^{\mathcal{I}})} \lesssim \|v\|_{\text{DG}}.$$

**Bound for  $T_4$ :** The green dots in Figure 2 with coordinates  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$  may be either: *i*) an internal point of some edge  $e \in \mathcal{F}_h^{\mathcal{I}}$ , or *ii*) a vertex of  $\mathcal{T}_h$ . Both situations are represented in Figure 3.



**Figure 3:** Illustration of the two types of points in the set  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$  used in the bound for  $T_4$ .

We consider each case separately.

**Case *i*)** Let  $e \subset \mathcal{F}_h^{\mathcal{I}}$  be the edge containing  $(\bar{x}, y_j)$ , and let  $\tilde{e}_{\max}$  be the largest segment of  $e$  having  $(\bar{x}, y_j)$  as a vertex. Let  $(x^*, y^*)$  be the remaining vertex. We deduce from  $h_{\tilde{e}_{\max}} \geq \frac{1}{2} h_e$  that  $h_{\tilde{e}_{\max}}^{-1} \leq 2h_e^{-1}$ . Set  $\Phi(t) := \llbracket v(\bar{x} + t(x^* - \bar{x}), y_j + t(y^* - y_j)) \rrbracket$ . Since  $\Phi \in \mathbb{P}^1((0, 1))$ , the inverse trace inequality  $\|v\|_{L^1(\partial D)} \lesssim h_D^{-1} \|v\|_{L^1(D)}$  shows that

$$|\Phi(0)| = \llbracket v(\bar{x}, y_j) \rrbracket \lesssim h_{\tilde{e}_{\max}}^{-1} \|\llbracket v \rrbracket\|_{L^1(\tilde{e}_{\max})} \lesssim \|h_e^{-1} \llbracket v \rrbracket\|_{L^1(e)}.$$

**Case *ii*)** Adding and subtracting the values of  $v$  at  $(\bar{x}, y_j)$  from all the elements having  $(\bar{x}, y_j)$  as a vertex, and using the triangle inequality, one can proceed as in case *i*).

**Conclusion of the bound for  $T_4$ :** Since the jumps at different points in  $\{(\bar{x}, y_j)\}_{j=1}^{\ell}$  are “lifted” to different edges, proceeding as in (4.19), we conclude that

$$T_4 = \sum_{j=1}^{\ell} \llbracket v(\bar{x}, y_j) \rrbracket \leq \sum_{j=1}^{\ell} \|\llbracket v(\bar{x}, y_j) \rrbracket\| \lesssim \|\mathbf{h}^{-1} \llbracket v \rrbracket\|_{L^1(\mathcal{F}_h^{\mathcal{I}})} \lesssim \|\mathbf{h}^{-\frac{3}{2}} \llbracket v \rrbracket\|_{L^2(\mathcal{F}_h^{\mathcal{I}})} = \|\mathbf{h}^{-\frac{3}{2}} \llbracket v \rrbracket_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \|v\|_{\text{DG}}.$$

**Bound for  $T_5$ :** Let  $(\hat{x}, \hat{y})$  be a vertex of  $\Omega$  such that the segment  $\tilde{\Gamma} := [(\hat{x}, \hat{y}), (\bar{x}, y^\partial)] \subset \partial\tilde{\Omega} \cap \partial\Omega$  has positive 1-dimensional measure<sup>4</sup>, and let  $\{(\hat{x}_i, \hat{y}_i)\}_{i=1}^k$ , with  $k \in \mathbb{N}$ , be the vertices of  $\mathcal{T}_h$  in the interior of  $\tilde{\Gamma}$ . Then, by the Fundamental Theorem of Calculus,

$$|v(\bar{x}, y^\partial)| \leq |v(\hat{x}, \hat{y})| + \|\nabla_{\tau, h} v\|_{L^1(\tilde{\Gamma})} + \sum_{i=1}^k \|\llbracket v(\hat{x}_i, \hat{y}_i) \rrbracket\|, \quad (4.24)$$

<sup>4</sup>The argument used to bound  $T_5$  is independent of whether  $(\bar{x}, y^\partial)$  is a mesh vertex or not.

where  $\nabla_{\tau,h}$  denotes the broken tangential derivative of  $v$ . Furthermore, applying Lemma 4.4 with  $r = 1$  along the side  $\Gamma$  of the boundary of  $\Omega$  containing  $\tilde{\Gamma}$ , we obtain

$$|v(\hat{x}, \hat{y})| \lesssim \|v\|_{L^1(\Gamma)} + \|\nabla_{\tau,h} v\|_{L^1(\Gamma)} + \sum_{i=1}^M |[[v(\hat{x}_i, \hat{y}_i)]]| \leq \|v\|_{L^1(\partial\Omega)} + \|\nabla_h v\|_{L^1(\partial\Omega)^d} + \sum_{i=1}^M |[[v(\hat{x}_i, \hat{y}_i)]]|,$$

where  $\{(\hat{x}_i, \hat{y}_i)\}_{i=1}^M$ , with  $M \geq k$ , are the vertices of  $\mathcal{T}_h$  in the interior of  $\Gamma$ . This, combined with (4.24), leads to

$$|v(\bar{x}, \bar{y}^\partial)| \lesssim \|v\|_{L^1(\partial\Omega)} + \|\nabla_h v\|_{L^1(\partial\Omega)^d} + \sum_{i=1}^M |[[v(\hat{x}_i, \hat{y}_i)]]| =: I_1 + I_2 + I_3.$$

Using Lemma 4.4 with  $r = 2$ , the terms  $I_1$  and  $I_2$  can be estimated as follows:

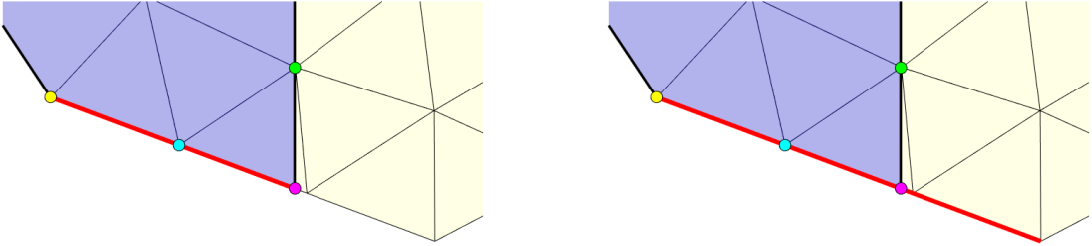
$$I_1 \lesssim \|v\|_{L^2(\Omega)} + \|\nabla_h v\|_{L^2(\Omega)^d} + \|\mathbf{h}^{-\frac{1}{2}}[[v]]_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \|v\|_{\text{DG}}, \quad (4.25a)$$

$$I_2 \lesssim \|\nabla_h v\|_{L^2(K)^d} + \|D_h^2 v\|_{L^2(K)^{d \times d}} + \|\mathbf{h}^{-\frac{1}{2}}[[\nabla_h v]]\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \|v\|_{\text{DG}}. \quad (4.25b)$$

Moreover, proceeding as for bound  $T_4$ , case *ii*), the term  $I_3$  can be estimated as

$$I_3 = \sum_{i=1}^M |[[v(\hat{x}_i, \hat{y}_i)]]| \lesssim \|\mathbf{h}^{-\frac{3}{2}}[[v]]_{\mathbf{N}}\|_{L^2(\mathcal{F}_h^{\mathcal{I}})^d} \lesssim \|v\|_{\text{DG}}. \quad (4.25c)$$

It follows from (4.25a), (4.25b), and (4.25c) that  $T_5 \lesssim \|v\|_{\text{DG}}$ , which completes the proof.



**Figure 4:** Example of the auxiliary segments  $\tilde{\Gamma}$  (left panel) and  $\Gamma$  (right panel) in the bound for  $T_5$ . The yellow dot has the coordinates  $(\hat{x}, \hat{y})$  and is a vertex of  $\Omega$ . The cyan dot has the coordinates  $(\hat{x}_1, \hat{x}_2)$

□

## 5 Numerical experiments

In this section, we assess the accuracy and entropy stability of the proposed method with some one- and two-dimensional test problems. The solutions to the nonlinear systems of equations stemming from the fully discrete method (2.20)–(2.21) are approximated using a quasi-Newton method, where the Jacobian of the nonlinear vector-valued function is evaluated on the approximation at the previous time. The tolerance ( $tol$ ) and the maximum number of linear iterations ( $s_{\max}$ ) of the nonlinear solver are specified in each test.

We use Gaussian elimination (for the one-dimensional problems) or a preconditioned BICG method (for the two-dimensional problems) to solve the linear system at each iteration of the nonlinear solver. In order to reduce the stencil of the gradient operator matrix  $B$ , we use directional numerical fluxes, i.e., we set  $\alpha_F = 1$  for all  $F \in \mathcal{F}_h^{\mathcal{I}}$  in (2.5); see [10].

## 5.1 One-dimensional porous medium equation

Given a real number  $m > 1$ , an initial datum  $\rho_0 : \Omega \rightarrow \mathcal{D}$ , and a Neumann boundary datum  $g_N : \partial\Omega \times (0, T) \rightarrow \mathbb{R}$ , we consider the following problem on a space-time cylinder  $Q_T = \Omega \times (0, T)$ :

$$\begin{cases} \partial_t \rho - \partial_{xx} \rho^m = 0 & \text{in } Q_T, \\ \partial_x(\rho^m) \underline{n}_\Omega = g_N & \text{on } \partial\Omega \times (0, T), \\ \rho = \rho_0 & \text{on } \Omega \times \{T\}, \end{cases} \quad (5.1)$$

where the first equation can be written as (1.1a) with  $N = 1$ ,  $A(\rho) = m\rho^{m-1}$ , and  $f(\rho) = 0$ .

We set  $\mathcal{D} = (0, 1)$  and define the entropy density  $s : \overline{\mathcal{D}} \rightarrow (0, \infty)$  as follows:

$$s(\rho) := \rho \log(\rho) + (1 - \rho) \log(1 - \rho) + \log(2),$$

whence,  $s'(\rho) = \log\left(\frac{\rho}{1-\rho}\right)$ ,  $s''(\rho) = \frac{1}{\rho(1-\rho)}$ , and  $u(w) = \frac{e^w}{1+e^w}$ . For this choice of  $s(\cdot)$ , Assumptions (H1)–(H2c) are satisfied with  $\gamma = m$  and  $C_f = 0$ , provided that  $m \in (1, 2]$ ; see [7, Prop. 4.2].

**$h$ -convergence.** In order to appraise the accuracy of the proposed method, we consider problem (5.1) with  $\Omega = (0, 1)$  and  $m = 2$ , and choose the initial datum  $\rho_0$  and the Neumann boundary datum  $g_N$  so that the exact solution is given by

$$\rho(x, t) = \left[ \frac{(m-1)(x-\alpha)^2}{2m(m+1)(\beta-t)} \right]^{\frac{1}{m-1}}, \quad (5.2)$$

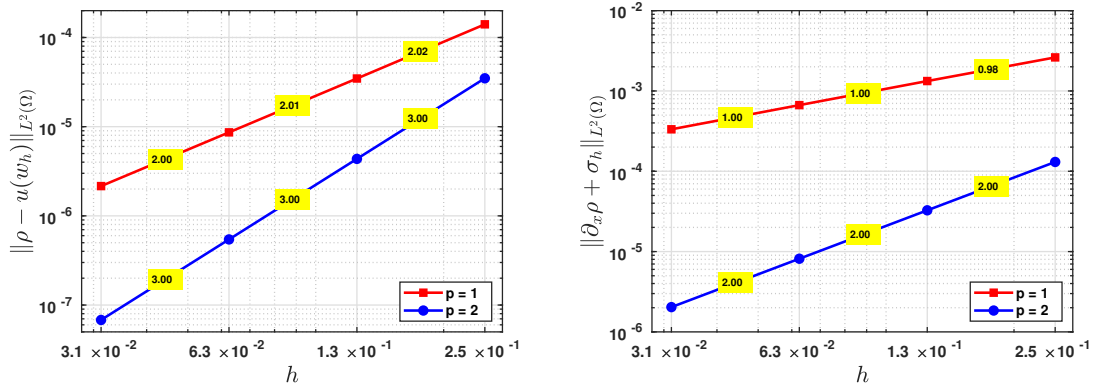
with  $\alpha = 2$  and  $\beta = 5$ ; cf. [7, §4.2].

We choose the parameters of the nonlinear solver as  $tol = 10^{-12}$  and  $s_{\max} = 50$ . We consider a set of meshes with uniformly distributed points for the spatial domain  $\Omega$ , and choose  $\tau = \mathcal{O}(h^{p+1})$  to balance the expected convergence rates in space with the first-order accuracy of the backward Euler time stepping scheme. Moreover, we set the regularization parameter to  $\varepsilon = 0$ .

In Figure 5, we show (in  $\log$ - $\log$  scale) the following errors obtained at the final time  $T = 1$ :

$$\|\rho - u(w_h)\|_{L^2(\Omega)} \quad \text{and} \quad \|\partial_x \rho + \sigma_h\|_{L^2(\Omega)}. \quad (5.3)$$

We observe, as expected, the convergence rates of order  $\mathcal{O}(h^{p+1})$  and  $\mathcal{O}(h^p)$ , respectively.



**Figure 5:**  $h$ -convergence of the errors in (5.3) at the final time  $T = 1$  for the porous medium equation with exact solution  $\rho$  in (5.2). The numbers in the yellow rectangles denote the experimental rates of convergence.

**Entropy stability.** We now consider problem (5.1) with  $\Omega = (-\pi/4, 5\pi/4)$ ,  $m = 2$ , homogeneous Neumann boundary conditions, and initial datum given by

$$\rho_0(x) = \begin{cases} \sin^{2/(m-1)}(\pi x) & \text{if } 0 \leq x \leq \pi, \\ 0 & \text{otherwise,} \end{cases} \quad (5.4)$$

whose exact solution keeps the support  $[0, \pi]$  of the initial condition until the waiting time  $t^* = (m-1)/(2m(m+1))$ ; see [2].

We choose the parameters for the nonlinear solver as  $tol = 10^{-6}$  and  $s_{\max} = 100$ , and consider  $T = 0.2$  as the final time. Moreover, we set the regularization parameter as  $\varepsilon = 10^{-6}$  and the bilinear form  $c_h(\cdot, \cdot)$  as in (4.20). In Figure 6 (first panel), we show the discrete approximation obtained for  $p = 5$ , a spatial mesh with uniformly distributed points and mesh size  $h \approx 0.04$ , and a fixed time step  $\tau = 10^{-3}$ . To represent the discrete solution, we have used linear interpolation in time, which preserves the uniform boundedness of the discrete approximation. In Figure 6 (second panel), we show the value of the discrete approximation at  $x = 0$ , where the expected behavior until  $t = t^*$  is observed; cf. [7, §4.2]. Since  $C_f = 0$ , we expect a (not necessarily monotonous) decreasing behavior of the discrete entropy values  $\{\mathcal{E}_n\}_{n=0}^{N_t}$ , where

$$\mathcal{E}_0 := \int_{\Omega} s(\rho_0) dx \quad \text{and} \quad \mathcal{E}_n := \int_{\Omega} s(u(w_h^{\varepsilon, n+1}(x))) dx \quad \text{for } n = 1, \dots, N_t. \quad (5.5)$$

Such an expected behaviour is numerically observed in Figure 6 (third panel).

Moreover, we define the discrete mass values  $\{\mathcal{M}_n\}_{n=0}^{N_t}$  as

$$\mathcal{M}_0 := \int_{\Omega} \rho_0 dx \quad \text{and} \quad \mathcal{M}_n := \int_{\Omega} u(w^{\varepsilon, n}) dx \quad \text{for } n = 1, \dots, N_t. \quad (5.6)$$

Since  $f(\rho) = 0$ , mass is conserved for analytical solutions. Standard arguments can be used to show that, for any solution  $\{w^{\varepsilon, n+1}\}_{n=0}^{N_t-1}$  to the fully discrete scheme (2.20)–(2.21), for  $n = 0, \dots, N_t - 1$ ,

$$\begin{aligned} |\mathcal{M}_{n+1} - \mathcal{M}_0| &\leq \varepsilon \sum_{m=1}^n \tau_{m+1} \int_{\Omega} |w^{\varepsilon, m+1}| dx \leq \varepsilon |\Omega|^{\frac{1}{2}} \sum_{m=1}^{N_t-1} \tau_{m+1} \|w^{\varepsilon, m+1}\|_{L^2(\Omega)} \\ &\leq \sqrt{\varepsilon} |\Omega|^{\frac{1}{2}} \left( \sum_{m=1}^{N_t-1} \tau_{m+1} \right)^{\frac{1}{2}} \left( \sum_{m=1}^{N_t-1} \varepsilon \tau_{m+1} \|w^{\varepsilon, m+1}\|_{L^2(\Omega)}^2 \right)^{\frac{1}{2}} \leq \sqrt{\varepsilon} |Q_T|^{\frac{1}{2}} \left( \int_{\Omega} s(\rho_0) dx \right)^{\frac{1}{2}}. \end{aligned}$$

In Figure 6 (fourth panel), we show (in *semilogy* scale) the error evolution of the mass values for different regularization parameters  $\varepsilon$ , where a mass loss of order  $\mathcal{O}(\varepsilon)$  is numerically observed.

## 5.2 Two-dimensional SKT model

We consider the two-dimensional Shigesada-Kawasaki-Teramoto (SKT) population system [22] with  $N = 2$  species. The diffusion matrix and Lotka–Volterra reaction terms (see (1.1)) read as

$$A_{ij}(\boldsymbol{\rho}) = \delta_{ij} \left( a_{i0} + \sum_{k=1}^2 a_{ik} \rho_k \right) + a_{ij} \rho_i, \quad i, j = 1, 2, \quad (5.7a)$$

$$\mathbf{f}_i(\boldsymbol{\rho}) = \rho_i \left( b_{i0} - \sum_{j=1}^2 b_{ij} \rho_j \right), \quad i = 1, 2, \quad (5.7b)$$

for some coefficients  $\{a_{ij}\}$  and  $\{b_{ij}\}$  satisfying  $a_{ii} > 0$ ,  $b_{ii} \geq 0$  for  $i = 1, 2$ , and  $a_{ij} \geq 0$ ,  $b_{ij} \geq 0$  for  $i \neq j$ .

We set  $\mathcal{D} = (0, \infty)^2$  and define the entropy density  $s : (0, \infty)^2 \rightarrow (0, \infty)$  as (see [18, Eq. (6)])

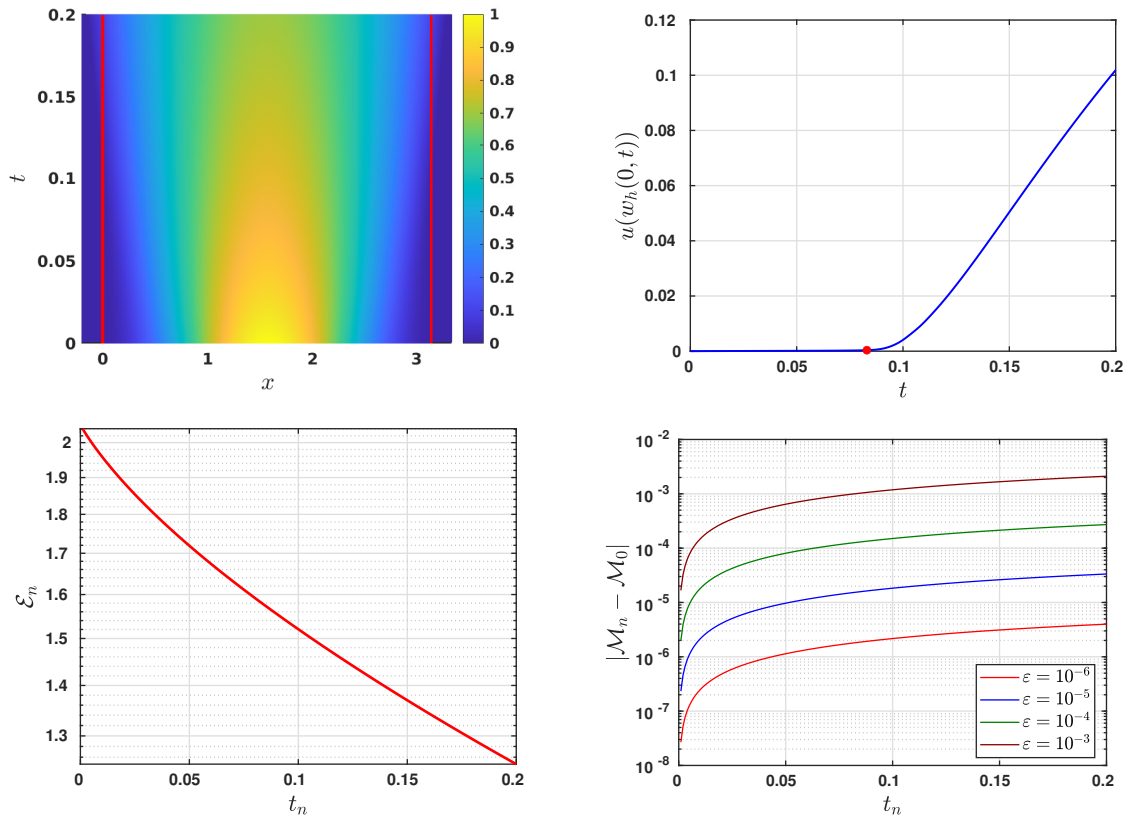
$$s(\boldsymbol{\rho}) := \sum_{i=1}^2 \pi_i (\rho_i (\log(\rho_i) - 1) + 1), \quad (5.8)$$

where  $\pi_1 = a_{21}$  and  $\pi_2 = a_{12}$ . Then  $s'(\boldsymbol{\rho}) = (\pi_1 \log \rho_1, \pi_2 \log \rho_2)$ ,  $s''(\boldsymbol{\rho}) = \text{diag}(\pi_i / \rho_i)$ , and  $\mathbf{u}(\mathbf{w}) = (\exp(w_1 / \pi_1), \exp(w_2 / \pi_2))$ .

Assumption (H2a) is satisfied with  $\gamma = \min_{i=1,2} \pi_i a_{ii} > 0$ ; see [18, §3.1]. Moreover, if the coefficients  $\{b_{ij}\}$  are all equal to zero, then Assumption (H2b) is trivially satisfied. For general coefficients  $\{b_{ij}\}$ , the reaction term satisfies the bound

$$\mathbf{f}(\boldsymbol{\rho}) \cdot s'(\boldsymbol{\rho}) \leq C_f (1 + s(\boldsymbol{\rho})) \quad \forall \boldsymbol{\rho} \in \mathcal{D}, \quad \text{with } C_f = \frac{2}{\log(2)} \max_{i=1,2} \left( b_{i0} + \frac{1}{e\pi_i} \sum_{j=1}^2 \pi_j b_{ji} \right),$$

which substitutes Assumption (H2b) in our theoretical results, by requiring that  $\tau < 1/C_f$ . Notice that the domain  $\mathcal{D} = (0, \infty)^2$  is not bounded, as required in Assumption (H1). As a consequence, we are not able to prove upper bounds for  $\rho_i$  but only the nonnegativity of  $\rho_i$ ; see [18].



**Figure 6:** Results obtained for the porous medium equation with initial condition (5.4). **First panel:** discrete approximation  $u(w_h^\varepsilon(\mathbf{x}, t))$  and support of the initial condition (red lines). **Second panel:** evolution of the value of the discrete approximation at the extreme  $x = 0$ . The theoretical waiting time has been highlighted with a red dot. **Third panel:** evolution of the entropy values in (5.5) in *semilogy* scale. **Fourth panel:** Error evolution of the mass values in (5.6) in *semilogy* scale for different values of the regularization parameter  $\varepsilon$ .

**$h$ -convergence.** We consider the SKT system with  $\Omega = (0, 1)^2$ , vanishing Lotka–Volterra terms, and the diffusion parameters (cf. [24, Example 5.1])

$$a_{i0} = 0 \quad \text{for } i = 1, 2, \quad a_{ij} = 1 \quad \text{for } i, j = 1, 2.$$

We choose the initial datum  $\rho_0$  and add a source term so that the exact solution is given by

$$\rho_1(x, y, t) = 0.25 \cos(2\pi x) \cos(\pi y) \exp(-t) + 0.5, \quad \rho_2(x, y, t) = 0.25 \cos(\pi x) \cos(2\pi y) \exp(-t) + 0.5. \quad (5.9)$$

We choose the parameters of the nonlinear solver as  $\text{tol} = 10^{-6}$  and  $s_{\max} = 50$ . We consider a set of structured simplicial meshes for the spatial domain  $\Omega$ , choose a fixed time step  $\tau = \mathcal{O}(h^{p+1})$  as in Section 5.1, and set the regularization parameter equal to  $\varepsilon = 0$ .

In Figure 7, we show (in *log-log* scale) the following errors obtained at the final time  $T = 0.5$ :

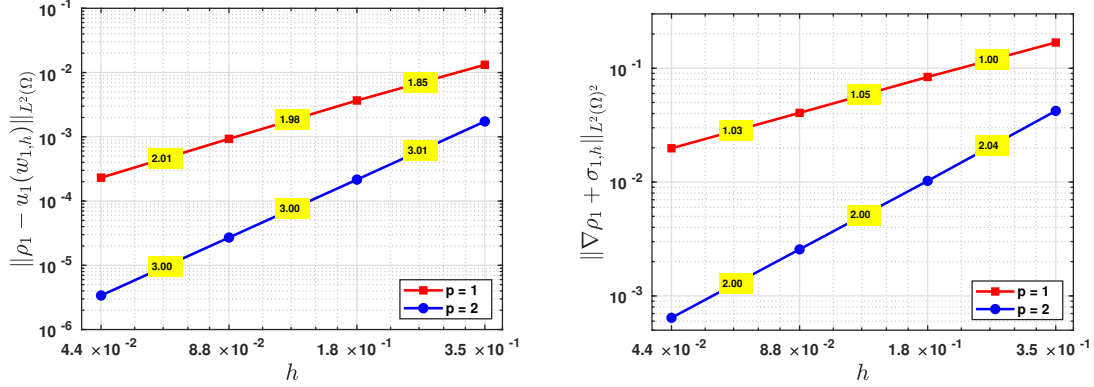
$$\|\rho_1 - u_1(w_{1,h})\|_{L^2(\Omega)} \quad \text{and} \quad \|\nabla \rho_1 + \sigma_{1,h}\|_{L^2(\Omega)^2}, \quad (5.10)$$

where convergence rates of order  $\mathcal{O}(h^{p+1})$  and  $\mathcal{O}(h^p)$  are observed, respectively. Similar results were obtained for the approximation of  $\rho_2$ , so they are omitted.

**Turing pattern.** We now consider a test from [18, §7.3]. More precisely, we choose  $\Omega = (0, 1)^2$ , and the coefficients for the diffusion matrix in (5.7a) and the reaction term in (5.7b) as follows:

$$\begin{aligned} a_{10} &= 0.05, & a_{11} &= 2.5 \times 10^{-5}, & a_{12} &= 1.025, & a_{20} &= 0.05, & a_{21} &= 0.075, & a_{22} &= 2.5 \times 10^{-5}, \\ b_{10} &= 59.7, & b_{11} &= 24.875, & b_{12} &= 19.9, & b_{20} &= 49.75, & b_{21} &= 19.9, & b_{22} &= 19.9. \end{aligned} \quad (5.11)$$





**Figure 7:**  $h$ -convergence of the errors in (5.10) at the final time  $T = 0.5$  for the SKT system with exact solution  $\rho$  in (5.9).

The initial datum is chosen as a perturbation of the equilibrium  $\rho^* = (2, 0.5)$ :

$$\rho_1(x, y, 0) = 2 + 0.31g(x - 0.25, y - 0.25) + 0.31g(x - 0.75, y - 0.75), \quad \rho_2(x, y, 0) = 0.5, \quad (5.12)$$

where  $g(x, y) = \max\{1 - 8^2x^2 - 8y^2, 0\}$ .

We select the parameters of the nonlinear solver as  $tol = 10^{-6}$  and  $s_{\max} = 50$ . We consider a rather coarse mesh with  $h \approx 1.41 \times 10^{-1}$  and use high-order approximations of degree  $p = 3$ . As for the time step, we use the adaptive strategy proposed in [18, §7.1], i.e., at the  $n$ th time step, if the desired tolerance has not been reached after 50 iterations, the time step  $\tau_{n+1}$  is reduced by a factor of 0.2 and the nonlinear solver is restarted, whereas, at the beginning of each time step, we increase the previous one by a factor of 1.1. The initial time step is set as  $\tau_1 = 10^{-4}$ . As in the previous experiment, we set the regularization parameter as  $\varepsilon = 0$ .

As discussed in [18, §7.3], due to the cross-diffusion, the equilibrium  $\rho^*$  is unstable for the SKT system (see [25, Thm. 3.1]), and the choice of the parameters  $\{b_{ij}\}$  leads to the coexistence of the two species (see [21, §6.2]). In Figure 8, we show the evolution of the approximations obtained for the densities  $\rho_1$  and  $\rho_2$  at times  $t = 0.5$  and  $t = 10$ , which exhibits the same Turing pattern formation obtained in [18, Fig. 1].

## 6 Conclusions

We designed and analyzed a structure-preserving backward Euler-LDG scheme for nonlinear cross-diffusion systems, which provides approximate solutions that respect the entropy structure of the system, and the positivity or boundedness of the physical unknown in a strong (pointwise) sense. The existence of discrete solutions and the asymptotic convergence to continuous weak solutions have been proven under some assumptions on the regularizing term and the discrete spaces, whose validity for different cases is verified. Moreover, high-order convergence rates are numerically observed for some  $L^2(\Omega)$ -errors at the final time.

## A Appendix

We present two examples of cross-diffusion systems satisfying Assumption (H2a) with  $s''A \in \mathcal{C}^0(\bar{\mathcal{D}}; \mathbb{R}^{N \times N})$ .

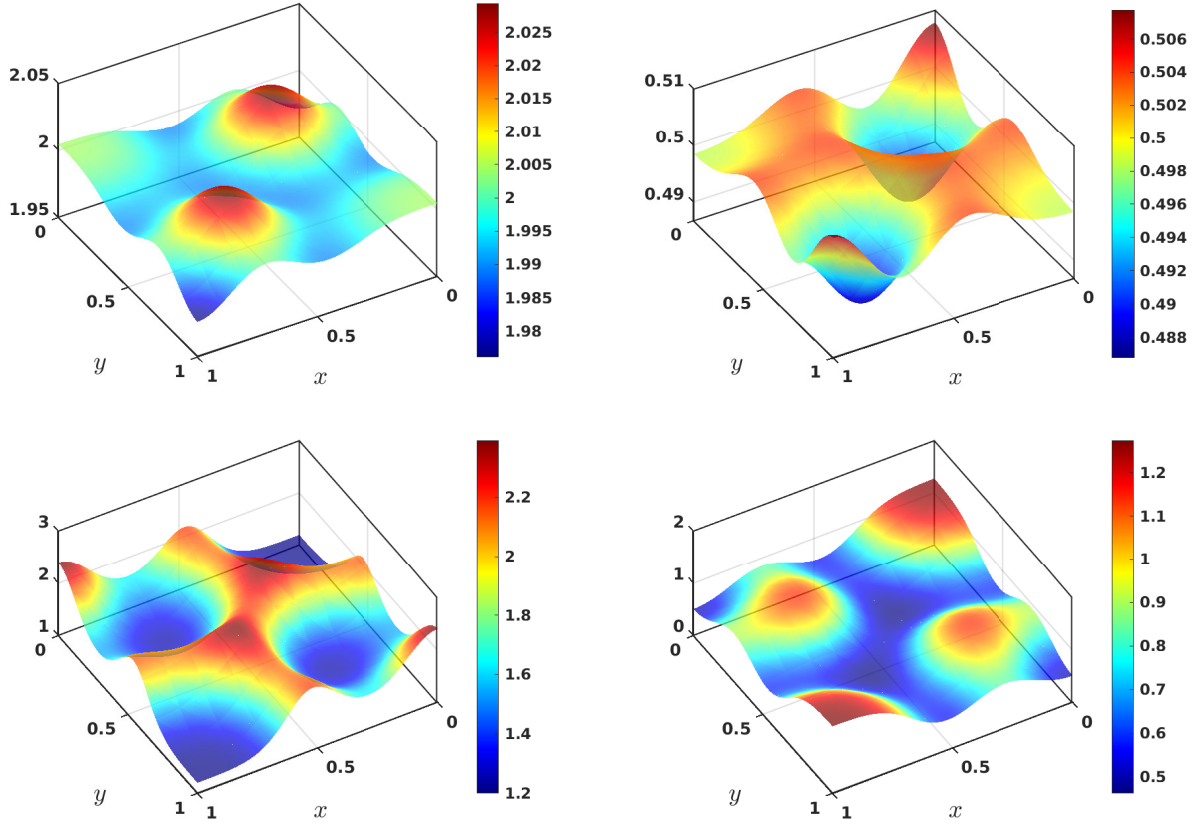
### A.1 Volume-filling diffusion model for fluid mixtures

We define the diffusion matrix  $A(\rho) = (A_{ij}(\rho))_{i,j=1}^N$  with  $N \in \mathbb{N}$  by

$$A_{ii}(\rho) = p_i \rho_i (1 - \rho_i), \quad A_{ji}(\rho) = -p_i \rho_i \rho_j \quad \text{for } i, j = 1, \dots, N,$$

where  $p_i > 0$  are pressure coefficients, and the entropy functional  $s(\rho)$  is defined by

$$s(\rho) = \sum_{i=1}^N \rho_i (\log \rho_i - 1) + \rho_0 (\log \rho_0 - 1) + N + 1, \quad \text{where } \rho_0 := 1 - \sum_{i=1}^N \rho_i,$$



**Figure 8:** Evolution of the approximations of the densities  $\rho_1$  (left panels) and  $\rho_2$  (right panels) for the SKT system with parameters (5.11) and initial datum (5.12) at times  $t = 0.5$  (first row) and  $t = 10$  (second row).

and  $\boldsymbol{\rho} \in \mathcal{D} := \{\boldsymbol{\rho} \in (0, 1)^N : \sum_{i=1}^N \rho_i < 1\}$ . The cross-diffusion system with this diffusion matrix can be derived from a multi-phase viscous fluid model in the diffusion limit similarly as in [15, §4.2], assuming that the partial pressures of the mixture are linear. The fluid mixture consists of  $N + 1$  components with the volume fractions  $\rho_0, \dots, \rho_N$ , which sum up to one.

The entropy  $s : \mathcal{D} \rightarrow (0, \infty)$  is convex and  $s'$  is invertible on  $\mathcal{D}$ . The Hessian  $s''(\boldsymbol{\rho}) = (H_{ij})_{i,j=1}^N$  of the entropy has the entries  $H_{ij} = \delta_{ij}/\rho_i + 1/\rho_0$  and therefore, for  $i \neq j$ ,

$$\begin{aligned} (A(\boldsymbol{\rho})^T s''(\boldsymbol{\rho}))_{ii} &= A_{ii}(\boldsymbol{\rho})H_{ii} + \sum_{k \neq i} A_{ik}(\boldsymbol{\rho})H_{ki} = p_i \rho_i (1 - \rho_i) \left( \frac{1}{\rho_i} + \frac{1}{\rho_0} \right) - p_i \rho_i \sum_{k \neq i} \frac{\rho_k}{\rho_0} = p_i, \\ (A(\boldsymbol{\rho})^T s''(\boldsymbol{\rho}))_{ij} &= \sum_{k=1}^N A_{ik}(\boldsymbol{\rho}) \frac{\delta_{jk}}{\rho_j} + \sum_{k=1}^N A_{ik}(\boldsymbol{\rho}) \frac{1}{\rho_0} = -p_i \rho_i + \frac{p_i}{\rho_0} \rho_i \left( 1 - \rho_i - \sum_{k \neq i} \rho_k \right) = 0. \end{aligned}$$

Thus,  $s''A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$  holds and Assumption (H2a) is satisfied with  $\gamma = \min\{p_1, \dots, p_N\} > 0$ .

## A.2 Tumor-growth model

The growth of an avascular tumor can be described by a cross-diffusion system with the diffusion matrix

$$A(\boldsymbol{\rho}) = \begin{pmatrix} 2r\rho_0(1 - \rho_1) - \beta\theta\rho_1\rho_2^2 & -2\beta\rho_1\rho_2(1 + \theta\rho_1) \\ -2\rho_1\rho_2 + \beta\theta(1 - \rho_2)\rho_2^2 & 2\beta\rho_2(1 - \rho_2)(1 + \theta\rho_1) \end{pmatrix},$$

where  $\rho_0$ ,  $\rho_1$ , and  $\rho_2$  denote the volume fractions of the interstitial fluid (water, nutrients), tumor cells, and extracellular matrix, respectively. The parameters  $\beta > 0$  and  $\theta > 0$  appear in the partial pressures for the extracellular fluid and tumor cells, respectively. We refer to [15, §4.2] for details about the modeling. We choose

the same entropy and domain  $\mathcal{D}$  as in the previous subsection. A straightforward computation shows that

$$A(\boldsymbol{\rho})^T s''(\boldsymbol{\rho}) = \begin{pmatrix} 2 & \beta\theta\rho_2 \\ 0 & 2\beta(\theta\rho_1 + 1) \end{pmatrix}$$

is positive definite if  $\theta < 4/\sqrt{\beta}$ . More precisely,  $s''A \in \mathcal{C}^0(\overline{\mathcal{D}}; \mathbb{R}^{N \times N})$  and there exists  $\gamma > 0$  such that Assumption (H2a) is satisfied. The constant  $\gamma$  vanishes if  $\theta = 4/\sqrt{\beta}$ , so the strict inequality is needed.

## Funding

The first author is member of the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM) and acknowledges the kind hospitality of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI), where part of this research was developed, and support from the Italian Ministry of University and Research through the project PRIN2020 “Advanced polyhedral discretizations of heterogeneous PDEs for multiphysics problems”. This research was funded in part by the Austrian Science Fund (FWF) projects 10.55776/F65 (AJ, IP), 10.55776/P33010 (AJ), and 10.55776/P33477 (IP). This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme, ERC Advanced Grant NEUROMORPH, no. 101018153.

## References

- [1] B. Andreianov, M. Bendahmane, and R. Ruiz-Baier. Analysis of a finite volume method for a cross-diffusion model in population dynamics. *Math. Models Methods Appl. Sci.*, 21(2):307–344, 2011.
- [2] D. G. Aronson. Regularity properties of flows through porous media: A counterexample. *SIAM J. Appl. Math.*, 19:299–307, 1970.
- [3] G. R. Barrenechea, E. H. Georgoulis, T. Pryer, and A. Veerer. A nodally bound-preserving finite element method. *IMA J. Numer. Anal.*, page drad055, 08 2023.
- [4] G. R. Barrenechea, V. John, and P. Knobloch. Finite element methods respecting the discrete maximum principle for convection-diffusion equations. *SIAM Rev.*, 66(1):3–88, 2024.
- [5] A. Bonito, D. Guignard, R. H. Nochetto, and S. Yang. Numerical analysis of the LDG method for large deformations of prestrained plates. *IMA J. Numer. Anal.*, 43(2):627–662, 2023.
- [6] F. Bonizzoni, M. Braukhoff, A. Jüngel, and I. Perugia. A structure-preserving discontinuous Galerkin scheme for the Fisher-KPP equation. *Numer. Math.*, 146(1):119–157, 2020.
- [7] M. Braukhoff, I. Perugia, and P. Stocker. An entropy structure preserving space-time formulation for cross-diffusion systems: analysis and Galerkin discretization. *SIAM J. Numer. Anal.*, 60(1):364–395, 2022.
- [8] A. Buffa and C. Ortner. Compact embeddings of broken Sobolev spaces and applications. *IMA J. Numer. Anal.*, 29(4):827–855, 2009.
- [9] C. Cancès and B. Gaudeul. A convergent entropy diminishing finite volume scheme for a cross-diffusion system. *SIAM J. Numer. Anal.*, 58(5):2684–2710, 2020.
- [10] P. Castillo. Stencil reduction algorithms for the local discontinuous Galerkin method. *Int. J. Numer. Methods Eng.*, 81(12):1475–1491, 2010.
- [11] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM J. Numer. Anal.*, 35(6):2440–2463, 1998.
- [12] M. Corti, F. Bonizzoni, and P.-F. Antonietti. Structure preserving polytopal discontinuous Galerkin methods for the numerical modeling of neurodegenerative diseases. *J. Sci. Comput.*, 100:Paper No. 39, 2024.
- [13] C. Dawson, V. Aizinger, and B. Cockburn. The local discontinuous Galerkin method for contaminant transport problems. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 309–314. Springer, Berlin, 2000.

- [14] A. Jüngel. The boundedness-by-entropy method for cross-diffusion systems. *Nonlinearity*, 28(6):1963–2001, 2015.
- [15] A. Jüngel. *Entropy methods for diffusive partial differential equations*. SpringerBriefs in Mathematics. Springer, 2016.
- [16] A. Jüngel and J.-P. Milišić. Entropy dissipative one-leg multistep time approximations of nonlinear diffusive equations. *Numer. Methods Partial Differential Equations*, 31(4):1119–1149, 2015.
- [17] A. Jüngel and M. Vetter. A Convergent Entropy-Dissipating BDF2 Finite-Volume Scheme for a Population Cross-Diffusion System. *Comput. Methods Appl. Math.*, 24(3):725–746, 2024.
- [18] A. Jüngel and A. Zurek. A convergent structure-preserving finite-volume scheme for the Shigesada–Kawasaki–Teramoto population system. *SIAM J. Num. Anal.*, 59(4):2286–2309, 2021.
- [19] S. Lemaire and J. Moatti. Structure preservation in high-order hybrid discretisations of potential-driven advection-diffusion: linear and nonlinear approaches. *Math. Eng.*, 6(1):100–136, 2024.
- [20] I. Perugia and D. Schötzau. An  $hp$ -analysis of the local discontinuous Galerkin method for diffusion problems. *J. Sci. Comput.*, 17(1-4):561–571, 2002.
- [21] N. Shigesada and K. Kawasaki. *Biological Invasions: Theory and Practice*. Oxford University Press, UK, 1997.
- [22] N. Shigesada, K. Kawasaki, and E. Teramoto. Spatial segregation of interacting species. *J. Theoret. Biol.*, 79(1):83–99, 1979.
- [23] D. R. Smart. *Fixed point theorems*, volume No. 66 of *Cambridge Tracts in Mathematics*. Cambridge University Press, London-New York, 1974.
- [24] Z. Sun, J. A. Carrillo, and C.-W. Shu. An entropy stable high-order discontinuous Galerkin method for cross-diffusion gradient flow systems. *Kinet. Relat. Models*, 12(4):885–908, 2019.
- [25] C. Tian, Z. Lin, and M. Pedersen. Instability induced by cross-diffusion in reaction-diffusion systems. *Nonlinear Anal., Real World Appl.*, 11(2):1036–1045, 2010.