



# Finite element approximation of time-dependent mean field games with nondifferentiable Hamiltonians

Yohance A. P. Osborne<sup>1</sup> · Iain Smears<sup>2</sup>

Received: 5 February 2024 / Revised: 7 November 2024 / Accepted: 27 November 2024 /

Published online: 11 December 2024

© The Author(s) 2024

## Abstract

The standard formulation of the PDE system of Mean Field Games (MFG) requires the differentiability of the Hamiltonian. However in many cases, the structure of the underlying optimal problem leads to a convex but nondifferentiable Hamiltonian. For time-dependent MFG systems, we introduce a generalization of the problem as a Partial Differential Inclusion (PDI) by interpreting the derivative of the Hamiltonian in terms of the subdifferential set. In particular, we prove the existence and uniqueness of weak solutions to the resulting MFG PDI system under standard assumptions in the literature. We propose a monotone stabilized finite element discretization of the problem, using conforming affine elements in space and an implicit Euler discretization in time with mass-lumping. We prove the strong convergence in  $L^2(H^1)$  of the value function approximations, and strong convergence in  $L^p(L^2)$  of the density function approximations, together with strong  $L^2$ -convergence of the value function approximations at the initial time.

**Mathematics Subject Classification** 65M12 · 65M60

## 1 Introduction

Mean Field Games (MFG), which were introduced by Lasry and Lions [26–28] and independently by Huang, Caines and Malhamé [23], describe the Nash equilibria of large numbers of players involved in a game of stochastic optimal control. This leads to a system of a Hamilton–Jacobi–Bellman (HJB) equation coupled with a Kolmogorov–Fokker–Planck (KFP) equation, where the unknowns are the value function of the

---

✉ Yohance A. P. Osborne  
yohance.a.osborne@durham.ac.uk

Iain Smears  
i.smears@ucl.ac.uk

<sup>1</sup> Department of Mathematical Sciences, Durham University, Stockton Road, Durham DH1 3LE, UK

<sup>2</sup> Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK

underlying stochastic optimal control problem, and the density function of the player distribution within the state space of the game. For extensive reviews on the theory and applications of MFG problems, we refer the reader to [3, 20–22].

We consider here the model problem

$$\begin{aligned} -\partial_t u - v\Delta u + H(t, x, \nabla u) &= F[m](t, x) && \text{in } (0, T) \times \Omega, \\ \partial_t m - v\Delta m - \operatorname{div} \left( m \frac{\partial H}{\partial p}(t, x, \nabla u) \right) &= G(t, x) && \text{in } (0, T) \times \Omega, \\ m(0, x) = m_0(x), \quad u(T, x) &= S[m(T, \cdot)](x) && \text{on } \Omega, \\ m = 0, \quad u = 0 &&& \text{on } (0, T) \times \partial\Omega, \end{aligned} \tag{1.1}$$

where  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , denotes a bounded polyhedral open connected set with Lipschitz boundary  $\partial\Omega$ ,  $T > 0$  is the time horizon, and  $v > 0$  is constant. The assumptions on the coupling  $F$ , the terminal cost  $S$ , the source term  $G$ , and the initial density  $m_0$  are specified in Sect. 2 below. We note from the onset that we allow for wide classes of both local and nonlocal operators  $F$  and  $S$ . The Dirichlet boundary conditions and the source term  $G$  arise in models where players may enter or exit the game. The Hamiltonian  $H$  in (1.1) is determined by the underlying optimal control problem, and is defined by

$$H(t, x, p) := \sup_{\alpha \in \mathcal{A}} (b(t, x, \alpha) \cdot p - f(t, x, \alpha)), \tag{1.2}$$

where  $\mathcal{A}$  is the control set, where  $b : [0, T] \times \overline{\Omega} \times \mathcal{A} \rightarrow \mathbb{R}^d$  is the opposite control-dependent drift, and  $f : [0, T] \times \overline{\Omega} \times \mathcal{A} \rightarrow \mathbb{R}$  is a control-dependent running-cost component. Note that  $p \mapsto H(t, x, p)$  is convex on  $\mathbb{R}^d$ . It is well-known in the optimal control literature that in many cases, the optimal controls are not necessarily unique, and are often of *bang-bang* type. In these cases the Hamiltonian  $H$  may be nondifferentiable, thus raising the question of how to interpret the KFP equation in (1.1) when  $\frac{\partial H}{\partial p}$  is not well-defined as a function. From a modelling perspective, this connects to the question as to how players of the game should choose among multiple optimal controls, and how this determines the evolution of the player density.

Whereas existing works on the analysis of MFG often assume that the Hamiltonian is differentiable or even  $C^1$ -regular with respect to the gradient, there are currently only a handful of works on the nondifferentiable case. In particular the special case of Hamiltonians of eikonal type, but also possibly dependent on  $m$ , is studied in [18]. In [32, 33], we analysed the steady-state analogue of (1.1), in the setting of Lipschitz continuous, but not necessarily differentiable Hamiltonians. We showed that the KFP equation in (1.1) can be generalized to a *Partial Differential Inclusion* (PDI) which, in the current setting, is

$$\partial_t m - v\Delta m - G \in \operatorname{div} (m \partial_p H(t, x, \nabla u)) \quad \text{in } (0, T) \times \Omega, \tag{1.3}$$

where  $\partial_p H$  denotes the Moreau–Rockafellar pointwise partial subdifferential of  $H$  w.r.t.  $p$ . From a modelling perspective, the PDI corresponds to the natural intuition that, in order to maintain a Nash equilibrium, players in the same state might be required to make distinct choices of optimal controls relative to one another. This entails that the drift term for the density is expected to be a (weighted) average of the chosen drift directions of the individual players, which can be shown to be an element of the subdifferential of the Hamiltonian, see also [35] for a heuristic derivation. In terms of the analysis, we showed in [32, 33] that, for steady-state problems, the MFG PDI system admits a notion of weak solution, with the existence of a solution guaranteed under very mild assumptions on the data, along with uniqueness of the solution guaranteed under a similar monotonicity condition on the coupling as made in [28]. Moreover, in [35] we analysed the connection between MFG PDI systems and their classical PDE counterparts via regularization of the Hamiltonians.

Our first main contribution in this work is to develop the analysis of the MFG PDI in the time-dependent setting, in extension of our work on the elliptic problem in [32, 33]. We show the existence of a weak solution to the problem, and we show that the uniqueness of the weak solution holds under standard monotonicity conditions on the problem data. Note that it is known from examples in [7] that uniqueness of the solution does not necessarily hold if the couplings are not monotone. The second main contribution of this work is the design and analysis of a monotone Finite Element Method (FEM) of the MFG PDI system. For problems with differentiable Hamiltonians, a range of numerical methods have been proposed and studied. This includes monotone finite difference methods on Cartesian grids [1, 2, 4, 10], methods based on optimization reformulations [11], as well as semi-Lagrangian scheme [14, 15]. In addition, in [34] we established the near and full quasi-optimality of a monotone stabilized FEM for stationary MFG PDE systems with  $C^{1,1}$  Hamiltonians and derived results on optimal convergence rates for the method. For problems with nondifferentiable Hamiltonians, the only work so far is [32, 33] which covers the steady-state case.

The monotone FEM that we propose here involves continuous piecewise affine finite element spaces for the spatial discretization, and an implicit Euler with mass-lumping for the temporal discretization. The discretized system uses the exact Hamiltonian and its subdifferential, which greatly helps the consistency and convergence analysis. The first-order terms of the operators are stabilized in order to satisfy a Discrete Maximum Principle (DMP) and ensure the positivity of the density approximations, and also to maintain stability in the small diffusion regime. Although there exist many choices of stabilization to achieve a DMP for FEM [8, 9, 12, 13, 17, 29, 40], we construct an original volume-based stabilization that has a simple analytical form. This helps to simplify the stability analysis in some discrete Bochner–Sobolev norms, and it is also easy to implement in practice. The analysis allows for unstructured shape-regular simplicial meshes in arbitrary dimensions, requiring only the standard Xu–Zikatanov condition [42]. This is not restrictive in practice, as it allows for many examples of meshes with nonacute elements, local refinements, etc. There is no restriction linking the time-step size to the mesh-size, and there is also no restriction on the spatial mesh-grading.

We show that the existence and uniqueness of the discrete approximations hold under the same assumptions as for their continuous counterparts, see Theorems 5.1 and 5.2 below. We aim here to show the basic convergence of the method without requiring stronger assumptions on the problem data or on the regularity of the exact solution. This is motivated by the fact that examples in [32, 33] show that the solution regularity can be quite limited, even in the interior of the domain. Our main convergence result is in Theorem 5.3, which shows the convergence of numerical approximations under the same assumptions on the data as required for the uniqueness results. In particular, we prove strong convergence of the value function approximations to  $u$  in  $L^2(0, T; H_0^1(\Omega))$  and the strong convergence of the density function approximations to  $m$  in  $L^p(0, T; L^2(\Omega))$  for any  $p \in [1, \infty)$ , along with weak convergence to  $m$  in  $L^2(0, T; H_0^1(\Omega))$ . These results do not require any additional regularity assumptions on the solution. The overall strategy of the analysis is to use the uniform boundedness of the numerical solutions in discrete Bochner–Sobolev norms, including negative-order norms for the time derivatives, to extract subsequences with suitable weak convergence properties, and with suitable strong convergence properties by discrete compact embedding results. Then we use the stability and consistency of the discretization to augment weak convergence to strong convergence in the appropriate norms. One of the key challenges that we address here is the problem of showing that the limit of the approximations satisfies the nonlinear final time coupling  $u(T, x) = S[m(T, \cdot)](x)$ , which we handle by showing strong convergence in  $L^2(\Omega)$  of the density approximations at the final time  $T$  through a duality argument and stability bounds in weighted discrete Bochner–Sobolev norms.

The paper is organized as follows. In Sect. 2 we describe the basic notation and setting of this work. The definition of the MFG PDI generalization and its notion of weak solution, along with the main theorems on existence and uniqueness, are presented in Sect. 3. Section 4 introduces the setting for discretization, including the mass-lumping and spatial stabilization. The FEM is then introduced in Sect. 5, along with the statements of the theorems on existence, uniqueness, and convergence of the discrete approximations. Section 6 gives the proof of the existence, uniqueness and stability of the numerical methods. Section 7 details the discrete functional analytic compactness tools in a self-contained exposition. The proof of convergence of the method is shown in Sect. 8. We present the results of a numerical experiment in Sect. 9. Appendices A and B give proofs of some of the supporting results, such as the proof of the DMP under the proposed stabilization.

## 2 Setting and notation

### 2.1 Basic notation

Let  $\Omega$  be a bounded, open connected subset of  $\mathbb{R}^d$  with Lipschitz boundary  $\partial\Omega$ . We use the following notation for  $L^2$ -norms and inner products: for a Lebesgue measurable set  $\omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , let  $\|\cdot\|_\omega$  and  $(\cdot, \cdot)_\omega$  denote respectively the standard  $L^2$ -norm and  $L^2$ -inner product for scalar- and vector-valued functions on  $\omega$ .

Given  $1 \leq p, q < \infty$ , let  $L^p(0, T; L^q(\Omega))$  denote the space of mappings  $w : (0, T) \rightarrow L^q(\Omega)$  that are strongly measurable (see [43, Ch. VI]) with  $\int_0^T \|w\|_{L^q(\Omega)}^p dt < \infty$ . In particular, we denote the norm on  $L^2(0, T; L^2(\Omega))$  by  $\|\cdot\|_{Q_T}$  where  $Q_T := (0, T) \times \Omega$ . When using notation for  $L^2$ -norms and the  $L^2$ -inner product, the dependence of functions on the state variable  $x \in \Omega$  will be suppressed. It is known that the spaces  $L^p(Q_T)$  and  $L^p(0, T; L^p(\Omega))$  are isometrically isomorphic to each other for every  $1 \leq p < \infty$  (see [19, Ex. 5.0.32]). As such, when  $1 \leq p < \infty$ , we identify the spaces  $L^p(Q_T)$  and  $L^p(0, T; L^p(\Omega))$ . Note that  $L^\infty(Q_T)$  is not isomorphic to  $L^\infty(0, T; L^\infty(\Omega))$ .

Let  $\langle \cdot, \cdot \rangle$  denote the duality pairing between  $H_0^1(\Omega)$  and its dual space  $H^{-1}(\Omega)$ . Let the norm  $\|\cdot\|_{H^{-1}(\Omega)}$  be defined by

$$\|w\|_{H^{-1}(\Omega)} := \sup_{v \in H_0^1(\Omega) \setminus \{0\}} \frac{\langle w, v \rangle}{\|\nabla v\|_\Omega} \quad \forall w \in H^{-1}(\Omega). \quad (2.1)$$

Let  $X := L^2(0, T; H_0^1(\Omega))$  and  $Y := L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ , with respective norms

$$\|v\|_X^2 := \int_0^T \|\nabla v\|_\Omega^2 dt, \quad \|w\|_Y^2 := \int_0^T \|\partial_t w\|_{H^{-1}(\Omega)}^2 + \|\nabla w\|_\Omega^2 dt, \quad (2.2)$$

for all  $v \in X$  and all  $w \in Y$ . Let  $Y_0 \subset Y$  denote the subspace of functions that vanish at  $t = 0$ .

Let the source term  $G \in L^2(0, T; H^{-1}(\Omega))$  and the initial distribution  $m_0 \in L^2(\Omega)$ . We will say that  $G$  is nonnegative in the sense of distributions if  $\int_0^T \langle G, v \rangle dt \geq 0$  whenever  $v \in L^2(0, T; H_0^1(\Omega))$  is such that  $v \geq 0$  a.e. in  $Q_T$ .

Next, let the coupling term  $F : L^2(0, T; L^2(\Omega)) \rightarrow L^2(0, T; H^{-1}(\Omega))$  be a continuous operator. Moreover, we assume that there exists a constant  $C_F > 0$  such that

$$\|F[v]\|_{L^2(0, T; H^{-1}(\Omega))} \leq C_F (\|v\|_{Q_T} + 1) \quad \forall v \in L^2(0, T; L^2(\Omega)). \quad (2.3)$$

We say that  $F$  is *strictly monotone* on  $L^2(0, T; H_0^1(\Omega))$  if

$$\int_0^T \langle F[w] - F[v], w - v \rangle dt \leq 0 \implies w = v \quad \text{in } L^2(0, T; H_0^1(\Omega))$$

whenever  $w, v \in L^2(0, T; H_0^1(\Omega))$ . Note that although  $F$  is defined on  $L^2(0, T; L^2(\Omega))$ , strict monotonicity will only be used later on the smaller space  $L^2(0, T; H_0^1(\Omega))$ . Finally, we suppose that the terminal cost operator  $S : L^2(\Omega) \rightarrow L^2(\Omega)$  is continuous and that there exists a constant  $C_S \geq 0$  such that

$$\|S[v]\|_\Omega \leq C_S (\|v\|_\Omega + 1) \quad \forall v \in L^2(\Omega). \quad (2.4)$$

We say that  $S$  is *monotone* on  $L^2(\Omega)$  if

$$(S[w] - S[v], w - v)_\Omega \geq 0 \quad \forall w, v \in L^2(\Omega).$$

**Example 2.1** The above hypotheses allow for a broad class of coupling operators  $F$ . This class includes for example local operators  $F : L^2(0, T; L^2(\Omega)) \rightarrow L^2(Q_T)$  of the form  $F[z](t, x) := f(z(t, x))$ , for a.e.  $(t, x) \in Q_T$ , where  $z \in L^2(0, T; L^2(\Omega))$  is given and the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  is strictly monotone and Lipschitz continuous. This class also includes nonlocal smoothing operators of the form  $F := (\partial_t - \nu \Delta)^{-1} : L^2(0, T; L^2(\Omega)) \rightarrow Y$  where, for each  $z \in L^2(0, T; L^2(\Omega))$ ,  $F[z]$  is defined as the unique function in  $Y$  satisfying  $\int_0^T \langle \partial_t F[z], \phi \rangle + \nu (\nabla F[z], \nabla \phi)_\Omega dt = \int_0^T (z, \phi)_\Omega dt$  for all  $\phi \in X$ . Moreover, this class admits examples of differential-type operators, e.g. operators  $F : L^2(0, T; L^2(\Omega)) \rightarrow L^2(0, T; H^{-1}(\Omega))$  which take the form  $\langle F[z](t), \phi \rangle := -(z \mathbf{v} \cdot \nabla \phi)_\Omega$  for  $z \in L^2(0, T; L^2(\Omega))$ ,  $\phi \in H_0^1(\Omega)$  and for a.e.  $t \in (0, T)$ , where  $\mathbf{v} \in C^1(\overline{\Omega}; \mathbb{R}^d)$  is some vector field that satisfies  $\operatorname{div}(\mathbf{v}) > 0$  in  $\Omega$ . In this case, for all  $z_1, z_2 \in X$ , we have  $\int_0^T \langle F[z_1](t) - F[z_2](t), z_1 - z_2 \rangle dt = \frac{1}{2} \int_0^T (\operatorname{div}(\mathbf{v}), (z_1 - z_2)^2)_\Omega dt$  so  $F$  is strictly monotone on  $L^2(0, T; H_0^1(\Omega))$ .

## 2.2 Subdifferential of the Hamiltonian

Recall that the Hamiltonian  $H$  is defined in (1.2). We assume that  $\mathcal{A}$  is a compact metric space, and that  $b : [0, T] \times \overline{\Omega} \times \mathcal{A} \rightarrow \mathbb{R}^d$  and  $f : [0, T] \times \overline{\Omega} \times \mathcal{A} \rightarrow \mathbb{R}$  are uniformly continuous functions. It then follows that the Hamiltonian  $H$  given by (1.2) is Lipschitz continuous with respect to its third argument, i.e.

$$|H(t, x, p_1) - H(t, x, p_2)| \leq L_H |p_1 - p_2| \quad \forall p_1, p_2 \in \mathbb{R}^d, \quad \forall (t, x) \in (0, T) \times \Omega, \quad (2.5)$$

where  $L_H := \|b\|_{C([0, T] \times \overline{\Omega} \times \mathcal{A}; \mathbb{R}^d)}$ .

Let the set-valued map  $\partial_p H : Q_T \times \mathbb{R}^d \rightrightarrows \mathbb{R}^d$  denote the pointwise Moreau-Rockafellar partial subdifferential of  $H$  with respect to  $p \in \mathbb{R}^d$ , which is defined by

$$\partial_p H(t, x, p) := \left\{ z \in \mathbb{R}^d : H(t, x, q) \geq H(t, x, p) + z \cdot (q - p) \quad \forall q \in \mathbb{R}^d \right\}. \quad (2.6)$$

Note that  $\partial_p H(t, x, p)$  is nonempty for all  $x \in \Omega$  and  $p \in \mathbb{R}^d$  because  $H$  is real-valued and convex in  $p$  for each fixed  $(t, x) \in Q_T$ .

Let  $w \in L^1(0, T; W^{1,1}(\Omega))$  be given. Following [32, 33], we give notation for a collection of measurable selections of  $\partial_p H(\cdot, \cdot, \nabla w)$ . We say that a vector field  $\tilde{b} \in L^\infty(Q_T; \mathbb{R}^d)$  is a *measurable selection* of  $\partial_p H(\cdot, \cdot, \nabla w)$  if  $\tilde{b}(t, x) \in \partial_p H(t, x, \nabla w(t, x))$  for a.e.  $(t, x) \in Q_T$ .

**Definition 2.1** Let  $H$  be the function given by (1.2). We define the set-valued map  $\mathcal{D}_p H : L^1(0, T; W^{1,1}(\Omega)) \rightrightarrows L^\infty(Q_T; \mathbb{R}^d)$  via

$$\mathcal{D}_p H[v] := \left\{ \tilde{b} \in L^\infty(Q_T; \mathbb{R}^d) : \tilde{b}(t, x) \in \partial_p H(t, x, \nabla v(t, x)) \text{ for a.e. } (t, x) \in Q_T \right\}.$$

We begin by showing that  $\mathcal{D}_p H$  has nonempty images for every argument in  $L^1(0, T; W^{1,1}(\Omega))$ . Define the set-valued map  $\Lambda : Q_T \times \mathbb{R}^d \rightrightarrows \mathcal{A}$  via

$$\Lambda(t, x, p) := \operatorname{argmax}_{\alpha \in \mathcal{A}} \{b(t, x, \alpha) \cdot p - f(t, x, \alpha)\} \quad \forall (t, x, p) \in Q_T \times \mathbb{R}^d. \quad (2.7)$$

Observe that  $\Lambda(t, x, p)$  is nonempty for all  $(t, x, p) \in Q_T \times \mathbb{R}^d$  since  $\mathcal{A}$  is compact and the functions  $b$  and  $f$  are uniformly continuous on  $[0, T] \times \bar{\Omega} \times \mathcal{A}$ . For a given  $v \in L^1(0, T; W^{1,1}(\Omega))$ , let  $\Lambda[v]$  denote the set of all Lebesgue measurable functions  $\alpha^* : Q_T \rightarrow \mathcal{A}$  that satisfy  $\alpha^*(t, x) \in \Lambda(t, x, \nabla v(t, x))$  for a.e.  $(t, x) \in Q_T$ . We will refer to each element of  $\Lambda[v]$  as a *measurable selection of  $\Lambda(\cdot, \cdot, \nabla v(\cdot, \cdot))$* . It can be shown that  $\Lambda[v]$  is nonempty for each  $v \in L^1(0, T; W^{1,1}(\Omega))$ , see e.g. [39, Theorem 10], where the proof of the existence of measurable selections ultimately rests upon the Kuratowski and Ryll–Nardzewski Selection Theorem [24].

**Lemma 2.1** For every  $v \in L^1(0, T; W^{1,1}(\Omega))$ , the set  $\mathcal{D}_p H[v]$  is a nonempty subset of  $L^\infty(Q_T; \mathbb{R}^d)$  and

$$\sup_{\tilde{b} \in \mathcal{D}_p H[v]} \|\tilde{b}\|_{L^\infty(Q_T; \mathbb{R}^d)} \leq L_H. \quad (2.8)$$

**Proof** Let  $v \in L^1(0, T; W^{1,1}(\Omega))$  be given. As explained above, the set  $\Lambda[v]$  is nonempty, so after choosing some  $\alpha^* \in \Lambda[v]$ , we deduce from (2.7) that

$$H(t, x, \nabla v(t, x)) = b(t, x, \alpha^*(t, x)) \cdot \nabla v(t, x) - f(t, x, \alpha^*(t, x)), \quad (2.9)$$

for a.e.  $(t, x) \in Q_T$ . Let  $b^* : (t, x) \mapsto b(t, x, \alpha^*(t, x))$  for a.e.  $(t, x) \in Q_T$ . It then follows that  $b^*(t, x)$  is Lebesgue measurable and  $b^*(t, x) \in \partial_p H(t, x, \nabla v(t, x))$  for a.e.  $(t, x) \in Q_T$ . Furthermore, we have that  $b^* \in L^\infty(Q_T; \mathbb{R}^d)$  since  $|b^*(t, x)| \leq \|b\|_{C(\bar{Q}_T \times \mathcal{A}; \mathbb{R}^d)} = L_H$  for a.e.  $(t, x) \in Q_T$ . Hence,  $b^* \in \mathcal{D}_p H[v]$ , thus establishing the nonemptiness of  $\mathcal{D}_p H[v]$  for each  $v \in L^1(0, T; W^{1,1}(\Omega))$ . To obtain the uniform bound (2.8), note that the Lipschitz continuity of  $H$ , c.f. (2.5), readily implies that, for all  $(t, x, p) \in (0, T) \times \Omega \times \mathbb{R}^d$ , the subdifferential  $\partial_p H(t, x, p)$  is contained in a closed ball in  $\mathbb{R}^d$  centred at the origin and of radius  $L_H$ . This implies that any  $\tilde{b} \in \mathcal{D}_p H[v]$  satisfies  $|\tilde{b}(t, x)| \leq L_H$  for a.e.  $(t, x) \in Q_T$ .  $\square$

The second preliminary result describes a closure property of the set-valued map  $\mathcal{D}_p H$  that will be key to the foregoing analysis. Recall that the space  $X = L^2(0, T; H_0^1(\Omega))$  was defined above.

**Lemma 2.2** *Let  $\{v_k\}_{k \in \mathbb{N}}$  be a sequence in  $X$  that converges strongly to a limit  $v \in X$ . If a sequence  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  satisfies  $\tilde{b}_k \in \mathcal{D}_p H[v_k]$  for all  $k \in \mathbb{N}$ , and if  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  converges weakly to some  $\tilde{b}$  in  $L^2(Q_T; \mathbb{R}^d)$  as  $k \rightarrow \infty$ , then  $\tilde{b} \in \mathcal{D}_p H[v]$ .*

This is proved similarly to the approach taken in [32, Lemma 3].

### 3 Continuous problem

We now give the definition of the weak solution of the MFG PDI problem.

**Definition 3.1** (Weak Solution) We say that a pair  $(u, m) \in X \times Y$  is a weak solution of (1.1) if  $m(0) = m_0$  in  $L^2(\Omega)$  and there exists a  $\tilde{b}_* \in \mathcal{D}_p H[u]$  such that, for all  $(\psi, \phi) \in Y_0 \times X$ ,

$$\begin{aligned} & \int_0^T \langle \partial_t \psi, u \rangle + v(\nabla u, \nabla \psi)_\Omega + (H[\nabla u], \psi)_\Omega dt \\ &= \int_0^T \langle F[m], \psi \rangle dt + (S[m(T)], \psi(T))_\Omega, \end{aligned} \quad (3.1a)$$

$$\int_0^T \langle \partial_t m, \phi \rangle + v(\nabla m, \nabla \phi)_\Omega + (m\tilde{b}_*, \nabla \phi)_\Omega dt = \int_0^T \langle G(t), \phi \rangle dt. \quad (3.1b)$$

**Remark 3.1** In order to see how Definition 3.1 leads to a PDI, define the set

$$\operatorname{div}(m\mathcal{D}_p H[u]) := \left\{ g \in L^2(0, T; H^{-1}(\Omega)); \exists \tilde{b} \in \mathcal{D}_p H[u] \text{ s.t. } g = \operatorname{div}(\tilde{b}m) \right\}, \quad (3.2)$$

where the equality above is understood in the sense of distributions. Then equation (3.1b) can be written as  $\partial_t m - v\Delta m - G \in \operatorname{div}(m\mathcal{D}_p H[u])$  in the sense of distributions in  $L^2(0, T; H^{-1}(\Omega))$ . We stress that Definition 3.1 does not require uniqueness of the term  $\tilde{b}_*$  appearing in (3.1), although in some special cases its uniqueness can be shown, see [32, 33]. It is furthermore known from examples in [32, 33] that the regularity of the solution can be quite limited even in the interior of the domain, in particular for the steady-state problem, there are examples where  $m \in H^{3/2-\epsilon}$  for arbitrarily small  $\epsilon > 0$  but  $m \notin H^{3/2}$ .

**Remark 3.2** The choice of test and trial spaces here, together with the condition on the transport vector  $\tilde{b}_*$ , are sufficient to ensure that each term in (3.1) is well-defined and finite. It is also possible to write other formulations of the problem that are ultimately equivalent, such as casting the time-derivative onto  $u$  in (3.1a), since if  $(u, m) \in X \times Y$  is a solution of (3.1), then  $u$  admits a distributional time derivative  $\partial_t u \in L^2(0, T; H^{-1}(\Omega))$  so that  $u \in Y$ , and moreover that  $u(T) = S[m(T)]$  a.e. in  $\Omega$ .



We now state the main result on the existence of a solution of (3.1).

**Theorem 3.1** (Existence of Solutions) *There exists at least one solution  $(u, m) \in Y \times Y$  of (3.1) in the sense of Definition 3.1.*

In this paper, we give a constructive proof of Theorem 3.1 in the case where the domain  $\Omega$  is a polytope by showing the convergence of finite element approximations of the problem. The proof of Theorem 3.1 is therefore given later in Sect. 8. We can also prove the existence of a solution of (3.1) for more general domains by a more direct analysis, although we omit this second approach for the sake of brevity; see the PhD thesis [31] of the first author for a complete analysis.

Under standard monotonicity assumptions, we recover the uniqueness of the solution in the sense of Definition 3.1. Crucially, by choosing the subdifferential set  $\mathcal{D}_p H$  as the natural generalization of the partial derivative of the Hamiltonian to the case of nondifferentiable Hamiltonians, we obtain a straightforward generalization of the uniqueness result of [28], see also [32, 33].

**Theorem 3.2** (Uniqueness of the Solution) *If the initial density  $m_0 \in L^2(\Omega)$  is non-negative a.e. in  $\Omega$ , the source term  $G$  is nonnegative in the sense of distributions in  $L^2(0, T; H^{-1}(\Omega))$ , the coupling term  $F$  is strictly monotone on  $X$  and the terminal cost  $S$  is monotone on  $L^2(\Omega)$ , then there is at most one weak solution  $(u, m)$  to (3.1) in the sense of Definition 3.1.*

**Proof** Suppose that there are two solutions  $(u_i, m_i)$ ,  $i \in \{1, 2\}$ , in the sense of Definition 3.1. Then, we have, for each  $i \in \{1, 2\}$ ,

$$\int_0^T \langle \partial_t m_i, \phi \rangle + v(\nabla m_i, \nabla \phi)_\Omega + (m_i \tilde{b}_i, \nabla \phi)_\Omega dt = \int_0^T \langle G(t), \phi \rangle dt, \quad (3.3)$$

for all  $\phi \in X$ , for some  $\tilde{b}_i \in \mathcal{D}_p H[u_i]$ , and

$$\begin{aligned} & \int_0^T \langle \partial_t \psi, u_i \rangle + v(\nabla u_i, \nabla \psi)_\Omega + (H[\nabla u_i], \psi)_\Omega dt \\ &= \int_0^T \langle F[m_i], \psi \rangle dt + (S[m_i(T)], \psi(T))_\Omega, \end{aligned} \quad (3.4)$$

for all  $\psi \in Y_0$ , with  $m_i(0) = m_0$  in  $L^2(\Omega)$ . Since  $m_0 \geq 0$  a.e. in  $\Omega$ , since  $G$  is nonnegative in the sense of distributions and since  $m_i|_{\partial\Omega} = 0$  for a.e.  $t \in (0, T)$ , the weak maximum principle (see, e.g. [5]) implies that  $m_i \geq 0$  a.e. in  $Q_T$ , for each  $i \in \{1, 2\}$ . After choosing test functions  $\psi = m_1 - m_2 \in Y_0$  and  $\phi = u_1 - u_2$  in (3.4) and (3.3), respectively, and subtracting the two equations, we deduce that

$$\int_0^T \int_{\Omega} m_1 \lambda_{12} + m_2 \lambda_{21} dx dt = \int_0^T \langle F[m_1] - F[m_2], m_1 - m_2 \rangle dt \\ + (S[m_1(T)] - S[m_2(T)], m_1(T) - m_2(T))_{\Omega}, \quad (3.5)$$

where  $\lambda_{ij} \in L^2(0, T; L^2(\Omega))$ ,  $i, j \in \{1, 2\}$ , is defined by

$$\lambda_{ij} := -H[\nabla u_j] + H[\nabla u_i] + \tilde{b}_i \cdot \nabla(u_j - u_i).$$

By definition of  $\mathcal{D}_p H$  in Definition 2.1, it follows that  $\lambda_{ij} \leq 0$  a.e. in  $Q_T$ . Using the nonnegativity of  $m_i$ ,  $i \in \{1, 2\}$ , we then deduce from (3.5) that

$$0 \geq \int_0^T \langle F[m_1] - F[m_2], m_1 - m_2 \rangle dt \\ + (S[m_1(T)] - S[m_2(T)], m_1(T) - m_2(T))_{\Omega}. \quad (3.6)$$

The strict monotonicity of  $F$  and the monotonicity of  $S$  then imply that  $m_1 = m_2$  in  $Q_T$ . Since  $Y$  is continuously embedded in  $C([0, T]; L^2(\Omega))$ , it follows that  $m_1(T) = m_2(T)$  a.e. in  $\Omega$ . Consequently, it follows that  $u_1$  and  $u_2$  solve (3.1a) with common right-hand side and common boundary conditions, and thus  $u_1 = u_2$  in  $Q_T$  by uniqueness of the solution of the corresponding HJB equation.  $\square$

## 4 Notation and setting for discretization

In this section we introduce a monotone finite element scheme for approximating solutions to the weak formulation (3.1). In the sequel, we shall further assume that  $\Omega$  is a polyhedron, in addition to the earlier assumption that it is a bounded connected open set with Lipschitz boundary.

### 4.1 Meshes

Let  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  be a shape-regular sequence of conforming simplicial meshes of the domain  $\Omega$ , c.f. [16, p. 51]. In addition we assume that the meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  are nested, i.e. every element in  $\mathcal{T}_{k+1}$  is either in  $\mathcal{T}_k$  or is a subdivision of an element in  $\mathcal{T}_k$ . For each  $k \in \mathbb{N}$ , define  $h_{\mathcal{T}_k} \in L^\infty(\Omega)$  the mesh-size function by  $h_{\mathcal{T}_k}|_K := \text{diam } K$  for each element  $K \in \mathcal{T}_k$ , where  $\text{diam } K$  denotes the diameter  $K$ . The maximum element size in  $\mathcal{T}_k$  is denoted by  $h_k := \|h_{\mathcal{T}_k}\|_{L^\infty(\Omega)}$ . We assume that  $h_k \rightarrow 0$  as  $k \rightarrow \infty$ . Shape-regularity of  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  refers to the condition that there exists a real-number  $\delta > 1$ , independent of  $k \in \mathbb{N}$ , such that  $h_{\mathcal{T}_k}|_K \leq \delta \rho_K$  for all  $K \in \mathcal{T}_k$ , for all  $k \in \mathbb{N}$ , where  $\rho_K$  denotes the radius of the largest inscribed ball in the element  $K$ .

**Notation for inequalities** For real numbers  $a, b$ , we write  $a \lesssim b$  if  $a \leq Cb$  for some constant  $C$  that depends only on the problem data  $\Omega, T, \mathcal{A}, b, v, H, F$  and  $S$ , and on the shape-regularity constant  $\delta$ , but is otherwise independent of the mesh  $\mathcal{T}_k$  and the mesh-size  $h_k$ . We write  $a \approx b$  if  $a \lesssim b$  and  $b \lesssim a$ .

**Sets of vertices and edges** For each  $k \in \mathbb{N}$ , let  $\mathcal{V}_k$  denote the set of all vertices of  $\mathcal{T}_k$  and let  $\mathcal{V}_{k,\Omega} = \mathcal{V}_k \cap \Omega$  denote the set of interior vertices of  $\mathcal{T}_k$ . Let  $\{x_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  be an enumeration of  $\mathcal{V}_k$ . Without loss of generality, we may choose the ordering such that  $x_i \in \mathcal{V}_{k,\Omega}$  if and only if  $i \leq M_k := \text{card } \mathcal{V}_{k,\Omega}$ . Two distinct vertices in  $\mathcal{V}_k$  are called neighbours if they belong to a common element of  $\mathcal{T}_k$ . For a vertex  $x_i \in \mathcal{V}_k$ , the set of neighbouring vertices of  $x_i$  is denoted by  $\mathcal{V}_{k,i}$ .

For each  $k \in \mathbb{N}$ , let  $\mathcal{E}_k$  denote the set of edges of the mesh  $\mathcal{T}_k$ , i.e. the set of all closed line segments formed by all pairs of neighbouring vertices. Given an edge  $E \in \mathcal{E}_k$ , let  $\mathcal{T}_{k,E} := \{K \in \mathcal{T}_k : E \subset K\}$  denote the set of elements of  $\mathcal{T}_k$  containing  $E$ . We say that an edge  $E \in \mathcal{E}_k$  is an internal edge if  $E$  contains at least one vertex in  $\mathcal{V}_{k,\Omega}$ . The set of all internal edges is denoted by  $\mathcal{E}_{k,\Omega}$ . For each vertex  $x_i \in \mathcal{V}_k$ , let  $\mathcal{E}_{k,i} := \{E \in \mathcal{E}_k : x_i \in E\}$  denote the set of edges containing  $x_i$ . For each  $K \in \mathcal{T}_k$ , let  $\mathcal{E}_K := \{E \subset K : E \in \mathcal{E}_{k,\Omega}\}$  denote the set of edges of the simplex  $K$  that are internal edges.

For each element  $K \in \mathcal{T}_k$  and a given vertex  $x_i \in K$ , let  $F_{K,i}$  denote the  $(d-1)$ -dimensional face of  $K$  that is opposite  $x_i$ , i.e.  $F_{K,i}$  is the convex hull of all vertices of  $K$  except  $x_i$ . Let  $\theta_{ij}^K$  denote the dihedral angle between the faces  $F_{K,i}$  and  $F_{K,j}$ . We assume that the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  satisfies the following hypothesis of Xu and Zikatanov (cf. [42]): for any  $k \in \mathbb{N}$  and for any internal edge  $E \in \mathcal{E}_k$  formed by neighbouring vertices  $x_i$  and  $x_j$ , there holds

$$\sum_{K \in \mathcal{T}_{k,E}} |F_{K,i} \cap F_{K,j}|_{d-2} \cot(\theta_{ij}^K) \geq 0, \quad (4.1)$$

where  $|\cdot|_{d-2}$  denotes the  $d-2$  dimensional Hausdorff measure (counting measure for  $d=2$ ). This condition ensures that the stiffness matrix for the Laplacian on  $V_k$  is an  $M$ -matrix [42]. In the case where  $d=2$ , this condition requires that the sum of the angles opposite to any edge should be less than or equal to  $\pi$ . Furthermore, meshes satisfying the condition (4.1) in the case  $d=2$  are equivalently Delaunay triangulations [42]. This includes, for instance, meshes with nonobtuse triangles and also some meshes with obtuse triangles. Note also that such meshes can be often be obtained by standard mesh refinement routines. For example, if the initial mesh  $\mathcal{T}_1$  consists of right-angled triangles then the newest vertex bisection method creates a sequence of meshes of only right-angled triangles, or if the initial mesh is acute then red refinement yields a sequence of meshes of acute triangles [41, p. 66].

## 4.2 Spatial finite element spaces

Given an element  $K \subset \mathbb{R}^d$ , we let  $\mathcal{P}_1(K)$  denote the vector space of  $d$ -variate real-valued polynomials on  $K$  of total degree at most one. For each  $k \in \mathbb{N}$ , let the spatial finite element space  $V_k$  be defined by

$$V_k := \{v \in H_0^1(\Omega), v|_K \in \mathcal{P}_1(K) \quad \forall K \in \mathcal{T}_k\}. \quad (4.2)$$

For each  $k \in \mathbb{N}$ , let  $\{\xi_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  denote the standard nodal Lagrange basis for the space of all continuous piecewise linear functions on  $\overline{\Omega}$  with respect to  $\mathcal{T}_k$ , where  $\xi_i(x_j) = \delta_{ij}$  for all  $i, j \in \{1, \dots, \text{card } \mathcal{V}_k\}$  for the chosen enumeration  $\{x_i\}_{i=1}^{\mathcal{V}_k}$  of  $\mathcal{V}_k$ . As there is no risk of confusion, we omit the dependence of the nodal basis on the index  $k$  of the mesh in the notation. Note that  $\{\xi_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  form a partition of unity on  $\Omega$ . Recalling that  $x_i \in \mathcal{V}_{k,\Omega}$  if and only if  $i \leq M_k := \text{card } \mathcal{V}_{k,\Omega}$ , we see that  $\{\xi_i\}_{i=1}^{M_k}$  is the standard nodal basis of  $V_k$ .

Let  $V_k^*$  denote the space of continuous linear functionals on  $V_k$ , where we let the duality pairing between  $V_k$  and  $V_k^*$  be denoted by  $\langle \cdot, \cdot \rangle_{V_k^* \times V_k}$ . We equip  $V_k^*$  with the standard dual norm  $\|\cdot\|_{V_k^*}$ . For any operator  $\mathcal{L} : V_k \rightarrow V_k^*$  we define the adjoint operator  $\mathcal{L}^* : V_k \rightarrow V_k^*$  by  $\langle \mathcal{L}^* w, v \rangle_{V_k^* \times V_k} := \langle \mathcal{L} v, w \rangle_{V_k^* \times V_k}$  for all  $w, v \in V_k$ .

### 4.3 Mass lumping

In order to obtain a numerical scheme satisfying a discrete maximum principle, we will use mass-lumping of the  $L^2$ -inner products in the discrete setting. We use the standard mass-lumping technique of defining a discrete inner-product by a quadrature approximation of the integral appearing in the  $L^2$ -inner product. This leads to a diagonal (lumped) mass matrix for the standard nodal basis of  $V_k$ . In particular, for each  $k \in \mathbb{N}$ , let the inner product  $(\cdot, \cdot)_{\Omega,k}$  on  $V_k$  be defined by

$$(w, v)_{\Omega,k} := \int_{\Omega} I_k(wv) dx \quad \forall w, v \in V_k, \quad (4.3)$$

where  $I_k : C(\overline{\Omega}) \cap H_0^1(\Omega) \rightarrow V_k$  denotes the canonical Lagrange interpolation operator, i.e.  $I_k v = \sum_{i=1}^{M_k} v(x_i) \xi_i$  for all  $v \in C(\overline{\Omega}) \cap H_0^1(\Omega)$ , where  $\mathcal{V}_{k,\Omega} = \{x_i\}_{i=1}^{M_k}$ . It is straightforward to check that the mass-lumped inner product satisfies

$$(w, v)_{\Omega,k} = \sum_{i=1}^{M_k} (\xi_i, 1)_{\Omega} w(x_i) v(x_i) \quad \forall w, v \in V_k. \quad (4.4)$$

Moreover,  $(\cdot, \cdot)_{\Omega,k}$  induces a norm on  $V_k$  that is given by  $\|w\|_{\Omega,k} := \sqrt{(w, w)_{\Omega,k}}$  for all  $w \in V_k$ . It is straightforward to show that

$$\|v\|_{\Omega} \leq \|v\|_{\Omega,k} \lesssim \|v\|_{\Omega} \quad \forall v \in V_k, \quad (4.5)$$

where the hidden constant in the second inequality depends only on the dimension  $d$  and the shape-regularity of  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$ . The first inequality follows immediately from the fact that  $v^2 \leq I_k(v^2)$  in  $\Omega$  for all  $v \in V_k$ . Indeed, since the basis functions  $\{\xi_i : 1 \leq i \leq M_k\}$  are nonnegative everywhere and satisfy  $\sum_{i=1}^{M_k} \xi_i|_K = 1$  for each

element  $K \in \mathcal{T}_k$ , we deduce from Jensen's inequality that

$$v^2|_K = \left( \sum_{i=1}^{M_k} v(x_i) \xi_i|_K \right)^2 \leq \sum_{i=1}^{M_k} v(x_i)^2 \xi_i|_K = I_k(v^2)|_K \quad \forall K \in \mathcal{T}_k,$$

for each  $v \in V_k$ . As such,  $v^2 \leq I_k(v^2)$  in  $\Omega$  holds for all  $v \in V_k$ , thereby showing the first inequality in (4.5). We stress that there is no generic constant in the first inequality in (4.5), which will be important for the weak compactness arguments that will follow in later sections. The second inequality in (4.5) is easily obtained by a standard scaling and equivalence-of-norms argument.

For each  $k \in \mathbb{N}$ , let  $R_k : L^2(\Omega) \rightarrow V_k$  be the linear operator defined by

$$R_k w := \sum_{i=1}^{M_k} \frac{(\xi_i, w)_\Omega}{(\xi_i, 1)_\Omega} \xi_i \quad \forall w \in L^2(\Omega). \quad (4.6)$$

It is clear that  $R_k$  is a Riesz-map between the  $L^2$ -inner product and  $(\cdot, \cdot)_{\Omega,k}$ , with

$$(v, R_k w)_{\Omega,k} = (v, w)_\Omega \quad \forall v \in V_k, \quad \forall w \in L^2(\Omega). \quad (4.7)$$

Furthermore,  $R_k$  is  $L^2$ -stable with the bound  $\|R_k w\|_\Omega \lesssim \|w\|_\Omega$  for all  $w \in L^2(\Omega)$ . We also have the well-known bound (see e.g. [41, p. 108])

$$\|h_{\mathcal{T}_k}^{-1}(w - R_k w)\|_\Omega + \|\nabla(w - R_k w)\|_\Omega \lesssim \|\nabla w\|_\Omega \quad \forall w \in H_0^1(\Omega). \quad (4.8)$$

It follows that  $R_k$  is also  $H_0^1(\Omega)$ -stable, i.e.  $\|\nabla R_k w\|_\Omega \lesssim \|\nabla w\|_\Omega$  for all  $w \in H_0^1(\Omega)$ .

We now introduce the following discrete dual norm  $\|\cdot\|_{k,*} : V_k \rightarrow \mathbb{R}_{\geq 0}$  defined by

$$\|w\|_{k,*} := \sup_{v \in V_k \setminus \{0\}} \frac{(w, v)_{\Omega,k}}{\|\nabla v\|_\Omega} \quad \forall w \in V_k. \quad (4.9)$$

**Lemma 4.1** *We have  $\|v\|_{H^{-1}(\Omega)} \approx \|v\|_{k,*}$  for all  $v \in V_k$ , where the hidden constants depend only on the dimension  $d$  and the shape-regularity of  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$ .*

**Proof** For any  $v \in V_k$ ,  $k \in \mathbb{N}$ , and  $w \in H_0^1(\Omega)$ , we have  $(v, w)_\Omega = (v, R_k w)_{\Omega,k}$  and thus the definition of  $\|\cdot\|_{k,*}$  and the  $H_0^1$ -stability of  $R_k$  imply that  $(v, w)_\Omega \leq \|v\|_{k,*} \|\nabla R_k w\|_\Omega \lesssim \|v\|_{k,*} \|\nabla w\|_\Omega$ . This shows that  $\|v\|_{H^{-1}(\Omega)} \lesssim \|v\|_{k,*}$ . To show the converse bound, now consider  $w \in V_k$ , and write  $(v, w)_{\Omega,k} = (v, w)_\Omega + (v, w - R_k w)_{\Omega,k}$ . The Cauchy–Schwarz inequality and standard inverse inequalities imply that  $(v, w - R_k w)_{\Omega,k} = \int_\Omega I_k(v(w - R_k w)) dx \lesssim \|h_{\mathcal{T}_k} v\|_\Omega \|h_{\mathcal{T}_k}^{-1}(w - R_k w)\|_\Omega$ . Note that  $\|h_{\mathcal{T}_k} v\|_\Omega \lesssim \|v\|_{H^{-1}(\Omega)}$ , which is deduced from the local bound  $\|h_{\mathcal{T}_k} v\|_K \lesssim \|v\|_{H^{-1}(K)}$  for all  $K \in \mathcal{T}_k$ , see e.g. [41, p. 113], and from  $\sum_{K \in \mathcal{T}_k} \|v\|_{H^{-1}(K)}^2 \leq \|v\|_{H^{-1}(\Omega)}^2$ . In conclusion, we then use (4.8) to find that  $(v, w - R_k w)_{\Omega,k} \leq \|v\|_{H^{-1}(\Omega)} \|\nabla w\|_\Omega$  for all  $w \in V_k$ , which implies that  $\|v\|_{k,*} \lesssim \|v\|_{H^{-1}(\Omega)}$ .  $\square$

#### 4.4 Spatial stabilization

We say that a linear operator  $L : V_k \rightarrow V_k^*$  satisfies the *discrete maximum principle* (DMP) provided that the following condition holds: if  $v \in V_k$  and  $\langle Lv, \xi_i \rangle_{V_k^* \times V_k} \geq 0$  for all  $i \in \{1, \dots, M_k\}$ , then  $v \geq 0$  in  $\Omega$ . Recall that under the condition (4.1), the stiffness matrix of the Laplace operator on  $V_k$  is an  $M$ -matrix, which implies monotonicity and the discrete maximum principle for the Laplace operator on  $V_k$ . In order to handle the advective terms of the problem, we construct in this section an original volume-based self-adjoint linear stabilization that maintains the adjoint relationship between the equations of the MFG system. For each  $k \in \mathbb{N}$  and each edge  $E \in \mathcal{E}_{k,\Omega}$ , let  $\omega_{k,E} \geq 0$  be a nonnegative weight to be chosen below, and let  $\mathbf{t}_E$  be a chosen unit tangent vector to the edge  $E$ . The choice of orientation of  $\mathbf{t}_E$  does not have any effect on the following. Then, we define the stabilization diffusion matrix  $D_k \in L^\infty(\Omega; \mathbb{R}^{d \times d})$  elementwise over  $\mathcal{T}_k$  by

$$D_k|_K := \sum_{E \in \mathcal{E}_K} \omega_{k,E} \mathbf{t}_E \otimes \mathbf{t}_E \quad \forall K \in \mathcal{T}_k, \quad (4.10)$$

where  $\otimes$  denotes the outer-product of vectors, i.e.  $\mathbf{t}_E \otimes \mathbf{t}_E := \mathbf{t}_E \mathbf{t}_E^T \in \mathbb{R}^{d \times d}$ , and where we recall that  $\mathcal{E}_K$  denotes the set of edges of  $K$  that are internal. Note that  $D_k$  is independent of the choice of orientations of the tangent vectors  $\mathbf{t}_E$ . Since the weights  $\omega_{k,E}$  are nonnegative, it follows that  $D_k|_K$  is a symmetric, positive semi-definite matrix in  $\mathbb{R}^{d \times d}$  for each  $K \in \mathcal{T}_k$ . Then, let  $W(V_k, D_k)$  denote the set of all linear operators  $L : V_k \rightarrow V_k^*$  of the form

$$\langle Lv, w \rangle_{V_k^* \times V_k} := \int_{\Omega} A_k \nabla v \cdot \nabla w + \tilde{b} \cdot \nabla v w \, dx \quad \forall w, v \in V_k, \quad (4.11)$$

where  $A_k := \nu \mathbb{I} + D_k$  and where  $\tilde{b} \in L^\infty(\Omega; \mathbb{R}^d)$  is some vector field that satisfies  $\|\tilde{b}\|_{L^\infty(\Omega; \mathbb{R}^d)} \leq L_H$ . Observe also that  $A_k \geq \nu \mathbb{I}$  in the sense of semi-definite matrices since  $D_k$  is positive semi-definite. Recall that  $\delta$  denotes the shape-regularity parameter of the meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$ .

**Theorem 4.2** (DMP) *Suppose that the weights in (4.10) are chosen such that*

$$\frac{\delta L_H \operatorname{diam} E}{2(d+1)} < \omega_{k,E} \lesssim L_H \operatorname{diam} E \quad \forall E \in \mathcal{E}_{k,\Omega}. \quad (4.12)$$

*Then, the following properties hold:*

- $\|D_k|_K\|_{L^\infty(K; \mathbb{R}^{d \times d})} \lesssim h_K$  for all  $K \in \mathcal{T}_k$ ,
- if  $L \in W(V_k, D_k)$  then  $L$  and its adjoint  $L^*$  satisfy the discrete maximum principle.

The proof of Theorem 4.2 is postponed to Appendix A. We remark that  $D_k$  can be chosen independently of the diffusion parameter  $\nu$  of the problem. Furthermore, it is clear that a monotone discretization of the parabolic operators of the MFG system can be obtained by combining the spatial stabilization above with a mass-lumped implicit

Euler discretization for the time derivative. Indeed, this is easily seen from the fact that if  $L \in W(V_k, D_k)$ , then the operator  $V_k \ni w \mapsto \langle Lw, \cdot \rangle_{V_k^* \times V_k} + \frac{1}{\tau_k}(w, \cdot)_{\Omega, k}$  also satisfies the discrete maximum principle for any  $\tau_k > 0$ . Finally, note that if  $L: V_k \rightarrow V_k^*$  satisfies the DMP, then  $L$  is necessarily a bijection since  $V_k$  is finite dimensional.

#### 4.5 Space-time finite element spaces

Given  $k \in \mathbb{N}$ , let  $\mathcal{J}_k := \{I_n\}_{n=1}^{N_k}$  denote the partition of  $[0, T]$  where  $I_n := (t_{n-1}, t_n)$  with  $t_n := n\tau_k$  for  $n \in \{1, \dots, N_k\}$ ,  $t_0 := 0$ , and the time-step  $\tau_k := T/N_k$ , with  $N_k \rightarrow \infty$  as  $k \rightarrow \infty$ . To simplify the analysis, we assume that the temporal meshes are nested, and thus  $\tau_{k+1} \leq \tau_k$  for all  $k \in \mathbb{N}$ , and we also assume that the initial time-step satisfies

$$\tau_1 < \nu L_H^{-2}. \quad (4.13)$$

The above assumption (4.13) on the initial time-step is not essential as we introduce it solely for the sake of brevity of the analysis. For treatment of the analysis in the case where this assumption is not satisfied see [31, Ch. 6].

For each  $k \in \mathbb{N}$ , we define  $\mathbb{V}_k$  the space of piecewise constant  $V_k$ -valued functions defined on  $[0, T]$  by

$$\mathbb{V}_k := \{v \in X, v|_{I_n} \in V_k \text{ is constant } \forall n \in \{1, \dots, N_k\}\}. \quad (4.14)$$

For a function  $v \in \mathbb{V}_k$  and  $n \in \{1, \dots, N_k - 1\}$ , we let  $v(t_n^-)$  and  $v(t_n^+)$  denote respectively the left- and right-limits of  $v$  at time  $t_n$ . A function  $v \in \mathbb{V}_k$  also admits a right-limit  $v(t_0^+) = v|_{I_1}$  at  $t_0 = 0$  and a left-limit  $v(t_{N_k}^-) = v|_{I_{N_k}}$  at  $t_{N_k} = T$ . Note that since  $v \in \mathbb{V}_k$  is piecewise constant, we have  $v(t_n^-) = v|_{I_n}$  for all  $n \in \{1, \dots, N_k\}$ , and that  $v(t_n^+) = v|_{I_{n+1}}$  for all  $n \in \{0, \dots, N_k - 1\}$ . In the following, it will help to also consider spaces of functions that are defined everywhere on  $[0, T]$  by left- or right-continuity. For a function  $v: [0, T] \rightarrow V_k$  with  $v|_{(0, T)} \in \mathbb{V}_k$ , we say that  $v$  is left-continuous if  $v(t_n) = v(t_n^-)$  for all  $n = 1, \dots, N_k$ , and that  $v$  is right-continuous if  $v(t_n) = v(t_n^+)$  for all  $n = 0, \dots, N_k - 1$ . We then define

$$\mathbb{V}_k^+ := \{v: [0, T] \rightarrow V_k, v|_{(0, T)} \in \mathbb{V}_k, \text{ and } v \text{ is left-continuous}\}, \quad (4.15a)$$

$$\mathbb{V}_k^- := \{v: [0, T] \rightarrow V_k, v|_{(0, T)} \in \mathbb{V}_k, \text{ and } v \text{ is right-continuous}\}. \quad (4.15b)$$

Note that here we adopt the point of view that functions in  $\mathbb{V}_k^\pm$  are defined at all points in  $[0, T]$ , and two functions in  $\mathbb{V}_k^\pm$  that agree up to a subset of measure zero of  $[0, T]$  are not identified. When  $v \in \mathbb{V}_k^+$  and  $w \in \mathbb{V}_k^-$ , we set  $v(t_0^-) := v(0)$  and  $w(t_{N_k}^+) := w(T)$ . For  $v \in \mathbb{V}_k$ , we define the jump  $\llbracket v \rrbracket_n$  of  $v$  at time  $t_n \in (0, T)$  by

$$\llbracket v \rrbracket_n := v(t_n^-) - v(t_n^+). \quad (4.16)$$

We define also the jump  $\llbracket v \rrbracket_0 = v(0) - v(0^+)$  for  $v \in \mathbb{V}_k^+$ , and  $\llbracket w \rrbracket_{N_k} = w(T^-) - w(T)$  for  $w \in \mathbb{V}_k^-$ .

**Reconstruction of time derivatives** Let the *forward-in-time* reconstruction operator  $\mathcal{I}_+ : \mathbb{V}_k^+ \rightarrow C[0, T; V_k]$  be defined by

$$(\mathcal{I}_+ v)(t) := v(t) + \frac{t_n - t}{\tau_k} \llbracket v \rrbracket_{n-1}, \quad t \in (t_{n-1}, t_n], \quad n \in \{1, \dots, N_k\}, \quad v \in \mathbb{V}_k^+, \quad (4.17)$$

where  $C[0, T; V_k]$  is the space of continuous functions from  $[0, T]$  to  $V_k$ . We also define the *backward-in-time* reconstruction operator  $\mathcal{I}_- : \mathbb{V}_k^- \rightarrow C[0, T; V_k]$  by

$$(\mathcal{I}_- w)(t) := w(t) - \frac{t - t_{n-1}}{\tau_k} \llbracket w \rrbracket_n, \quad t \in [t_{n-1}, t_n), \quad n \in \{1, \dots, N_k\}, \quad w \in \mathbb{V}_k^-. \quad (4.18)$$

As there is no risk of confusion, we omit the dependence of  $\mathcal{I}_+$  and  $\mathcal{I}_-$  on  $k$  to simplify the notation. Since  $v \in \mathbb{V}_k^+$  is by definition left-continuous, it is easy to check that  $\mathcal{I}_+ v$  is continuous on  $[0, T]$  and satisfies  $\mathcal{I}_+ v(t_n) = v(t_n) = v(t_n^-)$  for all  $v \in \mathbb{V}_k^+$  and  $n \in \{0, \dots, N_k\}$ . In particular,  $\mathcal{I}_+ v(0) = v(0)$ . Similarly,  $\mathcal{I}_- w$  is continuous on  $[0, T]$  and  $\mathcal{I}_- w(t_n) = w(t_n) = w(t_n^+)$  for all  $n \in \{0, \dots, N_k\}$  and  $\mathcal{I}_- w(T) = v(T)$ . It is also straightforward to verify that  $\mathcal{I}_+ v$  and  $\mathcal{I}_- w$  are piecewise continuously differentiable with respect to  $\mathcal{J}_k$  for any  $v \in \mathbb{V}_k^+$  and any  $w \in \mathbb{V}_k^-$ , and that

$$\partial_t \mathcal{I}_+ v|_{I_n} = -\frac{1}{\tau_k} \llbracket v \rrbracket_{n-1}, \quad \partial_t \mathcal{I}_- w|_{I_n} = -\frac{1}{\tau_k} \llbracket w \rrbracket_n, \quad n \in \{1, \dots, N_k\}. \quad (4.19)$$

This shows that  $\partial_t \mathcal{I}_+$  and  $\partial_t \mathcal{I}_-$  are respectively related to the first-order backward and forward difference operators. Furthermore, the operators  $\mathcal{I}_+$  and  $\mathcal{I}_-$  satisfy the discrete integration-by-parts formula

$$\int_0^T (\partial_t \mathcal{I}_+ v, w)_{\Omega, k} + (v, \partial_t \mathcal{I}_- w)_{\Omega, k} dt = (v(T), w(T))_{\Omega, k} - (v(0), w(0))_{\Omega, k}. \quad (4.20)$$

for all  $(v, w) \in \mathbb{V}_k^+ \times \mathbb{V}_k^-$ .

**Discrete norms** Let the norms  $\|\cdot\|_{\mathbb{V}_k^+}$  on  $\mathbb{V}_k^+$  and  $\|\cdot\|_{\mathbb{V}_k^-}$  on  $\mathbb{V}_k^-$  be defined by

$$\|v\|_{\mathbb{V}_k^+}^2 := \int_0^T \|\partial_t \mathcal{I}_+ v\|_{k,*}^2 + \|\nabla v\|_{\Omega}^2 dt + \|v(0)\|_{\Omega, k}^2 \quad \forall v \in \mathbb{V}_k^+, \quad (4.21a)$$



$$\|w\|_{\mathbb{V}_k^-}^2 := \int_0^T \|\partial_t \mathcal{I}_- w\|_{k,*}^2 + \|\nabla w\|_{\Omega}^2 dt + \|w(T)\|_{\Omega,k}^2 \quad \forall w \in \mathbb{V}_k^-. \quad (4.21b)$$

The following result shows that the norms  $\|\cdot\|_{\mathbb{V}_k^+}$  and  $\|\cdot\|_{\mathbb{V}_k^-}$  are upper bounds on the  $L^\infty(L^2)$  norm.

**Lemma 4.3** (Boundedness in  $L^\infty(L^2)$ ) *We have  $\|v(t)\|_{\Omega} \leq \|v\|_{\mathbb{V}_k^\pm}$  for all  $v \in \mathbb{V}_k^\pm$ , and all  $t \in [0, T]$ .*

**Proof** It is enough to prove the result in the case of  $\mathbb{V}_k^+$  since the case of  $\mathbb{V}_k^-$  follows the same argument up to a change of the time variable. Let  $v \in \mathbb{V}_k^+$  be arbitrary. The inequality  $\|v(0)\|_{\Omega} \leq \|v\|_{\mathbb{V}_k^+}$  follows immediately from (4.5) and (4.21a). It remains only to consider the case  $t \in (0, T]$ . Since  $v$  is piecewise constant in time and left-continuous, there exists a  $n \in \{1, \dots, N_k\}$  such that  $t \in (t_{n-1}, t_n]$  and  $v(t) = v|_{I_n}$ . The first inequality in (4.5) and the identity  $2 \int_0^{t_n} (\partial_t \mathcal{I}_+ v, v)_{\Omega,k} dt = \|v|_{I_n}\|_{\Omega,k}^2 + \sum_{i=0}^{n-1} \|\llbracket v \rrbracket_i\|_{\Omega,k}^2 - \|v(0)\|_{\Omega,k}^2$  for any  $n \in \{1, \dots, N_k\}$  imply that

$$\|v(t)\|_{\Omega}^2 = \|v|_{I_n}\|_{\Omega}^2 \leq \|v|_{I_n}\|_{\Omega,k}^2 \leq 2 \int_0^{t_n} (\partial_t \mathcal{I}_+ v, v)_{\Omega,k} dt + \|v(0)\|_{\Omega,k}^2 \leq \|v\|_{\mathbb{V}_k^+}^2,$$

where in the final inequality we have used  $(\partial_t \mathcal{I}_+ v, v)_{\Omega,k} \leq \|\partial_t \mathcal{I}_+ v\|_{k,*} \|\nabla v\|_{\Omega}$  for a.e.  $t \in (0, T)$  and Young's inequality.  $\square$

## 5 Finite element approximation and main results

We now introduce the finite element discretization of (3.1). For each  $k \in \mathbb{N}$ , let  $D_k \in L^\infty(\Omega; \mathbb{R}^{d \times d})$  of the form (4.10) be as in Theorem 4.2, and recall that  $A_k := v\mathbb{I} + D_k$  and that  $\|D_k\|_{L^\infty(\Omega; \mathbb{R}^{d \times d})} \lesssim h_k$ . For each  $k \in \mathbb{N}$ , the discrete problem is to find  $(u_k, m_k) \in \mathbb{V}_k \times \mathbb{V}_k^+$  such that, for some  $\tilde{b}_k \in \mathcal{D}_p H[u_k]$ ,

$$\int_0^T (\partial_t \mathcal{I}_+ \psi, u_k)_{\Omega,k} + (A_k \nabla u_k, \nabla \psi)_{\Omega} + (H[\nabla u_k], \psi)_{\Omega} dt \quad (5.1a)$$

$$= \int_0^T \langle F[m_k], \psi \rangle dt + (R_k S[m_k(T)], \psi(T))_{\Omega,k},$$

$$\int_0^T (\partial_t \mathcal{I}_+ m_k, \phi)_{\Omega,k} + (A_k \nabla m_k, \nabla \phi)_{\Omega} + (m_k \tilde{b}_k, \nabla \phi)_{\Omega} dt = \int_0^T \langle G, \phi \rangle dt, \quad (5.1b)$$

for all  $(\psi, \phi) \in \mathbb{V}_{k,0}^+ \times \mathbb{V}_k$ , and such that  $m_k(0) = R_k m_0 \in V_k$ .

**Remark 5.1** The FEM is presented above in a space-time variational form, as this reflects the coupled space-time nature of the problem, offers a succinct notation, and is also most convenient for the discrete functional analytic framework set out below. By choosing test functions supported on individual time-step intervals, it is easily seen that the discretization consists of a backward-in-time implicit Euler discretization for the HJB equation and a forward-in-time implicit Euler discretization for the KFP equation. Therefore, for practical computation the system can be re-written in a coupled forward-backward time-stepping form. The efficient solution of these coupled discrete systems remains an important challenge which is beyond the scope of this work, however one promising approach is to consider time- or space-time-parallel solvers, see [30] and the references therein.

Observe that if  $(u_k, m_k) \in \mathbb{V}_k \times \mathbb{V}_k^+$  solves (5.1), then we can adopt a different point of view and re-define  $u_k$  to a right-continuous function in  $\mathbb{V}_k^-$  (without change of notation) by setting  $u_k(T) := R_k S[m_k(T)]$  and  $u_k(t_n^+) := u_k|_{I_n}$  for all  $n \in \{0, \dots, M_k - 1\}$ . Using the discrete integration-by-parts identity (4.20), we can equivalently rewrite (5.1a) as

$$\int_0^T -(\psi, \partial_t \mathcal{I}_- u_k)_{\Omega,k} + (A_k \nabla u_k, \nabla \psi)_{\Omega} + (H[\nabla u_k], \psi)_{\Omega} dt = \int_0^T \langle F[m_k], \psi \rangle dt \quad (5.2)$$

for all  $\psi \in \mathbb{V}_{k,0}^+$ . Since the terms in (5.2) depend only on  $\psi|_{(0,T)}$ , it is clear that the set of test functions in (5.2) can then be extended to all  $\psi \in \mathbb{V}_k$ . In practical terms for computation, there is no difference between these two points of view.

## 5.1 Main results

The first main result states the existence of a solution to the discrete problem and gives a stability bound for any such discrete solution in terms of the data of the problem.

**Theorem 5.1** (Existence of discrete solutions and stability bound) *For every  $k \in \mathbb{N}$  there exists a discrete pair  $(u_k, m_k) \in \mathbb{V}_k^- \times \mathbb{V}_k^+$  satisfying (5.1). Any solution  $(u_k, m_k)$  of (5.1) satisfies*

$$\|u_k\|_{\mathbb{V}_k^-} + \|m_k\|_{\mathbb{V}_k^+} \lesssim 1 + \|G\|_{L^2(0,T;H^{-1}(\Omega))} + \|m_0\|_{\Omega}. \quad (5.3)$$

The proof of Theorem 5.1 is given in Sect. 6. The next main result shows that the uniqueness of the numerical solution holds under the same assumptions as those that were considered in the continuous setting of Theorem 3.2.

**Theorem 5.2** (Uniqueness) *Suppose that the initial density  $m_0 \in L^2(\Omega)$  is nonnegative a.e. in  $\Omega$ , the source term  $G$  is nonnegative in the sense of distributions in*

$L^2(0, T; H^{-1}(\Omega))$ , the coupling term  $F$  is strictly monotone on  $X$ , and that the terminal cost  $S$  is monotone on  $L^2(\Omega)$ . Then, for every  $k \in \mathbb{N}$ , there exists a unique solution  $(u_k, m_k) \in \mathbb{V}_k^- \times \mathbb{V}_k^+$  to (5.1).

The proof of Theorem 5.2 is given in Sect. 6.

Our main convergence result shows that strong convergence of the value function approximations in  $L^2(H_0^1)$ , and of the density approximations in  $L^p(L^2)$  for any  $1 \leq p < \infty$ , along with weak convergence of the density approximations in  $L^2(H_0^1)$ .

**Theorem 5.3** (Convergence of FEM) *Assume that the hypotheses of Theorem 5.2 hold. Then, there exists a unique solution  $(u, m) \in Y \times Y$  in the sense of Definition 3.1. Moreover, we have*

$$u_k \rightarrow u \text{ in } L^2(0, T; H_0^1(\Omega)), \quad u_k \rightarrow u \text{ in } L^p(0, T; L^2(\Omega)), \quad (5.4a)$$

$$m_k \rightarrow m \text{ in } L^2(0, T; H_0^1(\Omega)), \quad m_k \rightarrow m \text{ in } L^p(0, T; L^2(\Omega)), \quad (5.4b)$$

$$u_k(0) \rightarrow u(0) \text{ in } L^2(\Omega), \quad m_k(T) \rightarrow m(T) \text{ in } L^2(\Omega), \quad (5.4c)$$

as  $k \rightarrow \infty$ , for any  $1 \leq p < \infty$ .

The proof of Theorem 5.3 is given in Sect. 8 where we present a compactness argument that uses functional analytic tools developed in Sect. 7. We note that, if the conditions for uniqueness given in Theorem 5.2 are absent, then the proof we present still establishes the convergence (5.4) along subsequences.

Note that  $m_k \rightarrow m$  in  $L^2(0, T; H_0^1(\Omega))$  implies that  $\nabla m_k$  converges weakly to  $\nabla m$  in  $L^2(Q_T; \mathbb{R}^d)$ . At present, it is not currently known if the weak convergence of  $\nabla m_k$  to  $\nabla m$  is also strong in general. The difficulty lies in the fact that for nondifferentiable Hamiltonians, the convergence  $\nabla u_k \rightarrow \nabla u$  is not sufficient to prove convergence in sufficiently strong norms of  $\tilde{b}_k \in \mathcal{D}_p H[u_k]$  from (5.1) to  $\tilde{b}_* \in \mathcal{D}_p H[u]$  from (3.1). The next result shows that if some pre-compactness of  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is assumed, then we obtain strong convergence of approximations of the gradient of the density.

**Corollary 5.4** (Strong  $L^2(H_0^1)$ -Convergence for Density Approximations) *In addition to the hypotheses of Theorem 5.3, suppose that the sequence of transport vector fields  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  from (5.1) is pre-compact in  $L^1(Q_T; \mathbb{R}^d)$ . Then,  $m_k$  converges to  $m$  strongly in  $L^2(0, T; H_0^1(\Omega))$  as  $k \rightarrow \infty$ .*

The proof of Corollary 5.4 is in Sect. 8. Note that the pre-compactness of  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is assured, for instance, if the Hamiltonian  $H$  given by (1.2) is such that partial derivative  $\frac{\partial H}{\partial p}$  exists and is continuous in  $Q_T \times \mathbb{R}^d$ . Thus, in the case of continuously differentiable Hamiltonians, Corollary 5.4 shows that  $m_k \rightarrow m$  in norm in  $L^2(0, T; H_0^1(\Omega))$  as  $k \rightarrow \infty$ .

## 6 Proof of Theorems 5.1 and 5.2

In this section we give the proofs of the existence and uniqueness results for the discrete problem (5.1).

## 6.1 Well-posedness of discrete HJB and KFP equations

In Lemmas 6.1 and 6.2 below, we state well-posedness and stability results for the discretized KFP and HJB equations when considered separately. The proofs of these results are given in Appendix B. Recall that we make the assumption that  $\tau_k \leq \tau_1$  for all  $k \in \mathbb{N}$ , with  $\tau_1$  satisfying (4.13), in order to simplify the analysis.

**Lemma 6.1** (Well-posedness of discrete KFP equation) *Let  $k \in \mathbb{N}$ , let  $G \in L^2(0, T; H^{-1}(\Omega))$ , let  $g \in V_k$  and let  $\tilde{b} \in L^\infty(Q_T; \mathbb{R}^d)$  such that  $\|\tilde{b}\|_{L^\infty(Q_T; \mathbb{R}^d)} \leq L_H$ . Then, there exists a unique  $\bar{m} \in \mathbb{V}_k^+$  such that  $\bar{m}(0) = g$  and such that*

$$\int_0^T (\partial_t \mathcal{I}_+ \bar{m}, \phi)_{\Omega, k} + (A_k \nabla \bar{m}, \nabla \phi)_\Omega + (\bar{m}, \tilde{b} \cdot \nabla \phi)_\Omega dt = \int_0^T \langle G, \phi \rangle dt \quad \forall \phi \in \mathbb{V}_k. \quad (6.1)$$

We have the bound

$$\|\bar{m}\|_{\mathbb{V}_k^+} \lesssim \|G\|_{L^2(0, T; H^{-1}(\Omega))} + \|g\|_\Omega, \quad (6.2)$$

where the hidden constant depends on  $\tau_1$  but is otherwise independent of  $k$ . Furthermore, if  $g \geq 0$  a.e. in  $\Omega$  and if  $G$  is nonnegative in the sense of distributions, then  $\bar{m} \geq 0$  a.e. in  $Q_T$ .

**Lemma 6.2** (Well-posedness of discrete HJB equation) *Let  $k \in \mathbb{N}$ , let  $\tilde{F} \in L^2(0, T; H^{-1}(\Omega))$  and let  $\tilde{S} \in L^2(\Omega)$ . Then, there exists a unique  $\bar{u} \in \mathbb{V}_k^-$  such that  $\bar{u}(T) = R_k \tilde{S}$  and such that*

$$\begin{aligned} & \int_0^T (\partial_t \mathcal{I}_+ \psi, \bar{u})_{\Omega, k} + (A_k \nabla \bar{u}, \nabla \psi)_\Omega + (H[\nabla \bar{u}], \psi)_\Omega dt \\ &= \int_0^T \langle \tilde{F}, \psi \rangle dt + (R_k \tilde{S}, \psi(T))_{\Omega, k} \quad \forall \psi \in \mathbb{V}_{k,0}^+. \end{aligned} \quad (6.3)$$

Furthermore  $\bar{u}$  depends continuously on the data: if  $\bar{u}_1, \bar{u}_2 \in \mathbb{V}_k^-$  are respective solutions of (6.3) with data  $(\tilde{F}_1, \tilde{S}_1)$  and  $(\tilde{F}_2, \tilde{S}_2)$  respectively and if  $\bar{u}_1(T) = R_k \tilde{S}_1$  and  $\bar{u}_2(T) = R_k \tilde{S}_2$ , then

$$\|\bar{u}_1 - \bar{u}_2\|_{\mathbb{V}_k^-} \lesssim \|\tilde{F}_1 - \tilde{F}_2\|_{L^2(0, T; H^{-1}(\Omega))} + \|\tilde{S}_1 - \tilde{S}_2\|_\Omega, \quad (6.4)$$

where the hidden constant depends on  $\tau_1$  but is otherwise independent of  $k$ .

## 6.2 Proof of Theorem 5.1

We now prove Theorem 5.1 by an argument that is based on Kakutani's fixed point theorem [44, Ch. 9, Theorem 9.B].

**Theorem 6.3** (Kakutani's fixed point theorem [44]) *Suppose that*

1.  $\mathcal{B}$  is a nonempty, compact, convex set in a locally convex space  $X$ ;
2.  $\mathcal{V} : \mathcal{B} \rightrightarrows \mathcal{B}$  is a set-valued map such that  $\mathcal{V}[\tilde{b}]$  is nonempty, closed and convex for all  $\tilde{b} \in \mathcal{B}$ ; and
3.  $\mathcal{V}$  is upper semi-continuous.

*Then,  $\mathcal{V}$  has a fixed point: there exists  $\tilde{b} \in \mathcal{B}$  such that  $\tilde{b} \in \mathcal{V}[\tilde{b}]$ .*

To begin, recall that  $L_H$  is the Lipschitz constant of the Hamiltonian, c.f. (2.5). We equip the space  $L^\infty(Q_T; \mathbb{R}^d)$  with its weak-\* topology, noting that it is then a locally convex topological vector space. Let  $\mathcal{B}$  denote the ball

$$\mathcal{B} := \left\{ \tilde{b} \in L^\infty(Q_T; \mathbb{R}^d) : \|\tilde{b}\|_{L^\infty(Q_T; \mathbb{R}^d)} \leq L_H \right\}, \quad (6.5)$$

and note that  $\mathcal{B}$  is nonempty, convex, and closed in the weak-\* topology. We note further that the weak-\* topology on  $\mathcal{B}$  is metrisable since  $L^1(Q_T; \mathbb{R}^d)$  is separable [37, Ch. 15], and that  $\mathcal{B}$  is compact by Helly's theorem.

Let  $\mathcal{M}_k : \mathcal{B} \rightarrow \mathbb{V}_k^+$  be the map defined as follows: for each  $\tilde{b} \in \mathcal{B}$ , let  $\mathcal{M}_k[\tilde{b}]$  in  $\mathbb{V}_k^+$  be the unique solution of

$$\begin{aligned} & \int_0^T (\partial_t \mathcal{I}_+ \mathcal{M}_k[\tilde{b}], \phi)_{\Omega, k} + (A_k \nabla \mathcal{M}_k[\tilde{b}], \nabla \phi)_{\Omega} + (\mathcal{M}_k[\tilde{b}], \tilde{b} \cdot \nabla \phi)_{\Omega} dt \\ &= \int_0^T \langle G, \phi \rangle dt \quad \forall \phi \in \mathbb{V}_k \end{aligned} \quad (6.6)$$

with  $\mathcal{M}_k[\tilde{b}](0) = R_k m_0$  in  $V_k$ . The map  $\mathcal{M}_k$  is well-defined thanks to Lemma 6.1. Next, let  $\mathcal{U}_k : \mathbb{V}_k^+ \rightarrow \mathbb{V}_k^-$  be the map defined as follows: for each  $m_k \in \mathbb{V}_k^+$ , let  $\mathcal{U}_k[m_k] \in \mathbb{V}_k^-$  denote the unique solution of

$$\begin{aligned} & \int_0^T (\partial_t \mathcal{I}_+ \psi, \mathcal{U}_k[m_k])_{\Omega, k} + (A_k \nabla \mathcal{U}_k[m_k], \nabla \psi)_{\Omega} + (H[\nabla \mathcal{U}_k[m_k]], \psi)_{\Omega} dt \\ &= \int_0^T \langle F[m_k], \psi \rangle dt + (R_k S[m_k(T)], \psi(T))_{\Omega, k} \quad \forall \psi \in \mathbb{V}_{k,0}^+, \end{aligned} \quad (6.7)$$

with  $\mathcal{U}_k[m_k](T) = R_k S[m_k(T)]$  in  $V_k$ . The map  $\mathcal{U}_k$  is well-defined by Lemma 6.2.

Now, we define the set-valued map  $\mathcal{K}_k : \mathcal{B} \rightrightarrows L^\infty(Q_T; \mathbb{R}^d)$  as follows: for each  $\tilde{b} \in \mathcal{B}$ , let

$$\mathcal{K}_k[\tilde{b}] := \mathcal{D}_p H[\mathcal{U}_k[\mathcal{M}_k[\tilde{b}]]]. \quad (6.8)$$

Lemma 2.1 implies that  $\mathcal{K}_k[\tilde{b}] \subset \mathcal{B}$  for each  $\tilde{b} \in \mathcal{B}$ , so  $\mathcal{K}_k : \mathcal{B} \rightrightarrows \mathcal{B}$ . Moreover, for every  $\tilde{b} \in \mathcal{B}$ , the set  $\mathcal{K}_k[\tilde{b}]$  is nonempty and convex. Indeed, for each  $\tilde{b} \in \mathcal{B}$  the set  $\mathcal{K}_k[\tilde{b}]$  is nonempty by Lemma 2.1. Also,  $\mathcal{K}_k[\tilde{b}]$  is convex since  $\partial_p H$  has convex images. Furthermore,  $\mathcal{K}_k[\tilde{b}]$  is closed for all  $\tilde{b} \in \mathcal{B}$  thanks to Lemma 2.2 and the fact that  $\mathcal{B} \subset L^2(Q_T; \mathbb{R}^d)$  where  $|Q_T|_{d+1} < \infty$ .

The existence of a solution to the discrete problem (5.1) is equivalent to showing the existence of a fixed point of  $\mathcal{K}_k$ , i.e. that there exists a  $\tilde{b}_k \in \mathcal{B}$  such that  $\tilde{b}_k \in \mathcal{K}_k[\tilde{b}_k]$ . Indeed, if  $\tilde{b}_k \in \mathcal{B}$  satisfies  $\tilde{b}_k \in \mathcal{K}_k[\tilde{b}_k]$  then a solution pair  $(u_k, m_k)$  of the discrete problem (5.1) is given by  $m_k := \mathcal{M}_k[\tilde{b}_k]$  and  $u_k := \mathcal{U}_k[m_k]$  with  $\tilde{b}_k \in \mathcal{D}_p H[u_k]$ , while the converse is obvious.

We now verify that  $\mathcal{K}_k$  is upper semi-continuous. To this end, it suffices to prove that the graph of  $\mathcal{K}_k$  is closed; c.f. [6, Ch. 1, Corollary 1, p. 42]. Let  $\mathcal{W}_k$  denote the graph of  $\mathcal{K}_k$ , which is defined by

$$\mathcal{W}_k := \left\{ (\tilde{b}, \bar{b}) \in \mathcal{B} \times \mathcal{B} : \bar{b} \in \mathcal{K}_k[\tilde{b}] \right\}. \quad (6.9)$$

Since  $\mathcal{B}$  is metrisable, to show that the graph  $\mathcal{W}_k$  is a closed it is enough to show that whenever a sequence  $\{(\tilde{b}_i, \bar{b}_i)\}_{i \in \mathbb{N}} \subset \mathcal{W}_k$  converges weakly-\* in  $\mathcal{B} \times \mathcal{B}$  to a point  $(\tilde{b}, \bar{b})$  as  $i \rightarrow \infty$ , then  $(\tilde{b}, \bar{b}) \in \mathcal{W}_k$ , which is equivalent to  $\bar{b} \in \mathcal{K}_k[\tilde{b}]$ . Let us then suppose that we are given a sequence  $\{(\tilde{b}_i, \bar{b}_i)\}_{i \in \mathbb{N}} \subset \mathcal{W}_k$  that converges weakly-\* in  $\mathcal{B} \times \mathcal{B}$  to a point  $(\tilde{b}, \bar{b})$  as  $i \rightarrow \infty$ . To begin, we claim that  $\mathcal{M}_k[\tilde{b}_i] \rightarrow \mathcal{M}_k[\tilde{b}]$  in  $L^2(0, T; L^2(\Omega))$  and  $\mathcal{M}_k[\tilde{b}_i](T) \rightarrow \mathcal{M}_k[\tilde{b}](T)$  in  $L^2(\Omega)$  as  $i \rightarrow \infty$ . Indeed, since  $\{\tilde{b}_i\}_{i \in \mathbb{N}} \subset \mathcal{B}$ , for each  $i \in \mathbb{N}$  we apply Lemma 6.1 and the norm equivalence (4.5) to obtain the uniform bound

$$\sup_{i \in \mathbb{N}} \|\mathcal{M}_k[\tilde{b}_i]\|_{\mathbb{V}_k^+} \lesssim \|m_0\|_\Omega + \|G\|_{L^2(0, T; H^{-1}(\Omega))}. \quad (6.10)$$

Since  $\mathbb{V}_k^+$  is finite dimensional, this uniform bound implies that any given subsequence  $\{\mathcal{M}_k[\tilde{b}_{i_j}]\}_{j \in \mathbb{N}}$  contains a further subsequence  $\{\mathcal{M}_k[\tilde{b}_{i_{j_s}}]\}_{s \in \mathbb{N}}$  such that

$$\mathcal{M}_k[\tilde{b}_{i_{j_s}}] \rightarrow m_k \quad \text{in } \mathbb{V}_k^+$$

as  $s \rightarrow \infty$ , for some  $m_k \in \mathbb{V}_k^+$ . In particular, it follows from the  $L^\infty(L^2)$ -bound in Lemma 4.3 that  $m_k(T) = \lim_{s \rightarrow \infty} \mathcal{M}_k[\tilde{b}_{i_{j_s}}](T)$  in  $L^2(\Omega)$  and  $m_k(0) = R_k m_0$  in  $V_k$ . We then pass to the limit in the equation (6.6) satisfied by  $\mathcal{M}_k[\tilde{b}_{i_{j_s}}]$  to find that  $m_k \in \mathbb{V}_k^+$  satisfies

$$\int_0^T (\partial_t \mathcal{I}_+ m_k, \phi)_{\Omega, k} + (A_k \nabla m_k, \nabla \phi)_{\Omega} + (m_k, \tilde{b} \cdot \nabla \phi)_{\Omega} dt = \int_0^T \langle G, \phi \rangle dt \quad \forall \phi \in \mathbb{V}_k \quad (6.11)$$

with  $m_k(0) = R_k m_0$  in  $V_k$ . But we then see from the definition of  $\mathcal{M}_k[\tilde{b}]$  in (6.6) that  $m_k = \mathcal{M}_k[\tilde{b}]$  in  $\mathbb{V}_k^+$ . It follows that the entire sequence  $\{\mathcal{M}_k[\tilde{b}_i]\}_{i \in \mathbb{N}}$  satisfies  $\mathcal{M}_k[\tilde{b}_i] \rightarrow \mathcal{M}_k[\tilde{b}]$  in  $\mathbb{V}_k^+$  and  $\mathcal{M}_k[\tilde{b}_i](T) \rightarrow \mathcal{M}_k[\tilde{b}](T)$  in  $L^2(\Omega)$  as  $i \rightarrow \infty$ . In particular, the strong convergence of  $\{\mathcal{M}_k[\tilde{b}_i]\}_{i \in \mathbb{N}}$  to  $\mathcal{M}_k[\tilde{b}]$  in  $\mathbb{V}_k^+$ , together with the  $L^\infty(L^2)$ -bound in Lemma 4.3, implies that  $\mathcal{M}_k[\tilde{b}_i] \rightarrow \mathcal{M}_k[\tilde{b}]$  in  $L^2(0, T; L^2(\Omega))$  as  $i \rightarrow \infty$ . We deduce from continuity of  $S : L^2(\Omega) \rightarrow L^2(\Omega)$  that  $S[\mathcal{M}_k[\tilde{b}_i](T)] \rightarrow S[\mathcal{M}_k[\tilde{b}](T)]$  in  $L^2(\Omega)$  as  $i \rightarrow \infty$ , while the continuity of  $F : L^2(0, T; L^2(\Omega)) \rightarrow L^2(0, T; H^{-1}(\Omega))$  implies that  $F[\mathcal{M}_k[\tilde{b}_i]] \rightarrow F[\mathcal{M}_k[\tilde{b}]]$  in  $L^2(0, T; H^{-1}(\Omega))$  as  $i \rightarrow \infty$ . We then apply the continuous dependence result of Lemma 6.2 to conclude that  $\mathcal{U}_k[\mathcal{M}_k[\tilde{b}_i]] \rightarrow \mathcal{U}_k[\mathcal{M}_k[\tilde{b}]]$  in  $\mathbb{V}_k^-$  as  $i \rightarrow \infty$  with  $\mathcal{U}_k[\mathcal{M}_k[\tilde{b}]](T) = R_k S[\mathcal{M}_k[\tilde{b}](T)]$  in  $V_k$ . By hypothesis,  $\bar{b}_i \in \mathcal{K}_k[\tilde{b}_i] = \mathcal{D}_p H[\mathcal{U}_k[\mathcal{M}_k[\tilde{b}_i]]]$  for  $i \in \mathbb{N}$  and  $\bar{b}_i \rightharpoonup^* \bar{b}$  in  $L^\infty(Q_T; \mathbb{R}^d)$  as  $i \rightarrow \infty$ . In particular,  $\bar{b}_i \rightharpoonup \bar{b}$  in  $L^2(Q_T; \mathbb{R}^d)$  as  $i \rightarrow \infty$  since  $\{\bar{b}_i\}_{i \in \mathbb{N}} \subset \mathcal{B} \subset L^2(Q_T; \mathbb{R}^d)$  and  $|\mathcal{Q}_T|_{d+1} < \infty$ . We conclude from Lemma 2.2 that  $\bar{b} \in \mathcal{D}_p H[\mathcal{U}_k[\mathcal{M}_k[\tilde{b}]]]$ , i.e.  $\bar{b} \in \mathcal{K}_k[\tilde{b}]$ . We have therefore shown that  $\mathcal{W}_k$  is closed, so  $\mathcal{K}_k$  is upper semi-continuous.

We have thus shown that the map  $\mathcal{K}_k : \mathcal{B} \rightrightarrows \mathcal{B}$  satisfies the conditions of Kakutani's fixed point theorem, so  $\mathcal{K}_k$  admits a fixed point and therefore there exists a solution to the discrete problem (5.1), as required.

Finally, to deduce the stability bound (5.3) satisfied by solutions  $(u_k, m_k)$  to the FEM (5.1), we first observe that Lemma 6.1 and the discrete KFP equation satisfied by  $m_k$  with drift in  $\mathcal{B}$  imply that

$$\sup_{k \in \mathbb{N}} \|m_k\|_{\mathbb{V}_k^+} \lesssim \|m_0\|_{\Omega} + \|G\|_{L^2(0, T; H^{-1}(\Omega))}.$$

This bound, the linear growth of  $F$  and  $S$ , together with the  $L^\infty(L^2)$ -bound in Lemma 4.3, allow us to apply (6.4) to obtain

$$\sup_{k \in \mathbb{N}} \|u_k\|_{\mathbb{V}_k^-} \lesssim \|m_0\|_{\Omega} + \|G\|_{L^2(0, T; H^{-1}(\Omega))} + 1.$$

We then deduce the stability bound (5.3), as desired.  $\square$

### 6.3 Proof of Theorem 5.2

Suppose that  $(u_{k,i}, m_{k,i}) \in \mathbb{V}_k \times \mathbb{V}_k^+$ ,  $i \in \{1, 2\}$  are solutions of (5.1), for some respective  $\tilde{b}_{k,i} \in \mathcal{D}_p H[u_{k,i}]$ . It follows from the nonnegativity of  $m_0$  and of  $G$  and from the discrete maximum principle (see Theorem 4.2), that  $m_{k,i} \geq 0$  a.e. in  $Q_T$  for each  $i \in \{1, 2\}$ . Observe also that  $m_{k,1}(0) = m_{k,2}(0)$  so that  $m_{k,1} - m_{k,2} \in \mathbb{V}_{k,0}^+$ . Note that  $(R_k S[m_{k,i}(T)], m_{k,j}(T))_{\Omega, k} = (S[m_{k,i}(T)], m_k^{(j)}(T))_{\Omega}$  for all  $i, j \in \{1, 2\}$

by (4.7). Then, proceeding similarly to the proof of Theorem 3.2, we test (5.1a) with  $\psi = m_k^{(1)} - m_k^{(2)} \in \mathbb{V}_{k,0}^+$  and (5.1b) with  $\phi = u_{k,1} - u_{k,2}$  to find that

$$\begin{aligned} \int_0^T \int_{\Omega} m_1 \lambda_{k,1,2} + m_2 \lambda_{k,2,1} dx dt &= \int_0^T \langle F[m_{k,1}] - F[m_{k,2}], m_{k,1} - m_{k,2} \rangle dt \\ &\quad + (S[m_{k,1}(T)] - S[m_{k,2}(T)], m_{k,1}(T) - m_{k,2}(T))_{\Omega}, \end{aligned} \quad (6.12)$$

where

$$\lambda_{k,i,j} := -H[\nabla u_{k,j}] + H[\nabla u_{k,i}] + \tilde{b}_{k,i} \cdot \nabla(u_{k,j} - u_{k,i}), \quad i, j \in \{1, 2\}. \quad (6.13)$$

Note that  $\lambda_{k,i,j} \leq 0$  a.e. in  $Q_T$  for all  $i, j \in \{1, 2\}$  since  $\tilde{b}_{k,i} \in \mathcal{D}_p H[u_{k,i}]$ . Therefore  $\int_0^T \int_{\Omega} m_1 \lambda_{k,1,2} + m_2 \lambda_{k,2,1} dx dt \leq 0$ , and so we conclude from strict monotonicity of  $F$  and monotonicity of  $S$  applied to (6.12) that  $m_{k,1} = m_{k,2}$  a.e. in  $Q_T$ . Since functions in  $\mathbb{V}_k^+$  are left-continuous, it follows that  $m_{k,1}(T) = m_{k,2}(T)$ . Therefore  $u_{k,1}$  and  $u_{k,2}$  both satisfy (5.1a) with common right hand side, so Lemma 6.2 implies that  $u_{k,1} = u_{k,2}$ .  $\square$

## 7 Discrete functional analysis

In this section we establish auxiliary compactness results that will be used in the convergence analysis of the finite element scheme (5.1). We begin by collecting some fundamental weak convergence results for the spaces  $\mathbb{V}_k^{\pm}$ .

**Theorem 7.1** (Weak convergence properties) *Let  $\{v_k\}_{k \in \mathbb{N}}$  be a sequence such that  $v_k \in \mathbb{V}_k^+$  for each  $k \in \mathbb{N}$  and  $\sup_{k \in \mathbb{N}} \|v_k\|_{\mathbb{V}_k^+} < \infty$ . Then, there exists a  $v \in Y$  and a subsequence  $\{v_{k_j}\}_{j \in \mathbb{N}}$  such that*

$$\begin{aligned} v_{k_j} &\rightharpoonup v \text{ in } X, & \partial_t \mathcal{I}_+ v_{k_j} &\rightharpoonup \partial_t v \text{ in } L^2(0, T; H^{-1}(\Omega)), \\ v_{k_j}(T) &\rightharpoonup v(T) \text{ in } L^2(\Omega), & v_{k_j}(0) &\rightharpoonup v(0) \text{ in } L^2(\Omega), \end{aligned} \quad (7.1)$$

as  $j \rightarrow \infty$ . Furthermore, for any fixed  $w_{\ell} \in \mathbb{V}_{\ell}$ ,  $\ell \in \mathbb{N}$ , there holds

$$\lim_{j \rightarrow \infty} \int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_{\ell})_{\Omega, k_j} dt = \int_0^T \langle \partial_t v, w_{\ell} \rangle dt. \quad (7.2)$$

The properties (7.1) and (7.2) hold analogously for sequences  $\{\tilde{v}_k\}_{k \in \mathbb{N}}$  such that  $\tilde{v}_k \in \mathbb{V}_k^-$  for each  $k \in \mathbb{N}$  and  $\sup_{k \in \mathbb{N}} \|\tilde{v}_k\|_{\mathbb{V}_k^-} < \infty$ , with  $\mathcal{I}_-$  in place of  $\mathcal{I}_+$ .



**Proof** We give the proof only for the case of a sequence in  $\mathbb{V}_k^+$  as the proof for the case of  $\mathbb{V}_k^-$  is very similar. The hypothesis  $\sup_{k \in \mathbb{N}} \|v_k\|_{\mathbb{V}_k^+} < \infty$  implies that the sequence  $\{v_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $X$ , and also that  $\{\partial_t \mathcal{I}_+ v_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^2(0, T; H^{-1}(\Omega))$  by Lemma 4.1. Furthermore, Lemma 4.3 implies that  $\{v_k(T)\}_{k \in \mathbb{N}}$  and  $\{v_k(0)\}_{k \in \mathbb{N}}$  are uniformly bounded sequences in  $L^2(\Omega)$ . So, there exists a subsequence such that  $v_{k_j} \rightharpoonup v$  in  $X$ , that  $\partial_t \mathcal{I}_+ v_{k_j} \rightharpoonup \varphi$  in  $L^2(0, T; H^{-1}(\Omega))$ , that  $v_{k_j}(T) \rightharpoonup v_T$  and  $v_{k_j}(0) \rightharpoonup v_0$  as  $j \rightarrow \infty$ , for some  $v \in X$ , some  $\varphi \in L^2(0, T; H^{-1}(\Omega))$ , and some  $v_T, v_0 \in L^2(\Omega)$ . To show that  $\varphi = \partial_t v$  so that  $v \in Y$ , and that  $v(T) = v_T$  and  $v(0) = v_0$ , let  $\phi \in C^1([0, T]; H_0^1(\Omega))$  be arbitrary. Then, using integration-by-parts for the piecewise constant function  $v_k \in \mathbb{V}_k^+$  and the identities in (4.19), we find that

$$\begin{aligned} \int_0^T \langle \partial_t \mathcal{I}_+ v_{k_j}, \phi \rangle + (v_{k_j}, \partial_t \phi)_\Omega dt &= (v_{k_j}(T), \phi(T))_\Omega - (v_{k_j}(0), \phi(0))_\Omega \\ &\quad + \sum_{n=1}^{N_{k_j}} \int_{I_n} (\partial_t \mathcal{I}_+ v_{k_j}, \phi(t) - \phi(t_{n-1}))_\Omega dt. \end{aligned} \quad (7.3)$$

It is straightforward to show that the last term on the right-hand side of (7.3) vanishes in the limit  $j \rightarrow \infty$  since  $\{\partial_t \mathcal{I}_+ v_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^2(0, T; H^{-1}(\Omega))$  and since  $\phi \in C^1([0, T]; H_0^1(\Omega))$ . Therefore, passing to the limit in (7.3), we get

$$\int_0^T \langle \varphi, \phi \rangle + (v, \partial_t \phi)_\Omega dt = (v_T, \phi(T))_\Omega - (v_0, \phi(0))_\Omega \quad \forall \phi \in C^1([0, T]; H_0^1(\Omega)). \quad (7.4)$$

It is then easy to deduce from (7.4) that  $v \in Y$  with  $\partial_t v = \varphi \in L^2(0, T; H^{-1}(\Omega))$ , and moreover that  $v(T) = v_T$  and that  $v(0) = v_0$ .

To show (7.2), let  $v_\ell \in \mathbb{V}_\ell$ ,  $\ell \in \mathbb{N}$  be fixed but arbitrary. Using (4.7), we have

$$\begin{aligned} \int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell)_{\Omega, k_j} dt &= \int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell)_\Omega dt \\ &\quad + \int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell - R_{k_j} w_\ell)_{\Omega, k_j} dt. \end{aligned} \quad (7.5)$$

It follows from the weak convergence properties above that  $\int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell)_\Omega dt \rightarrow \int_0^T \langle \partial_t v, w_\ell \rangle dt$  as  $j \rightarrow \infty$ . Recall that the function  $v_\ell$  is piecewise constant-in-time with respect to the partition  $\mathcal{J}_\ell$  of  $[0, T]$ . In a slight abuse of notation, let  $\{t_q\}_{q=0}^{M_\ell}$

denote the nodes of  $\mathcal{J}_\ell$ , i.e.  $\mathcal{J}_\ell = \{[t_{q-1}, t_q]\}_{q=0}^{M_\ell}$ . Therefore,

$$\int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell - R_{k_j} w_\ell)_{\Omega, k_j} dt = \sum_{q=1}^{M_\ell} (\mathcal{I}_+ v_{k_j}(t_q) - \mathcal{I}_+ v_{k_j}(t_{q-1}), z_{j, \ell, q})_{\Omega, k_j}, \quad (7.6)$$

where  $z_{j, \ell, q} := w_\ell|_{(t_{q-1}, t_q)} - R_{k_j}(w_\ell|_{(t_{q-1}, t_q)})$ . The  $L^2$ -approximation bound in (4.8) and (4.5) together imply that  $\|z_{j, \ell, q}\|_{\Omega, k} \lesssim h_{k_j} \|\nabla w_\ell|_{(t_{q-1}, t_q)}\|_\Omega$  for all  $q = 1, \dots, M_j$  and all  $j \in \mathbb{N}$ . Also, Lemma 4.3 and (4.5) imply that  $\|\mathcal{I}_+ v_{k_j}(t_q)\|_{\Omega, k} \lesssim \|v_{k_j}\|_{\mathbb{V}_k^+}$ , so there exists a constant  $C_*$  independent of  $k \in \mathbb{N}$  such that  $\|\mathcal{I}_+ v_{k_j}(t_q)\|_{\Omega, k_j} \leq C_*$  for all  $j \in \mathbb{N}$  and all  $q \in \{0, \dots, N_\ell\}$ . Applying these bounds in (7.6), we see that

$$\lim_{j \rightarrow \infty} \left| \int_0^T (\partial_t \mathcal{I}_+ v_{k_j}, w_\ell - R_{k_j} w_\ell)_{\Omega, k_j} dt \right| \lesssim \lim_{j \rightarrow \infty} C_* h_{k_j} \sum_{q=1}^{M_\ell} \|\nabla w_\ell|_{(t_{q-1}, t_q)}\|_\Omega = 0. \quad (7.7)$$

Therefore we obtain (7.2).  $\square$

The following Lemma generalizes a well-known inequality for real-valued functions of bounded variation to the case of functions in  $\mathbb{V}_k^\pm$ .

**Lemma 7.2** *Let  $\|\cdot\| : V_k \rightarrow \mathbb{R}_{\geq 0}$  be a norm on  $V_k$ . Then, for all  $v \in \mathbb{V}_k^\pm$ , we have*

$$\sup_{s \in (0, T)} \int_0^{T-s} s^{-1} \|v(t+s) - v(t)\| dt \leq \int_0^T \|\partial_t \mathcal{I}_\pm v\| dt. \quad (7.8)$$

**Proof** We show the proof for the case of  $v \in \mathbb{V}_k^+$ , as the argument for  $\mathbb{V}_k^-$  is quite similar. Let  $s \in (0, T)$  be arbitrary, and let  $\underline{\sigma}, \bar{\sigma} \in \mathbb{R}$  be any positive real numbers such that  $0 < \underline{\sigma} < \bar{\sigma} < T$  and  $\bar{\sigma} - s > \underline{\sigma}$ . For each  $\epsilon > 0$  such that  $\epsilon < \min(\underline{\sigma}, T - \bar{\sigma})$ , let  $\eta_\epsilon = \frac{1}{\epsilon} \eta_0(\frac{\cdot}{\epsilon})$  where  $\eta_0$  is the standard mollifier on  $\mathbb{R}$ , so  $\eta_\epsilon \in C_0^\infty(\mathbb{R})$  is a smooth nonnegative function with  $\text{supp } \eta_\epsilon \subset [-\epsilon, \epsilon]$  and  $\int_{\mathbb{R}} \eta_\epsilon(t) dt = 1$ . For  $v \in \mathbb{V}_k^+$ , define  $v_\epsilon : [\underline{\sigma}, \bar{\sigma}] \rightarrow V_k$  by  $v_\epsilon(t) := (v * \eta_\epsilon)(t) = \int_0^T v(s) \eta_\epsilon(t-s) ds$ . Note that  $v_\epsilon$  is well defined on  $[\underline{\sigma}, \bar{\sigma}]$  and depends only on  $v|_{(0, T)}$ . Since  $\eta(t - \cdot) \in C_0^\infty(0, T)$  for all  $t \in [\underline{\sigma}, \bar{\sigma}]$ , integration-by-parts shows that  $\partial_t v_\epsilon(t) = \sum_{n=1}^{N_k-1} \llbracket v \rrbracket_n \eta_\epsilon(t - t_n)$  for all  $t \in [\underline{\sigma}, \bar{\sigma}]$ . Therefore, the triangle inequality leads to

$$\int_{\underline{\sigma}}^{\bar{\sigma}} \|\partial_t v_\epsilon\| dt \leq \sum_{n=1}^{N_k-1} \|\llbracket v \rrbracket_n\| \int_{\underline{\sigma}}^{\bar{\sigma}} \eta_\epsilon(t - t_n) dt \leq \sum_{n=1}^{N_k-1} \|\llbracket v \rrbracket_n\| \leq \int_0^T \|\partial_t \mathcal{I}_+ v\| dt. \quad (7.9)$$

Furthermore,  $\int_{\underline{\sigma}}^{\bar{\sigma}} \|v - v_\epsilon\| dt \rightarrow 0$  as  $\epsilon \rightarrow 0$ . Therefore,

$$\begin{aligned} \int_{\underline{\sigma}}^{\bar{\sigma}-s} \|v(t+s) - v(t)\| dt &= \lim_{\epsilon \rightarrow 0} \int_{\underline{\sigma}}^{\bar{\sigma}-s} \|v_\epsilon(t+s) - v_\epsilon(t)\| dt \\ &\leq \limsup_{\epsilon \rightarrow 0} \int_{\underline{\sigma}}^{\bar{\sigma}-s} \int_t^{t+s} \|\partial_t v_\epsilon(r)\| dr dt \leq \limsup_{\epsilon \rightarrow 0} s \int_{\underline{\sigma}}^{\bar{\sigma}} \|\partial_t v_\epsilon\| dr \leq s \int_0^T \|\partial_t \mathcal{I}_+ v\| dr, \end{aligned} \quad (7.10)$$

where we invoke Fubini's Theorem and (7.9) in the second line above. We then readily conclude (7.8) by letting  $\underline{\sigma} \rightarrow 0$  and  $\bar{\sigma} \rightarrow T$ , and recalling that  $s \in (0, T)$  is arbitrary.  $\square$

A direct consequence of Lemma 7.2 and the definition of  $\|\cdot\|_{\mathbb{V}_k^\pm}$  in (4.21) is that

$$\sup_{s \in (0, T)} \int_0^{T-s} s^{-1} \|v(t+s) - v(t)\|_{k,*} dt \leq \sqrt{T} \|v\|_{\mathbb{V}_k^\pm} \quad \forall v \in \mathbb{V}_k^\pm. \quad (7.11)$$

**Theorem 7.3** (Compactness in  $L^p(L^2)$ ) *Let  $\{v_k\}_{k \in \mathbb{N}} \subset X$  be a sequence such that  $v_k \in \mathbb{V}_k^+$  for each  $k \in \mathbb{N}$  and  $\sup_{k \in \mathbb{N}} \|v_k\|_{\mathbb{V}_k^+} < \infty$ . Then,  $\{v_k\}_{k \in \mathbb{N}}$  is pre-compact in  $L^p(0, T; L^2(\Omega))$  for any  $1 \leq p < \infty$ . This result holds analogously for sequences  $\{\tilde{v}_k\}_{k \in \mathbb{N}}$  such that  $\tilde{v}_k \in \mathbb{V}_k^-$  for each  $k \in \mathbb{N}$  and  $\sup_{k \in \mathbb{N}} \|\tilde{v}_k\|_{\mathbb{V}_k^-} < \infty$ .*

**Proof** It is enough to prove the result in the case of  $\mathbb{V}_k^+$  by symmetry of  $\mathbb{V}_k^+$  and  $\mathbb{V}_k^-$  under reversal of the time variable. We show that  $\{v_k\}_{k \in \mathbb{N}}$  satisfies the hypotheses of [38, Theorem 3], which asserts that  $\{v_k\}_{k \in \mathbb{N}}$  is pre-compact in  $L^p(0, T; L^2(\Omega))$  provided that  $\{v_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^1(0, T; B_1)$  for some Banach space  $B_1$  that is compactly embedded in  $L^2(\Omega)$ , and that for all  $\epsilon > 0$ , there is an  $s_0 > 0$  such that for all  $s < s_0$ ,  $\|v_k(\cdot + s) - v_k(\cdot)\|_{L^p(0, T-s; L^2(\Omega))} < \epsilon$  for all  $k \in \mathbb{N}$ . The first hypothesis holds immediately since  $\{v_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $X = L^2(0, T; H_0^1(\Omega))$  with  $H_0^1(\Omega)$  compactly embedded in  $L^2(\Omega)$ . To show the second hypothesis, let  $M_* := \sup_{k \in \mathbb{N}} \|v_k\|_{\mathbb{V}_k^+}$ , and let  $r_{k,s} := v_k(\cdot + s) - v_k$  for each  $k \in \mathbb{N}$  and  $s \in (0, T)$ . Then, we find that

$$\begin{aligned} \int_0^{T-s} \|r_{k,s}\|_{\Omega}^p dt &\leq 2^{p-1} M_*^{p-1} \int_0^{T-s} \|r_{k,s}\|_{\Omega} dt \leq 2^{p-1} M_*^{p-1} \int_0^{T-s} \|r_{k,s}\|_{\Omega,k} dt \\ &\leq 2^{p-1} M_*^{p-1} \left( \int_0^{T-s} \|\nabla r_{k,s}\|_{\Omega} dt \right)^{\frac{1}{2}} \left( \int_0^{T-s} \|r_{k,s}\|_{k,*} dt \right)^{\frac{1}{2}} \leq 2^{p-\frac{1}{2}} \sqrt{T} M_*^p s^{\frac{1}{2}}, \end{aligned} \quad (7.12)$$

where we have applied first Lemma 4.3, then the first inequality of (4.5), followed by the Cauchy–Schwarz inequality and finally (7.11). Thus we obtain from (7.12) that  $\|v_k(\cdot + s) - v_k(\cdot)\|_{L^p(0, T-s, L^2(\Omega))} \lesssim s^{\frac{1}{2p}} M_*$  for all  $k \in \mathbb{N}$ , and all  $s \in (0, T)$ . This verifies the second hypothesis above and thus completes the proof.  $\square$

## 8 Proofs of Theorems 3.1 and 5.3 and of Corollary 5.4

In this section, we assume that the hypotheses of Theorem 3.2 hold, i.e. that  $m_0$  is nonnegative a.e. in  $\Omega$ , that  $G$  is nonnegative in the sense of distributions in  $L^2(0, T; H^{-1}(\Omega))$ , that  $F$  is strictly monotone, and that  $S$  is monotone. We now prove convergence of the finite element approximations  $\{(u_k, m_k)\}_{k \in \mathbb{N}}$  that are determined by the scheme (5.1).

### 8.1 KFP equation

For each  $k \in \mathbb{N}$ , let  $\tilde{b}_k$  be an element of  $\mathcal{D}_p H[u_k]$  for which (5.1) holds; in cases where  $\tilde{b}_k$  may be nonunique, we may choose one such element.

**Lemma 8.1** (Weak convergence of discrete KFP approximations) *There exists an  $m \in Y$  and a  $\tilde{b}_* \in L^\infty(Q_T; \mathbb{R}^d)$  such that, after passing to a subsequence without change of notation,  $\partial_t \mathcal{I}_+ m_k \rightharpoonup \partial_t m$  in  $L^2(0, T; H^{-1}(\Omega))$ ,  $m_k \rightharpoonup m$  in  $X$ ,  $m_k(T) \rightharpoonup m(T)$  in  $L^2(\Omega)$ ,  $m_k \rightarrow m$  in  $L^p(0, T; L^2(\Omega))$  for all  $p \in [1, \infty)$ , and  $\tilde{b}_k \rightharpoonup \tilde{b}_*$  in  $L^2(Q_T; \mathbb{R}^d)$  as  $k \rightarrow \infty$ . Furthermore, we have  $m(0) = m_0$  in  $L^2(\Omega)$  and*

$$\int_0^T \langle \partial_t m, \phi \rangle + \nu(\nabla m, \nabla \phi)_\Omega + (m \tilde{b}_*, \nabla \phi)_\Omega dt = \int_0^T \langle G, \phi \rangle dt \quad (8.1)$$

for all  $\phi \in X$ .

**Proof** Theorem 7.1 shows the existence of a  $m \in Y$  and a subsequence of  $\{m_k\}_{k \in \mathbb{N}}$ , to which we pass without change of notation, that satisfies the weak convergence properties stated above. In particular, we have  $m_k(0) = R_k m_0 \rightharpoonup m(0)$  as  $k \rightarrow \infty$ , and since  $R_k m_0 \rightarrow m_0$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$ , we conclude that  $m(0) = m_0$ . Furthermore, the strong convergence  $m_k \rightarrow m$  in  $L^p(0, T; L^2(\Omega))$  for all  $p \in [1, \infty)$  as  $k \rightarrow \infty$  follows from Theorem 7.3. Since  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^\infty(Q_T; \mathbb{R}^d)$  by Lemma 2.1, and hence also in  $L^2(Q_T; \mathbb{R}^d)$ , we may pass to a further subsequence without change of notation to find that  $\tilde{b}_k \rightharpoonup \tilde{b}_*$  for some  $\tilde{b}_* \in L^2(Q_T; \mathbb{R}^d)$ . Mazur's Theorem and the Riesz–Fischer Theorem further imply that  $\tilde{b}_* \in L^\infty(Q_T; \mathbb{R}^d)$  with  $\|\tilde{b}_*\|_{L^\infty(Q_T; \mathbb{R}^d)} \leq L_H$ . It remains only to show (8.1). Let  $\ell \in \mathbb{N}$  be fixed and let  $\phi \in \mathbb{V}_\ell$  be fixed but arbitrary. Recall that  $\mathbb{V}_\ell \subset \mathbb{V}_k$  for all  $\ell \leq k$ . Then, we have  $m_k \nabla \phi \rightarrow m \nabla \phi$  in  $L^2(Q_T; \mathbb{R}^d)$  as  $k \rightarrow \infty$  since  $\nabla \phi \in L^\infty(Q_T; \mathbb{R}^d)$ . This implies that  $\int_0^T (m_k, \tilde{b}_k \cdot \nabla \phi)_\Omega dt \rightarrow \int_0^T (m, \tilde{b}_* \cdot \nabla \phi)_\Omega dt$  by weak-times-strong convergence. Next, we have  $\int_0^T (\partial_t \mathcal{I}_+ m_k, \phi)_\Omega dt \rightarrow \int_0^T (\partial_t m, \phi)_\Omega dt$  by (7.2). We also have  $A_k \rightarrow$

$v\mathbb{I}$  as  $k \rightarrow \infty$  in  $L^\infty(\Omega; \mathbb{R}^{d \times d})$  by Theorem 4.2. Therefore we may pass to the limit in (5.1b) to obtain (8.1) for all  $\phi \in \mathbb{V}_\ell$ . Since  $\bigcup_{\ell \in \mathbb{N}} \mathbb{V}_\ell$  is dense in  $X$ , we deduce that (8.1) holds for all  $\phi \in X$ .  $\square$

To establish the strong  $L^2$ -convergence of the density approximations at terminal time, we find it helpful to first establish some compactness properties associated to a class of related discrete dual problems.

**Discrete dual problems** For each  $k \in \mathbb{N}$ , we define  $z_k \in \mathbb{V}_k^-$  as the unique solution of the discrete dual problem

$$\int_0^T -(\partial_t \mathcal{I} z_k, v)_{\Omega, k} + (A_k \nabla z_k, \nabla v)_\Omega + (\tilde{b}_k \cdot \nabla z_k, v)_\Omega dt = 0 \quad \forall v \in \mathbb{V}_k, \quad (8.2)$$

with the condition  $z_k(T) = m_k(T)$  in  $V_k$ .

**Lemma 8.2** (Weak convergence of discrete dual solutions) *After passing to a subsequence without change of notation, we have  $z_k \rightharpoonup z$  in  $X$ ,  $\partial_t \mathcal{I} z_k \rightharpoonup \partial_t z$  in  $L^2(0, T; H^{-1}(\Omega))$ ,  $z_k \rightarrow z$  in  $L^2(0, T; L^2(\Omega))$  and  $z_k(0) \rightarrow z(0)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$  where  $z \in Y$  satisfies*

$$\int_0^T -\langle \partial_t z, \phi \rangle + v(\nabla z, \nabla \phi)_\Omega + (\tilde{b}_* \cdot \nabla z, \phi)_\Omega dt = 0 \quad \forall \phi \in X, \quad (8.3)$$

with  $z(T) = m(T)$  in  $L^2(\Omega)$ , where  $\tilde{b}_* \in L^\infty(Q_T; \mathbb{R}^d)$  is as in Lemma 8.1.

**Proof** The fact that  $\|m_k(T)\|_\Omega \lesssim \|m_k\|_{\mathbb{V}_k^+}$ , see Lemma 4.3, and the fact that  $\sup_{k \in \mathbb{N}} \|m_k\|_{\mathbb{V}_k^+} < \infty$  together imply that  $\sup_{k \in \mathbb{N}} \|z_k\|_{\mathbb{V}_k^-} < \infty$ . Theorem 7.1 implies that, after passing to a subsequence without change of notation, there exists a  $z \in Y$  such that  $z_k \rightharpoonup z$  in  $X$ ,  $\partial_t \mathcal{I} z_k \rightharpoonup \partial_t z$  in  $L^2(0, T; H^{-1}(\Omega))$ ,  $z_k(T) \rightarrow z(T)$  and  $z_k(0) \rightarrow z(0)$  as  $k \rightarrow \infty$ . Moreover, Theorem 7.3 implies that  $z_k \rightarrow z$  in  $L^2(0, T; L^2(\Omega))$ . Note that  $z(T) = m(T)$  since  $z_k(T) = m_k(T) \rightarrow m(T)$  as  $k \rightarrow \infty$  by Lemma 8.1. Since  $\{\tilde{b}_k \cdot \nabla z_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^2(0, T; L^2(\Omega))$ , after possibly passing to a further subsequence without change of notation, there exists a  $g \in L^2(0, T; L^2(\Omega))$  such that  $\tilde{b}_k \cdot \nabla z_k \rightarrow g$  as  $k \rightarrow \infty$ . After fixing a test function  $v \in \mathbb{V}_j$  in (8.2) for some fixed  $j \leq k$ , we use the weak convergence properties of  $z_k$  above, along with (7.2) and  $\|A_k - v\mathbb{I}\|_{L^\infty(\Omega; \mathbb{R}^{d \times d})} \lesssim h_k$ , to pass to the limit  $k \rightarrow \infty$  in (8.2) to find that  $z \in Y$  satisfies

$$\int_0^T -\langle \partial_t z, v \rangle + v(\nabla z, \nabla v)_\Omega dt = - \int_0^T (g, v)_\Omega dt \quad \forall v \in \mathbb{V}_j. \quad (8.4)$$

Since  $\bigcup_{j \in \mathbb{N}} \mathbb{V}_j$  is dense in  $X$ , we see that (8.4) further holds for all  $v \in X$ .

In order to show that  $g = \tilde{b}_* \cdot \nabla z$  so that (8.3) holds, it is enough to show that

$$\lim_{k \rightarrow \infty} \int_0^T (T-t) \|\nabla(z - z_k)\|_{\Omega}^2 dt = 0. \quad (8.5)$$

Indeed, supposing momentarily that (8.5) is known, we conclude first that  $z_k \rightarrow z$  in  $L^2(0, T - \epsilon, H_0^1(\Omega))$  for all  $\epsilon > 0$ . Weak-times-strong convergence then implies that  $\tilde{b}_k \cdot \nabla z_k \rightarrow \tilde{b}_* \cdot \nabla z$  in  $L^2(0, T - \epsilon, L^2(\Omega))$  for all  $\epsilon \in (0, T)$ , so that  $g = \tilde{b}_* \cdot \nabla z$  on  $(0, T - \epsilon)$  for arbitrarily small  $\epsilon$ , and thus  $g = \tilde{b}_* \cdot \nabla z$  on  $(0, T)$ .

In order to prove (8.5), define

$$w_k|_{I_j} := (T - t_j)z_k|_{I_j}, \quad \tilde{w}_k|_{I_j} := (T - t_j)^{\frac{1}{2}} z_k|_{I_j} \quad \forall j \in \{1, \dots, N_k\}. \quad (8.6)$$

Since  $z_k$  converges to  $z$  strongly in  $L^2(0, T; L^2(\Omega))$  and weakly in  $X = L^2(0, T, H_0^1(\Omega))$ , it is easy to show that  $w_k \rightarrow w$  strongly in  $L^2(0, T; L^2(\Omega))$  where  $w(t) := (T - t)z \in X$ , and also that  $\tilde{w}_k \rightarrow \tilde{w}$  in  $X$  where  $\tilde{w} := (T - t)^{\frac{1}{2}} w$ . Note that  $\tilde{w}_k \rightarrow \tilde{w}$  in  $X$  and  $z_k(0) \rightarrow z(0)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$  respectively imply that

$$\|\tilde{w}\|_X^2 \leq \liminf_{k \rightarrow \infty} \|\tilde{w}_k\|_X^2, \quad \|z(0)\|_{\Omega}^2 \leq \liminf_{k \rightarrow \infty} \|z_k(0)\|_{\Omega}^2. \quad (8.7)$$

We also deduce from strong convergence of  $z_k \rightarrow z$  in  $L^2(0, T; L^2(\Omega))$  that

$$\lim_{k \rightarrow \infty} \int_0^T \|z_k\|_{\Omega, k}^2 dt = \int_0^T \|z\|_{\Omega}^2 dt. \quad (8.8)$$

Indeed (8.8) follows from  $\|z_k\|_{\Omega, k}^2 = \|z_k\|_{\Omega}^2 + (z_k, z_k - R_k z_k)_{\Omega, k}$  a.e. on  $(0, T)$  by (4.7), and from  $|(z_k, z_k - R_k z_k)_{\Omega, k}| \lesssim h_k \|\nabla z_k\|_{\Omega}^2$  a.e. on  $(0, T)$  by (4.5), (4.8) and the Poincaré inequality. Then, a straightforward calculation gives

$$\frac{T}{2} \|z_k(0)\|_{\Omega, k}^2 + \sum_{j=1}^{N_k} \frac{T - t_j}{2} \|\llbracket z_k \rrbracket_j\|_{\Omega, k}^2 = \int_0^T -(\partial_t \mathcal{I}_- z_k, w_k)_{\Omega, k} + \frac{1}{2} \|z_k\|_{\Omega, k}^2 dt. \quad (8.9)$$

Thus, testing (8.2) with  $w_k \in \mathbb{V}_k$  and re-arranging terms shows that

$$\begin{aligned} \frac{T}{2} \|z_k(0)\|_{\Omega}^2 + \nu \|\tilde{w}_k\|_X^2 &\leq \frac{T}{2} \|z_k(0)\|_{\Omega, k}^2 + \int_0^T (A_k \nabla z_k, \nabla w_k)_{\Omega} dt \\ &\leq \int_0^T -(\partial_t \mathcal{I}_- z_k, w_k)_{\Omega, k} + (A_k \nabla z_k, \nabla w_k)_{\Omega} + \frac{1}{2} \|z_k\|_{\Omega, k}^2 dt \end{aligned}$$

$$= \int_0^T -(\tilde{b}_k \cdot \nabla z_k, w_k)_\Omega + \frac{1}{2} \|z_k\|_{\Omega, k}^2 dt. \quad (8.10)$$

Note that in the first inequality above we have used (4.5) as well as the inequality  $\nu \|\nabla \tilde{w}_k\|_\Omega^2 \leq (A_k \nabla z_k, \nabla w_k)_\Omega$  since  $A_k \geq \nu \mathbb{I}$  in the partial ordering of positive semi-definite matrices. We then combine the strong convergence  $w_k \rightarrow w = (T - t)z$  in  $L^2(0, T; L^2(\Omega))$  with weak convergence  $\tilde{b}_k \cdot \nabla z_k \rightarrow g$  in  $L^2(0, T; L^2(\Omega))$ , and (8.8) to pass to the limit in the last line of (8.10) to find that

$$\begin{aligned} \lim_{k \rightarrow \infty} \int_0^T -(\tilde{b}_k \cdot \nabla z_k, w_k)_\Omega + \frac{1}{2} \|z_k\|_{\Omega, k}^2 dt &= \int_0^T -(g, (T - t)z)_\Omega + \frac{1}{2} \|z\|_\Omega^2 dt \\ &= \frac{T}{2} \|z(0)\|_\Omega^2 + \int_0^T (T - t) \nu \|\nabla z\|_\Omega^2 dt \\ &= \frac{T}{2} \|z(0)\|_\Omega^2 + \nu \|\tilde{w}\|_X^2, \end{aligned} \quad (8.11)$$

where the second equality is obtained by using the test function  $w = (T - t)z$  in (8.4). This shows that

$$\limsup_{k \rightarrow \infty} \left[ \frac{T}{2} \|z_k(0)\|_\Omega^2 + \nu \|\tilde{w}_k\|_X^2 \right] \leq \frac{T}{2} \|z(0)\|_\Omega^2 + \nu \|\tilde{w}\|_X^2. \quad (8.12)$$

We then conclude from (8.7) and (8.12) that  $\|z_k(0)\|_\Omega \rightarrow \|z(0)\|_\Omega$  and  $\|\tilde{w}\|_X \rightarrow \|\tilde{w}\|_X$  as  $k \rightarrow \infty$ , which combined with the weak convergence properties above imply that  $\|z(0) - z_k(0)\|_\Omega \rightarrow 0$  and  $\|\tilde{w} - \tilde{w}_k\|_X \rightarrow 0$  as  $k \rightarrow \infty$ . The strong convergence  $\tilde{w}_k \rightarrow \tilde{w}$  in  $X$  and the triangle inequality imply that

$$\lim_{k \rightarrow \infty} \int_0^T (T - t) \|\nabla(z - z_k)\|_\Omega^2 dt \lesssim \lim_{k \rightarrow \infty} \left[ \|\tilde{w} - \tilde{w}_k\|_X^2 + \tau_k \|z_k\|_X^2 \right] = 0,$$

where we have used the fact that  $\sup_{k \in \mathbb{N}} \|z_k\|_X \lesssim \sup_{k \in \mathbb{N}} \|z_k\|_{V_k^-} < \infty$ . This completes the proof of (8.5) and thus of the lemma.  $\square$

We can now prove the strong  $L^2$ -convergence of the density function approximations at terminal time.

**Theorem 8.3** (Terminal-time  $L^2$ -convergence of discrete KFP approximations) *After possibly passing to a subsequence without change of notation, we have  $m_k(T) \rightarrow m(T)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$ , where  $m$  solves (8.1).*

**Proof** Since  $m_k(T) \rightharpoonup m(T)$  in  $L^2(\Omega)$  by Lemma 8.1, it is enough to show that  $\limsup_{k \rightarrow \infty} \|m_k(T)\|_{\Omega} \leq \|m(T)\|_{\Omega}$ . The discrete integration-by-parts identity (4.20) implies that

$$\begin{aligned} \|m_k(T)\|_{\Omega,k}^2 &= (m_k(T), z_k(T))_{\Omega,k} \\ &= (m_k(0), z_k(0))_{\Omega,k} + \int_0^T (\partial_t \mathcal{I}_+ m_k, z_k)_{\Omega,k} + (m_k, \partial_t \mathcal{I}_- z_k)_{\Omega,k} dt. \end{aligned}$$

Then, since  $z_k \in \mathbb{V}_k^-$  and  $m_k \in \mathbb{V}_k^+$  naturally embed into  $\mathbb{V}_k$ , we use  $z_k$  as a test function in (5.1b) with  $m_k$  as a test function in (8.2) to obtain

$$\|m_k(T)\|_{\Omega,k}^2 = (m_k(0), z_k(0))_{\Omega,k} + \int_0^T \langle G, z_k \rangle dt. \quad (8.13)$$

Similarly, since  $Y$  embeds continuously in  $X$ , we use  $z$  as a test function in (8.1) and  $m$  as test function in (8.3) to obtain

$$\|m(T)\|_{\Omega}^2 = (m_0, z(0))_{\Omega} + \int_0^T \langle G, z \rangle dt. \quad (8.14)$$

The convergence  $z_k(0) \rightarrow z(0)$  in  $L^2(\Omega)$  and  $z_k \rightharpoonup z$  in  $X = L^2(0, T; H_0^1(\Omega))$  from Lemma 8.2, together with the fact that  $m_k(0) = R_k m_0$  in  $V_k$ , then implies that

$$(m_k(0), z_k(0))_{\Omega,k} = (m_0, z_k(0))_{\Omega} \rightarrow (m_0, z(0))_{\Omega}, \quad \int_0^T \langle G, z_k \rangle dt \rightarrow \int_0^T \langle G, z \rangle dt, \quad (8.15)$$

as  $k \rightarrow \infty$ . Therefore, passing to the limit in (8.13), we see that

$$\limsup_{k \rightarrow \infty} \|m_k(T)\|_{\Omega}^2 \leq \limsup_{k \rightarrow \infty} \|m_k(T)\|_{\Omega,k}^2 \leq \|m(T)\|_{\Omega}^2. \quad (8.16)$$

where we recall that the first inequality above follows from (4.5). This implies that  $m_k(T)$  converges in norm to  $m(T)$  in  $L^2(\Omega)$ .  $\square$

## 8.2 HJB equation

With the terminal time compactness result Theorem 8.3 established for the density approximations, we can now prove convergence of the value function approximations to a weak solution to the HJB equation.

**Lemma 8.4** (Convergence to the HJB equation) *After passing to a subsequence without change of notation, we have  $u_k \rightarrow u$  in  $X$ ,  $u_k \rightarrow u$  in  $L^p(0, T; L^2(\Omega))$  for any  $1 \leq$*



$p < \infty$ , and  $u_k(0) \rightarrow u(0)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$  where  $u \in Y$  solves  $u(T) = S[m(T)]$  in  $L^2(\Omega)$  and

$$\int_0^T -\langle \partial_t u, \psi \rangle + v(\nabla u, \nabla \psi)_\Omega + (H[\nabla u], \psi)_\Omega dt = \int_0^T \langle F[m], \psi \rangle dt \quad \forall \psi \in X, \quad (8.17)$$

where  $m \in Y$  satisfies (8.1).

Note that an integration-by-parts argument on the time derivative in (8.17) shows that  $u$  equivalently solves (3.1a) for all  $\psi \in Y_0$ .

**Proof** Since  $\sup_{k \in \mathbb{N}} \|u_k\|_{V_k^-} < \infty$  by (5.3), Theorems 7.1 and 7.3 imply that, after passing to a subsequence without change of notation, there exists a  $u \in Y$  such that  $\partial_t \mathcal{I}_- u_k \rightarrow \partial_t u$  in  $L^2(0, T; H^{-1}(\Omega))$ ,  $u_k \rightarrow u$  in  $X$ ,  $u_k(0) \rightarrow u(0)$  in  $L^2(\Omega)$ ,  $u_k(T) \rightarrow u(T)$  in  $L^2(\Omega)$  and  $u_k \rightarrow u$  in  $L^p(0, T; L^2(\Omega))$  for all  $p \in [1, \infty)$ . Furthermore, since  $m_k(T) \rightarrow m(T)$  by Theorem 8.3, the continuity of  $S$  on  $L^2(\Omega)$  and the  $L^2$ -stability of  $R_k$  imply that  $u_k(T) = R_k S[m_k(T)] \rightarrow S[m(T)]$  in norm in  $L^2(\Omega)$ , hence  $u(T) = S[m(T)]$ . It follows from (2.5) that  $\{H[\nabla u_k]\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^2(0, T; L^2(\Omega))$  and thus, after passing to a further subsequence without change of notation, there exists a  $H_* \in L^2(0, T; L^2(\Omega))$  such that  $H[\nabla u_k] \rightarrow H_*$  in  $L^2(0, T; L^2(\Omega))$ . Furthermore, the continuity of  $F$  on  $L^2(0, T; L^2(\Omega))$  and the strong convergence  $m_k \rightarrow m$  in  $L^2(0, T; L^2(\Omega))$  from Lemma 8.1 imply that  $F[m_k] \rightarrow F[m]$  in  $L^2(0, T; H^{-1}(\Omega))$ . We may then use density of  $\bigcup_{\ell \in \mathbb{N}} V_\ell$  in  $X$  and (7.2) to pass to the limit as  $k \rightarrow \infty$  in (5.2) to find that  $u$  solves

$$\int_0^T -\langle \partial_t u, \psi \rangle + v(\nabla u, \nabla \psi)_\Omega + (H_*, \psi)_\Omega dt = \int_0^T \langle F[m], \psi \rangle dt \quad \psi \in X. \quad (8.18)$$

We now claim that  $u_k \rightarrow u$  in norm in  $X$  as  $k \rightarrow \infty$  which will imply  $H_* = H[\nabla u]$ . First, note that

$$\begin{aligned} \frac{1}{2} \|u_k(0)\|_\Omega^2 &\leq \frac{1}{2} \|u_k(0)\|_{\Omega, k}^2 + \frac{1}{2} \sum_{n=1}^{N_k} \| [u_k]_n \|_{\Omega, k}^2 \\ &= \int_0^T -(\partial_t \mathcal{I}_- u_k, u_k)_{\Omega, k} dt + \frac{1}{2} \|u_k(T)\|_{\Omega, k}^2. \end{aligned} \quad (8.19)$$

where the first inequality above follows from (4.5). Observe also that  $v \|u_k\|_X^2 = \int_0^T v \|\nabla u_k\|_\Omega^2 dt \leq \int_0^T (A_k \nabla u_k, \nabla u_k)_\Omega dt$  since  $A_k \geq v \mathbb{I}$  a.e. in  $\Omega$  in the sense of semi-definite matrices. Next, note that  $\|u_k(T)\|_{\Omega, k}^2 = \|R_k S[m_k(T)]\|_{\Omega, k}^2 \rightarrow$

$\|S[m(T)]\|_{\Omega}^2$  as  $k \rightarrow \infty$  since  $\|R_k S[m_k(T)]\|_{\Omega,k}^2 = (S[m_k(T)], R_k S[m_k(T)])_{\Omega}$  and both  $R_k S[m_k(T)]$  and  $S[m_k(T)]$  converge to  $S[m(T)]$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$ . Furthermore, weak convergence  $H[\nabla u_k] \rightharpoonup H_*$  and strong convergence  $u_k \rightarrow u$  in  $L^2(0, T; L^2(\Omega))$  imply that  $\int_0^T (H[\nabla u_k], u_k)_{\Omega} dt \rightarrow \int_0^T (H_*, u)_{\Omega} dt$ . Since  $F[m_k] \rightarrow F[m]$  in  $L^2(0, T; H^{-1}(\Omega))$  we also have  $\int_0^T \langle F[m_k], u_k \rangle dt \rightarrow \int_0^T \langle F[m], u \rangle dt$  as  $k \rightarrow \infty$ . Therefore, testing (5.2) with  $u_k$  gives

$$\begin{aligned}
 & \limsup_{k \rightarrow \infty} \left[ \frac{1}{2} \|u_k(0)\|_{\Omega}^2 + \nu \|u_k\|_X^2 \right] \\
 & \leq \lim_{k \rightarrow \infty} \left[ \int_0^T -(\partial_t \mathcal{I} u_k, u_k)_{\Omega,k} + (A_k \nabla u_k, \nabla u_k)_{\Omega} dt + \frac{1}{2} \|u_k(T)\|_{\Omega,k}^2 \right] \\
 & = \lim_{k \rightarrow \infty} \left[ \int_0^T \langle F[m_k], u_k \rangle - (H[\nabla u_k], u_k)_{\Omega} dt + \frac{1}{2} \|R_k S[m_k(T)]\|_{\Omega,k}^2 \right] \\
 & = \int_0^T \langle F[m], u \rangle - (H_*, u)_{\Omega} dt + \frac{1}{2} \|S[m(T)]\|_{\Omega}^2 \\
 & = \frac{1}{2} \|u(0)\|_{\Omega}^2 + \nu \|u\|_X^2,
 \end{aligned} \tag{8.20}$$

where the last line follows from testing (8.18) with  $u$ . In conjunction with the inequalities  $\|u(0)\|_{\Omega} \leq \liminf_{k \rightarrow \infty} \|u_k(0)\|_{\Omega}$  and  $\|u\|_X \leq \liminf_{k \rightarrow \infty} \|u_k\|_X$  owing to weak convergence, we see that (8.20) implies that  $u_k(0) \rightarrow u(0)$  in norm in  $L^2(\Omega)$  and that  $u_k \rightarrow u$  in norm in  $X$  as  $k \rightarrow \infty$ . In turn, this implies that  $H_* = H[\nabla u]$  since  $H[\nabla u_k] \rightarrow H[\nabla u]$  in  $L^2(0, T; L^2(\Omega))$  by Lipschitz continuity of  $H$ , and thus  $u$  solves (8.17).  $\square$

### 8.3 Proof of Theorems 5.3 and 3.1

To begin, we observe that since  $\tilde{b}_k \in \mathcal{D}_p H[u_k]$  for all  $k \in \mathbb{N}$ ,  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is uniformly bounded in  $L^\infty(Q_T; \mathbb{R}^d)$  (see Lemma 2.1). We may then pass to a subsequence, without change of notation such that  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  converges weakly in  $L^2(Q_T; \mathbb{R}^d)$  to some  $\tilde{b}_* \in L^\infty(Q_T; \mathbb{R}^d)$ . We first obtain from Lemma 8.1 and Theorem 8.3 that  $m_k \rightarrow m$  in  $L^p(0, T; L^2(\Omega))$  for any  $1 \leq p < \infty$ , and  $m_k(T) \rightarrow m(T)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$  where  $m \in Y$  uniquely satisfies the KFP equation (8.1) and the initial condition  $m(0) = m_0$ .

Consequently, Lemma 8.4 implies that  $u_k \rightarrow u$  in  $X$ ,  $u_k \rightarrow u$  in  $L^p(0, T; L^2(\Omega))$  for any  $1 \leq p < \infty$ , and  $u_k(0) \rightarrow u(0)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$  where  $u \in Y$  uniquely satisfies the HJB equation (8.17).

Since  $\tilde{b}_k \in \mathcal{D}_p H[u_k]$  for all  $k \in \mathbb{N}$ , while  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  converges weakly in  $L^2(Q_T; \mathbb{R}^d)$  to  $\tilde{b}_*$  and  $u_k \rightarrow u$  in  $X$ , we can apply Lemma 2.2 to obtain  $\tilde{b}_* \in \mathcal{D}_p H[u]$ . This shows that  $(u, m) \in Y$  is a weak solution of (3.1), and thus proves Theorem 3.1.

In summary, we have shown that a subsequence of the numerical approximations converges to a solution  $(u, m)$  of (3.1) in the sense of (5.4). Then, uniqueness of the solution, c.f. Theorem 3.2, implies that the whole sequence  $\{(u_k, m_k)\}_{k \in \mathbb{N}}$  converges to  $(u, m)$  in the sense of (5.4). This completes the proof of Theorem 5.3.  $\square$

## 8.4 Proof of Corollary 5.4

Under the hypotheses of Corollary 5.4, we now prove strong convergence in  $L^2(0, T; H_0^1(\Omega))$  of the density approximations.

We know by hypothesis that  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is pre-compact in  $L^1(Q_T; \mathbb{R}^d)$ . Moreover, as a consequence of Lemma 2.1 and the definition of the scheme (5.1), the sequence is uniformly bounded in  $L^\infty(Q_T; \mathbb{R}^d)$ . Since  $\text{meas}(\Omega) < \infty$  and  $T < \infty$ , it can then be shown that  $\{\tilde{b}_k\}_{k \in \mathbb{N}}$  is pre-compact in  $L^q(Q_T; \mathbb{R}^d)$  for any  $q \in [1, \infty)$ . Since the density  $m \in Y$ , we have by the embedding  $Y \rightarrow C([0, T]; L^2(\Omega))$  that  $m \in X \cap C([0, T]; L^2(\Omega))$ . It then follows from the parabolic embedding [25, Equation (6.39), p. 466] that  $m \in L^{2+\frac{4}{d}}(Q_T)$ . Passing to a suitable subsequence, without change of notation, we get that  $\tilde{b}_k \rightarrow \tilde{b}_*$  in  $L^{d+2}(Q_T; \mathbb{R}^d)$  as  $k \rightarrow \infty$  for some  $\tilde{b}_* \in L^{d+2}(Q_T; \mathbb{R}^d)$ . It then follows that  $m_k \tilde{b}_k \rightarrow m \tilde{b}_*$  in  $L^2(Q_T; \mathbb{R}^d)$  as  $k \rightarrow \infty$ .

We now show that  $m_k \rightarrow m$  in  $X$  as  $k \rightarrow \infty$ , along a subsequence. For given  $k \in \mathbb{N}$ , testing the equation (5.1b) with  $\phi = m_k \in \mathbb{V}_k^+$  leads to

$$\begin{aligned} & \limsup_{k \rightarrow \infty} \left[ \frac{1}{2} \|m_k(T)\|_\Omega^2 + \nu \|m_k\|_X^2 \right] \\ & \leq \lim_{k \rightarrow \infty} \left[ \int_0^T \langle G, m_k \rangle - (m_k \tilde{b}_k, \nabla m_k)_\Omega dt + \frac{1}{2} \|R_k m_0\|_{\Omega, k}^2 \right] \\ & = \int_0^T \langle G, m \rangle - (m \tilde{b}_*, \nabla m)_\Omega dt + \frac{1}{2} \|m_0\|_\Omega^2 \\ & = \frac{1}{2} \|m(T)\|_{L^2(\Omega)}^2 + \nu \|m\|_X^2, \end{aligned} \quad (8.21)$$

where the last line above is obtained by testing (3.1b) with  $m$ . Since  $m_k(T) \rightarrow m(T)$  in  $L^2(\Omega)$  as  $k \rightarrow \infty$ , it follows from (8.21) that  $\limsup_{k \rightarrow \infty} \|m_k\|_X \leq \|m\|_X$ . Together with the fact that  $m_k \rightarrow m$  in  $X$  as  $k \rightarrow \infty$ , we obtain that  $m_k \rightarrow m$  in  $X$  as  $k \rightarrow \infty$ . This completes the proof.  $\square$

## 9 Numerical experiment

In this section we test the performance of the finite element scheme (5.1) in approximating a smooth solution to a MFG PDI of the form (3.1). Let  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$  be the unit square, and let the time horizon  $T = 1$ . We set  $\nu = 1$ , and we set

$$H(t, x, p) := \max_{\alpha \in B_1(0)} (\alpha \cdot p) = |p| \quad \forall (t, x, p) \in \overline{Q_T} \times \mathbb{R}^2, \quad (9.1)$$

where  $\overline{B_1(0)}$  denotes the closed unit ball in  $\mathbb{R}^2$  and  $L_H = 1$ . We choose the data such that the exact solution to (3.1) is

$$u(x, y, t) := (2 \arctan(t) + 1)xy(e^{1.2} - e^{1.2x})(e^{0.7} - e^{0.7y}), \quad (9.2a)$$

$$m(x, y, t) := \left(\frac{1}{4} \tanh(t) + 1\right) \sinh(x) \sinh(1-x)y \ln(2-y), \quad (9.2b)$$

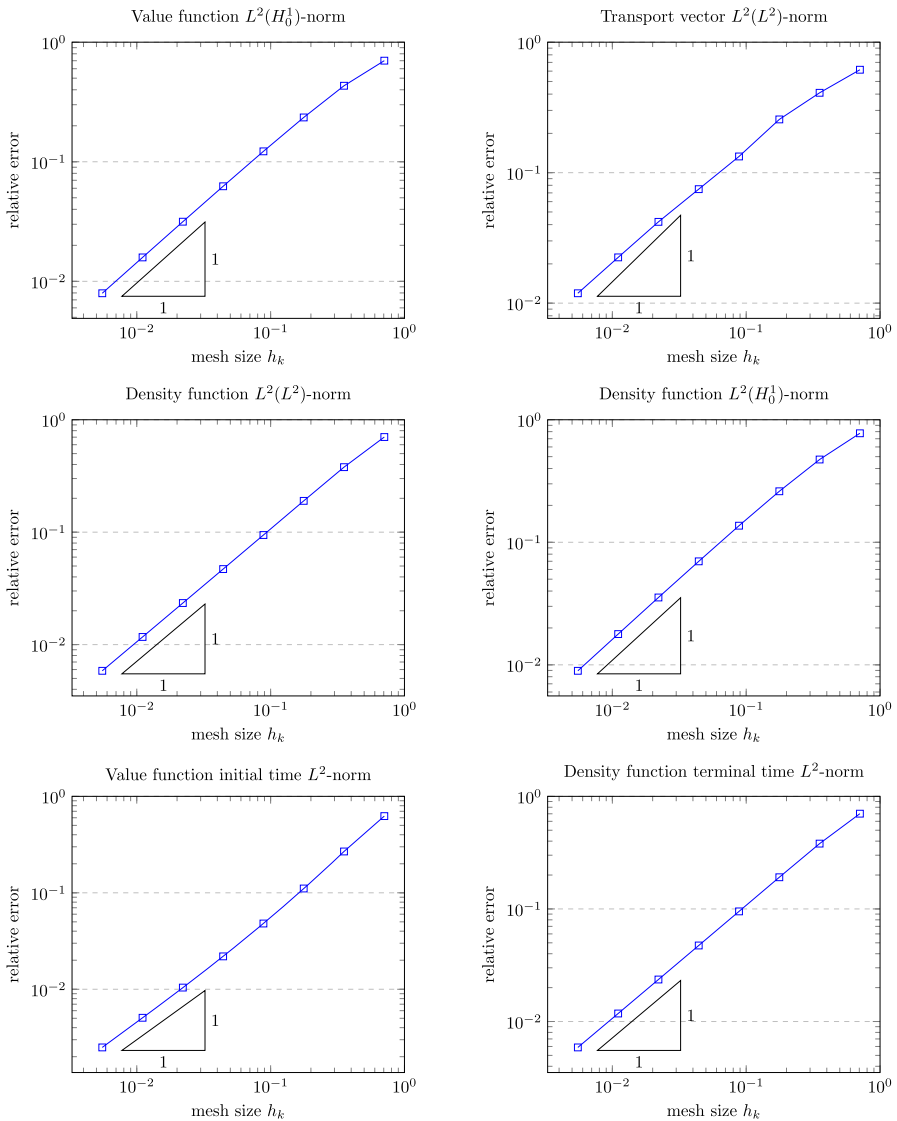
for all  $(x, y, t) \in \overline{Q_T}$ , where the vector field  $\tilde{b}_* \in \mathcal{D}_p H[u]$  in (3.1) is unique and given by  $\tilde{b}_* = |\nabla u|^{-1} \nabla u$  a.e. in  $Q_T$ . In particular, we choose of  $F : L^2(Q_T) \rightarrow L^2(Q_T)$  and  $S : L^2(\Omega) \rightarrow L^2(\Omega)$  to be given by  $F[m] = m + F_0$  and  $S[m] := \tanh(m) + S_0$  respectively, where  $F_0 \in L^2(Q_T)$  and  $S_0 \in L^2(\Omega)$  are determined from  $u$  and  $m$  above. The source term  $G \in L^2(0, T; H^{-1}(\Omega))$  is also determined from  $m$  above and can be shown to be nonnegative in the sense of distributions (see [31, Sections 3.3 & 5.3]). Overall, the data satisfy the hypotheses of Theorem 3.2, so the solution of the problem is unique. We use a shape-regular sequence of uniform conforming meshes of  $\Omega$  where the elements are right-angled triangles. The weights in the diffusion tensor (4.10) with are chosen as  $\omega_{k,E} := L_H \text{diam}(E)$ , for all  $E \in \mathcal{E}_k, k \in \mathbb{N}$ . We take the time-step to be comparable to the mesh-size, namely  $\tau_k = \frac{h_k}{(1+2^{-k})\sqrt{2}} \approx h_k$  for  $k \in \mathbb{N}$ . The computations are performed using Firedrake [36].

The results of the experiment are displayed in Fig. 1. We observe that the approximations for the value function converge to  $u$  in the  $L^2(H_0^1)$ -norm with a rate of order one, and likewise we observe a first-order rate of convergence for the approximations of  $m$  in the  $L^2(L^2)$ -norm. These observed rates are the best possible to be expected since the discretization spaces  $\{\mathbb{V}_k^\pm\}_{k=1}^8$  consist of functions that are piecewise affine in space and piecewise constant in time, and the stabilization term in the scheme (5.1) is consistent to first order in the mesh-size.

Furthermore, we observe strong convergence of the transport vector fields to  $\tilde{b}_*$  in the  $L^2(L^2)$ -norm. This strong convergence is due to the fact that the value function satisfies  $\nabla u \neq 0$  a.e. in  $Q_T$  along with uniqueness of  $\tilde{b}_*$  (see [31, Chapter 6] for further details). In light of Corollary 5.4, the strong convergence of the transport vector fields implies the strong convergence of the density function approximations in the  $L^2(H_0^1)$ -norm, which we observe in Fig. 1 occurs with a convergence rate of order one. Finally, we observe strong convergence of the value function approximations at the initial time and of the density function approximations at the terminal time in the  $L^2$ -norm, which is in agreement with Theorem 5.3.

## Appendix A Proof of Theorem 4.2

In this section we prove Theorem 4.2. For this, we will need the following key formula associated with the diffusion tensor  $D_k$  that was defined in (4.10). Recall that for a given



**Fig. 1** Convergence plots for approximations of the value function, density function, and transport vector. Optimal rates of convergence are observed for the value function and density function errors in the  $H^1$ -norm

vertex  $x_i \in \mathcal{V}_k$ , the corresponding piecewise affine nodal basis function is denoted by  $\xi_i$ .

**Lemma A.1** Let  $k \in \mathbb{N}$  be given. Let  $\{x_i, x_j\} \subset \mathcal{V}_k$  be a pair of distinct neighbouring vertices of  $\mathcal{T}_k$ , and let  $E \in \mathcal{E}_k$  denote the edge between  $x_i$  and  $x_j$ . Then

$$\int_{\Omega} D_k \nabla \xi_j \cdot \nabla \xi_i dx = -\frac{\omega_{k,E}}{(\text{diam } E)^2} \sum_{K \in \mathcal{T}_{k,E}} |K|_d, \quad (\text{A1})$$

where  $|\cdot|_d$  denotes the  $d$ -dimensional Lebesgue measure.

**Proof** Consider an element  $K \in \mathcal{T}_k$  containing both vertices  $x_i$  and  $x_j$ . Note that  $\nabla \xi_i|_K$  is orthogonal to all vectors that are tangent to  $F_{K,i}$  the face of  $K$  that is opposite  $x_i$ . Furthermore, any edge  $\tilde{E} \in \mathcal{E}_K$  that does not contain  $x_i$  is contained in  $F_{K,i}$ . Therefore  $\nabla \xi_i|_K \cdot \mathbf{t}_{\tilde{E}} = 0$  for all edges  $\tilde{E} \in \mathcal{E}_K$  that do not contain  $x_i$ . Likewise,  $\nabla \xi_j|_K \cdot \mathbf{t}_{\tilde{E}} = 0$  for all edges  $\tilde{E} \in \mathcal{E}_K$  that do not contain  $x_j$ . Since the only edge of  $\mathcal{E}_K$  that contains both  $x_i$  and  $x_j$  is  $E$ , we see that

$$\begin{aligned} \int_K D_k \nabla \xi_i \cdot \nabla \xi_j dx &= \sum_{\tilde{E} \in \mathcal{E}_K} \omega_{k,\tilde{E}} |K|_d (\nabla \xi_i \cdot \mathbf{t}_{\tilde{E}}) (\nabla \xi_j \cdot \mathbf{t}_{\tilde{E}}) \\ &= \omega_{k,E} |K|_d (\nabla \xi_i \cdot \mathbf{t}_E) (\nabla \xi_j \cdot \mathbf{t}_E), \end{aligned}$$

i.e. the sum above simplifies to a single term. It is easy to show that  $\nabla \xi_i|_K \cdot \mathbf{t}_E = \pm \frac{1}{\text{diam } E}$ , where the sign depends on the choice of orientation of the vector  $\mathbf{t}_E$ . Since  $\xi_i + \xi_j = 1$  identically on  $E$ , we also have  $\nabla \xi_j|_K \cdot \mathbf{t}_E = -\nabla \xi_i|_K \cdot \mathbf{t}_E$ . We therefore conclude that  $\int_K D_k \nabla \xi_i \cdot \nabla \xi_j dx = -\frac{\omega_{k,E} |K|_d}{(\text{diam } E)^2}$ . The identity (A1) then follows by summation over all elements of the mesh.  $\square$

**Proof of Theorem 4.2** It is clear that (4.12) implies that  $\|D_k\|_{L^\infty(K; \mathbb{R}^{d \times d})} \lesssim h_K$  for all  $K \in \mathcal{T}_k$ , so it remains only to prove the DMP property for arbitrary elements of  $W(V_k, D_k)$ . Fix  $k \in \mathbb{N}$  and let  $L \in W(V_k, D_k)$  be an operator taking the form (4.11) for some vector field  $\tilde{b} : \Omega \rightarrow \mathbb{R}^d$  that satisfies the bound  $\|\tilde{b}\|_{L^\infty(\Omega; \mathbb{R}^d)} \leq L_H$ . Note that  $L$  satisfies the DMP if and only if the adjoint operator  $L^*$  satisfies the DMP. We therefore restrict our attention to showing that  $L$  satisfies the DMP. Suppose  $v \in V_k$  is such that  $\langle Lv, \xi_i \rangle_{V_k^* \times V_k} \geq 0$  for all  $i \in \{1, \dots, M_k\}$ . Since  $v$  is piecewise affine and continuous on  $\Omega$ , there exists a vertex  $x_i$  of the mesh  $\mathcal{T}_k$  where  $v$  achieves a global minimum value over  $\overline{\Omega}$ . In the case where  $x_i \in \partial\Omega$  then the result is immediate since  $v(x_i) = 0$ . Consider now the case where  $x_i \in \Omega$  is an interior vertex. Let  $\mathbb{S}_i$  denote the index set of all vertices  $x_j \in \mathcal{V}_k$  that are neighbours to  $x_i$ . Note that by definition  $\mathbb{S}_i$  does not include  $x_i$ , but  $\mathbb{S}_i$  possibly includes neighbouring vertices on the boundary  $\partial\Omega$ .

Since the nodal basis functions form a partition of unity, we have  $\nabla v = \sum_{j \in \mathbb{S}_i} \nabla \xi_j (v(x_j) - v(x_i))$  in  $\text{supp } \xi_i$ . Therefore, we find that

$$\langle Lv, \xi_i \rangle \leq \sum_{j \in \mathbb{S}_i} \left[ (D_k \nabla \xi_j, \nabla \xi_i)_\Omega + (\tilde{b} \cdot \nabla \xi_j, \xi_i)_\Omega \right] (v(x_j) - v(x_i)), \quad (\text{A2})$$

where we have used the fact that  $\int_{\Omega} v \nabla v \cdot \nabla \xi_i dx \leq 0$  as a result of the condition (4.1), c.f. [42]. We will show below that the condition on the weights  $\omega_{k,E}$  in (4.12) implies that, for all  $i \in \{1, \dots, M_k\}$ ,

$$(D_k \nabla \xi_j, \nabla \xi_i)_{\Omega} + (\tilde{b} \cdot \nabla \xi_j, \xi_i)_{\Omega} < 0 \quad \forall j \in \mathbb{S}_i, \quad (\text{A3})$$

for any  $\tilde{b} \in L^{\infty}(\Omega; \mathbb{R}^d)$  satisfying  $\|\tilde{b}\|_{L^{\infty}(\Omega; \mathbb{R}^d)} \leq L_H$ . Assuming (A3) for the moment, we then deduce from  $\langle Lv, \xi_i \rangle \geq 0$  and from (A2) that  $v(x_j) = v(x_i)$  for all  $j \in \mathbb{S}_i$ , i.e.  $v$  achieves a minimum at all neighbouring vertices.

Repeating the above argument, we can eventually deduce that the global minimum of  $v$  is also attained on the boundary  $\partial\Omega$ , where  $v$  vanishes. This shows that  $v \geq 0$  in  $\Omega$  and thus  $L$  satisfies the DMP.

Returning to the proof of (A3), let  $x_i$  and  $x_j$  be a pair of distinct neighbouring vertices, and let  $E$  be the edge between  $x_i$  and  $x_j$ . Then, it is straightforward to show that  $\|\nabla \xi_j\|_{L^{\infty}(K)} \leq \frac{\delta}{2 \text{diam } E}$  for each element  $K$  containing the edge  $E$ . Therefore, using the bounds  $\|\tilde{b}\|_{L^{\infty}(\Omega; \mathbb{R}^d)} \leq L_H$  and  $\|\xi_i\|_{L^1(K)} \leq \frac{|K|_d}{d+1}$ , we find that  $(\tilde{b} \cdot \nabla \xi_j, \xi_i)_{\Omega} \leq \sum_{K \in \mathcal{T}_{k,E}} \frac{\delta L_H |K|_d}{2(d+1) \text{diam } E}$ . Using Lemma A.1, we therefore obtain

$$(D_k \nabla \xi_j, \nabla \xi_i)_{\Omega} + (\tilde{b} \cdot \nabla \xi_j, \xi_i)_{\Omega} \leq \sum_{K \in \mathcal{T}_{k,E}} \left( \frac{\delta L_H \text{diam } E}{2(d+1)} - \omega_{k,E} \right) \frac{|K|_d}{(\text{diam } E)^2}. \quad (\text{A4})$$

We then see that (A3) follows from (A4) and (4.12).  $\square$

## Appendix B Proofs of Lemmas 6.1 and 6.2

In order to prove Lemmas 6.1 and 6.2, we introduce some weighted norms on the discrete spaces that will help to provide a sharper analysis. Recall that the matrix-valued function  $A_k \in L^{\infty}(\Omega; \mathbb{R}^{d \times d})$  is defined by  $A_k := \nu \mathbb{I} + D_k \in L^{\infty}(\Omega; \mathbb{R}^{d \times d})$  with  $D_k$  defined in (4.10). Let the norm  $\|\cdot\|_{A_k} := L^2(\Omega; \mathbb{R}^d) \rightarrow \mathbb{R}_{\geq 0}$  be defined by  $\|w\|_{A_k} = \sqrt{(A_k w, w)_{\Omega}}$  for all  $w \in L^2(\Omega; \mathbb{R}^d)$ . Note that  $\|\cdot\|_{A_k}$  is equivalent to the standard norm on  $L^2(\Omega; \mathbb{R}^d)$ , i.e.  $\|w\|_{A_k} \approx \|w\|_{\Omega}$  for all  $w \in L^2(\Omega; \mathbb{R}^d)$ , since  $A_k \in L^{\infty}(\Omega; \mathbb{R}^{d \times d})$  and since  $A_k \geq \nu \mathbb{I}$ . Next we define the dual norm of  $\|\cdot\|_{A_k} : V_k \rightarrow \mathbb{R}_{\geq 0}$  by

$$\|w\|_{A_k,*} := \sup_{v \in V_k \setminus \{0\}} \frac{(w, v)_{\Omega,k}}{\|\nabla v\|_{A_k}} \quad \forall w \in V_k. \quad (\text{B1})$$

Observe that  $\|w\|_{A_k,*} \approx \|w\|_{k,*}$  for all  $w \in V_k^*$  where  $\|\cdot\|_{k,*}$  is defined in (4.9). Define the bilinear form  $B_k : \mathbb{V}_k^+ \times \mathbb{V}_k \rightarrow \mathbb{R}$  by

$$B_k(u, v) := \sum_{n=1}^{N_k} \int_{I_n} (\partial_t \mathcal{I}_+ u, v)_{\Omega, k} + (A_k \nabla u, \nabla v)_{\Omega} dt \quad \forall (u, v) \in \mathbb{V}_k^+ \times \mathbb{V}_k. \quad (\text{B2})$$

Given  $k \in \mathbb{N}$ , let  $\tilde{a}_k \in L^\infty(0, T)$  denote the piecewise constant time-dependent weight function that is defined by

$$\tilde{a}_k|_{I_n} := \tilde{a}_{k,n} = \frac{1}{(1 + v^{-1} L_H^2 \tau_k)^n} \quad \forall n \in \{1, \dots, N_k\}. \quad (\text{B3})$$

Since the time-steps are assumed to be uniform, it is straight-forward to show that  $a_k \in L^\infty(0, T)$  is strictly positive and satisfies the uniform bounds

$$\exp(-v^{-1} L_H^2 T) \leq \tilde{a}_k \leq 1 \quad \text{a.e. in } (0, T), \quad \forall k \in \mathbb{N}. \quad (\text{B4})$$

We now define the following weighted discrete norms on the spaces  $\mathbb{V}_k$  and  $\mathbb{V}_k^+$  that will help to simplify the analysis. First, let the norm  $\|\cdot\|_{\mathbb{V}_k, A_k} : \mathbb{V}_k \rightarrow \mathbb{R}_{\geq 0}$  be defined by

$$\|v\|_{\mathbb{V}_k, A_k}^2 := \int_0^T \tilde{a}_k^{-1} \|\nabla v\|_{A_k}^2 dt \quad \forall v \in \mathbb{V}_k. \quad (\text{B5})$$

Note that  $\|\cdot\|_{\mathbb{V}_k, A_k}$  is equivalent to  $\|\cdot\|_X$  on  $\mathbb{V}_k$ , where we recall that  $\|\cdot\|_X$  is defined in (2.2). Second, let the norm  $\|\cdot\|_{\mathbb{V}_k^+, A_k} : \mathbb{V}_k^+ \rightarrow \mathbb{R}_{\geq 0}$  be defined by

$$\begin{aligned} \|w\|_{\mathbb{V}_k^+, A_k}^2 &:= \int_0^T \tilde{a}_k \left( \|\partial_t \mathcal{I}_+ w\|_{A_k, *}^2 + \|\nabla w\|_{A_k}^2 + \frac{v^{-1} L_H^2}{1 + v^{-1} L_H^2 \tau_k} \|w\|_{\Omega, k}^2 \right) dt \\ &\quad + \frac{\tilde{a}_{k, N_k} \|w(T)\|_{\Omega, k}^2}{1 + v^{-1} L_H^2 \tau_k} + \frac{1}{1 + v^{-1} L_H^2 \tau_k} \sum_{n=0}^{N_k-1} \tilde{a}_{k,n} \|[w]\|_{\Omega, k}^2, \end{aligned} \quad (\text{B6})$$

for all  $w \in \mathbb{V}_k^+$ . The next result gives parabolic inf-sup identities that show the relation between the weighted norms  $\|\cdot\|_{\mathbb{V}_k, A_k}$  and  $\|\cdot\|_{\mathbb{V}_k^+, A_k}$ . Recall that  $\mathbb{V}_{k,0}^+ = \{v \in \mathbb{V}_k^+, v(0) = 0\}$ .

**Theorem B.1** *For all  $w \in \mathbb{V}_k^+$  we have*

$$\|w\|_{\mathbb{V}_k^+, A_k}^2 = \left[ \sup_{v \in \mathbb{V}_k \setminus \{0\}} \frac{B_k(w, v)}{\|v\|_{\mathbb{V}_k, A_k}} \right]^2 + \frac{\|w(0)\|_{\Omega, k}^2}{1 + v^{-1} L_H^2 \tau_k}, \quad (\text{B7})$$



and for all  $v \in \mathbb{V}_k$

$$\|v\|_{\mathbb{V}_k, A_k} = \sup_{w \in \mathbb{V}_{k,0}^+ \setminus \{0\}} \frac{B_k(w, v)}{\|w\|_{\mathbb{V}_k^+, A_k}}. \quad (\text{B8})$$

**Proof** Let  $w \in \mathbb{V}_k^+$  be arbitrary. It is clear that the supremum on the right-hand side in (B7) is achieved by the test function  $v_{\dagger} = \tilde{a}_k(z+w) \in \mathbb{V}_k$ , where  $z \in \mathbb{V}_k$  denotes the unique element of  $\mathbb{V}_k$  such that  $(A_k \nabla z|_{I_n}, \nabla v) = (\partial_t \mathcal{I}_+ w|_{I_n}, v)_{\Omega, k}$  for all  $v \in V_k$ , for all  $n \in \{1, \dots, N_k\}$ . Observe that  $v_{\dagger} \in \mathbb{V}_k$  is justified by the fact that  $\tilde{a}_k$  is piecewise constant in time. We therefore deduce that

$$\begin{aligned} \left[ \sup_{v \in \mathbb{V}_k \setminus \{0\}} \frac{B_k(w, v)}{\|v\|_{\mathbb{V}_k, A_k}} \right]^2 &= \int_0^T \tilde{a}_k \|\nabla(z+w)\|_{A_k}^2 dt \\ &= \int_0^T \tilde{a}_k \left( \|\partial_t \mathcal{I}_+ w\|_{A_{k,*}}^2 + \|\nabla w\|_{A_k}^2 + 2(\partial_t \mathcal{I}_+ w, w)_{\Omega, k} \right) dt, \end{aligned} \quad (\text{B9})$$

where we have expanded the square in the second-line above. A straightforward calculation shows that

$$\begin{aligned} &2 \int_0^T \tilde{a}_k (\partial_t \mathcal{I}_+ w, w)_{\Omega, k} dt \\ &= \frac{\nu^{-1} L_H^2}{1 + \nu^{-1} L_H^2 \tau_k} \int_0^T \tilde{a}_k \|w\|_{\Omega, k}^2 dt + \frac{\tilde{a}_{k, N_k} \|w(T)\|_{\Omega, k}^2}{1 + \nu^{-1} L_H^2 \tau_k} \\ &\quad + \frac{1}{1 + \nu^{-1} L_H^2 \tau_k} \sum_{n=0}^{N_k-1} \tilde{a}_{k, n} \|[w]_n\|_{\Omega, k}^2 - \frac{\|w(0)\|_{\Omega, k}^2}{1 + \nu^{-1} L_H^2 \tau_k}. \end{aligned} \quad (\text{B10})$$

Thus (B9) and (B10) imply (B7). The second identity (B8) then follows by duality.  $\square$

**Corollary B.2** We have  $\|w\|_{\mathbb{V}_k^+, A_k} \approx \|w\|_{\mathbb{V}_k^+}$  for any  $w \in \mathbb{V}_k^+$ .

**Proof** Let  $w \in \mathbb{V}_k^+$  be arbitrary. The upper bound  $\|w\|_{\mathbb{V}_k^+, A_k} \lesssim \|w\|_{\mathbb{V}_k^+}$  is a consequence of the uniform bounds on  $\tilde{a}_k$  in (B4), the maximum norm bound of Lemma 4.3, and the bound for the jumps  $\sum_{n=0}^{N_k-1} \|[w]_n\|_{\Omega, k}^2 \leq \|w\|_{\mathbb{V}_k^+}^2$ . The converse bound follows again from (B4), and from (B7) which implies that  $\|w(0)\|_{\Omega, k} \lesssim \|w\|_{\mathbb{V}_k^+, A_k}$ .  $\square$

**Lemma B.3** For each  $k \in \mathbb{N}$ , let the constant  $\gamma_k > 0$  be defined by

$$\gamma_k := \sqrt{\frac{1 + v^{-1} L_H^2 \tau_k}{2}}. \quad (\text{B11})$$

Then, for any  $w \in \mathbb{V}_k^+$ , we have

$$\int_0^T v^{-1} L_H^2 \tilde{a}_k \|w\|_{\Omega}^2 dt \leq \gamma_k^2 \|w\|_{\mathbb{V}_k^+, A_k}^2 + \frac{1}{2} \|w(0)\|_{\Omega, k}^2. \quad (\text{B12})$$

**Proof** Let  $w \in \mathbb{V}_k^+$  be arbitrary. The definition of the dual norm  $\|\cdot\|_{A_k, *}$  and the Cauchy–Schwarz inequality, applied to the left-hand side of (B10), imply that

$$\begin{aligned} & \frac{v^{-1} L_H^2}{1 + v^{-1} L_H^2 \tau_k} \int_0^T \tilde{a}_k \|w\|_{\Omega, k}^2 dt - \frac{\|w(0)\|_{\Omega, k}^2}{1 + v^{-1} L_H^2 \tau_k} \\ & \leq \int_0^T \tilde{a}_k \left( \|\partial_t \mathcal{I}_+ w\|_{A_k, *}^2 + \|\nabla w\|_{A_k}^2 \right) dt. \end{aligned}$$

Therefore, we add  $\frac{v^{-1} L_H^2}{1 + v^{-1} L_H^2 \tau_k} \int_0^T \tilde{a}_k \|w\|_{\Omega, k}^2 dt$  to both sides of the equation above to find that

$$\frac{2v^{-1} L_H^2}{1 + v^{-1} L_H^2 \tau_k} \int_0^T \tilde{a}_k \|w\|_{\Omega, k}^2 dt - \frac{\|w(0)\|_{\Omega, k}^2}{1 + v^{-1} L_H^2 \tau_k} \leq \|w\|_{\mathbb{V}_k^+, A_k}^2. \quad (\text{B13})$$

We then obtain (B12) from (B13) after noting that  $\|w\|_{\Omega} \leq \|w\|_{\Omega, k}$  by (4.5).  $\square$

**Proof of Lemma 6.1** This result is essentially well-known so we sketch only the main ideas. Note that (6.1) can be re-written in time-stepping form as the sequence of equations

$$(\bar{m}|_{I_n}, v|_{I_n})_{\Omega, k} + \tau_k \langle L_n \bar{m}|_{I_n}, v|_{I_n} \rangle_{V_k^* \times V_k} = (\bar{m}|_{I_{n-1}}, v|_{I_n})_{\Omega, k} + \int_{I_n} \langle G, v|_{I_n} \rangle dt, \quad (\text{B14})$$

for each  $n \in \{1, \dots, N_k\}$  for some operator  $L_n \in W(V_k, D_k)$  of the form of (4.11). Therefore, the existence and uniqueness of the solution  $\bar{m}$  along with its nonnegativity in the case of nonnegative data is a consequence of the DMP, c.f. Theorem 4.2. It remains only to show the bound (6.2). Taking the square-roots to both sides of the identity in (B7) and applying the triangle inequality implies that

$$\|\bar{m}\|_{\mathbb{V}_k^+, A_k} \leq \sup_{v \in \mathbb{V}_k \setminus \{0\}} \frac{B_k(\bar{m}, v)}{\|v\|_{\mathbb{V}_k, A_k}} + \frac{\|g\|_{\Omega, k}}{\sqrt{1 + v^{-1} L_H^2 \tau_k}},$$

where we have used the fact  $\bar{m}(0) = g$ , and where  $B_k(\bar{m}, v)$  is found by rewriting (6.1) equivalently as

$$B_k(\bar{m}, v) = \int_0^T \langle G, v \rangle - (\bar{m}, \tilde{b} \cdot \nabla v)_{\Omega} dt \quad \forall v \in \mathbb{V}_k. \quad (\text{B15})$$

Note that  $\sup_v \frac{B_k(\bar{m}, v)}{\|v\|_{\mathbb{V}_k, A_k}} \leq \sup_v \frac{\int_0^T \langle G, v \rangle dt}{\|v\|_{\mathbb{V}_k, A_k}} + \sup_v \frac{\int_0^T (\bar{m}, \tilde{b} \cdot \nabla v)_{\Omega} dt}{\|v\|_{\mathbb{V}_k, A_k}}$ , where the suprema are over all functions in  $\mathbb{V}_k \setminus \{0\}$ . Since  $\int_0^T \tilde{a}_k^{-1} \nu L_H^{-2} \|\tilde{b} \cdot \nabla v\|_{\Omega}^2 dt \leq \|v\|_{\mathbb{V}_k, A_k}^2$  for all  $v \in \mathbb{V}_k$ , we have

$$\begin{aligned} \sup_{v \in \mathbb{V}_k \setminus \{0\}} \frac{\int_0^T (\bar{m}, \tilde{b} \cdot \nabla v)_{\Omega} dt}{\|v\|_{\mathbb{V}_k, A_k}} &\leq \left( \int_0^T \nu^{-1} L_H^2 \tilde{a}_k \|\bar{m}\|_{\Omega}^2 dt \right)^{\frac{1}{2}} \\ &\leq \gamma_k \|\bar{m}\|_{\mathbb{V}_k^+, A_k} + \frac{1}{\sqrt{2}} \|g\|_{\Omega, k}, \end{aligned} \quad (\text{B16})$$

where the second inequality follows from (B12) of Lemma B.3. Therefore, we find that

$$\|\bar{m}\|_{\mathbb{V}_k^+, A_k} \leq \gamma_k \|\bar{m}\|_{\mathbb{V}_k^+, A_k} + C (\|G\|_{L^2(0, T; H^{-1}(\Omega))} + \|g\|_{\Omega}),$$

where the constant  $C$  depends only on  $\nu$ ,  $L_H$ ,  $d$ ,  $T$ , and the shape-regularity of the meshes. We therefore obtain (6.2) after noting that  $\gamma_k \leq \gamma_1 < 1$  by (4.13) and using the equivalence of norms of Corollary B.2.  $\square$

**Proof of Lemma 6.2** Let  $k \in \mathbb{N}$ ,  $\tilde{F} \in L^2(0, T; H^{-1}(\Omega))$  and let  $\tilde{S} \in L^2(\Omega)$  be fixed. For each  $w \in \mathbb{V}_k$ , let  $\Gamma_k(w) \in \mathbb{V}_k$  be defined as the unique solution of

$$B_k(\psi, \Gamma_k(w)) = \int_0^T \langle \tilde{F}, \psi \rangle - (H[\nabla w], \psi)_{\Omega} dt + (R_k \tilde{S}, \psi(T))_{\Omega, k} \quad \forall \psi \in \mathbb{V}_{k,0}^+, \quad (\text{B17})$$

where  $B_k$  is defined in (B2) above. It is clear that  $\Gamma_k: \mathbb{V}_k \rightarrow \mathbb{V}_k$  is well-defined, and moreover it is clear that  $\bar{u} \in \mathbb{V}_k$  solves (6.3) if and only if  $\bar{u}$  is a fixed point of  $\Gamma_k$ , i.e.  $\bar{u} = \Gamma_k(\bar{u})$ . We now show that  $\Gamma_k$  is a contraction on  $\mathbb{V}_k$  with respect to the norm  $\|\cdot\|_{\mathbb{V}_k, A_k}$ , which implies existence and uniqueness of the the solution  $\bar{u}$  of (6.3) by the Banach Fixed Point Theorem. To see that  $\Gamma_k$  is a contraction map on  $\mathbb{V}_k$ , let  $w_1, w_2 \in \mathbb{V}_k$  be arbitrary, and note that (B8) implies that

$$\|\Gamma_k(w_1) - \Gamma_k(w_2)\|_{\mathbb{V}_k, A_k} = \sup_{\psi \in \mathbb{V}_{k,0}^+} \frac{\int_0^T (H[\nabla w_2] - H[\nabla w_1], \psi)_{\Omega} dt}{\|\psi\|_{\mathbb{V}_k^+, A_k}}. \quad (\text{B18})$$

The Cauchy–Schwarz inequality then implies that, for any  $\psi \in \mathbb{V}_{k,0}^+$ ,

$$\begin{aligned} & \int_0^T (H[\nabla w_2] - H[\nabla w_1], \psi)_{\Omega} dt \\ & \leq \left( \int_0^T \nu L_H^{-2} \tilde{a}_k^{-1} \|H[\nabla w_2] - H[\nabla w_1]\|_{\Omega}^2 dt \right)^{\frac{1}{2}} \left( \int_0^T \nu^{-1} L_H^2 \tilde{a}_k \|\psi\|_{\Omega}^2 dt \right)^{\frac{1}{2}}. \end{aligned} \quad (\text{B19})$$

We then see that  $\left( \int_0^T \nu L_H^{-2} \tilde{a}_k \|H[\nabla w_2] - H[\nabla w_1]\|_{\Omega}^2 dt \right)^{\frac{1}{2}} \leq \|w_1 - w_2\|_{\mathbb{V}_k, A_k}$  since  $H$  is Lipschitz continuous by (2.5) and since  $\nu \mathbb{I} \leq A_k$ . Furthermore, Lemma B.3 applied to  $\psi \in \mathbb{V}_{k,0}^+$  implies that  $\left( \int_0^T \nu^{-1} L_H^2 \tilde{a}_k \|\psi\|_{\Omega}^2 dt \right)^{\frac{1}{2}} \leq \gamma_k \|\psi\|_{\mathbb{V}_k^+, A_k}$ . Thus we find that

$$\|\Gamma_k(w_1) - \Gamma_k(w_2)\|_{\mathbb{V}_k, A_k} \leq \gamma_k \|w_1 - w_2\|_{\mathbb{V}_k, A_k}, \quad (\text{B20})$$

which shows that  $\Gamma_k$  is a contraction on  $\mathbb{V}_k$ , since  $\gamma_k \leq \gamma_1 < 1$  by (4.13).

It remains only to show the continuous dependence bound (6.4), where  $\bar{u}_1$  and  $\bar{u}_2$  are respective solutions of (6.3) with respective data  $(\tilde{F}_1, \tilde{S}_1)$  and  $(\tilde{F}_2, \tilde{S}_2)$ . Then, we have

$$\begin{aligned} B_k(\psi, u_1 - u_2) &= \int_0^T (H[\nabla u_2] - H[\nabla u_1], \psi)_{\Omega} + \langle \tilde{F}_1 - \tilde{F}_2, \psi \rangle dt \\ &\quad + (R_k \tilde{S}_1 - R_k \tilde{S}_2, \psi(T))_{\Omega, k}, \end{aligned}$$

for all  $\psi \in \mathbb{V}_{k,0}^+$ . Combining Lemma 4.3, the equivalence of norms  $\|\cdot\|_{\mathbb{V}_k^+, A_k} \approx \|\cdot\|_{\mathbb{V}_k^+}$ , the  $L^2$ -stability of  $R_k$ , and Lemma B.3, we obtain

$$\|u_1 - u_2\|_{\mathbb{V}_k, A_k} \leq \gamma_k \|u_1 - u_2\|_{\mathbb{V}_k, A_k} + C \left( \|\tilde{F}_1 - \tilde{F}_2\|_{L^2(0,T;H^{-1}(\Omega))} + \|\tilde{S}_1 - \tilde{S}_2\|_{\Omega} \right), \quad (\text{B21})$$

where  $C$  is a constant that depends on  $d, \nu, L_H, T$  and the shape-regularity of the meshes. Therefore,  $\|u_1 - u_2\|_{\mathbb{V}_k, A_k} \lesssim \|\tilde{F}_1 - \tilde{F}_2\|_{L^2(0,T;H^{-1}(\Omega))} + \|\tilde{S}_1 - \tilde{S}_2\|_{\Omega}$  since  $\gamma_k \leq \gamma_1 < 1$  by (4.13).

Next, since  $u_1$  and  $u_2$  are extended to functions in  $\mathbb{V}_k^-$  with  $u_1(T) = R_k \tilde{S}_1$  and  $u_2(T) = R_k \tilde{S}_2$ , the discrete integration-by-parts formula (4.20) implies that

$$\int_0^T -(\partial_t \mathcal{I}_-(u_1 - u_2), \psi)_{\Omega, k} dt = \int_0^T (H[\nabla u_2] - H[\nabla u_1], \psi)_{\Omega} + \langle \tilde{F}_1 - \tilde{F}_2, \psi \rangle dt$$

$$- \int_0^T (A_k \nabla(u_1 - u_2), \nabla \psi)_{\Omega} dt, \quad (\text{B22})$$

which extends to all test functions  $\psi \in \mathbb{V}_k$ . Therefore  $\int_0^T \|\partial_t \mathcal{I}_-(u_1 - u_2)\|_{k,*}^2 dt \lesssim \|u_1 - u_2\|_{\mathbb{V}_k, A_k}^2 + \|\tilde{F}_1 - \tilde{F}_2\|_{L^2(0,T;H^{-1}(\Omega))}^2$ . Combining this bound with  $\|u_1(T) - u_2(T)\|_{\Omega} \lesssim \|\tilde{S}_1 - \tilde{S}_2\|_{\Omega}$ , we conclude that  $\|u_1 - u_2\|_{\mathbb{V}_k^-} \lesssim \|\tilde{F}_1 - \tilde{F}_2\|_{L^2(0,T;H^{-1}(\Omega))} + \|\tilde{S}_1 - \tilde{S}_2\|_{\Omega}$  which completes the proof of (6.4).  $\square$

**Acknowledgements** This work was supported by the Engineering and Physical Sciences Research Council [Grant Number EP/Y008758/1]. This work was also completed with the support of The Royal Society Career Development Fellowship. We would like to acknowledge the use of the UCL Myriad High Performance Computing Facility (Myriad@UCL), and associated support services, in the completion of this work.

**Funding** This work was supported by The Royal Society Career Development Fellowship of the first author and the Engineering and Physical Sciences Research Council grant [Grant Number EP/Y008758/1] of the second author.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Achdou, Y., Camilli, F., Capuzzo-Dolcetta, I.: Mean field games: convergence of a finite difference method. *SIAM J. Numer. Anal.* **51**(5), 2585–2612 (2013). <https://doi.org/10.1137/120882421>
2. Achdou, Y., Capuzzo-Dolcetta, I.: Mean field games: numerical methods. *SIAM J. Numer. Anal.* **48**(3), 1136–1162 (2010). <https://doi.org/10.1137/090758477>
3. Achdou, Y., Cardaliaguet, P., Delarue, F., Porretta, A., Santambrogio, F.: *Mean Field Games*, vol. 2281. Springer, Switzerland (2019). <https://doi.org/10.1007/978-3-030-59837-2>
4. Achdou, Y., Porretta, A.: Convergence of a finite difference scheme to weak solutions of the system of partial differential equations arising in mean field games. *SIAM J. Numer. Anal.* **54**(1), 161–186 (2016). <https://doi.org/10.1137/15M1015455>
5. Aronson, D.G., Serrin, J.: Local behavior of solutions of quasilinear parabolic equations. *Arch. Ration. Mech. Anal.* **25**(2), 81–122 (1967). <https://doi.org/10.1007/BF00281291>
6. Aubin, J.-P., Cellina, A.: Differential inclusions. In: *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, vol. 264, p. 342. Springer, Berlin (1984). Set-valued maps and viability theory. <https://doi.org/10.1007/978-3-642-69512-4>
7. Bardi, M., Fischer, M.: On non-uniqueness and uniqueness of solutions in finite-horizon mean field games. *ESAIM Control Optim. Calc. Var.* **25**, 44–33 (2019). <https://doi.org/10.1051/cocv/2018026>
8. Barenechea, G.R., Burman, E., Karakatsani, F.: Edge-based nonlinear diffusion for finite element approximations of convection-diffusion equations and its relation to algebraic flux-correction schemes. *Numer. Math.* **135**(2), 521–545 (2017). <https://doi.org/10.1007/s00211-016-0808-z>
9. Barenechea, G.R., John, V., Knobloch, P., Rankin, R.: A unified analysis of algebraic flux correction schemes for convection-diffusion equations. *SeMA J.* **75**(4), 655–685 (2018). <https://doi.org/10.1007/s40324-018-0160-6>

10. Bonnans, J.F., Liu, K., Pfeiffer, L.: Error estimates of a theta-scheme for second-order mean field games. *ESAIM Math. Model. Numer. Anal.* **57**(4), 2493–2528 (2023). <https://doi.org/10.1051/m2an/2023059>
11. Briceño-Arias, L.M., Kalise, D., Silva, F.J.: Proximal methods for stationary mean field games with local couplings. *SIAM J. Control. Optim.* **56**(2), 801–836 (2018). <https://doi.org/10.1137/16M1095615>
12. Burman, E., Ern, A.: Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the convection-diffusion-reaction equation. *Comput. Methods Appl. Mech. Eng.* **191**(35), 3833–3855 (2002). [https://doi.org/10.1016/S0045-7825\(02\)00318-3](https://doi.org/10.1016/S0045-7825(02)00318-3)
13. Burman, E., Ern, A.: Stabilized Galerkin approximation of convection-diffusion-reaction equations: discrete maximum principle and convergence. *Math. Comp.* **74**(252), 1637–1652 (2005). <https://doi.org/10.1090/S0025-5718-05-01761-8>
14. Carlini, E., Silva, F.J.: A fully discrete semi-Lagrangian scheme for a first order mean field game problem. *SIAM J. Numer. Anal.* **52**(1), 45–67 (2014). <https://doi.org/10.1137/120902987>
15. Carlini, E., Silva, F.J.: A semi-Lagrangian scheme for a degenerate second order mean field game system. *Discrete Contin. Dyn. Syst.* **35**(9), 4269–4292 (2015). <https://doi.org/10.3934/dcds.2015.35.4269>
16. Ciarlet, P.G.: The finite element method for elliptic problems. In: *Classics in Applied Mathematics*, vol. 40, p. 530. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, (2002). <https://doi.org/10.1137/1.9780898719208>
17. Ciarlet, P.G., Raviart, P.-A.: Maximum principle and uniform convergence for the finite element method. *Comput. Methods Appl. Mech. Eng.* **2**(1), 17–31 (1973). [https://doi.org/10.1016/0045-7825\(73\)90019-4](https://doi.org/10.1016/0045-7825(73)90019-4)
18. Ducasse, R., Mazanti, G., Santambrogio, F.: Second order local minimal-time mean field games. *Nonlinear Differ. Equ. Appl.* **29**, 37 (2022). <https://doi.org/10.1007/s00030-022-00767-2>
19. Fattorini, H.O.: *Infinite Dimensional Optimization and Control Theory*, vol. 54. Cambridge University Press, London (1999). <https://doi.org/10.1017/CBO9780511574795>
20. Gomes, D.A., Pimentel, E.A., Voskanyan, V.: *Regularity Theory for Mean-Field Game Systems*. Springer, Switzerland (2016). <https://doi.org/10.1007/978-3-319-38934-9>
21. Gomes, D.A., Saúde, J.A.: Mean field games models—a brief survey. *Dyn. Games Appl.* **4**(2), 110–154 (2014). <https://doi.org/10.1007/s13235-013-0099-2>
22. Guéant, O., Lasry, J.-M., Lions, P.-L.: Mean field games and applications. In: *Paris-Princeton Lectures on Mathematical Finance 2010. Lecture Notes in Math.*, vol. 2003, pp. 205–266. Springer, Berlin (2011). [https://doi.org/10.1007/978-3-642-14660-2\\_3](https://doi.org/10.1007/978-3-642-14660-2_3)
23. Huang, M., Malhamé, R.P., Caines, P.E.: Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle. *Commun. Inf. Syst.* **6**(3), 221–252 (2006). <https://doi.org/10.4310/CIS.2006.v6.n3.a5>
24. Kuratowski, K., Ryll-Nardzewski, C.: A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **13**(6), 397–403 (1965)
25. Ladyženskaja, O.A., Solonnikov, V.A., Uralceva, N.N.: *Linear and quasilinear equations of parabolic type*. p. 648. American Mathematical Society, Providence: Translated from the Russian by Smith, S. (1968). <https://doi.org/10.1090/mmono/023>
26. Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. I—Le cas stationnaire. *Comptes Rendus Mathématique* **343**(9), 619–625 (2006). <https://doi.org/10.1016/j.crma.2006.09.019>
27. Lasry, J.-M., Lions, P.-L.: Jeux à champ moyen. II—Horizon fini et contrôle optimal. *Comptes Rendus Mathématique* **343**(10), 679–684 (2006). <https://doi.org/10.1016/j.crma.2006.09.018>
28. Lasry, J.-M., Lions, P.-L.: Mean field games. *Jpn. J. Math.* **2**(1), 229–260 (2007). <https://doi.org/10.1007/s11537-007-0657-8>
29. Mizukami, A., Hughes, T.J.R.: A Petrov–Galerkin finite element method for convection-dominated flows: an accurate upwinding technique for satisfying the maximum principle. *Comput. Methods Appl. Mech. Eng.* **50**(2), 181–193 (1985). [https://doi.org/10.1016/0045-7825\(85\)90089-1](https://doi.org/10.1016/0045-7825(85)90089-1)
30. Neumüller, M., Smears, I.: Time-parallel iterative solvers for parabolic evolution equations. *SIAM J. Sci. Comput.* **41**(1), 28–51 (2019). <https://doi.org/10.1137/18M1172466>
31. Osborne, Y.A.P.: *Analysis and Numerical Approximation of Mean Field Game Partial Differential Inclusions*. PhD thesis, University College London (2024). <https://discovery.ucl.ac.uk/id/eprint/10194241>

32. Osborne, Y.A.P., Smears, I.: Analysis and numerical approximation of stationary second-order mean field game partial differential inclusions. *SIAM J. Numer. Anal.* **62**(1), 138–166 (2024). <https://doi.org/10.1137/22M1519274>
33. Osborne, Y.A.P., Smears, I.: Erratum: analysis and numerical approximation of stationary second-order mean field game partial differential inclusions. *SIAM J. Numer. Anal.* **62**(5), 2415–2417 (2024). <https://doi.org/10.1137/24M165123X>
34. Osborne, Y.A.P., Smears, I.: Near and full Quasi-optimality of finite element approximations of stationary second-order mean field games. *arXiv preprint arXiv:2402.00685* (2024). <https://doi.org/10.48550/arXiv.2402.00685>
35. Osborne, Y.A.P., Smears, I.: Regularization of stationary second-order mean field game partial differential inclusions. *arXiv preprint arXiv:2408.10810* (2024). <https://doi.org/10.48550/arXiv.2408.10810>
36. Rathgeber, F., Ham, D.A., Mitchell, L., Lange, M., Luporini, F., McRae, A.T., Bercea, G.-T., Markall, G.R., Kelly, P.H.: Firedrake: automating the finite element method by composing abstractions. *ACM Trans. Math. Softw. (TOMS)* **43**(3), 1–27 (2016). <https://doi.org/10.1145/2998441>
37. Royden, H., Fitzpatrick, P.: Real analysis. In: Pearson Modern Classics for Advanced Mathematics Series. Pearson, Indianapolis (2018)
38. Simon, J.: Compact sets in the space  $L^p(0, T; B)$ . *Ann. Mat. Pura Appl.* **146**, 65–96 (1987). <https://doi.org/10.1007/BF01762360>
39. Smears, I., Süli, E.: Discontinuous Galerkin finite element approximation of Hamilton–Jacobi–Bellman equations with Cordes coefficients. *SIAM J. Numer. Anal.* **52**(2), 993–1016 (2014). <https://doi.org/10.1137/130909536>
40. Tabata, M.: A finite element approximation corresponding to the upwind finite differencing. *Mem. Numer. Math.* **4**, 47–63 (1977)
41. Verfürth, R.: A posteriori error estimation techniques for finite element methods. In: Numerical Mathematics and Scientific Computation, p. 393. Oxford University Press, Oxford (2013). <https://doi.org/10.1093/acprof:oso/9780199679423.001.0001>
42. Xu, J., Zikatanov, L.: A monotone finite element scheme for convection-diffusion equations. *Math. Comput.* **68**(228), 1429–1446 (1999). <https://doi.org/10.1090/S0025-5718-99-01148-5>
43. Yosida, K.: Functional Analysis, 6th edn., p. 501. Springer, Berlin (1980). <https://doi.org/10.1007/978-3-642-61859-8>
44. Zeidler, E.: Nonlinear functional analysis and its applications. In: Fixed-Point Theorems, vol. I, p. 897. Springer, New York, Translated from the German by Peter R. Wadsack (1986)

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