

# Near and full quasi-optimality of finite element approximations of stationary second-order mean field games

Yohance A. P. Osborne<sup>‡</sup> and Iain Smears<sup>†</sup>

January 28, 2025

## Abstract

We establish *a priori* error bounds for monotone stabilized finite element discretizations of stationary second-order mean field games (MFG) on Lipschitz polytopal domains. Under suitable hypotheses, we prove that the approximation is asymptotically nearly quasi-optimal in the  $H^1$ -norm in the sense that, on sufficiently fine meshes, the error between exact and computed solutions is bounded by the best approximation error of the corresponding finite element space, plus possibly an additional term, due to the stabilization, that is of optimal order with respect to the mesh size. We thereby deduce optimal rates of convergence of the error with respect to the mesh-size for solutions with sufficient regularity. We further show full asymptotic quasi-optimality of the approximation error in the more restricted case of sequences of strictly acute meshes. Our third main contribution is to further show, in the case where the domain is convex, that the convergence rate for the  $H^1$ -norm error of the value function approximation remains optimal even if the density function only has minimal regularity in  $H^1$ .

## 1 Introduction

Mean field games (MFG), introduced by Lasry & Lions [23, 24, 25] and independently by Huang, Caines & Malhamé [19], model Nash equilibria of rational games of stochastic optimal control in which there are infinitely many players. The equilibria are characterized by a system of partial differential equations (PDE) that consist of a Hamilton–Jacobi–Bellman (HJB) equation for the the value function  $u$  of the underlying optimal control problem faced by the players, and a Kolmogorov–Fokker–Planck (KFP) equation for the the density  $m$  of players in the state space of the game. We consider as a model problem a quasilinear elliptic MFG system of the form

$$-\nu\Delta u + H(x, \nabla u) = F[m](x) \quad \text{in } \Omega, \quad (1.1a)$$

$$-\nu\Delta m - \operatorname{div} \left( m \frac{\partial H}{\partial p}(x, \nabla u) \right) = G(x) \quad \text{in } \Omega, \quad (1.1b)$$

along with homogeneous Dirichlet boundary conditions  $u = 0$  and  $m = 0$  on  $\partial\Omega$ . Here  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , denotes the state space of the game, which is assumed to be a bounded,

---

<sup>†</sup>Department of Mathematical Sciences, Durham University, Stockton Road, DH1 3LE Durham, United Kingdom (yohance.a.osborne@durham.ac.uk).

<sup>‡</sup>Department of Mathematics, University College London, Gower Street, WC1E 6BT London, United Kingdom (i.smears@ucl.ac.uk).

connected, polytopal open set with Lipschitz boundary  $\partial\Omega$ . Whereas many works treat MFG with periodic boundary conditions and with zero source term in (1.1b), we consider here Dirichlet conditions and nonzero source terms. This model problem can be motivated as a steady-state solution for a game where the players have an infinite horizon zero-discounted control problem with stopping at the first-exit time from the domain  $\Omega$ . For more details on how the Dirichlet boundary conditions arise from the first-exit time stopping of the stochastic processes see for instance [16, Chapter 4]. The first-exit stopping time models players exiting the game upon reaching the boundary, which occurs for instance in applications to traffic flow and evacuation planning. The source term  $G$  in (1.1b) also arises in situations where new players enter/exit the game, or possibly after transformation of the problem from one with an inhomogeneous Dirichlet condition to one with a homogeneous one. Note that for MFG systems with Dirichlet boundary conditions, the player density  $m$  is usually not required to be a probability density. The real-valued Hamiltonian  $H = H(x, p)$ ,  $x \in \bar{\Omega}$ ,  $p \in \mathbb{R}^d$ , in (1.1) originates from the underlying optimal control problem, c.f. (2.3) below, and is generally a convex function in its second variable. In many applications, it is also globally Lipschitz continuous in the second variable, which results for instance from compactness of the underlying set of controls and continuity of the stochastic dynamics with respect to the controls. In this work, we assume in addition that  $H$  is everywhere differentiable with respect to  $p$  with a partial derivative  $\frac{\partial H}{\partial p}$  that is Lipschitz continuous with respect to  $p \in \mathbb{R}^d$ , uniformly in  $x \in \Omega$ . The diffusion parameter  $\nu > 0$  in (1.1) is a constant. The coupling operator  $F : L^2(\Omega) \rightarrow H^{-1}(\Omega)$  is a component of the running cost in the players' optimal control problems, and our analysis allows it to be either a local or nonlocal operator. Note that in contrast to many works on MFG, we do not require that  $F$  be *smoothing*; in fact  $F$  can even be of differential type.

In this work, we are primarily interested in the *a priori* error analysis of numerical approximations of the solutions of MFG systems. Existing results on error analysis for MFG in the literature so far have primarily been of a *qualitative* kind, such as proving that the error (in some norm) between exact and computed solutions vanishes in the small mesh/grid limit; this is often called *plain* convergence. For problems with differentiable Hamiltonians, there are results on the plain convergence of methods on Cartesian grids, such as monotone finite difference methods [2, 1, 3, 20] and semi-Lagrangian methods [10, 11, 12, 13, 5]. Convergence of semi-Lagrangian methods has also been shown for fractional and nonlocal MFG in [14]. For problems with nondifferentiable Hamiltonians, we recently showed in [27, 28, 29] that the MFG system can be generalized as a *Partial Differential Inclusion* (PDI), and we proved the convergence of a monotone stabilized Finite Element Method (FEM) with piecewise affine elements. Although the analysis in these works was of a more qualitative nature, the quantitative performance of the FEM for MFG has been shown through extensive numerical experiments in [26, 29, 27, 28]. Furthermore, in [30], we analysed the connection between the MFG PDI for nondifferentiable Hamiltonians and its classical PDE counterpart via regularization of the Hamiltonians.

In contrast to the range of qualitative results in the literature, *quantitative* results on the error analysis, such as bounds on the error on a given computational mesh/grid, are rare. Some of the main challenges of the MFG system in this regard include the quasilinear coupling of the system and the general lack of coercivity properties of the differential operators. To the best of our knowledge, quantitative error bounds have so far only been obtained by Bonnans, Liu & Pfeiffer in [6]. In particular, in the context of a fully discrete finite difference scheme for some second-order parabolic MFG systems, they showed that if the exact solution  $u$  and  $m$  have  $C^{1+\tau/2, 2+\tau}$  regularity over the space-time domain for some

$r \in (0, 1)$ , then the error, as measured in the discrete maximum norm for the value function and a discrete  $L^\infty(L^1)$ -norm for the density, is of order  $h^r$  (the time-step size  $\Delta t$  does not appear explicitly in the bound owing to their assumption that  $\Delta t \lesssim h^2$ ). Note that [6] allows also for some Hamiltonians with superlinear growth. Besides the results in [6], quantitative error bounds for the numerical approximation of solutions of MFG systems appear to be largely untouched.

As our main contributions in this work, we show error bounds for the stabilized piecewise affine FEM from [27, 28, 29] applied to the model problem (1.1). The motivation for considering stabilized FEM is, in a first instance, that the stabilization leads to a discrete maximum principle and nonnegativity of the approximations in the density, which is then used in the proof of well-posedness of the discrete problems, see [27, 28, Theorem 5.3]. It is thus natural that the nonnegativity of the density approximations also plays an important role in the analysis of the error bounds. The stabilization also plays a role in uniform stability bounds for the discretized HJB and KFP equations, which are essential for the analysis of the MFG system.

Our first main result, stated in Theorem 4.1 below, is an  $H^1$ -norm error bound for a class of stabilized FEM which covers the stabilizations of [29, 27, 28]. In particular, when applied to the method of [29], our bound implies that for a shape-regular sequence of computational meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  that satisfies the Xu–Zikatanov condition [34], the finite element approximation  $(u_k, m_k) \in V_k \times V_k$  satisfies

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} + \|h_{\mathcal{T}_k} \nabla m\|_{L^2(\Omega)} + \|h_{\mathcal{T}_k} \nabla u\|_{L^2(\Omega)}, \end{aligned} \quad (1.2)$$

for all  $k$  is sufficiently large, where  $V_k$  denotes the  $H_0^1$ -conforming finite element space on the mesh  $\mathcal{T}_k$ , where  $h_{\mathcal{T}_k}$  denotes the (elementwise) mesh-size function of the mesh, and where the constant hidden in the inequality is independent of  $k \in \mathbb{N}$  (see Section 2 for details on the notation). The additional terms  $\|h_{\mathcal{T}_k} \nabla m\|_{L^2(\Omega)}$  and  $\|h_{\mathcal{T}_k} \nabla u\|_{L^2(\Omega)}$  in (1.2) stem from the stabilization terms of the method that are included to satisfy a discrete maximum principle. The bound (1.2) holds for solutions  $(u, m)$  with the minimal  $H^1$ -regularity that is guaranteed by the well-posedness theory in [27, 28]. The main assumptions for the analysis are that the Hamiltonian  $H$  has a Lipschitz continuous partial derivative w.r.t. the gradient variable uniformly in space, that the coupling term  $F : L^2(\Omega) \rightarrow H^{-1}(\Omega)$  is strongly monotone, and that the source term  $G$  is nonnegative.

Recall that a numerical method is said to be *quasi-optimal* in a given norm if the error between exact and computed solutions is bounded by a constant times the best-approximation error from the same approximation space. It is asymptotically quasi-optimal if the mesh/grid is additionally required to be sufficiently fine. Since the additional terms in (1.2) are typically of same or higher order than the  $H^1$ -norm best-approximation error for piecewise affine elements, we therefore say that the stabilized FEM considered here is *nearly* asymptotically quasi-optimal in the  $H^1$ -norm. If a stronger condition is placed on the meshes, namely a strict acuteness condition as in [27, 28], then these additional terms can be removed and the resulting method is *fully* asymptotically quasi-optimal with

$$\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)}, \quad (1.3)$$

for all  $k$  sufficiently large. However, the fact that the Xu–Zikatanov condition is less restrictive in practice than the strict acuteness condition means that the bound (1.2) is likely the more relevant one in computational practice.

As a consequence of the bound (1.2), for sufficiently regular solutions, in Corollary 4.2 below, we prove that the numerical approximations have optimal rates of convergence with respect to the mesh-size. For instance, if both  $u, m$  are in  $H^2(\Omega)$ , then

$$\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \lesssim h_k (\|m\|_{H^2(\Omega)} + \|u\|_{H^2(\Omega)}), \quad (1.4)$$

where  $h_k = \|h_{\mathcal{T}_k}\|_{L^\infty(\Omega)}$  is the maximum mesh-size across the domain, see Corollary 4.2 below covering the more general case where the solution has regularity in Sobolev spaces of fractional order. To the best of our knowledge, this is the first proof of optimal rates of convergence for any numerical method for MFG systems in the literature so far.

The second main contribution of this work is to analyse the rates of convergence of the errors for the separate components of the solution. More precisely, we aim to explain why it was observed in the numerical experiments of [27, 28] that the convergence rate of the error  $\|u - u_k\|_{H^1(\Omega)}$  for the value function approximations was of optimal order  $h_k$  even in cases where the density  $m$  has low regularity. This observation is not explained by the bounds above, i.e. (1.4) requires  $m \in H^2(\Omega)$ . In order to illuminate this matter, we prove in Theorem 4.3 below that

$$\|m - m_k\|_{L^2(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \lesssim \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} + h_k (\|u\|_{H^1(\Omega)} + \|m\|_{H^1(\Omega)}), \quad (1.5)$$

for all  $k$  sufficiently large, which holds even if  $u$  and  $m$  have minimal regularity in  $H^1(\Omega)$ , under the additional assumption that  $\Omega$  is convex. Note crucially that the left-hand side of (1.5) involves only the  $L^2$ -norm of the error for the density, rather than the  $H^1$ -norm as in (1.2), which is key for obtaining a sharper bound on  $\|u - u_k\|_{H^1(\Omega)}$ . The bound (1.5) implies that the convergence rate of the value function approximations does not generally depend on the regularity of  $m$ . For instance, if  $u \in H^2(\Omega)$ , then (1.5) implies that  $\|m - m_k\|_{L^2(\Omega)} + \|u - u_k\|_{H^1(\Omega)}$  is of order  $h_k$ , even for  $m$  with minimal regularity in  $H^1$ . This result offers some insight into earlier computational results, despite the difference that the experiments in [27, 28] involved nondifferentiable Hamiltonians whereas the proof of (1.5) assumes that  $H$  has a Lipschitz continuous derivative.

This paper is organized as follows. In Section 2 we set the notation and specify the assumptions on the problem. In Section 3 we present the framework for discretization, including the general class of stabilized FEM covered by the analysis, as well as two concrete instances from [27, 28, 29]. The statement of the main results mentioned above are given in Section 4. The proofs of the main results are then the subjects of Sections 5 to 8. In particular, Section 5 shows key stability results for the discretized HJB and KFP equations. In Section 6, we consider the near quasi-optimality of nonnegative approximations of the density function, and the  $L^2$ -norm stability of the discrete MFG system. We then apply these results in Section 7 to prove Theorem 4.1 and in Section 8 to prove Theorem 4.3. We conclude the paper in Section 9 with numerical experiments that illustrate Corollary 4.2 of Theorem 4.1 and the conclusion of Theorem 4.3 for problems involving nonsmooth solutions.

## 2 Setting and Notation

**Notation for inequalities.** In order to avoid the proliferation of generic constants, in the following, we write  $a \lesssim b$  for real numbers  $a$  and  $b$  if  $a \leq Cb$  for some constant  $C$  that may depend on the problem data and some fixed parameters for the sequence of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  appearing below, such as the shape-regularity parameter, but is otherwise independent of

the mesh  $\mathcal{T}_k$  and the mesh-size  $h_k$ . Throughout this work we will regularly specify the particular dependencies of the hidden constants.

**General notation.** We denote  $\mathbb{N} := \{1, 2, 3, \dots\}$ . For a Lebesgue measurable set  $\omega \subset \mathbb{R}^d$ ,  $d \in \mathbb{N}$ , the inner product on  $L^2(\omega)$  and  $L^2(\omega; \mathbb{R}^d)$  is denoted commonly by  $(\cdot, \cdot)_\omega$ , with induced norm  $\|\cdot\|_\omega$ . There is no risk of confusion here since the scalar and vector cases of the inner-product and norm can then be determined from their arguments. Moreover, we equip the space  $L^\infty(\omega; \mathbb{R}^{d \times d})$  of bounded matrix-valued functions on  $\omega$  with the essential supremum norm  $\|\cdot\|_{L^\infty(\omega; \mathbb{R}^{d \times d})}$  that is induced by the Frobenius norm on  $\mathbb{R}^{d \times d}$ . We let  $|\cdot|_{d-2}$  denote the  $(d-2)$ -dimensional Hausdorff measure, which we note reduces to the counting measure when  $d = 2$ .

**Problem data.** Let  $\Omega$  be a bounded, connected, polytopal open subset of  $\mathbb{R}^d$  with Lipschitz boundary  $\partial\Omega$ . Note that although we focus here on the case of polytopal domains, a common procedure for the case of domains with curved boundaries is to approximate the geometry via the meshes; this technique has been studied extensively in the literature on numerics for surface PDEs [9]. For an integer  $s \geq 0$ , let  $H^s(\Omega)$  denote the Sobolev space of order  $s$ , which consists of functions in  $L^2(\Omega)$  that have weak partial derivatives of order up to and including  $s$  also in  $L^2(\Omega)$ . Let  $H_0^1(\Omega)$  denote the space of functions in  $H^1(\Omega)$  with vanishing trace on  $\Omega$ , and let  $H^{-1}(\Omega)$  denote the dual space of  $H_0^1(\Omega)$ . Since  $\Omega$  is assumed to have a Lipschitz boundary, the Sobolev space  $H^s(\Omega)$  can be defined for real  $s \geq 0$ , for instance, by interpolation of spaces, see [4, Chapter 7] and [7, Chapter 14].

We make the following assumptions on the data appearing in (1.1). Let the diffusion  $\nu > 0$  be constant, and let  $G \in H^{-1}(\Omega)$  be of the form  $G = g_0 - \nabla \cdot \tilde{g}$  with  $g_0 \in L^{q/2}(\Omega)$  and  $\tilde{g} \in L^q(\Omega; \mathbb{R}^d)$  for some  $q > d$ . We assume that  $G \in H^{-1}(\Omega)$  is nonnegative in the distributional sense, i.e. that  $\langle G, \phi \rangle_{H^{-1} \times H_0^1} \geq 0$  for all functions  $\phi \in H_0^1(\Omega)$  that are nonnegative a.e. in  $\Omega$ . Next, let  $F : L^2(\Omega) \rightarrow H^{-1}(\Omega)$  be a Lipschitz continuous operator, i.e. that satisfies

$$\|F[w] - F[v]\|_{H^{-1}(\Omega)} \leq L_F \|w - v\|_\Omega \quad \forall w, v \in L^2(\Omega) \quad (2.1)$$

for some constant  $L_F \geq 0$ . Note that  $F$  is allowed to be either a local or nonlocal operator; and also that  $F$  is not necessarily of smoothing type; in fact  $F$  can even be of differential type. We assume that  $F$  is *strongly monotone* in the sense that there exists a constant  $c_F > 0$  such that

$$c_F \|w - v\|_\Omega^2 \leq \langle F[w] - F[v], w - v \rangle_{H^{-1} \times H_0^1} \quad \forall w, v \in H_0^1(\Omega). \quad (2.2)$$

Note that although the domain of  $F$  is  $L^2(\Omega)$ , the monotonicity condition (2.2) is needed only for arguments in the smaller space  $H_0^1(\Omega)$ . We refer the reader to [27, 28, Example 1] for several examples of operators  $F$  that satisfy the above conditions. The strong monotonicity condition (2.2) can be seen as a quantitative version of the strict monotonicity condition on  $F$  that is employed to prove uniqueness of the solution of the problem, see e.g. [25], and to prove plain convergence of numerical approximations, see e.g. [27, 28].

In applications, the Hamiltonian  $H$  is usually specified in terms of the controlled drift and running cost of the associated optimal control problem, so we assume that

$$H(x, p) := \sup_{\alpha \in \mathcal{A}} (b(x, \alpha) \cdot p - f(x, \alpha)), \quad \forall (x, p) \in \bar{\Omega} \times \mathbb{R}^d, \quad (2.3)$$

where we let the control set  $\mathcal{A}$  denote a compact metric space, and we assume that both the control-dependent drift  $b : \Omega \times \mathcal{A} \rightarrow \mathbb{R}^d$  and the control-dependent component of the running cost  $f : \Omega \times \mathcal{A} \rightarrow \mathbb{R}$  are uniformly continuous. The Hamiltonian  $H$  is then convex w.r.t.  $p$  uniformly in  $x \in \bar{\Omega}$  and satisfies the following bounds

$$|H(x, p)| \leq C_H (|p| + 1) \quad \forall (x, p) \in \bar{\Omega} \times \mathbb{R}^d, \quad (2.4a)$$

$$|H(x, p) - H(x, q)| \leq L_H |p - q| \quad \forall (x, p, q) \in \bar{\Omega} \times \mathbb{R}^d \times \mathbb{R}^d, \quad (2.4b)$$

with  $C_H := \max \left\{ \|b\|_{C(\bar{\Omega} \times \mathcal{A}; \mathbb{R}^d)}, \|f\|_{C(\bar{\Omega} \times \mathcal{A})} \right\}$  and  $L_H := \|b\|_{C(\bar{\Omega} \times \mathcal{A}; \mathbb{R}^d)}$ . We assume that  $H$  is differentiable w.r.t.  $p$  with continuous derivative  $\frac{\partial H}{\partial p} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  satisfying the Lipschitz condition

$$\left| \frac{\partial H}{\partial p}(x, p) - \frac{\partial H}{\partial p}(x, q) \right| \leq L_{H_p} |p - q| \quad \forall (x, p, q) \in \Omega \times \mathbb{R}^d \times \mathbb{R}^d, \quad (2.5)$$

for some constant  $L_{H_p} \geq 0$ . Furthermore, we note that (2.4b) implies that  $\frac{\partial H}{\partial p}$  is uniformly bounded with

$$\left| \frac{\partial H}{\partial p}(x, p) \right| \leq L_H \quad \forall (x, p) \in \Omega \times \mathbb{R}^d. \quad (2.6)$$

It is clear that the mappings  $H^1(\Omega) \ni v \mapsto H(\cdot, \nabla v) \in L^2(\Omega)$  and  $H^1(\Omega) \ni v \mapsto \frac{\partial H}{\partial p}(\cdot, \nabla v) \in L^2(\Omega; \mathbb{R}^d)$  are Lipschitz continuous. Furthermore, (2.6) implies that  $\frac{\partial H}{\partial p}(\cdot, \nabla v) \in L^\infty(\Omega; \mathbb{R}^d)$  for each  $v \in H^1(\Omega)$ . For given  $v \in H^1(\Omega)$ , we will often abbreviate these compositions by writing  $H[\nabla v] := H(\cdot, \nabla v)$  and  $\frac{\partial H}{\partial p}[\nabla v] := \frac{\partial H}{\partial p}(\cdot, \nabla v)$  a.e. in  $\Omega$ .

We conclude this section with a bound on the linearizations of the Hamiltonians when composed with gradients of functions in Sobolev spaces.

**Lemma 2.1.** *For any  $\epsilon > 0$ , there exists a  $R > 0$ , depending only on  $\epsilon$ ,  $\Omega$ ,  $L_H$  and  $L_{H_p}$ , such that*

$$\left\| H[\nabla v] - H[\nabla w] - \frac{\partial H}{\partial p}[\nabla w] \cdot \nabla(v - w) \right\|_{H^{-1}(\Omega)} \leq \epsilon \|v - w\|_{H^1(\Omega)}, \quad (2.7)$$

whenever  $v, w \in H^1(\Omega)$  satisfy  $\|v - w\|_{H^1(\Omega)} \leq R$ .

*Proof.* To alleviate the notation, let  $R_H[\nabla v, \nabla w] := H[\nabla v] - H[\nabla w] - \frac{\partial H}{\partial p}[\nabla w] \cdot \nabla(v - w)$  for any  $v, w \in H^1(\Omega)$ . We note that, for every  $v, w \in H^1(\Omega)$ ,  $R_H[\nabla v, \nabla w]$  is well-defined as a Lebesgue measurable function on  $\Omega$  since  $H : \bar{\Omega} \times \mathbb{R}^d \rightarrow \mathbb{R}$  and  $\frac{\partial H}{\partial p} : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  are both continuous, and the linear growth property (2.4a) and the uniform bound (2.6) imply that  $R_H[\nabla v, \nabla w] \in L^2(\Omega)$ . Define the real-number  $s_* \in (1, 2)$  by  $s_* = 3/2$  if  $d = 2$  and  $s_* := \frac{2d}{d+2}$  if  $d \geq 3$ . Note that the Sobolev embedding theorem shows that  $H_0^1(\Omega)$  is continuously embedded in  $L^{s'_*}(\Omega)$  where  $1/s_* + 1/s'_* = 1$ . It then follows from the Hölder inequality that  $L^{s_*}(\Omega)$  is continuously embedded in  $H^{-1}(\Omega)$  and thus

$$\|R_H[\nabla v, \nabla w]\|_{H^{-1}(\Omega)} \leq C_\Omega \|R_H[\nabla v, \nabla w]\|_{L^{s_*}(\Omega)} \quad (2.8)$$

for all  $v$  and  $w$  in  $H^1(\Omega)$ , where we note that the embedding constant  $C_\Omega$  depends only on  $\Omega$  and  $R_H[\nabla v, \nabla w] \in L^{s_*}(\Omega)$  since  $L^2(\Omega) \subset L^{s_*}(\Omega)$ . We now derive an upper bound for the term  $\|R_H[\nabla v, \nabla w]\|_{L^{s_*}(\Omega)}$  as follows.

Let  $\gamma \in [0, 1]$ ,  $x \in \Omega$  and  $p, q \in \mathbb{R}^d$  be given. The Mean Value Theorem implies the existence of  $\theta_{x,p,q} \in (0, 1)$  such that

$$\begin{aligned} H(x, p) - H(x, q) - \frac{\partial H}{\partial p}(x, q) \cdot (p - q) \\ = \left( \frac{\partial H}{\partial p}(x, \theta_{x,p,q}p + (1 - \theta_{x,p,q})q) - \frac{\partial H}{\partial p}(x, q) \right) \cdot (p - q). \end{aligned} \quad (2.9)$$

Since  $\theta_{x,p,q} \in (0, 1)$ , the above identity (2.9), the Lipschitz condition (2.5) and the uniform bound (2.6) together imply that

$$\begin{aligned} \left| H(x, p) - H(x, q) - \frac{\partial H}{\partial p}(x, q) \cdot (p - q) \right| \\ \leq (2L_H)^{1-\gamma} L_{H_p}^\gamma |p - q|^{\gamma+1} \quad \forall (p, q) \in \mathbb{R}^d \times \mathbb{R}^d, \quad \forall x \in \Omega \end{aligned} \quad (2.10)$$

and for any  $\gamma \in [0, 1]$ . Consequently, for any  $v, w \in H^1(\Omega)$  and  $\gamma \in [0, 1]$ , we have

$$|R_H[\nabla v, \nabla w]| \leq (2L_H)^{1-\gamma} L_{H_p}^\gamma |\nabla(v - w)|^{\gamma+1} \quad (2.11)$$

a.e. in  $\Omega$ . Hence, we obtain from (2.11) the bound

$$\|R_H[\nabla v, \nabla w]\|_{L^{s_*}(\Omega)} \leq (2L_H)^{1-\gamma} L_{H_p}^\gamma \|\nabla(v - w)\|_{L^{(1+\gamma)s_*}(\Omega)}^{1+\gamma}. \quad (2.12)$$

We now choose  $\gamma = 1/9$  if  $d = 2$  and  $0 < \gamma < 2/d$  if  $d \geq 3$ , which ensures that  $(\gamma + 1)s_* \in (1, 2)$ . Thus, by applying the Hölder inequality we obtain

$$\begin{aligned} \|\nabla(v - w)\|_{L^{(1+\gamma)s_*}(\Omega)}^{1+\gamma} &= \left( \int_{\Omega} |\nabla(v - w)|^{(1+\gamma)s_*} dx \right)^{\frac{1}{s_*}} \\ &\leq \left( \int_{\Omega} 1 dx \right)^{\frac{1}{s_*} - \frac{1+\gamma}{2}} \|\nabla(v - w)\|_{L^2(\Omega)}^{1+\gamma}, \end{aligned} \quad (2.13)$$

which, together with (2.8) and (2.12), implies

$$\|R_H[\nabla v, \nabla w]\|_{H^{-1}(\Omega)} \leq \tilde{C}_{\Omega, d} (2L_H)^{1-\gamma} L_{H_p}^\gamma \|v - w\|_{H^1(\Omega)}^{\gamma+1}, \quad (2.14)$$

for some constant  $\tilde{C}_{\Omega}$  depending only on  $\Omega$  and  $d$ . Then, for any  $\epsilon > 0$ , it is then clear that (2.7) holds whenever  $\|v - w\|_{H^1(\Omega)} \leq R$  for  $R := \left( \epsilon^{-1} \tilde{C}_{\Omega, d} (2L_H)^{1-\gamma} L_{H_p}^\gamma \right)^{-1/\gamma}$ .  $\square$

**Remark 2.1** (Semismoothness of the Hamiltonian). Lemma 2.1 can be seen as a special case (with a simpler proof) of a more general result on the semismoothness of Hamilton–Jacobi–Bellman operators on function spaces, which was first shown in [33, Theorem 13], where differentiability of the Hamiltonian is not required.

## 2.1 Continuous Problem

The weak form of (1.1) is to find  $(u, m) \in H_0^1(\Omega) \times H_0^1(\Omega)$  that satisfies

$$\int_{\Omega} \nu \nabla u \cdot \nabla \psi + H[\nabla u] \psi \, dx = \langle F[m], \psi \rangle_{H^{-1} \times H_0^1}, \quad (2.15a)$$

$$\int_{\Omega} \nu \nabla m \cdot \nabla \phi + m \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi \, dx = \langle G, \phi \rangle_{H^{-1} \times H_0^1}, \quad (2.15b)$$

for all  $\psi, \phi \in H_0^1(\Omega)$ .

**Remark 2.2** (Existence and uniqueness of weak solutions). By [27, 28, Theorem 3.3, Theorem 3.4], there exists a unique solution  $(u, m)$  to the continuous problem (2.15), and we have the bounds

$$\|m\|_{H^1(\Omega)} \lesssim \|G\|_{H^{-1}(\Omega)}, \quad (2.16)$$

$$\|u\|_{H^1(\Omega)} \lesssim 1 + \|G\|_{H^{-1}(\Omega)} + \|f\|_{C(\bar{\Omega} \times \mathcal{A})}, \quad (2.17)$$

where the hidden constants depend only on  $d$ ,  $\Omega$ ,  $\nu$ ,  $L_H$ , and  $L_F$ . Furthermore, the density function  $m$  is nonnegative a.e. in  $\Omega$  owing to the assumption that  $G$  is nonnegative in the sense of distributions.

**Remark 2.3** (Essential boundedness of the density function). We have assumed above that the source term  $G$  has the form  $G = g_0 - \nabla \cdot \tilde{g}$  with  $g_0 \in L^{q/2}(\Omega)$  and  $\tilde{g} \in L^q(\Omega; \mathbb{R}^d)$  for some  $q > d$ . This allows us to apply [17, Theorem 8.15] which shows that the density function  $m$  is essentially bounded in  $\Omega$  and satisfies  $\|m\|_{L^\infty(\Omega)} \lesssim \|m\|_\Omega + \|g_0\|_{L^{q/2}(\Omega)} + \|\tilde{g}\|_{L^q(\Omega; \mathbb{R}^d)}$  where the hidden constant depends only on  $d, \Omega, L_H, q$  and  $\nu$ . In fact, the  $H^1$ -norm bound (2.16) given in Remark 2.2 further implies that there exists a constant  $M_\infty$  of the form  $M_\infty = C(\|G\|_{H^{-1}(\Omega)} + \|g_0\|_{L^{q/2}(\Omega)} + \|\tilde{g}\|_{L^q(\Omega; \mathbb{R}^d)})$ , with a constant  $C$  depending only on  $d, \Omega, L_H, \nu, q$  and  $L_F$ , such that

$$\|m\|_{L^\infty(\Omega)} \leq M_\infty. \quad (2.18)$$

We will use (2.18) in the analysis that follows.

### 3 Finite Element Discretization

**Meshes.** We let  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  denote a sequence of conforming simplicial meshes of the domain  $\Omega$  (c.f. [15, p. 51]) that are nested, which is to say that each element in  $\mathcal{T}_{k+1}$  is either in  $\mathcal{T}_k$  or is a subdivision of an element in  $\mathcal{T}_k$ . Note that the nestedness assumption will not be used explicitly in the analysis in this paper, but rather only implicitly in order to apply some existing results from earlier works, see Remark 3.2 below. For each  $k \in \mathbb{N}$  and each element  $K \in \mathcal{T}_k$ , let  $\text{diam } K$  denote the diameter of  $K$ . We define the mesh-size function  $h_{\mathcal{T}_k} \in L^\infty(\Omega)$  by  $h_{\mathcal{T}_k}|_K := \text{diam } K$  for each element  $K \in \mathcal{T}_k$ . For each  $K \in \mathcal{T}_k$ , let  $\rho_K$  denote the radius of the largest inscribed ball in  $K$ . We assume that the sequence of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  is shape-regular, i.e. that there exists a  $\delta > 1$ , independent of  $k \in \mathbb{N}$ , such that  $h_{\mathcal{T}_k}|_K \leq \delta \rho_K$  for all  $K \in \mathcal{T}_k$  and for all  $k \in \mathbb{N}$ . We let  $h_k := \|h_{\mathcal{T}_k}\|_{L^\infty(\Omega)}$  denote the maximum element diameter in  $\mathcal{T}_k$ , which we assume satisfies  $h_k \rightarrow 0$  as  $k \rightarrow \infty$ .

**Vertices and edges.** For each  $k \in \mathbb{N}$ , we let  $\mathcal{V}_k$  denote the set of all vertices of the mesh  $\mathcal{T}_k$ . We denote the set of vertices of  $\mathcal{T}_k$  that lie in the interior of the domain  $\Omega$  by  $\mathcal{V}_{k,\Omega} = \mathcal{V}_k \cap \Omega$ . Let  $\{x_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  denote an enumeration of  $\mathcal{V}_k$  which, without loss of generality, we assume to be ordered in such a way that  $x_i \in \mathcal{V}_{k,\Omega}$  if and only if  $1 \leq i \leq M_k := \text{card } \mathcal{V}_{k,\Omega}$ . We call two distinct vertices in  $\mathcal{V}_k$  neighbours if they belong to a common element of  $\mathcal{T}_k$ . The set of neighbouring vertices of a given vertex  $x_i$  is denoted by  $\mathcal{V}_{k,i}$ .

Given  $k \in \mathbb{N}$ , the collection of all closed line segments formed by all pairs of neighbouring vertices of the mesh  $\mathcal{T}_k$  will be denoted by  $\mathcal{E}_k$ . Thus, for each  $k \in \mathbb{N}$ ,  $\mathcal{E}_k$  is simply the set of edges of the mesh  $\mathcal{T}_k$ . Note that in two space dimensions, edges and faces of the mesh coincide, whereas they are distinct in three space dimensions and above. For each edge  $E \in \mathcal{E}_k$ , we let the set of elements of  $\mathcal{T}_k$  containing  $E$  be denoted by  $\mathcal{T}_{k,E} := \{K \in \mathcal{T}_k : E \subset K\}$ . An edge  $E \in \mathcal{E}_k$  is called an internal edge if there exists at least one vertex  $x_* \in \mathcal{V}_{k,\Omega}$  such



that  $x_* \in E$ . We denote the set of all internal edges by  $\mathcal{E}_{k,\Omega}$ . For each  $1 \leq i \leq \text{card } \mathcal{V}_k$ , we let  $\mathcal{E}_{k,i} := \{E \in \mathcal{E}_k : x_i \in E\}$  denote the set of edges containing the vertex  $x_i \in \mathcal{V}_k$ . Given a simplex  $K \in \mathcal{T}_k$  we define the set  $\mathcal{E}_K := \{E \subset K : E \in \mathcal{E}_{k,\Omega}\}$  which is the collection of edges of  $K$  that are internal edges.

**Approximation space.** Let  $k \in \mathbb{N}$  be given. We let  $\{\xi_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  denote the standard nodal Lagrange basis for the space of all continuous piecewise affine functions on  $\bar{\Omega}$  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  $\{x_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  of  $\mathcal{V}_k$ . Note that  $\{\xi_i\}_{i=1}^{\text{card } \mathcal{V}_k}$  form a partition of unity on  $\Omega$ . 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.

For each element  $K \subset \mathbb{R}^d$ , let the vector space of  $d$ -variate real-valued polynomials on  $K$  of total degree at most one be denoted by  $\mathcal{P}_1(K)$ . For each  $k \in \mathbb{N}$ , we let the finite element space  $V_k$  of  $H_0^1$ -conforming piecewise affine functions be defined by

$$V_k := \{v_k \in H_0^1(\Omega) : v_k|_K \in \mathcal{P}_1(K) \quad \forall K \in \mathcal{T}_k\}. \quad (3.1)$$

We let the space  $V_k$  inherit the standard norm on  $H_0^1(\Omega)$ , and we denote the norm by  $\|\phi_k\|_{V_k} := \|\phi_k\|_{H^1(\Omega)}$  for  $\phi_k \in V_k$ . Note that  $\{\xi_i\}_{i=1}^{M_k}$  is the standard nodal basis of  $V_k$  since  $x_i \in \mathcal{V}_{k,\Omega}$  if and only if  $1 \leq i \leq M_k := \text{card } \mathcal{V}_{k,\Omega}$ . Note also that the union  $\cup_{k \in \mathbb{N}} V_k$  is a dense subspace of  $H_0^1(\Omega)$ . Furthermore, the space of continuous linear functionals on  $V_k$  is denoted by  $V_k^*$ , 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 dual norm  $\|\cdot\|_{V_k^*}$  defined by

$$\|g\|_{V_k^*} := \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \langle g, \phi_k \rangle_{V_k^* \times V_k} \quad \forall g \in V_k^*. \quad (3.2)$$

Given an operator  $L : V_k \rightarrow V_k^*$  we define its adjoint operator  $L^* : V_k \rightarrow V_k^*$  by the pairing  $\langle L^* w_k, v_k \rangle_{V_k^* \times V_k} := \langle L v_k, w_k \rangle_{V_k^* \times V_k}$  for all  $w_k, v_k \in V_k$ .

### 3.1 General stabilized methods

In order to accommodate a range of stabilized methods such as those in [29, 27, 28], we consider the class of discretizations of (2.15) of the following form: *given  $k \in \mathbb{N}$ , find  $(u_k, m_k) \in V_k \times V_k$  such that*

$$\int_{\Omega} A_k \nabla u_k \cdot \nabla \psi_k + H[\nabla u_k] \psi_k \, dx = \langle F[m_k], \psi_k \rangle_{H^{-1} \times H_0^1}, \quad (3.3a)$$

$$\int_{\Omega} A_k \nabla m_k \cdot \nabla \phi_k + m_k \frac{\partial H}{\partial p}[\nabla u_k] \cdot \nabla \phi_k \, dx = \langle G, \phi_k \rangle_{H^{-1} \times H_0^1}, \quad (3.3b)$$

for all  $\psi_k, \phi_k \in V_k$ . Here  $A_k \in L^\infty(\Omega; \mathbb{R}_{\text{sym}}^{d \times d})$ , where  $\mathbb{R}_{\text{sym}}^{d \times d}$  denotes the space of symmetric matrices in  $\mathbb{R}^{d \times d}$ , is a numerical diffusion tensor defined as

$$A_k := \nu \mathbb{I}_d + D_k, \quad (3.4)$$

where  $\mathbb{I}_d$  denotes the  $d \times d$  identity matrix, and  $D_k \in L^\infty(\Omega; \mathbb{R}_{\text{sym},+}^{d \times d})$ , where  $\mathbb{R}_{\text{sym},+}^{d \times d}$  denotes the cone of positive semi-definite matrices in  $\mathbb{R}_{\text{sym}}^{d \times d}$ , is a stabilization term determined by the choice of method. In this work, we will consider two prime examples of choices of  $D_k$ , namely the construction from [27, 28] for the case of strictly acute meshes, and the construction from [29] for more general meshes that satisfy the Xu–Zikatanov (XZ) condition, see Sections 3.2 and 3.3 below.

In order to give a unified analysis of the different stabilization methods, we formulate the main features required for the analysis in the following two assumptions, which we verify in the two examples of meshes satisfying the XZ condition and of strictly acute meshes in the following subsections. Our first main assumption ensures that the stabilization term is of optimal order:

(H1) The matrix-valued function  $D_k \in L^\infty(\Omega; \mathbb{R}_{\text{sym},+}^{d \times d})$  for all  $k \in \mathbb{N}$  and there exists a constant  $C_D$ , independent of  $k$ , such that  $|D_k| \leq C_D h_{\mathcal{T}_k}$  a.e. in  $\Omega$ , for all  $k \in \mathbb{N}$ .

Our second main assumption ensures that the scheme (3.3) is monotone, due to the satisfaction of a discrete maximum principle by the discretizations. Recall that a linear operator  $L : V_k \rightarrow V_k^*$  is said to satisfy the *discrete maximum principle* (DMP) provided that the following condition holds: if  $v_k \in V_k$  and  $\langle Lv_k, \xi_i \rangle_{V_k^* \times V_k} \geq 0$  for all  $i \in \{1, \dots, M_k\}$ , then  $v_k \geq 0$  in  $\Omega$ . For each  $k \in \mathbb{N}$ , let  $W(V_k, D_k)$  denote the set of all linear operators  $L : V_k \rightarrow V_k^*$  of the form

$$\langle Lv_k, w_k \rangle_{V_k^* \times V_k} := \int_{\Omega} A_k \nabla v_k \cdot \nabla w_k + \tilde{b} \cdot \nabla v_k w_k \, dx \quad \forall w_k, v_k \in V_k, \quad (3.5)$$

where  $A_k$  is given by (3.4), and where  $\tilde{b} \in L^\infty(\Omega; \mathbb{R}^d)$  is a Lebesgue measurable vector field that satisfies  $\|\tilde{b}\|_{L^\infty(\Omega; \mathbb{R}^d)} \leq L_H$ . Then, the second main assumption ensures a discrete maximum principle:

(H2) For every  $k \in \mathbb{N}$ , each  $L \in W(V_k, D_k)$  and its adjoint  $L^*$  satisfy the discrete maximum principle.

This assumption leads to two important consequences which show that the scheme (3.3) is monotone. Firstly, the DMP ensured by (H2) provides a discrete counterpart of the Weak Maximum Principle (see e.g. [17, Theorem 8.1]), which allows one to deduce the uniqueness of solutions to the discrete HJB equation (3.3a) for each arbitrary source term in  $H^{-1}(\Omega)$ . The uniqueness of solutions to the discrete HJB equation is shown by study of its linearizations, which yield operators of the form (3.5) (c.f. proof of [27, 28, Lemma 6.3]). We note that finite element discretizations of HJB equations with DMP have been studied in the earlier works [22, 21], where the convergence of the approximations to the viscosity solution was shown for classes of degenerate fully nonlinear HJB and Isaacs equations.

Secondly, (H2) is useful in ensuring that the approximations for the density function given by (3.3) are nonnegative everywhere without requiring *a priori* knowledge of the vector field  $\frac{\partial H}{\partial p}[\nabla u_k]$  beyond its satisfaction of a suitable  $L^\infty$ -bound. Indeed, we can rewrite (3.3b) in the form  $\langle L_k^* m_k, \phi \rangle_{V_k^* \times V_k} = \langle G, \phi \rangle_{H^{-1} \times H_0^1}$  where  $L_k$  is the operator of the form (3.5), with  $\tilde{b} = \frac{\partial H}{\partial p}[\nabla u_k]$ , because  $\|\tilde{b}\|_{L^\infty(\Omega; \mathbb{R}^d)} \leq L_H$  holds by (2.6). Since the nonnegativity of  $G$  in  $H^{-1}(\Omega)$  also holds by hypothesis, we conclude from the DMP that  $m_k \geq 0$  in  $\Omega$ .

The assumptions (H1) and (H2) are verified below in Lemmas 3.2 and 3.3 for the stabilizations of [27, 28, 29]. Among the main consequences of these assumptions, there is a uniform stability bound for the class operators  $W(V_k, D_k)$  of the form given in (3.5).

**Lemma 3.1.** *Assume (H1) and (H2). Then, for every  $k \in \mathbb{N}$ , each operator  $L \in W(V_k, D_k)$  and its adjoint  $L^*$  are invertible. Moreover, there exists a constant  $C_{\text{stab}}$  independent of  $k \in \mathbb{N}$  such that*

$$\sup_{k \in \mathbb{N}} \sup_{L \in W(V_k, D_k)} \max \left\{ \|L^{-1}\|_{\mathcal{L}(V_k^*, V_k)}, \|L^{*-1}\|_{\mathcal{L}(V_k^*, V_k)} \right\} \leq C_{\text{stab}}. \quad (3.6)$$

*Proof.* The result is proved in [27, 28, Lemma 6.2] for the case of strictly acute meshes. The argument in the proof of [27, 28, Lemma 6.2] only requires that  $D_k$  is positive semi-definite a.e. in  $\Omega$  for all  $k \in \mathbb{N}$ , that (H2) holds, and that  $\{D_k\}_{k \in \mathbb{N}}$  vanishes in  $L^\infty(\Omega; \mathbb{R}^{d \times d})$  in the limit as  $k \rightarrow \infty$ . It is then clear that the result extends to the current setting under the assumptions (H1) and (H2).  $\square$

In the analysis later on, we will often employ the fact that the uniform stability bound (3.6) can be written equivalently as inf-sup stability bounds for operators in  $W(V_k, D_k)$ ,  $k \in \mathbb{N}$ , as follows: for each  $k \in \mathbb{N}$  and  $L \in W(V_k, D_k)$  there hold

$$C_{\text{stab}}^{-1} \leq \inf_{v_k \in V_k \setminus \{0\}} \sup_{\substack{\psi_k \in V_k: \\ \|\psi_k\|_{H^1(\Omega)}=1}} \frac{\langle Lv_k, \psi_k \rangle_{V_k^* \times V_k}}{\|v_k\|_{H^1(\Omega)}}, \quad (3.7a)$$

$$C_{\text{stab}}^{-1} \leq \inf_{w_k \in V_k \setminus \{0\}} \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \frac{\langle L^* w_k, \phi_k \rangle_{V_k^* \times V_k}}{\|w_k\|_{H^1(\Omega)}}, \quad (3.7b)$$

or equivalently

$$\|v_k\|_{H^1(\Omega)} \leq C_{\text{stab}} \sup_{\substack{\psi_k \in V_k: \\ \|\psi_k\|_{H^1(\Omega)}=1}} \langle Lv_k, \psi_k \rangle_{V_k^* \times V_k} \quad \forall v_k \in V_k, \quad (3.8a)$$

$$\|w_k\|_{H^1(\Omega)} \leq C_{\text{stab}} \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \langle L^* w_k, \phi_k \rangle_{V_k^* \times V_k} \quad \forall w_k \in V_k. \quad (3.8b)$$

**Remark 3.1** (Dependencies of  $C_{\text{stab}}$ ). It is clear that the constant  $C_{\text{stab}}$  appearing in (3.6) will generally depend on  $d, \Omega, \delta, \nu, L_H, C_D$ , etc. However, since the constant  $C_{\text{stab}}$  appears frequently in the following analysis, it will often be simpler and clearer for us to state how other constants depend on  $C_{\text{stab}}$ , rather than on these other more primitive quantities.

**Remark 3.2** (Well-posedness of discrete approximations and convergence). The existence and uniqueness of the numerical solutions is shown in [27, 28, Theorems 5.2 & 5.3] for the case of strictly acute meshes. Following the arguments of the proofs in these results, it is straightforward to show that existence and uniqueness of the discrete solution also holds for the abstract class of methods considered here of the form (3.3) under the assumptions (H1) and (H2). Moreover, in the current setting where  $H$  is continuously differentiable, it is also straightforward to extend [27, 28, Theorem 5.4, Corollary 5.5] on the convergence of the method to deduce that  $\{u_k\}_{k \in \mathbb{N}}, \{m_k\}_{k \in \mathbb{N}}$  are uniformly bounded in the  $H^1$ -norm with  $u_k \rightarrow u$  and  $m_k \rightarrow m$  in  $H_0^1(\Omega)$  as  $k \rightarrow \infty$ . We note that we will use these results on plain convergence of the method in the analysis that comes below. Since the analysis from [27, 28] assumes that the finite element spaces are nested, we therefore make the same assumption here.

**Vanishing stabilization.** In some cases, it is possible to design the numerical methods such that stabilization becomes unnecessary once the mesh is sufficiently fine, i.e. one can satisfy (H2) whilst also satisfying

$$D_k \equiv 0 \quad \text{in } \Omega \quad \forall k \geq k^*, \quad (3.9)$$

for some  $k^* \in \mathbb{N}$  sufficiently large. Note that (3.9) trivially implies (H1). An important example where (3.9) holds is the case of sequences of strictly acute meshes, see Section 3.3

below. Situations where the stabilization vanishes are of special interest since we will show below that these lead to the full quasi-optimality of error without additional terms in the error bounds originating from stabilization, see Remark 4.2 below.

### 3.2 Xu–Zikatanov (XZ) condition [34]

**Formulation of the condition.** Let  $K \in \mathcal{T}_k$  be given. For each vertex  $x_i \in K$ , we let  $F_{K,i}$  denote the convex hull of all vertices of  $K$  except  $x_i$ , i.e.  $F_{K,i}$  is the  $(d-1)$ -dimensional face of  $K$  that is opposite  $x_i$ . The dihedral angle between the faces  $F_{K,i}$  and  $F_{K,j}$  is denoted by  $\theta_{ij}^K$ . We say that the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  satisfies the *XZ condition* [34] if the following holds: 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

$$\sum_{K' \in \mathcal{T}_{k,E}} |F_{K',i} \cap F_{K',j}|_{d-2} \cot(\theta_{ij}^{K'}) \geq 0. \quad (3.10)$$

For instance, when  $d = 2$ , the condition (3.10) requires that the sum of the angles opposite to any edge in the mesh should be less than or equal to  $\pi$ . More generally, the XZ condition allows for some unstructured meshes with nonacute elements. Note that under condition (3.10), the stiffness matrix for the Laplacian on  $V_k$  is an  $M$ -matrix, see [34].

**Construction of the stabilization.** Given an interior edge  $E \in \mathcal{E}_{k,\Omega}$ , let  $\omega_{k,E} \geq 0$  denote a nonnegative weight that satisfies

$$\frac{\delta L_H \text{diam } E}{2(d+1)} < \omega_{k,E} \leq C L_H \text{diam } E, \quad (3.11)$$

for some fixed constant  $C$  independent of  $E$  and  $k$ . Recall that  $\delta$  denotes here the shape-regularity parameter of the sequence of meshes. For each  $E \in \mathcal{E}_{k,\Omega}$ , let  $\mathbf{t}_E$  denote a fixed choice of unit tangent vector to  $E$ ; the orientation of  $\mathbf{t}_E$  has no effect on what is to follow. For sequences of meshes satisfying the XZ condition, the stabilization matrix  $D_k \in L^\infty(\Omega; \mathbb{R}^{d \times d})$  introduced in [29] is defined element-wise over the mesh  $\mathcal{T}_k$  by

$$D_k|_K := \sum_{E \in \mathcal{E}_K} \omega_{k,E} \mathbf{t}_E \otimes \mathbf{t}_E \quad \forall K \in \mathcal{T}_k, \quad (3.12)$$

where we recall that  $\mathcal{E}_K \subset \mathcal{E}_{k,\Omega}$  denotes the edges of  $K$  that are internal edges, and  $\otimes$  denotes the outer-product of vectors, i.e.  $\mathbf{t}_E \otimes \mathbf{t}_E := \mathbf{t}_E \mathbf{t}_E^T \in \mathbb{R}^{d \times d}$ . It is clear that  $D_k$  is well-defined since each summand in the right-hand-side of (3.12) is independent of the choice of orientations of the tangent vectors  $\mathbf{t}_E$ . Moreover, for each  $K \in \mathcal{T}_k$ ,  $D_k|_K$  is a symmetric, positive semi-definite matrix in  $\mathbb{R}^{d \times d}$  since the weights  $\omega_{k,E}$  are nonnegative.

**Lemma 3.2** ([29]). *Suppose that the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  satisfies the XZ condition (3.10) and that, for each  $k \in \mathbb{N}$ , the matrix-valued function  $D_k$  is given by (3.12) with the weights satisfying (3.11). Then assumptions (H1) and (H2) are satisfied.*

*Proof.* See [29, Proof of Theorem 4.2, pp. 28–29].  $\square$

### 3.3 Strict acuteness condition

**Formulation of the condition.** We say that the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  is *strictly acute* [8] if the following condition holds: there exists  $\theta \in (0, \pi/2)$ , independent of  $k \in \mathbb{N}$ , such that, for each  $k \in \mathbb{N}$ , the nodal basis  $\{\xi_1, \dots, \xi_{M_k}\}$  of  $V_k$  satisfies

$$\nabla \xi_i \cdot \nabla \xi_j|_K \leq -\sin(\theta) |\nabla \xi_i|_K |\nabla \xi_j|_K \quad \forall 1 \leq i, j \leq M_k, i \neq j, \forall K \in \mathcal{T}_k. \quad (3.13)$$

It is well-known that when  $d = 2$  the strict acuteness condition (3.13) indicates that the largest angle of a given triangle  $K \in \mathcal{T}_k$  is at most  $\frac{\pi}{2} - \theta$ , while in the case  $d = 3$  this condition indicates that each angle formed by the six pairs of faces of any tetrahedron  $K \in \mathcal{T}_k$  is at most  $\frac{\pi}{2} - \theta$  (see [8]). Furthermore, the strict acuteness condition implies the XZ condition.

**Construction of the stabilization.** In this case, we consider the stabilization method given by artificial diffusion as in [27, 28]. More precisely, let  $\mu > 1$  be a fixed constant and let  $k \in \mathbb{N}$  be given. For  $K \in \mathcal{T}_k$ , we denote by  $\{\psi_{k,0}^K, \dots, \psi_{k,d}^K\} \subset V_k$  the set of nodal basis functions associated with the  $d + 1$  nodes of  $K$ . Let  $\sigma_K^k := (\text{diam } K) \min_{0 \leq i \leq d} |\nabla \psi_{k,i}^K|$  and  $\sigma^k := \min_{K \in \mathcal{T}_k} \sigma_K^k$ . We note that, owing to the shape regularity of the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$ , there exist constants  $\underline{\sigma}, \bar{\sigma} > 0$ , that are independent of  $k \in \mathbb{N}$ , such that  $\underline{\sigma} \leq \sigma^k \leq \bar{\sigma}$  for all  $k \in \mathbb{N}$ . We then define  $D_k$  element-wise over  $\mathcal{T}_k$  by

$$D_k|_K := \max \left( \frac{\mu L_H h_{\mathcal{T}_k}|_K}{\sigma^k \sin(\theta)} - \nu, 0 \right) \mathbb{I}_d \quad \forall K \in \mathcal{T}_k. \quad (3.14)$$

It is clear that  $D_k$  is in  $L^\infty(\Omega; \mathbb{R}_{\text{sym},+}^{d \times d})$ . Although the strict acuteness condition is less general than the XZ condition, it is nonetheless a case of special interest. Indeed, note if  $k \in \mathbb{N}$  is sufficiently large, in particular if the mesh-size is sufficiently small, then  $\frac{\mu L_H h_{\mathcal{T}_k}}{\sigma^k \sin(\theta)} \leq \nu$  almost everywhere in  $\Omega$ , and therefore (3.14) implies that (3.9) holds whenever  $k$  is sufficiently large.

**Lemma 3.3** ([27, 28]). *Suppose that the family of meshes  $\{\mathcal{T}_k\}_{k \in \mathbb{N}}$  satisfies the strict acuteness condition (3.13) and that, for each  $k \in \mathbb{N}$ , the matrix-valued function  $D_k$  is given by (3.14). Then assumptions (H1) and (H2) are satisfied. Moreover, there exists  $k^* \in \mathbb{N}$  such that (3.9) holds for all  $k \geq k^*$ .*

*Proof.* See [27, 28, Lemma 6.1], and also [8, Theorem 4.2] and [22, Section 8].  $\square$

## 4 Main results

### 4.1 Bounds for the $H^1$ -norm error of the density and value functions

We now present our main result on *a priori* bounds for the  $H^1$ -norm of the error for the approximations of the solution of the MFG system (2.15). First, in Theorem 4.1 below, we give the main bound for the general abstract method of Section 3.1. Then, in Remarks 4.1 and 4.2 below, we show how the bound of Theorem 4.1 applies to the stabilized methods of [29, 27, 28]. Then, in Corollary 4.2, we show how the bound of Theorem 4.1 implies optimal rates of convergence of the error with respect to the mesh-size for sufficiently smooth solutions.

**Theorem 4.1** (A priori bound). *Assume (H1) and (H2). Then, there exists a  $k_0 \in \mathbb{N}$  such that*

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_\Omega + \|D_k \nabla u\|_\Omega, \end{aligned} \quad (4.1)$$

for all  $k \geq k_0$ . The hidden constant in (4.1) depends only on  $d$ ,  $\Omega$ ,  $c_F$ ,  $L_F$ ,  $L_{H_p}$ ,  $L_H$ ,  $\nu$ ,  $C_D$ ,  $C_{\text{stab}}$ , and  $M_\infty$ .

We stress that Theorem 4.1 does not require any additional regularity on the solution. The proof of Theorem 4.1 is postponed to Section 7 below.

**Remark 4.1** (Near quasi-optimality). As an immediate consequence of (4.1) and of the bound  $|D_k| \leq C_D h_{\mathcal{T}_k}$  in  $\Omega$  from (H1), we obtain

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} + \|h_{\mathcal{T}_k} \nabla m\|_{\Omega} + \|h_{\mathcal{T}_k} \nabla u\|_{\Omega}. \end{aligned} \quad (4.2)$$

We then see from (4.2) that, for sufficiently refined meshes, the error of the stabilized FEM is bounded by the best approximation error plus an additional term that is of first-order with respect to the mesh-size function  $h_{\mathcal{T}_k}$ . Recall that a numerical method is said to be *quasi-optimal* if the error between exact and computed solutions is bounded by a constant times the best approximation error from the same approximation space. Note that for piecewise affine elements, in general the convergence rate of  $\inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)}$  and  $\inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)}$  with respect to the mesh-size is generally at best of first-order, and may be lower for nonsmooth solutions. Therefore, the additional terms  $\|h_{\mathcal{T}_k} \nabla m\|_{\Omega}$  and  $\|h_{\mathcal{T}_k} \nabla u\|_{\Omega}$  appearing in (4.2) are of optimal order with respect to the mesh-size. This is why we shall say that the method is asymptotically *nearly* quasi-optimal for the  $H^1$ -norm of the error. In particular, this applies to the stabilization for XZ meshes from Section 3.2.

**Remark 4.2** (Full quasi-optimality for vanishing stabilization). In the case where the stabilization terms vanish identically once the meshes are sufficiently refined, i.e. if (3.9) holds, then Theorem 4.1 implies that

$$\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)}, \quad (4.3)$$

for all  $k$  sufficiently large. The bound (4.3) shows that the method is then asymptotically fully quasi-optimal in the  $H^1$ -norm, i.e. without any additional terms. In particular, this stronger result holds for the case of the stabilization for strictly acute meshes from Section 3.3.

As an immediate corollary of Theorem 4.1, we now show that the approximations have optimal rates of convergence if the solution has sufficient regularity. Recall that  $h_k := \|h_{\mathcal{T}_k}\|_{L^\infty(\Omega)}$  denotes the maximum element size of  $\mathcal{T}_k$ .

**Corollary 4.2** (Optimal convergence rates). *Assume (H1) and (H2). Suppose also that  $m$  and  $u$  are in  $H^{1+s}(\Omega)$  for some  $s \in [0, 1]$ . Then, there exists  $k_0 \in \mathbb{N}$  such that*

$$\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \lesssim h_k^s (\|m\|_{H^{1+s}(\Omega)} + \|u\|_{H^{1+s}(\Omega)}), \quad (4.4)$$

for all  $k \geq k_0$ .

*Proof.* To begin, we show that any  $v \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$ , with  $s \in [0, 1]$ , satisfies

$$\inf_{\bar{v}_k \in V_k} \|v - \bar{v}_k\|_{H^1(\Omega)} \lesssim h_k^s \|v\|_{H^{1+s}(\Omega)} \quad (4.5)$$

Proving this bound in the case  $s = 0$  is trivial, while proof for the case  $s = 1$  follows from well-known results for piecewise affine finite element approximation of  $H^2$ -regular functions by quasi-interpolation (see e.g. [32]). For the case where  $s \in (0, 1)$ , the bound (4.5) is a

standard consequence of the theory of interpolation between Banach spaces (see e.g. [7, Chapter 14]). We include the proof of (4.5) in the case  $s \in (0, 1)$  for completeness.

For each  $k \in \mathbb{N}$ , let  $\mathcal{I}_k : H^1(\Omega) \rightarrow H^1(\Omega)$  denote the Scott–Zhang linear quasi-interpolation operator [32] associated with the shape-regular mesh  $\mathcal{T}_k$  that is defined in [7, Theorem (4.8.3.8)] and which satisfies  $\mathcal{I}_k w = 0$  on  $\partial\Omega$  if  $w \in H_0^1(\Omega)$ . We then have that  $\mathcal{I}_k w \in V_k$  for each  $w \in H_0^1(\Omega)$ , and so, for each  $v \in H_0^1(\Omega) \cap H^{1+s}(\Omega)$  we get

$$\inf_{\bar{v}_k \in V_k} \|v - \bar{v}_k\|_{H^1(\Omega)} \leq \|v - \mathcal{I}_k v\|_{H^1(\Omega)} \quad \forall k \in \mathbb{N}. \quad (4.6)$$

For each  $k \in \mathbb{N}$  define the linear operator  $Q_k : H^1(\Omega) \rightarrow H^1(\Omega)$  via  $Q_k w := w - \mathcal{I}_k w$  for  $w \in H^1(\Omega)$ . By applying [7, Theorem (4.8.12)] we obtain that, for all  $w \in H^2(\Omega)$ ,

$$\|Q_k w\|_{H^1(\Omega)} \lesssim h_k |v|_{H^2(\Omega)}, \quad (4.7)$$

so  $Q_k$  maps  $H^2(\Omega)$  into  $H^1(\Omega)$  with

$$\|Q_k\|_{\mathcal{L}(H^2(\Omega), H^1(\Omega))} \lesssim h_k \quad \forall k \in \mathbb{N}. \quad (4.8)$$

We also have by [7, Corollary (4.8.15)] that  $\|Q_k w\|_{H^1(\Omega)} \lesssim \|w\|_{H^1(\Omega)}$  for all  $w \in H^1(\Omega)$ , so

$$\|Q_k\|_{\mathcal{L}(H^1(\Omega), H^1(\Omega))} \lesssim 1 \quad \forall k \in \mathbb{N}. \quad (4.9)$$

We then apply the Banach space interpolation result [7, Proposition (14.1.5)] with  $A_0 = B_0 = B_1 = H^1(\Omega)$ , and  $A_1 = H^2(\Omega)$  to deduce that  $Q_k$  maps  $H^{1+s}(\Omega)$  to  $H^1(\Omega)$  with

$$\|Q_k\|_{\mathcal{L}(H^{1+s}(\Omega), H^1(\Omega))} \leq \|Q_k\|_{\mathcal{L}(H^1(\Omega), H^1(\Omega))}^{1-s} \|Q_k\|_{\mathcal{L}(H^2(\Omega), H^1(\Omega))}^s \lesssim h_k^s \quad (4.10)$$

after using (4.8) and (4.9). This then implies that

$$\|Q_k v\|_{H^1(\Omega)} = \|v - \mathcal{I}_k v\|_{H^1(\Omega)} \lesssim h_k^s \|v\|_{H^{1+s}(\Omega)} \quad \forall v \in H^{1+s}(\Omega) \quad (4.11)$$

This bound and (4.6) then give desired bound (4.5).

The regularity assumptions on  $u$  and  $m$ , together with the bound (4.5), imply

$$\inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} \lesssim h_k^s \|m\|_{H^{1+s}(\Omega)}, \quad (4.12a)$$

$$\inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} \lesssim h_k^s \|u\|_{H^{1+s}(\Omega)}. \quad (4.12b)$$

Applying bounds (4.12a) and (4.12b) to (4.2) gives

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim h_k^s (\|m\|_{H^{1+s}(\Omega)} + \|u\|_{H^{1+s}(\Omega)}) + h_k (\|m\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)}). \end{aligned} \quad (4.13)$$

We then obtain (4.4) upon noting that  $h_k (\|m\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)}) \lesssim h_k^s (\|m\|_{H^{1+s}(\Omega)} + \|u\|_{H^{1+s}(\Omega)})$  since  $s \in [0, 1]$ .  $\square$

We note that the rate of convergence with respect to the mesh-size that is shown in Corollary 4.2 is optimal in the sense that it cannot be improved for general  $u$  and  $m$  in  $H^{1+s}(\Omega)$  when considering the  $H^1$ -norm of the error. In particular, a first-order convergence rate in the  $H^1$ -norm is achieved for the case of  $s = 1$ , i.e. when  $u, m \in H^2(\Omega)$ , which is the maximum rate of convergence that is generally possible for piecewise affine finite element approximations.

**Remark 4.3** (Time-dependent problems). Although this work treats the case of steady-state problems, the numerical method considered here has been extended already to the case of time-dependent problems in [29], which employs a stabilized finite element method in space coupled with implicit Euler discretizations of the time derivatives with mass lumping. The plain convergence of the resulting method is proved in [29]. The analysis of rates of convergence of finite element approximations of time-dependent problems will be the subject of future work. Recall also that [6] obtains a rate of convergence for a finite difference approximation of time-dependent MFG systems.

## 4.2 Bounds for the $L^2$ -norm error of the density function and $H^1$ -norm error of the value function

In the numerical experiments of [27, 28], it was observed that the convergence rate of  $\|u - u_k\|_{H^1(\Omega)}$  can attain the maximum possible convergence rate of order one, even for problems where  $m$  had low regularity, e.g.  $m \notin H^2(\Omega)$ . We now give theoretical support to those experimental observations by proving that the error  $\|u - u_k\|_{H^1(\Omega)}$  has optimal rates of convergence even when  $m$  has minimal regularity in  $H^1(\Omega)$ , at least in the case where some additional elliptic regularity is available, such as on convex domains. Note that  $m \notin H^2(\Omega)$  is still possible when  $\Omega$  is convex, owing to the distributional datum  $G \in H^{-1}(\Omega)$ . We also remark that the analysis here assumes differentiability of the Hamiltonian, in contrast to the computations in [27, 28], which involved nondifferentiable Hamiltonians.

The main idea to show this is to seek bounds for a composite norm on the error that involves the  $L^2$ -norm error for the density and  $H^1$ -norm error for the value function, rather than the stronger  $H^1$ -norm in both components that was considered in the previous subsection. Our main result on this subject is in Theorem 4.3 below, which gives an error bound in this composite norm for the general class of methods of Section 3.1, without any regularity assumptions on the solution  $(u, m)$ .

**Theorem 4.3.** *Assume (H1) and (H2). Suppose additionally that the domain  $\Omega$  is convex. Then, there exists  $k_0 \in \mathbb{N}$  such that*

$$\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)} \lesssim \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} + h_k (\|u\|_{H^1(\Omega)} + \|m\|_{H^1(\Omega)}), \quad (4.14)$$

for all  $k \geq k_0$ . The hidden constant in the above bound depends only on  $d$ ,  $\Omega$ ,  $c_F$ ,  $L_F$ ,  $L_{H_p}$ ,  $L_H$ ,  $\nu$ ,  $\delta$ ,  $C_D$ ,  $C_{\text{stab}}$ , and  $M_{\infty}$ .

We prove Theorem 4.3 in Section 8. We now show that Theorem 4.3 implies an optimal asymptotic rate of convergence of the value function approximations that is independent of the regularity of  $m$ .

**Corollary 4.4.** *In addition to the hypotheses of Theorem 4.3, suppose that  $u \in H^{1+s}(\Omega)$  for some  $s \in [0, 1]$ . Then, there exists  $k_0 \in \mathbb{N}$  such that*

$$\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)} \lesssim h_k^s (\|m\|_{H^1(\Omega)} + \|u\|_{H^{1+s}(\Omega)}), \quad (4.15)$$

for all  $k \geq k_0$ .

*Proof.* As explained in the proof of Corollary 4.2, if  $u \in H^{1+s}(\Omega)$  for  $s \in [0, 1]$  then  $\inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)} \lesssim h_k^s \|u\|_{H^{1+s}(\Omega)}$ . Moreover, it is clear that  $h_k (\|m\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)}) \lesssim h_k^s (\|m\|_{H^1(\Omega)} + \|u\|_{H^{1+s}(\Omega)})$ . Applying these bounds to (4.14) then yields (4.15).  $\square$



We note that Corollary 4.4 implies a rate of convergence in the total  $L^2$ - $H^1$  norm in (4.15) that requires a regularity hypothesis only on  $u$  but not on  $m$ , in contrast to Corollary 4.2. However, Corollary 4.2 implies a rate of convergence in the total  $H^1$ -norm in (4.4) without requiring  $\Omega$  to be convex.

## 5 Stability of discrete HJB and KFP equations

In this section, we start the analysis of the approximations defined by discrete MFG system (3.3) by showing the stability of the discrete HJB and KFP equations considered separately. The stability of these discrete problems can be expressed in terms of discrete residuals of the equations, which we now define. For each  $k \in \mathbb{N}$ , introduce the discrete residual operators  $R_k^1, R_k^2 : V_k \times V_k \rightarrow V_k^*$  given by

$$\langle R_k^1(\bar{m}_k, \bar{u}_k), \psi_k \rangle_{V_k^* \times V_k} := \langle F[\bar{m}_k], \psi_k \rangle_{H^{-1} \times H_0^1} - \int_{\Omega} (A_k \nabla \bar{u}_k \cdot \nabla \psi_k + H[\nabla \bar{u}_k] \psi_k) dx, \quad (5.1a)$$

$$\langle R_k^2(\bar{m}_k, \bar{u}_k), \phi_k \rangle_{V_k^* \times V_k} := \langle G, \phi_k \rangle_{H^{-1} \times H_0^1} - \int_{\Omega} \left( A_k \nabla \bar{m}_k \cdot \nabla \phi_k + \bar{m}_k \frac{\partial H}{\partial p} [\nabla \bar{u}_k] \cdot \nabla \phi_k \right) dx, \quad (5.1b)$$

for all  $\bar{m}_k, \bar{u}_k, \psi_k, \phi_k \in V_k$ . Since the pair  $(m_k, u_k) \in V_k \times V_k$  satisfies the discrete problem (3.3), it is clear that  $R_k^1(m_k, u_k) = 0$  and  $R_k^2(m_k, u_k) = 0$  in  $V_k^*$ .

We now show that the discrete residuals are bounded by the approximation error plus some additional terms coming from the stabilization.

**Lemma 5.1.** *Assume (H1). For all  $k \in \mathbb{N}$ , and all  $\bar{u}_k, \bar{m}_k \in V_k$ , we have*

$$\|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*} \lesssim \|m - \bar{m}_k\|_{\Omega} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla u\|_{\Omega}, \quad (5.2a)$$

$$\|R_k^2(\bar{m}_k, \bar{u}_k)\|_{V_k^*} \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega}, \quad (5.2b)$$

where the hidden constants depend only on  $d, \Omega, L_F, L_{H_p}, L_H, \nu, C_D$ , and  $M_{\infty}$ .

*Proof.* Since  $(u, m)$  solves (2.15), we have

$$\begin{aligned} \langle R_k^1(\bar{m}_k, \bar{u}_k), \psi_k \rangle_{V_k^* \times V_k} &= \langle F[\bar{m}_k] - F[m], \psi_k \rangle_{H^{-1} \times H_0^1} + \int_{\Omega} A_k \nabla (u - \bar{u}_k) \cdot \nabla \psi_k dx \\ &\quad + \int_{\Omega} (H[\nabla u] - H[\nabla \bar{u}_k]) \psi_k dx - \int_{\Omega} D_k \nabla u \cdot \nabla \psi_k dx, \end{aligned} \quad (5.3)$$

for all  $\psi_k \in V_k$ . The Lipschitz continuity of  $F$  in (2.1) implies that  $\langle F[\bar{m}_k] - F[m], \psi_k \rangle_{H^{-1} \times H_0^1} \leq L_F \|\bar{m}_k - m_k\|_{\Omega} \|\psi_k\|_{H^1(\Omega)}$  for all  $\psi_k \in V_k$ . The definition of  $A_k$  in (3.4) and the bound on  $D_k$  in from (H1) then imply that  $|\int_{\Omega} A_k \nabla (u - \bar{u}_k) \cdot \nabla \psi_k dx| \lesssim \|u - \bar{u}_k\|_{H^1(\Omega)} \|\psi_k\|_{H^1(\Omega)}$  for all  $\psi_k \in V_k$ , for some hidden constant that depends only on  $d, \nu, C_D$  and on  $\Omega$ . The Lipschitz continuity of  $H$  in (2.4b) implies that  $|\int_{\Omega} (H[\nabla u] - H[\nabla \bar{u}_k]) \psi_k dx| \lesssim \|u - \bar{u}_k\|_{H^1(\Omega)} \|\psi_k\|_{H^1(\Omega)}$  for all  $\psi_k \in V_k$ . The Cauchy–Schwarz inequality implies that  $|\int_{\Omega} D_k \nabla u \cdot \nabla \psi_k dx| \leq \|D_k \nabla u\|_{\Omega} \|\psi_k\|_{H^1(\Omega)}$  for all  $\psi_k \in V_k$ . Combining these bounds then yields (5.2a).

Next, we have

$$\begin{aligned} \langle R_k^2(\bar{m}_k, \bar{u}_k), \phi_k \rangle_{V_k^* \times V_k} &= \int_{\Omega} A_k \nabla (m - \bar{m}_k) \cdot \nabla \phi_k + (m - \bar{m}_k) \frac{\partial H}{\partial p} [\nabla \bar{u}_k] \cdot \nabla \phi_k dx \\ &\quad - \int_{\Omega} D_k \nabla m \cdot \nabla \phi_k dx + \int_{\Omega} m \left( \frac{\partial H}{\partial p} [\nabla u] - \frac{\partial H}{\partial p} [\nabla \bar{u}_k] \right) \cdot \nabla \phi_k dx, \end{aligned} \quad (5.4)$$

for all  $\phi_k \in V_k$ . We then obtain (5.2b) similarly as above, where we additionally use the bound  $\|\frac{\partial H}{\partial p}[\nabla \bar{u}_k]\|_{L^\infty(\Omega, \mathbb{R}^d)} \leq L_H$  which follows from (2.6), along with the Lipschitz continuity of  $\frac{\partial H}{\partial p}$  in (2.5) and the  $L^\infty$ -bound on  $m$  in (2.18).  $\square$

We now turn to the (local) stability of the discrete HJB equation (3.3a).

**Lemma 5.2** (Stability of discrete HJB equation). *Assume (H1) and (H2). There exists a constant  $R > 0$  that is independent of  $k \in \mathbb{N}$  such that*

$$\|\bar{u}_k - u_k\|_{H^1(\Omega)} \lesssim \|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*} + \|\bar{m}_k - m_k\|_\Omega, \quad (5.5)$$

for any  $\bar{m}_k \in V_k$ , and for any  $\bar{u}_k \in V_k$  satisfying  $\|\bar{u}_k - u_k\|_{H^1(\Omega)} \leq R$ . The hidden constant in (5.5) depends only on  $L_F$  and  $C_{\text{stab}}$ .

*Proof of Lemma 5.2.* Applying Lemma 3.1 to the operator  $L \in W(V_k, D_k)$  for the case  $\tilde{b} = \frac{\partial H}{\partial p}[\nabla \bar{u}_k]$ , we employ the equivalent inf-sup stability bound (3.8a) with  $v_k := \bar{u}_k - u_k$  to get

$$\begin{aligned} & \|\bar{u}_k - u_k\|_{H^1(\Omega)} \\ & \leq C_{\text{stab}} \sup_{\substack{\psi_k \in V_k: \\ \|\psi_k\|_{H^1(\Omega)}=1}} \left[ \int_\Omega A_k \nabla(\bar{u}_k - u_k) \cdot \nabla \psi_k - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \cdot \nabla(u_k - \bar{u}_k) \psi_k dx \right]. \end{aligned}$$

Using the discrete HJB equation of (3.3), we find that

$$\begin{aligned} & \|\bar{u}_k - u_k\|_{H^1(\Omega)} \\ & \leq C_{\text{stab}} \sup_{\substack{\psi_k \in V_k: \\ \|\psi_k\|_{H^1(\Omega)}=1}} \left[ \langle F[\bar{m}_k] - F[m_k], \psi_k \rangle_{H^{-1} \times H_0^1} - \langle R_k^1(\bar{m}_k, \bar{u}_k), \psi_k \rangle_{V_k^* \times V_k} \right. \\ & \quad \left. + \int_\Omega \left( H[\nabla u_k] - H[\nabla \bar{u}_k] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \cdot \nabla(u_k - \bar{u}_k) \right) \psi_k dx \right]. \end{aligned}$$

Hence, the Lipschitz continuity of  $F$ , c.f. (2.1), and the triangle inequality imply that

$$\begin{aligned} \|\bar{u}_k - u_k\|_{H^1(\Omega)} & \leq C_{\text{stab}} \|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*} + C_{\text{stab}} L_F \|\bar{m}_k - m_k\|_\Omega \\ & \quad + C_{\text{stab}} \left\| H[\nabla u_k] - H[\nabla \bar{u}_k] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \cdot \nabla(u_k - \bar{u}_k) \right\|_{H^{-1}(\Omega)}. \quad (5.6) \end{aligned}$$

Next, we apply Lemma 2.1 with  $\epsilon = (2C_{\text{stab}})^{-1}$  in order to absorb the last term on the right-hand side of (5.6) into the left-hand side, and thereby deduce that there exists  $R > 0$  such that (5.5) holds whenever  $\|\bar{u}_k - u_k\|_{H^1(\Omega)} \leq R$ .  $\square$

We now consider the stability of the discrete KFP equation (3.3b).

**Lemma 5.3** (Stability of discrete KFP equation). *Assume (H1) and (H2). Then for all  $k \in \mathbb{N}$  and all  $\bar{m}_k \in V_k$ , we have*

$$\|m_k - \bar{m}_k\|_{H^1(\Omega)} \lesssim \|R_k^2(\bar{m}_k, u_k)\|_{V_k^*}, \quad (5.7)$$

where the constant depends only on  $C_{\text{stab}}$ .

*Proof.* Let  $k \in \mathbb{N}$  and let  $\bar{m}_k \in V_k$  be arbitrary. Then, the equivalent inf-sup stability bound (3.8b) implied by Lemma 3.1, with  $w_k := m_k - \bar{m}_k$ , gives

$$\begin{aligned}
& \|m_k - \bar{m}_k\|_{H^1(\Omega)} \\
& \lesssim \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \left[ \int_{\Omega} A_k \nabla(m_k - \bar{m}_k) \cdot \nabla \phi_k + (m_k - \bar{m}_k) \frac{\partial H}{\partial p}[\nabla u_k] \cdot \nabla \phi_k dx \right] \\
& = \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \left[ \langle G, \phi_k \rangle_{H^{-1} \times H_0^1} - \int_{\Omega} A_k \nabla \bar{m}_k \cdot \nabla \phi_k + \bar{m}_k \frac{\partial H}{\partial p}[\nabla u_k] \cdot \nabla \phi_k dx \right] \\
& = \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \langle R_k^2(\bar{m}_k, u_k), \phi_k \rangle_{V_k^* \times V_k} = \|R_k^2(\bar{m}_k, u_k)\|_{V_k^*},
\end{aligned}$$

where we have used the discrete KFP equation of (3.3) in the second line above.  $\square$

## 6 Nonnegative approximations of the density

It is well-known that the nonnegativity of the density and its discrete approximation respectively play important roles in the proofs of uniqueness of the solution of the MFG system (2.15) and its discrete approximation (3.3), see e.g. [25] and [27, 28]. It is therefore not surprising that nonnegative approximations of the density will play an important part in the error analysis.

Let  $V_{k,+}$  denote the set of functions in  $V_k$  that are nonnegative in  $\Omega$ . First, we show in Lemma 6.1 below that the density  $m$  can be approximated nearly quasi-optimally, up to stabilization terms, by discrete functions in  $V_{k,+}$ . To this end, let us assume (H1) and (H2). We define  $m_{k,+} \in V_k$  to be the unique solution of

$$\int_{\Omega} A_k \nabla m_{k,+} \cdot \nabla \phi_k + m_{k,+} \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx = \langle G, \phi_k \rangle_{H^{-1} \times H_0^1} \quad \forall \phi_k \in V_k. \quad (6.1)$$

Note that Lemma 3.1 implies that  $m_{k,+}$  is well-defined. Moreover, under the assumption (H2), we further have  $m_{k,+} \in V_{k,+}$  since  $G$  is nonnegative in the sense of distributions. It is also straightforward to show that  $m_{k,+}$  satisfies  $\|m_{k,+}\|_{H^1(\Omega)} \lesssim \|m\|_{H^1(\Omega)}$  for some constant that depends only on  $\Omega$ ,  $\nu$ ,  $L_H$ , and  $C_{\text{stab}}$ . Indeed, using the weak KFP equation (2.15b), we can rewrite (6.1) as

$$\begin{aligned}
& \int_{\Omega} A_k \nabla m_{k,+} \cdot \nabla \phi_k + m_{k,+} \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx \\
& = \int_{\Omega} \nu \nabla m \cdot \nabla \phi_k + m \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx \quad \forall \phi_k \in V_k. \quad (6.2)
\end{aligned}$$

It is clear by (2.6) that

$$\sup_{\phi_k \in V_k: \|\phi_k\|_{H^1(\Omega)}=1} \int_{\Omega} \nu \nabla m \cdot \nabla \phi_k + m \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx \lesssim \|m\|_{H^1(\Omega)} \quad (6.3)$$

where the hidden constant depends only on  $\nu$  and  $L_H$ . We then immediately conclude from (6.2) and Lemma 3.1 that  $\|m_{k,+}\|_{H^1(\Omega)} \lesssim \|m\|_{H^1(\Omega)}$  for some constant that depends only on  $\Omega$ ,  $\nu$ ,  $L_H$ , and  $C_{\text{stab}}$ .

**Lemma 6.1.** Assume (H1) and (H2). Then, for all  $k \in \mathbb{N}$ ,

$$\inf_{\bar{m}_k \in V_{k,+}} \|m - \bar{m}_k\|_{H^1(\Omega)} \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega}. \quad (6.4)$$

The hidden constant in (6.4) depends only on  $d, \Omega, \nu, L_H, C_{\text{stab}},$  and  $C_D$ .

*Proof.* Since  $m_{k,+} \in V_{k,+}$ , it is clearly sufficient to show that  $\|m - m_{k,+}\|_{H^1(\Omega)}$  is bounded by the right-hand side of (6.4) above. Let  $\bar{m}_k \in V_k$  be arbitrary. Then (6.1) implies that

$$\begin{aligned} & \int_{\Omega} A_k \nabla(m_{k,+} - \bar{m}_k) \cdot \nabla \phi_k + (m_{k,+} - \bar{m}_k) \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx \\ &= \int_{\Omega} A_k \nabla(m - \bar{m}_k) \cdot \nabla \phi_k + (m - \bar{m}_k) \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx - \int_{\Omega} D_k \nabla m \cdot \nabla \phi_k dx \end{aligned} \quad (6.5)$$

for all  $\phi_k \in V_k$ . Using Lemma 3.1, for the case  $\tilde{b} = \frac{\partial H}{\partial p}[\nabla u]$ , we apply the equivalent inf-sup stability bound (3.8b) with  $w_k := m_{k,+} - \bar{m}_k$  to find that

$$\begin{aligned} & \|m_{k,+} - \bar{m}_k\|_{H^1(\Omega)} \\ & \lesssim \sup_{\substack{\phi_k \in V_k: \\ \|\phi_k\|_{H^1(\Omega)}=1}} \left[ \int_{\Omega} A_k \nabla(m - \bar{m}_k) \cdot \nabla \phi_k + (m - \bar{m}_k) \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla \phi_k dx \right. \\ & \quad \left. - \int_{\Omega} D_k \nabla m \cdot \nabla \phi_k dx \right] \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega}, \end{aligned} \quad (6.6)$$

where the second inequality is obtained by applying the Cauchy–Schwarz inequality to the various terms, along with (2.6) and the bound on  $D_k$  from (H2). The constant in (6.6) depends only on  $C_{\text{stab}}, d, \Omega, \nu, C_D$  and  $L_H$ . Since  $\bar{m}_k \in V_k$  is arbitrary, we then deduce from the triangle inequality that  $\|m - m_{k,+}\|_{H^1(\Omega)} \lesssim \inf_{\bar{m}_k \in V_k} \|m - \bar{m}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega}$ , which completes the proof of (6.4).  $\square$

We now prove the  $L^2$ -norm stability in the density component for the discrete MFG system (3.3), when restricted to the set of nonnegative functions in  $V_{k,+}$ . This bound can be seen as a quantitative analogue to some of the inequalities used in the well-known proofs of uniqueness of solutions due to Lasry and Lions in [25].

**Lemma 6.2.** Assume (H2). For each  $k \in \mathbb{N}$ , any  $\bar{u}_k \in V_k$  and  $\bar{m}_k \in V_{k,+}$ , we have

$$c_F \|\bar{m}_k - m_k\|_{\Omega}^2 \leq \langle R_k^1(\bar{m}_k, \bar{u}_k), \bar{m}_k - m_k \rangle_{V_k^* \times V_k} - \langle R_k^2(\bar{m}_k, \bar{u}_k), \bar{u}_k - u_k \rangle_{V_k^* \times V_k}. \quad (6.7)$$

*Proof of Lemma 6.2.* Let  $\bar{u}_k \in V_k$  and  $\bar{m}_k \in V_{k,+}$  be fixed but arbitrary. To abbreviate the notation, let

$$\mathcal{R}_k := \langle R_k^1(\bar{m}_k, \bar{u}_k), \bar{m}_k - m_k \rangle_{V_k^* \times V_k} - \langle R_k^2(\bar{m}_k, \bar{u}_k), \bar{u}_k - u_k \rangle_{V_k^* \times V_k}. \quad (6.8)$$

Since  $R_k^1(m_k, u_k) = 0$  and  $R_k^2(m_k, u_k) = 0$  in  $V_k^*$ , we have that

$$\begin{aligned} \mathcal{R}_k &= \langle R_k^1(\bar{m}_k, \bar{u}_k) - R_k^1(m_k, u_k), \bar{m}_k - m_k \rangle_{V_k^* \times V_k} \\ & \quad - \langle R_k^2(\bar{m}_k, \bar{u}_k) - R_k^2(m_k, u_k), \bar{u}_k - u_k \rangle_{V_k^* \times V_k}. \end{aligned} \quad (6.9)$$

Recall that  $(\cdot, \cdot)_\Omega$  denotes the inner product on  $L^2(\Omega)$ . Using the discrete HJB equation (3.3a) and the discrete KFP equation (3.3b) we find that

$$\begin{aligned} & \langle R_k^1(\bar{m}_k, \bar{u}_k) - R_k^1(m_k, u_k), \bar{m}_k - m_k \rangle_{V_k^* \times V_k} \\ &= \langle F[\bar{m}_k] - F[m_k], \bar{m}_k - m_k \rangle_{H^{-1} \times H_0^1} + \int_\Omega (H[\nabla u_k] - H[\nabla \bar{u}_k])(\bar{m}_k - m_k) dx \\ & \quad + \int_\Omega A_k \nabla(u_k - \bar{u}_k) \cdot \nabla(\bar{m}_k - m_k) dx, \end{aligned} \quad (6.10)$$

and

$$\begin{aligned} & \langle R_k^2(\bar{m}_k, \bar{u}_k) - R_k^2(m_k, u_k), \bar{u}_k - u_k \rangle_{V_k^* \times V_k} \\ &= \int_\Omega \left( m_k \frac{\partial H}{\partial p}[\nabla u_k] - \bar{m}_k \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \right) \cdot \nabla(\bar{u}_k - u_k) dx + \int_\Omega A_k \nabla(\bar{m}_k - m_k) \cdot \nabla(u_k - \bar{u}_k) dx. \end{aligned} \quad (6.11)$$

Using the symmetry of  $A_k$  we then obtain from (6.9), (6.10) and (6.11) that

$$\begin{aligned} \mathcal{R}_k &= \langle F[\bar{m}_k] - F[m_k], \bar{m}_k - m_k \rangle_{H^{-1} \times H_0^1} \\ & \quad + (m_k, R_H[\nabla \bar{u}_k, \nabla u_k])_\Omega + (\bar{m}_k, R_H[\nabla u_k, \nabla \bar{u}_k])_\Omega, \end{aligned} \quad (6.12)$$

where we define  $R_H[\nabla v, \nabla w] := H[\nabla v] - H[\nabla w] - \frac{\partial H}{\partial p}[\nabla w] \cdot \nabla(v - w)$  for any  $v, w \in H^1(\Omega)$ . Since  $H$  is convex and differentiable w.r.t.  $p$ , we have  $R_H[\nabla v, \nabla w] \geq 0$  a.e. for any  $v, w \in H^1(\Omega)$ . Also, we have  $m_k \geq 0$  in  $\Omega$  from (H2) and  $\bar{m}_k \geq 0$  in  $\Omega$  by hypothesis. Therefore the terms  $(m_k, R_H[\nabla \bar{u}_k, \nabla u_k])_\Omega$  and  $(\bar{m}_k, R_H[\nabla u_k, \nabla \bar{u}_k])_\Omega$  in (6.12) are both nonnegative, and thus the strong monotonicity of  $F$ , c.f. (2.2), implies that

$$c_F \|\bar{m}_k - m_k\|_\Omega^2 \leq \langle F[\bar{m}_k] - F[m_k], \bar{m}_k - m_k \rangle_{H^{-1} \times H_0^1} \leq \mathcal{R}_k, \quad (6.13)$$

which shows (6.7).  $\square$

## 7 Proof of Theorem 4.1

We now prove Theorem 4.1. The main ingredients of the proof are the separate stability bounds for the discrete HJB and KFP equations given in Section 5 above, along with the near quasi-optimality of nonnegative approximations and the  $L^2$ -stability bound for the discrete MFG system as shown in Section 6 above.

**Proof of Theorem 4.1** Let  $R > 0$  be the constant given in Lemma 5.2. As explained in Remark 3.2, we have  $u_k \rightarrow u$  in  $H_0^1(\Omega)$  as  $k \rightarrow \infty$ . Therefore, there exists  $k_0 \in \mathbb{N}$  such that for all  $k \geq k_0$ , we have  $u_k \in B_k(u, R/2)$  where  $B_k(u, R/2) := \{v_k \in V_k; \|v_k - u\|_{H^1(\Omega)} < R/2\}$  is the restriction to  $V_k$  of the  $H^1$ -norm ball of radius  $R/2$  centred on  $u$ .

Consider now  $k \geq k_0$ , an arbitrary  $\bar{m}_k \in V_{k,+}$  and an arbitrary  $\bar{u}_k \in B_k(u, R/2)$ , noting that  $B_k(u, R/2)$  is nonempty for all  $k \geq k_0$ . The triangle inequality then implies that  $\|\bar{u}_k - u_k\|_{H^1(\Omega)} \leq R$ , so that each  $\bar{u}_k \in B_k(u, R/2)$ , with  $k \geq k_0$ , satisfies the hypotheses of Lemma 5.2.

We start by combining Lemma 5.3 and (5.2b) to obtain

$$\|m_k - \bar{m}_k\|_{H^1(\Omega)} \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_\Omega. \quad (7.1)$$

Therefore, the triangle inequality implies that

$$\begin{aligned} & \|m_k - \bar{m}_k\|_{H^1(\Omega)} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|u_k - \bar{u}_k\|_{H^1(\Omega)}. \end{aligned} \quad (7.2)$$

Since  $\|u_k - \bar{u}_k\|_{H^1(\Omega)} \leq R$ , we then apply Lemma 5.2 to the last term on the right-hand side of (7.2) above to find that

$$\begin{aligned} & \|m_k - \bar{m}_k\|_{H^1(\Omega)} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*} + \|\bar{m}_k - m_k\|_{\Omega} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega} + \|\bar{m}_k - m_k\|_{\Omega}, \end{aligned} \quad (7.3)$$

where in passing to the last line above, we have applied (5.2a) from Lemma 5.1 to bound  $\|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*}$ . Then, since by hypothesis  $\bar{m}_k \in V_{k,+}$ , we use the bound for  $\|\bar{m}_k - m_k\|_{\Omega}$  in (6.7) (after taking square-roots) from Lemma 6.2 which implies that

$$\begin{aligned} & \|m_k - \bar{m}_k\|_{H^1(\Omega)} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega} \\ & \quad + \|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*}^{\frac{1}{2}} \|m_k - \bar{m}_k\|_{H^1(\Omega)}^{\frac{1}{2}} + \|R_k^2(\bar{m}_k, \bar{u}_k)\|_{V_k^*}^{\frac{1}{2}} \|u_k - \bar{u}_k\|_{H^1(\Omega)}^{\frac{1}{2}}. \end{aligned} \quad (7.4)$$

Applying Young's inequality to the last two terms on the right-hand side of (7.4) and simplifying, we deduce that

$$\begin{aligned} & \|m_k - \bar{m}_k\|_{H^1(\Omega)} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega} \\ & \quad + \|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*} + \|R_k^2(\bar{m}_k, \bar{u}_k)\|_{V_k^*} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega}, \end{aligned} \quad (7.5)$$

where the last inequality is obtained by applying Lemma 5.1 to bound  $\|R_k^1(\bar{m}_k, \bar{u}_k)\|_{V_k^*}$  and  $\|R_k^2(\bar{m}_k, \bar{u}_k)\|_{V_k^*}$ . It follows from (7.5) and the triangle inequality that

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim \|m - \bar{m}_k\|_{H^1(\Omega)} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega}. \end{aligned} \quad (7.6)$$

Since  $\bar{m}_k$  was arbitrary in  $V_{k,+}$  and  $\bar{u}_k$  was arbitrary in  $B_k(u, R/2)$ , we consider the infima in (7.6) over all such  $\bar{m}_k$  and  $\bar{u}_k$  to obtain

$$\begin{aligned} & \|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)} \\ & \lesssim \inf_{\bar{m}_k \in V_{k,+}} \|m - \bar{m}_k\|_{H^1(\Omega)} + \inf_{\bar{u}_k \in B_k(u, R/2)} \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla m\|_{\Omega} + \|D_k \nabla u\|_{\Omega}. \end{aligned} \quad (7.7)$$

Since  $B_k(u, R/2)$  is nonempty, it is clear that

$$\inf_{\bar{u}_k \in B_k(u, R/2)} \|u - \bar{u}_k\|_{H^1(\Omega)} = \inf_{\bar{u}_k \in V_k} \|u - \bar{u}_k\|_{H^1(\Omega)}. \quad (7.8)$$

Thus, using this identity and using Lemma 6.1 in (7.7) yields (4.1), thereby completing the proof.  $\square$

## 8 Proof of Theorem 4.3

As explained in Section 4.1, we seek bounds for a composite norm on the error that involves the  $L^2$ -norm of the error for the density and  $H^1$ -norm of the error for the value function. The starting point for the analysis is in Lemma 8.1, which shows that, when considering the specific approximation  $m_{k,+} \in V_{k,+}$  defined in (6.1) above, we obtain a sharper bound on the residual in the KFP equation which involves the approximation error for the density function only in the  $L^2$ -norm rather than the  $H^1$ -norm.

**Lemma 8.1.** *Assume (H1) and (H2). Then, for any  $\bar{u}_k \in V_k$ ,*

$$\|R_k^2(m_{k,+}, \bar{u}_k)\|_{V_k^*} \lesssim \|m - m_{k,+}\|_{\Omega} + \|u - \bar{u}_k\|_{H^1(\Omega)}, \quad (8.1)$$

where the hidden constant depends only on  $L_H$ ,  $L_{H_p}$  and  $M_{\infty}$ .

Note that the right-hand side of (8.1) features only the  $L^2$ -norm term  $\|m - m_{k,+}\|_{\Omega}$ , in contrast to the right-hand side of (5.2b) which includes an  $H^1$ -norm approximation error term for the density.

*Proof.* Fix  $k \in \mathbb{N}$  and let  $\bar{u}_k \in V_k$  be given. For any  $\phi_k \in V_k$ , it follows from the definition of  $m_{k,+}$  given in (6.1) that

$$\begin{aligned} & \langle R_k^2(m_{k,+}, \bar{u}_k), \phi_k \rangle_{V_k^* \times V_k} \\ &= \langle G, \phi_k \rangle_{H^{-1} \times H_0^1} - \int_{\Omega} A_k \nabla m_{k,+} \cdot \nabla \phi_k + m_{k,+} \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \cdot \nabla \phi_k \, dx \\ &= \int_{\Omega} m_{k,+} \left( \frac{\partial H}{\partial p}[\nabla u] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \right) \cdot \nabla \phi_k \, dx \\ &= \int_{\Omega} (m_{k,+} - m) \left( \frac{\partial H}{\partial p}[\nabla u] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \right) \cdot \nabla \phi_k + m \left( \frac{\partial H}{\partial p}[\nabla u] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \right) \cdot \nabla \phi_k \, dx. \end{aligned}$$

Consequently, we have for all  $\phi_k \in V_k$  that

$$\begin{aligned} & |\langle R_k^2(m_{k,+}, \bar{u}_k), \phi_k \rangle_{V_k^* \times V_k}| \\ & \leq \left( 2L_H \|m_{k,+} - m\|_{\Omega} + M_{\infty} \left\| \frac{\partial H}{\partial p}[\nabla u] - \frac{\partial H}{\partial p}[\nabla \bar{u}_k] \right\|_{\Omega} \right) \|\nabla \phi_k\|_{\Omega} \quad (8.2) \end{aligned}$$

where we have applied (2.6) and the fact that  $m \in L^{\infty}(\Omega)$  by (2.18). Then the Lipschitz condition (2.5) implies that

$$\|R_k^2(m_{k,+}, \bar{u}_k)\|_{V_k^*} \leq 2L_H \|m - m_{k,+}\|_{\Omega} + M_{\infty} L_{H_p} \|u - \bar{u}_k\|_{H^1(\Omega)}, \quad (8.3)$$

which shows (8.1).  $\square$

Assuming that the domain  $\Omega$  is convex, we use a standard duality argument to give a bound on the  $L^2$  norm of the approximation error  $\|m - m_{k,+}\|_{\Omega}$ .

**Lemma 8.2.** *Assume (H1) and (H2). Suppose also that  $\Omega$  is convex. Then*

$$\|m - m_{k,+}\|_{\Omega} \lesssim h_k \|m\|_{H^1(\Omega)}, \quad (8.4)$$

where the hidden constant depends only on  $d$ ,  $\Omega$ ,  $\nu$ ,  $L_H$ ,  $\delta$ ,  $C_{\text{stab}}$ , and  $C_D$ .

*Proof.* To begin, for each  $k \in \mathbb{N}$ , let  $z_k$  be the unique function in  $H_0^1(\Omega)$  such that

$$\int_{\Omega} \nu \nabla z_k \cdot \nabla \phi + \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla z_k \phi dx = \int_{\Omega} (m - m_{k,+}) \phi dx \quad \forall \phi \in H_0^1(\Omega). \quad (8.5)$$

The existence and uniqueness of such  $z_k$  for each  $k \in \mathbb{N}$  is guaranteed by [17, Theorem 8.3] and moreover  $\|z_k\|_{H^1(\Omega)} \lesssim \|m - m_{k,+}\|_{\Omega}$  with a constant depending only on  $d, \nu, L_H$ , and on  $\Omega$ , see also [27, 28, Lemma 4.5]. Since  $\Omega$  is convex, and since  $\frac{\partial H}{\partial p}[\nabla u] \in L^\infty(\Omega; \mathbb{R}^d)$ , the  $H^2$ -regularity of solutions of Poisson's equation on bounded convex domains, see [18, Theorem 3.2.1.2], implies that  $z_k \in H^2(\Omega)$  with  $\|z_k\|_{H^2(\Omega)} \lesssim \|m - m_{k,+}\|_{\Omega}$ . Now let  $\bar{z}_k \in V_k$  be any quasi-interpolant of  $z_k$  that satisfies the following approximation and stability bounds

$$\|z_k - \bar{z}_k\|_{H^1(\Omega)} \lesssim h_k \|z_k\|_{H^2(\Omega)}, \quad \|\bar{z}_k\|_{H^1(\Omega)} \lesssim \|z_k\|_{H^1(\Omega)}. \quad (8.6)$$

For instance, one can use  $\bar{z}_k = \mathcal{I}_k z_k$  where  $\mathcal{I}_k$  denotes the Scott–Zhang quasi-interpolation operator [32, 7], see e.g. (4.8) and (4.9) above.

Then, using  $\phi = m - m_{k,+}$  in (8.5), and using the definition of  $m_{k,+}$  from (6.1), we eventually find that

$$\begin{aligned} \|m - m_{k,+}\|_{\Omega}^2 &= \int_{\Omega} \nu \nabla z_k \cdot \nabla (m - m_{k,+}) + (m - m_{k,+}) \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla z_k dx \\ &= \int_{\Omega} \nu \nabla (z_k - \bar{z}_k) \cdot \nabla (m - m_{k,+}) + (m - m_{k,+}) \frac{\partial H}{\partial p}[\nabla u] \cdot \nabla (z_k - \bar{z}_k) dx \\ &\quad + \int_{\Omega} D_k \nabla m_{k,+} \cdot \nabla \bar{z}_k dx. \end{aligned} \quad (8.7)$$

Combining (8.7) with (8.6), we obtain

$$\begin{aligned} \|m - m_{k,+}\|_{\Omega}^2 &\lesssim (h_k \|m - m_{k,+}\|_{H^1(\Omega)} + \|D_k \nabla m_{k,+}\|_{\Omega}) \|z_k\|_{H^2(\Omega)} \\ &\lesssim h_k \|m\|_{H^1(\Omega)} \|m - m_{k,+}\|_{\Omega}, \end{aligned} \quad (8.8)$$

where in the last line we used the bound on  $D_k$  from (H1), the stability bound  $\|m_{k,+}\|_{H^1(\Omega)} \lesssim \|m\|_{H^1(\Omega)}$ , and the elliptic regularity of  $z_k$  above. We then simplify (8.8) to obtain (8.4).  $\square$

We now prove Theorem 4.3, with the main idea being to compare the discrete density approximation  $m_k$  to the specific approximation  $m_{k,+}$  defined in (6.1) above. Combining the  $L^2$ -norm bounds above with the results in the previous sections, we are then able to obtain a bound on the relevant composite norm for the error.

**Proof of Theorem 4.3** Similar to the proof of Theorem 4.1, we consider  $k_0$  sufficiently large so that  $u_k \in B_k(u, R/2)$  for all  $k \geq k_0$ , where  $R$  is the constant given in Lemma 5.2. For each  $k \geq k_0$ , let  $\bar{u}_k \in B_k(u, R/2)$  be arbitrary, noting that  $B_k(u, R/2)$  is nonempty. In the following, let  $E_k \geq 0$ ,  $k \in \mathbb{N}$ , be the quantity defined by

$$E_k := \|m - m_{k,+}\|_{\Omega} + \|u - \bar{u}_k\|_{H^1(\Omega)} + \|D_k \nabla u\|_{\Omega} + \|D_k \nabla m\|_{\Omega}. \quad (8.9)$$

We stress that it is the  $L^2$ -norm of  $m - m_{k,+}$  appearing in the definition of  $E_k$  above. We start by noting that Lemmas 5.1 and 8.1 imply that

$$\|R_k^1(m_{k,+}, \bar{u}_k)\|_{V_k^*} + \|R_k^2(m_{k,+}, \bar{u}_k)\|_{V_k^*} \lesssim E_k. \quad (8.10)$$



Next, Lemmas 5.3 and 8.1 combined with (8.10) imply that

$$\begin{aligned} \|m_k - m_{k,+}\|_{H^1(\Omega)} &\lesssim \|R_k^2(m_{k,+}, u_k)\|_{V_k^*} \\ &\lesssim \|m - m_{k,+}\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)} \leq E_k + \|u_k - \bar{u}_k\|_{H^1(\Omega)}, \end{aligned} \quad (8.11)$$

where we have applied the triangle inequality in passing to the last inequality above. Applying Lemma 5.2 to the last term on the right-hand side of (8.11) and combining with (8.10), we thus obtain

$$\|m_k - m_{k,+}\|_{H^1(\Omega)} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \lesssim E_k + \|m_k - m_{k,+}\|_{\Omega}. \quad (8.12)$$

Then, Lemma 6.2 and (8.10) imply that

$$\begin{aligned} \|m_k - m_{k,+}\|_{\Omega}^2 &\lesssim \|R_k^1(m_{k,+}, \bar{u}_k)\|_{V_k^*} \|m_k - m_{k,+}\|_{H^1(\Omega)} + \|R_k^2(m_{k,+}, \bar{u}_k)\|_{V_k^*} \|u_k - \bar{u}_k\|_{H^1(\Omega)} \\ &\leq E_k^2 + E_k \|m_k - m_{k,+}\|_{\Omega}. \end{aligned} \quad (8.13)$$

Applying Young's inequality with a parameter, we deduce that

$$\|m_k - m_{k,+}\|_{\Omega} \lesssim E_k. \quad (8.14)$$

Therefore, using (8.12), (8.14) and the triangle inequality, we get

$$\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)} \leq E_k + \|m_k - m_{k,+}\|_{\Omega} + \|u_k - \bar{u}_k\|_{H^1(\Omega)} \lesssim E_k. \quad (8.15)$$

Finally, we apply Lemma 8.2 and (H1) to conclude that

$$\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)} \lesssim E_k \lesssim \|u - \bar{u}_k\|_{H^1(\Omega)} + h_k(\|m\|_{H^1(\Omega)} + \|u\|_{H^1(\Omega)}). \quad (8.16)$$

Since  $\bar{u}_k$  was arbitrary in  $B_k(u, R/2)$ , we may then consider the infimum in (8.16) and, recalling here (7.8), we conclude (4.14).  $\square$

## 9 Numerical Experiments

In this section, we present numerical results that verify the optimal asymptotic rates of convergence that are implied by Theorems 4.1 and 4.3, respectively, for some concrete examples of the system (1.1) with explicit weak solutions  $(u, m)$ . Recall that the rate of convergence of the approximations for a given norm is called optimal if it equals the rate of convergence of a sequence of best approximations from the same sequence of approximation spaces and in the same norm. In particular, for the first experiment where  $u$  is not  $H^2$ -regular and  $m$  is smooth, we showcase Corollary 4.2 of Theorem 4.1 on the rate of convergence of the total  $H^1$ -norm error  $\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)}$ . In the second experiment, we illustrate the conclusion of Theorem 4.3 on the convergence rate of the composite  $L^2$ - $H^1$ -norm error  $\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)}$ . Both our experiments involve examples where at least one of the solution components  $u$  or  $m$  has limited fractional Sobolev regularity.

### 9.1 Set-up

For each experiment, we consider the system (1.1) where  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$  is the unit square, the diffusion  $\nu = 1$ , and the Hamiltonian  $H : \bar{\Omega} \times \mathbb{R}^2 \rightarrow \mathbb{R}$  is defined by

$$H(x, p) = \sup_{\alpha \in B_1(0)} (\alpha \cdot p + \sqrt{1 - |\alpha|^2}) - 1 = \sqrt{|p|^2 + 1} - 1 \quad \forall (x, p) \in \bar{\Omega} \times \mathbb{R}^2. \quad (9.1)$$

It is clear that  $L_H = 1$ . In each experiment we specify the choices of the operator  $F$  and the source term  $G$ . For each experiment we apply the FEM (3.3) along a dyadic, nested sequence  $\{\mathcal{T}_k\}_{k=1}^9$  of shape-regular uniform meshes on  $\Omega$  consisting of right-angled simplicial elements. We note that the sequence  $\{\mathcal{T}_k\}_{k=1}^9$  satisfies the XZ condition (3.10). The stabilization is given by the diffusion tensor (3.12) where the weights are taken to be  $\omega_{k,E} = L_H \text{diam}(E)$  for all  $E \in \mathcal{E}_k$ ,  $k \in \{1, \dots, 9\}$ . Therefore, hypotheses (H1) and (H2) are satisfied according to Lemma 3.2, which ensures that the FEM (3.3) is monotone in these experiments. The computations for each experiment were performed in Firedrake [31].

## 9.2 Experiment 1

(Test with nonsmooth value function)

The purpose of this experiment is to verify numerically the asymptotic rate of convergence of the total  $H^1$ -norm error  $\|m - m_k\|_{H^1(\Omega)} + \|u - u_k\|_{H^1(\Omega)}$  predicted by Theorem 4.1, via Corollary 4.2, for a problem where at least one of  $u$  or  $m$  is not  $H^2$ -regular. To this end, we consider the problem (2.15) with model data that yields a unique weak solution  $(u, m)$ , where the value function  $u$  has limited fractional order Sobolev regularity.

Let the operator  $F : L^2(\Omega) \rightarrow H^{-1}(\Omega)$  take the form

$$\langle F[m], \psi \rangle_{H^{-1} \times H_0^1} := \int_{\Omega} m \psi dx + \langle F_1, \psi \rangle_{H^{-1} \times H_0^1} \quad \forall \psi \in H_0^1(\Omega), \quad (9.2)$$

where  $F_1 \in H^{-1}(\Omega)$  is given by

$$\langle F_1, \psi \rangle := \int_{\Omega} \left( \sqrt{|\tilde{v}_1|^2 + 1} - 1 - xy \ln(2-x) \ln(2-y) \right) \psi + \tilde{v}_1 \cdot \nabla \psi dx \quad \forall \psi \in H_0^1(\Omega), \quad (9.3)$$

with the vector field  $\tilde{v}_1 : \Omega \rightarrow \mathbb{R}^2$  defined by

$$\tilde{v}_1(x, y) := \frac{12.8}{(xy(1-x)(1-y))^{0.2}} \begin{pmatrix} (1-2x)y(1-y) \\ x(1-x)(1-2y) \end{pmatrix}, \quad (9.4)$$

for  $(x, y) \in \Omega$ . We take the source term  $G \in H^{-1}(\Omega)$  of the form

$$\langle G, \phi \rangle_{H^{-1} \times H_0^1} := \int_{\Omega} \tilde{v}_2 \cdot \nabla \phi dx \quad \forall \phi \in H_0^1(\Omega), \quad (9.5)$$

where  $\tilde{v}_2 : \Omega \rightarrow \mathbb{R}^2$  is defined by

$$\tilde{v}_2(x, y) := \begin{pmatrix} \left( \ln(2-x) - \frac{x}{2-x} \right) y \ln(2-y) \\ x \ln(2-x) \left( \ln(2-y) - \frac{y}{2-y} \right) \end{pmatrix} + \frac{xy \ln(2-x) \ln(2-y)}{\sqrt{|\tilde{v}_1(x, y)|^2 + 1}} \tilde{v}_1(x, y), \quad (9.6)$$

for  $(x, y) \in \Omega$ . It is clear that  $F$  defined by (9.2) via (9.3) and (9.4) is strongly monotone in the sense of condition (2.2). By following the proof of [26, Lemma 3.3.1], one can show that  $G$  defined by (9.5) via (9.6) is nonnegative in the sense of distributions in  $H^{-1}(\Omega)$ , and that the unique solution to the weak formulation (2.15) is given by

$$u(x, y) := 16(xy(1-x)(1-y))^{4/5}, \quad m(x, y) := xy \ln(2-x) \ln(2-y), \quad (9.7)$$

for all  $(x, y) \in \Omega$ . It can be shown that the value function  $u$  above lies in the fractional order Sobolev space  $H^{13/10-\epsilon}(\Omega)$ , for all  $0 < \epsilon \leq 13/10$  while the density function  $m$  above

is clearly  $H^2$ -regular. Theorem 4.1 and Corollary 4.2 imply a rate of convergence for the total  $H^1$ -norm error of order  $3/10$ . Figure 1 confirms this theoretical prediction, where we plot the total error in the  $H^1$ -norm, as well as its constituent parts, against the mesh size  $h$ . We observe that the convergence rate for the total error is of order  $3/10$ , as predicted. In addition, the convergence rates in the errors  $\|u - u_k\|_{H^1(\Omega)}$  and  $\|m - m_k\|_{H^1(\Omega)}$ , being of order  $3/10$  and 1 respectively, are also optimal in view of the corresponding regularity of  $u$  and  $m$ .

### 9.3 Experiment 2

(Test with nonsmooth density function)

The aim of this experiment is to illustrate the conclusion of Theorem 4.3 on the convergence rate of the total  $L^2$ - $H^1$ -norm error  $\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)}$ . As such, we consider an example where  $u$  is smooth, but  $m$  has limited fractional order Sobolev regularity. Let the operator  $F : L^2(\Omega) \rightarrow H^{-1}(\Omega)$  take the form

$$\langle F[m], \psi \rangle_{H^{-1} \times H_0^1} := \int_{\Omega} m \psi dx + \langle F_2, \psi \rangle_{H^{-1} \times H_0^1} \quad \forall \psi \in H_0^1(\Omega), \quad (9.8)$$

where  $F_2 \in H^{-1}(\Omega)$  is given by

$$\langle F_2, \psi \rangle := \int_{\Omega} \left( \sqrt{|\tilde{r}_1|^2 + 1} - 1 - xy \ln(x) \ln(y) \right) \psi + \tilde{r}_1 \cdot \nabla \psi dx \quad \forall \psi \in H_0^1(\Omega), \quad (9.9)$$

with the vector field  $\tilde{r}_1 : \Omega \rightarrow \mathbb{R}^2$  defined by

$$\tilde{r}_1(x, y) := \begin{pmatrix} 16(1-2x)y(1-y) \\ 16x(1-x)(1-2y) \end{pmatrix}, \quad (9.10)$$

for  $(x, y) \in \Omega$ . We take the source term  $G \in H^{-1}(\Omega)$  of the form

$$\langle G, \phi \rangle_{H^{-1} \times H_0^1} := \int_{\Omega} \tilde{r}_2 \cdot \nabla \phi dx \quad \forall \phi \in H_0^1(\Omega), \quad (9.11)$$

where  $\tilde{r}_2 : \Omega \rightarrow \mathbb{R}^2$  is defined by

$$\tilde{r}_2(x, y) := \begin{pmatrix} (1 + \ln(x)) y \ln(y) \\ x \ln(x) (1 + \ln(y)) \end{pmatrix} + \frac{xy \ln(x) \ln(y)}{\sqrt{|\tilde{r}_1(x, y)|^2 + 1}} \tilde{r}_1(x, y), \quad (9.12)$$

for  $(x, y) \in \Omega$ . It is clear that  $F$  defined by (9.8) via (9.9) and (9.10) is strongly monotone in the sense of condition (2.2). By following the proof of [26, Lemma 3.3.1], one can show that  $G$  defined by (9.11) via (9.12) is nonnegative in the sense of distributions in  $H^{-1}(\Omega)$ , and that the unique solution to the weak formulation (2.15) is given by

$$u(x, y) := 16xy(1-x)(1-y), \quad m(x, y) := xy \ln(x) \ln(y), \quad (9.13)$$

for all  $(x, y) \in \Omega$ . It is clear that the value function  $u$  above is  $H^2$ -regular, while the density function  $m$  above can be shown to lie in the fractional order Sobolev space  $H^{3/2-\epsilon}(\Omega)$ , for all  $0 < \epsilon \leq 3/2$ . Recall that the domain  $\Omega$  in this experiment is convex. We therefore deduce from Theorem 4.3 and Corollary 4.4, that the total  $L^2$ - $H^1$ -norm error  $\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)}$  should converge with a rate of order 1, despite the lower regularity of  $m$  relative to  $u$ . In Figure 2, we see that the results of the experiment are in agreement with

the theoretical predictions. In particular, the composite  $L^2$ - $H^1$  error converges with a rate of order 1, which is optimal due to the  $H^2$ -regularity of  $u$ . It is also seen that the density function approximations converge to  $m$  in the  $H^1$ -norm at the optimal asymptotic rate of order  $1/2$ . The results of this experiment thus illustrate how in particular the  $H^1$ -norm error for the value function approximations can converge with optimal rates, independently of the lower regularity of the density function  $m$  of the weak solution.

## Acknowledgements

Yohance A. P. Osborne was supported by The Royal Society Career Development Fellowship programme. Iain Smears was supported by the Engineering and Physical Sciences Research Council [grant number EP/Y008758/1].

## References

- [1] Y. Achdou, F. Camilli, and I. Capuzzo-Dolcetta, *Mean field games: convergence of a finite difference method*, SIAM J. Numer. Anal., **51** (2013), no. 5, 2585–2612. <https://doi.org/10.1137/120882421>.
- [2] Y. Achdou and I. Capuzzo-Dolcetta, *Mean field games: numerical methods*, SIAM J. Numer. Anal., **48** (2010), no. 3, 1136–1162. <https://doi.org/10.1137/090758477>.
- [3] Y. Achdou and A. Porretta, *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** (2016), no. 1, 161–186. <https://doi.org/10.1137/15M1015455>.
- [4] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003.
- [5] Y. Ashrafyan and D. Gomes, *A fully-discrete semi-Lagrangian scheme for a price formation MFG model*, arXiv preprint arXiv:2403.02785, (2024).
- [6] J. F. Bonnans, K. Liu, and L. Pfeiffer, *Error estimates of a theta-scheme for second-order mean field games*, ESAIM Math. Model. Numer. Anal., **57** (2023), no. 4, 2493–2528. <https://doi.org/10.1051/m2an/2023059>.
- [7] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008, <https://doi.org/10.1007/978-0-387-75934-0>.
- [8] E. Burman and A. Ern, *Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the convection–diffusion–reaction equation*, Computer Methods in Applied Mechanics and Engineering, **191** (2002), no. 35, 3833–3855. [https://doi.org/10.1016/S0045-7825\(02\)00318-3](https://doi.org/10.1016/S0045-7825(02)00318-3).
- [9] E. Burman, P. Hansbo, and M. G. Larson, *A cut finite element method with boundary value correction*, Math. Comp., **87** (2018), no. 310, 633–657. <https://doi.org/10.1090/mcom/3240>.
- [10] E. Carlini and F. J. Silva, *A fully discrete semi-Lagrangian scheme for a first order mean field game problem*, SIAM J. Numer. Anal., **52** (2014), no. 1, 45–67. <https://doi.org/10.1137/120902987>.

- [11] E. Carlini and F. J. Silva, *A semi-Lagrangian scheme for a degenerate second order mean field game system*, *Discrete Contin. Dyn. Syst.*, **35** (2015), no. 9, 4269–4292. <https://doi.org/10.3934/dcds.2015.35.4269>.
- [12] E. Carlini and F. J. Silva, *On the discretization of some nonlinear Fokker–Planck–Kolmogorov equations and applications*, *SIAM Journal on Numerical Analysis*, **56** (2018), no. 4, 2148–2177.
- [13] E. Carlini, F. J. Silva, and A. Zorkot, *A Lagrange–Galerkin scheme for first order mean field game systems*, *SIAM Journal on Numerical Analysis*, **62** (2024), no. 1, 167–198. <https://doi.org/10.1137/23M1561762>.
- [14] I. Chowdhury, O. Ersland, and E. R. Jakobsen, *On numerical approximations of fractional and nonlocal mean field games*, *Found. Comput. Math.*, **23** (2023), no. 4, 1381–1431. <https://doi.org/10.1007/s10208-022-09572-w>.
- [15] P. G. Ciarlet, *The finite element method for elliptic problems*, vol. 40 of *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, <https://doi.org/10.1137/1.9780898719208>.
- [16] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, vol. 25 of *Stochastic Modelling and Applied Probability*, Springer, New York, second ed., 2006.
- [17] D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, 2001, <https://doi.org/10.1007/978-3-642-61798-0>.
- [18] P. Grisvard, *Elliptic problems in nonsmooth domains*, vol. 69 of *Classics in Applied Mathematics*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011, Reprint of the 1985 original, with a foreword by Susanne C. Brenner.
- [19] M. Huang, R. P. Malhamé, and P. E. Caines, *Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle*, *Commun. Inf. Syst.*, **6** (2006), no. 3, 221–251. <https://doi.org/10.4310/cis.2006.v6.n3.a5>
- [20] D. Inoue, Y. Ito, T. Kashiwabara, N. Saito, and H. Yoshida, *A fictitious-play finite-difference method for linearly solvable mean field games*, *ESAIM: M2AN*, **57** (2023), no. 4, 1863–1892. <https://doi.org/10.1051/m2an/2023026>.
- [21] M. Jensen,  $L^2(H_\gamma^1)$  finite element convergence for degenerate isotropic Hamilton–Jacobi–Bellman equations, *IMA J. Numer. Anal.*, **37** (2017), no. 3, 1300–1316. <https://doi.org/10.1093/imanum/drw055>.
- [22] M. Jensen and I. Smears, *On the convergence of finite element methods for Hamilton–Jacobi–Bellman equations*, *SIAM J. Numer. Anal.*, **51** (2013), no. 1, 137–162. <https://doi.org/10.1137/110856198>.
- [23] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. I. Le cas stationnaire*, *C. R. Math. Acad. Sci. Paris*, **343** (2006), no. 9, 619–625. <https://doi.org/10.1016/j.crma.2006.09.019>.

- [24] J.-M. Lasry and P.-L. Lions, *Jeux à champ moyen. II. Horizon fini et contrôle optimal*, C. R. Math. Acad. Sci. Paris, **343** (2006), no. 10, 679–684. <https://doi.org/10.1016/j.crma.2006.09.018>.
- [25] J.-M. Lasry and P.-L. Lions, *Mean field games*, Jpn. J. Math., **2** (2007), no. 1, 229–260. <https://doi.org/10.1007/s11537-007-0657-8>.
- [26] Y. A. P. Osborne, *Analysis and Numerical Approximation of Mean Field Game Partial Differential Inclusions*, PhD thesis, University College London, UK, 2024.
- [27] Y. A. P. Osborne and I. Smears, *Analysis and numerical approximation of stationary second-order mean field game partial differential inclusions*, SIAM Journal on Numerical Analysis, **62** (2024), no. 1, 138–166. <https://doi.org/10.1137/22M1519274>.
- [28] Y. A. P. Osborne and I. Smears, *Erratum: Analysis and numerical approximation of stationary second-order mean field game partial differential inclusions*, SIAM Journal on Numerical Analysis, **62** (2024), no. 5, 2415–2417. <https://doi.org/10.1137/24M165123X>.
- [29] Y. A. P. Osborne and I. Smears, *Finite element approximation of time-dependent mean field games with nondifferentiable Hamiltonians*, Numerische Mathematik, (2024), 1–47. <https://doi.org/10.1007/s00211-024-01447-2>.
- [30] Y. A. P. Osborne and I. Smears, *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>.
- [31] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. McRae, G.-T. Bercea, G. R. Markall, and P. H. Kelly, *Firedrake: automating the finite element method by composing abstractions*, ACM Transactions on Mathematical Software (TOMS), **43** (2016), no. 3, 1–27. <https://doi.org/10.1145/2998441>.
- [32] L. R. Scott and S. Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., **54** (1990), no. 190, 483–493. <https://doi.org/10.2307/2008497>.
- [33] I. Smears and E. Süli, *Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients*, SIAM J. Numer. Anal., **52** (2014), no. 2, 993–1016. <https://doi.org/10.1137/130909536>.
- [34] J. Xu and L. Zikatanov, *A monotone finite element scheme for convection-diffusion equations*, Math. Comp., **68** (1999), no. 228, 1429–1446. <https://doi.org/10.1090/S0025-5718-99-01148-5>.

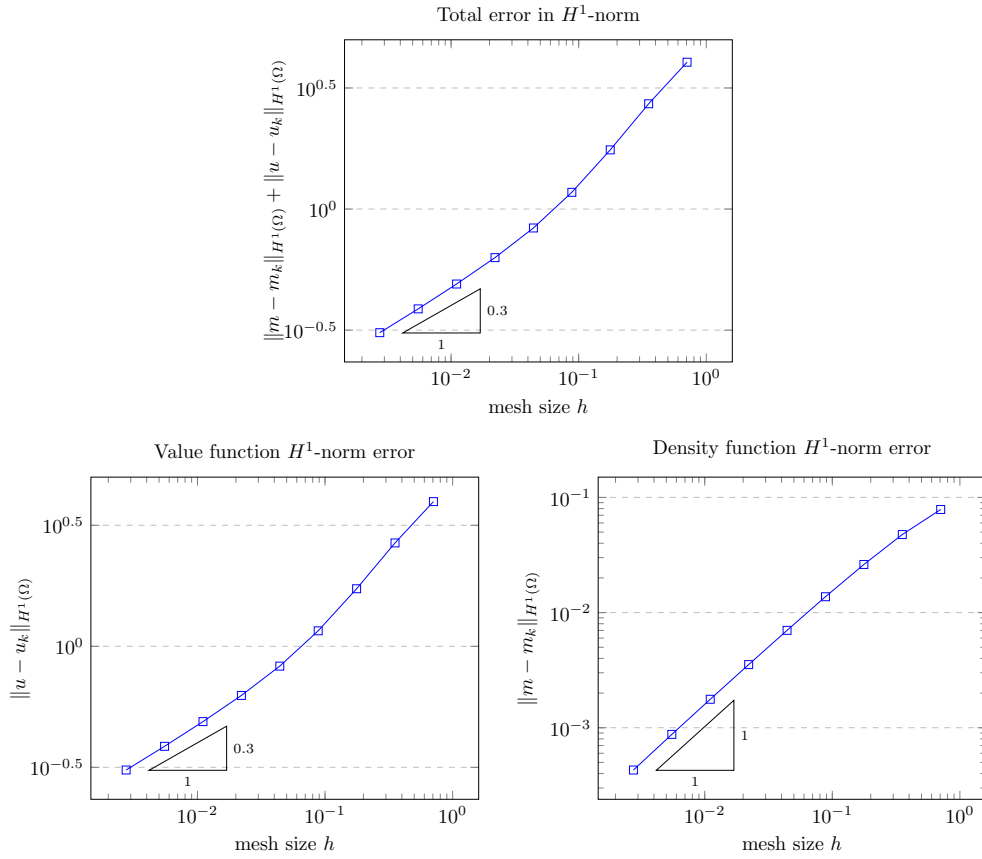


Figure 1: Experiment 1 – convergence plots for approximations of the value function and density function. The asymptotic rate of convergence for the total error in the  $H^1$ -norm is close to order  $3/10$ . This is due to the observed asymptotic rate of convergence in the  $H^1$ -norm for the approximations of the value function being close to the optimal value of  $3/10$ . The rate of convergence in the  $H^1$ -norm of the density function approximations is of optimal order 1.

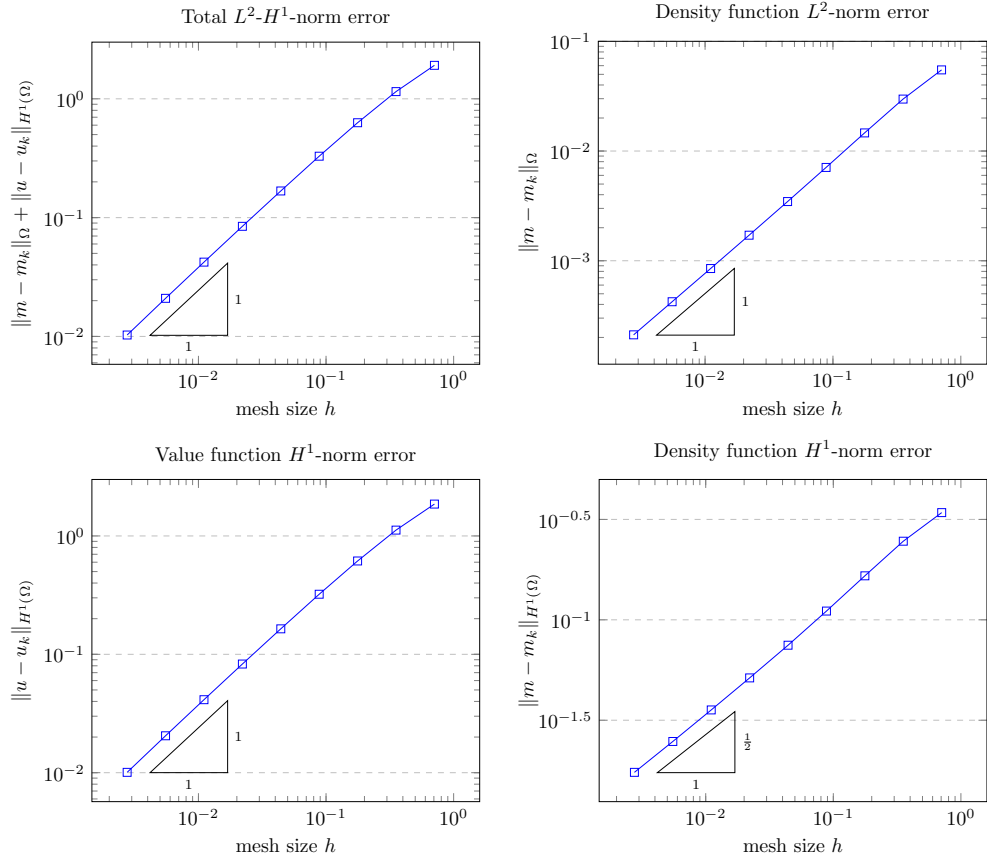


Figure 2: Experiment 2 – convergence plots for approximations of the value function and density function. The asymptotic rate of convergence in the total error  $\|m - m_k\|_{\Omega} + \|u - u_k\|_{H^1(\Omega)}$  is of optimal order 1. The convergence rate in the  $H^1$ -norm for the approximations of the density function is of order 1/2, which is also optimal given the lower regularity of the density function.