ESAIM: M2AN 53 (2019) 351–374
https://doi.org/10.1051/m2an/2018067

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

FINITE ELEMENT APPROXIMATION OF THE ISAACS EQUATION

Abner J. Salgado[1,*] and Wujun Zhang[2]

Abstract. We propose and analyze a two-scale finite element method for the Isaacs equation. The fine scale is given by the mesh size h whereas the coarse scale ε is dictated by an integro-differential approximation of the partial differential equation. We show that the method satisfies the discrete maximum principle provided that the mesh is weakly acute. This, in conjunction with weak operator consistency of the finite element method, allows us to establish convergence of the numerical solution to the viscosity solution as $\varepsilon, h \to 0$, and $\varepsilon \gtrsim (h|\log h|)^{1/2}$. In addition, using a discrete Alexandrov Bakelman Pucci estimate we deduce rates of convergence, under suitable smoothness assumptions on the exact solution.

Mathematics Subject Classification. 65N12, 65N15, 65N30, 35J60, 35D40, 35Q91.

Received January 30, 2018. Accepted October 28, 2018.

1. Introduction

Fully nonlinear elliptic partial differential equations (PDE) arise naturally from differential geometry, optimal mass transportation, stochastic optimal control and other fields of science and engineering. In spite of their wide range of applications, the numerical methods for this type of PDEs is still under development and this is particularly the case if one wants to apply the finite element method (FEM). A major difficulty in their numerical approximation is that, for fully nonlinear PDEs, the correct notion of solution is the so-called viscosity solution, which is based on the maximum principle instead of a variational one. This is reflected in the fact that, in contrast to an extensive literature for linear and quasilinear elliptic PDEs in divergence form, the numerical approximation via finite elements of fully nonlinear PDEs reduces to a few papers; we refer the reader to [14,31] for an overview. The situation is somewhat more satisfactory for finite difference approximations, where convergence to the viscosity solution without rates was studied in the early works [2,24]. Rates of convergence for convex/concave fully nonlinear elliptic equations have been established, *e.g.* by Krylov [21,22]; Barles and Jakobsen [1] and Debrabant and Jakobsen [12]. However, rates of convergence for the nonconvex/nonconcave case remained an open problem as the techniques used in the convex case could not be generalized to this scenario. For many years the existing results were rather specialized. For instance, [17] considered a one dimensional problem and obtained rates of convergence but its arguments do not extend to more dimensions. A particular nonconvex equation: an obstacle problem for a Hamilton Jacobi Bellman equation was studied in [18], where this particular structure is exploited to obtain convergence rates. This changed when, in 2008, Caffarelli and

Keywords and phrases. Fully nonlinear equations, discrete maximum principle, finite elements.

[1] Department of Mathematics, University of Tennessee, Knoxville, TN 37996, USA.

[2] Department of Mathematics, Rutgers University, Piscataway, NJ, USA.

[*]Corresponding author: asalgad1@utk.edu

Souganidis [9] established a rate of convergence for finite difference approximations of elliptic equations of the form

$$F[\mathrm{D}^2 u](x) = f(x),$$

with Dirichlet boundary conditions. This result was extended by Turanova [37] to the case where F is also dependent on x and by Krylov [23] to the Isaacs equation, where he allows the operator F to depend on x, ∇u and u. We finally comment that, to our knowledge, no rates of convergence are available for semi-Lagrangian schemes. Reference [12] provides convergence, but no rates are available for nonconvex operators of order two. It is remarkable, however, to note that explicit rates of convergence for nonlocal operators of order $\mu \in (1,2)$ are known. Namely, [3] shows that the rate of convergence is $\mathcal{O}(h^{\frac{2-\mu}{\mu}})$. Note that this rate degenerates as the order of the operators approaches two and that the authors of this work state that this is optimal and expected.

In all the works mentioned above, convergence hinges on operator consistency and monotonicity, which are nontrivial properties to be satisfied, especially in the FEM. To overcome this, the recent work [32] introduced a two scale FEM for linear elliptic PDEs in nondivergence form and, on the basis of monotonicity and *weak* operator consistency, the authors were able to prove convergence of the method with rates. The purpose of this work is to extend these ideas to a particular kind of uniformly elliptic Isaacs equations on convex domains, namely,

$$\mathfrak{F}[\mathsf{u}](x) := \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left[A^{\alpha,\beta}(x) : \mathrm{D}^2\mathsf{u}(x)\right] = \mathsf{f}(x) \text{ in } \Omega, \quad \mathsf{u} = 0 \text{ on } \partial\Omega, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^d$ $(d \geq 2)$ is an open, bounded and convex domain, $\mathsf{f} \in \mathrm{C}^{0,1}(\bar{\Omega})$, the sets $\mathcal{A}$ and $\mathcal{B}$ are arbitrary finite sets and the matrices $A^{\alpha,\beta} \in \mathrm{C}^{0,1}(\bar{\Omega}, \mathrm{GL}_d(\mathbb{R}))$ are symmetric and uniformly elliptic, in the sense that there are constants $\lambda, \Lambda \in \mathbb{R}$, with $0 < \lambda \leq \Lambda$, such that

$$\lambda I \leq A^{\alpha,\beta} \leq \Lambda I \quad \forall \alpha \in \mathcal{A}, \beta \in \mathcal{B}. \tag{1.2}$$

We will introduce, following [32], a two-scale FEM, show its convergence to the viscosity solution of (1.1) and provide rates of convergence. The method is based on the approximation of (1.1) proposed by Caffarelli and Silvestre [8]: we formally rewrite

$$\mathfrak{F}[u](x) = \frac{\lambda}{2}\Delta u(x) + \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} \left[\left(A^{\alpha,\beta}(x) - \frac{\lambda}{2}I\right) : \mathrm{D}^2 u(x)\right]$$

and we approximate the operator above by the *integro-differential* operator

$$\mathfrak{F}^\varepsilon[u](x) := \frac{\lambda}{2}\Delta u(x) + \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u](x) \tag{1.3}$$

where

$$I_\varepsilon^{\alpha,\beta}[u](x) = \frac{1}{\varepsilon^{d+2}\det M^{\alpha,\beta}(x)} \int_{\mathbb{R}^d} \mathfrak{d}u(x,y)\varphi\left(\frac{M^{\alpha,\beta}(x)^{-1}y}{\varepsilon}\right)\mathrm{d}y,$$

with

$$M^{\alpha,\beta}(x) := \left(A^{\alpha,\beta}(x) - \frac{\lambda}{2}I\right)^{1/2}.$$

Hereafter, φ is a radially symmetric function with support in the unit ball B_1 of $\mathbb{R}^d$, where $d \geq 1$ is the dimension, that verifies

$$\int_{\mathbb{R}^d} |y|^2\varphi(y)\,\mathrm{d}y = d$$

and

$$\mathfrak{d}u(x,y) := u(x+y) - 2u(x) + u(x-y) \tag{1.4}$$

is the centered second difference operator. The operator $I_\varepsilon^{\alpha,\beta}[\cdot]$ is a consistent approximation of $\left(A^{\alpha,\beta}(x) - \frac{\lambda}{2}I\right) : \mathrm{D}^2u(x)$ in the sense that if u is a quadratic polynomial, then

$$I_\varepsilon^{\alpha,\beta}[u](x) = \left(A^{\alpha,\beta}(x) - \frac{\lambda}{2}I\right) : \mathrm{D}^2u(x) \qquad \forall \varepsilon > 0, \quad \forall u \in \mathbb{P}_2, \tag{1.5}$$

see Lemma 2.1.

We discretize (1.1) by using this approximation as follows: Introduce a triangulation $\mathscr{T}_h$ of Ω and let $\mathscr{N}_h$ be the set of internal nodes. We introduce a finite element space consisting of piecewise linear functions and denote by $\{\phi_z\}_{z\in\mathscr{N}_h}$ its Lagrange nodal basis. Now, multiply (1.3) by ϕ_z and integrate over Ω. Integrate the first term by parts and apply mass lumping to the second to finally obtain that the finite element approximation, u_h^ε, satisfies

$$\mathfrak{F}_h^\varepsilon[u_h^\varepsilon](z) := -\frac{\lambda}{2}\int_\Omega \nabla u_h^\varepsilon(x)\cdot\nabla\phi_z(x)\,\mathrm{d}x + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z)\int_\Omega \phi_z(x)\,\mathrm{d}x = \int_\Omega \phi_z(x)\mathsf{f}(x)\,\mathrm{d}x \qquad \forall z \in \mathscr{N}_h,$$

We show that the FEM is monotone provided that meshes are weakly acute. To show existence and uniqueness, we employ a *discrete* version of Perron's method, which seems to not have been considered before, especially in the finite element literature. Exploiting monotonicity, and using the notion of *weak consistency* introduced by Jensen and Smears [19] we show convergence to the viscosity solution following [2, 24].

We also derive rates of convergence for the method. The main difficulty to derive a rate of convergence is to establish a suitable notion of stability for the FEM applied to this fully nonlinear PDE. To address this issue, we resort to the discrete Alexandrov Bakelman Pucci (ABP) estimate of [32], which reads

$$\sup_\Omega (u_h^\varepsilon)^- \lesssim \left(\sum_{\left\{z\in\mathscr{N}_h : u_h^\varepsilon(z) = \Gamma(u_h^\varepsilon)(z)\right\}} |f_z|^d\omega_z\right)^{1/d}. \tag{1.6}$$

Here $\{z \in \mathscr{N}_h : u_h^\varepsilon(z) = \Gamma(u_h^\varepsilon)(z)\}$ denotes the *(lower) nodal contact set*, $\Gamma(u_h^\varepsilon)$ is the convex envelope of u_h^ε (see (3.17)),

$$f_z = \left(\int_\Omega \mathsf{f}(x)\phi_z(x)\,\mathrm{d}x\right)\left(\int_\Omega \phi_z(x)\,\mathrm{d}x\right)^{-1},$$

and, for each node $z \in \mathscr{N}_h$, we set $\omega_z = \int_\Omega \phi_z(x)\,\mathrm{d}x$. Note that the nodal contact set is just a finite collection of nodes. With (1.6) we obtain control of the negative part of u_h^ε. If we consider the concave envelope and corresponding (upper) contact set, we can estimate the positive part. A combination of these bounds yields stability, in the L^∞-norm, of u_h^ε in terms of the L^d-norm of the right hand side f. Suitable notions of consistency, monotonicity and stability yield rates of convergence. Moreover, by combining them with the regularity estimates for the Isaacs equation of [8], we show that, for some $\sigma \in (0, 1]$,

$$\|\mathsf{u} - u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim h^{\frac{\sigma}{\sigma+2}}|\log h|^{\frac{\sigma}{\sigma+2}}\|\mathsf{f}\|_{\mathrm{C}^{0,1}(\bar{\Omega})},$$

which, at best, gives us a $\mathcal{O}(h^{1/3}|\log h|^{1/3})$ error estimate.

We also discuss how to practically realize the method in question. We study, following [6], a variant of Howard's algorithm to solve the ensuing discrete (nonlinear) systems.

Before proceeding any further, we must also comment that our ideas and methods also apply to a version of the Hamilton Jacobi Bellman equation. Namely, setting $\#\mathcal{A} = 1$ or $\#\mathcal{B} = 1$ we arrive at a convex or concave operator. In this case much more could be said for this problem, and it is possible that our analysis yields suboptimal error estimates. In particular, the rate of convergence σ, which comes form (2.6) below, might be known. Since our main interest here is to deal with the Isaacs equation, we will not explore this refinement any further.

We must also mention that the linear case, *i.e.* $\#\mathcal{A} = \#\mathcal{B} = 1$, was already treated in [32]. Our analysis, however, departs from theirs in several respects:

- To apply the regularity and approximation results of [8], as opposed to [32], the second difference operator $\mathfrak{d}$ must be of the form (1.4), and should not be modified near the boundary. This forces us to consider *extensions of functions outside the domain* Ω, and our analysis must take this into account.
- Since our problem is nonlinear existence is not guaranteed by uniqueness, as it is the case in [32]. Thus, although uniqueness is obtained by similar (monotonicity and maximum principle based) arguments, we must develop a *discrete Perron's method* to assert existence of solutions.
- Due to the fact that solutions to the Isaacs equation (1.1) exhibit *very low regularity*, the convergence arguments that we employ must take this into account. In particular this means that the relation between the two scales, ε and h, of our method must be modified, and that suitable barrier functions must be constructed.

The rest of this paper is organized as follows. In Section 2, we recall the approximation of elliptic problems by integro-differential equations and the main regularity results that follow from it. Section 3 presents our discretization and proves convergence to the viscosity solution. We recall the discrete ABP estimate for finite element methods in Section 4. The discrete ABP estimate allows us to show stability and rates of convergence for our discretization. We discuss some implementation details in Section 5, where we present a convergent iterative scheme.

We will follow standard notation concerning differential operators and function spaces. The relation $A \lesssim B$ means that there is a nonessential constant C such that $A \leq CB$. The value of this constant might change at each occurrence. By $A \gtrsim B$ we mean $B \lesssim A$.

2. Approximation of elliptic problems by integro-differential operators

Let us review the approximation, proposed by Caffarelli and Silvestre [8], of fully nonlinear elliptic PDEs by an integro-differential equation and the convergent finite element scheme of [32] for the approximation of elliptic problems in nondivergence form based on this idea. In this work, we exploit this approximation to discretize the fully nonlinear problem (1.1).

2.1. Integral operator

Let us begin by fixing some notation. Recall that, owing to (1.2), for all $x \in \Omega$ the matrices $A^{\alpha,\beta}(x)$ are uniformly positive definite. Thus, we can define

$$M^{\alpha,\beta}(x) := \left(A^{\alpha,\beta}(x) - \frac{\lambda}{2} I \right)^{1/2}.$$

Let $\varepsilon > 0$ and φ be a radially symmetric function with compact support in the unit ball B_1 and such that $\int |x|^2 \varphi(x) \,\mathrm{d}x = d$. Let $Q = \sqrt{2/\lambda}$ and notice that, again because of (1.2), for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ we have that if

$$\varphi^{\alpha,\beta}_{x,\varepsilon} : y \mapsto \varphi\left(\frac{1}{\varepsilon} M^{\alpha,\beta}(x)^{-1} y \right),$$

then $\operatorname{supp} \varphi^{\alpha,\beta}_{x,\varepsilon} \subset B_{Q\varepsilon}$. For this reason, we define

$$(\partial\Omega)_\varepsilon = \left\{ x \in \mathbb{R}^d \setminus \Omega : \operatorname{dist}(x, \partial\Omega) < Q\varepsilon \right\}. \tag{2.1}$$

a $Q\varepsilon$-neighborhood of $\partial\Omega$ and

$$\Omega_\varepsilon := \Omega \cup (\partial\Omega)_\varepsilon. \tag{2.2}$$

Given a function $w \in \mathrm{C}^0(\overline{\Omega_\varepsilon})$ we define, for each α and β the function $I_\varepsilon^{\alpha,\beta}[w] : \Omega \to \mathbb{R}$ by

$$I_\varepsilon^{\alpha,\beta}[w](x) = \frac{1}{\varepsilon^{d+2} \det M^{\alpha,\beta}(x)} \int_{\mathbb{R}^d} \mathfrak{d}w(x,y) \varphi_{x,\varepsilon}^{\alpha,\beta}(y) \,\mathrm{d}y, \tag{2.3}$$

where we denoted by $\mathfrak{d}w(x,y)$ the second order difference of the function w at the point x in the direction y which is given in (1.4). Notice that, for $x \in \Omega$ and $y \in \operatorname{supp} \varphi_{x,\varepsilon}^{\alpha,\beta}$, we have that $x \pm y \in \overline{\Omega_\varepsilon}$ so that (2.3) is well defined. It is easy to check that if w is a quadratic polynomial, then we have $\mathfrak{d}w(x,y) = \mathrm{D}^2 w(x) : (y \otimes y)$ for all $x \in \Omega$.

The integral operator $I_\varepsilon^{\alpha,\beta}[w]$ is a consistent approximation of the differential operator $\left(A^{\alpha,\beta} - \frac{\lambda}{2} I\right) : \mathrm{D}^2 w$ in the following sense [8,32].

Lemma 2.1 (Approximation properties of $I_\varepsilon^{\alpha,\beta}[\cdot]$)**.** *Let $I_\varepsilon^{\alpha,\beta}[\cdot]$ be the integral operator defined by (2.3) and assume that $\varepsilon > 0$ and $x \in \Omega$. Then,*

(1) *If $p \in \mathbb{P}_2$,* i.e. *it is a quadratic polynomial, then*

$$I_\varepsilon^{\alpha,\beta}[p](x) = \left(A^{\alpha,\beta}(x) - \frac{\lambda}{2} I \right) : \mathrm{D}^2 p(x).$$

(2) *Let $\varepsilon_0 > 0$ be fixed. If $w \in \mathrm{C}^2(\Omega_{\varepsilon_0})$ then, as $\varepsilon \to 0$, we have that*

$$I_\varepsilon^{\alpha,\beta}[w](x) \to \left(A^{\alpha,\beta}(x) - \frac{\lambda}{2} I \right) : \mathrm{D}^2 w(x).$$

Let us, from now on, assume that the matrices $M^{\alpha,\beta}$ have a uniform modulus of continuity ϖ. In other words, there is a nondecreasing function ϖ such that $\lim_{t \downarrow 0} \varpi(t) = \varpi(0) = 0$ and

$$\sup_{x_1, x_2 \in \bar\Omega, |x_1 - x_2| \le t} \| M^{\alpha,\beta}(x_1) - M^{\alpha,\beta}(x_2) \| \le \varpi(t), \quad \forall \alpha \in \mathcal{A}, \beta \in \mathcal{B}. \tag{2.4}$$

Under this assumption, we can show that, for every $\varepsilon > 0$, the integral operator maps continuously $\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$ into $\mathrm{C}^0(\bar\Omega)$.

Lemma 2.2 (Continuity of $I_\varepsilon^{\alpha,\beta}[\cdot]$)**.** *If $w \in \mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$, then, for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ and all $x, z \in \bar\Omega$*

$$\left| I_\varepsilon^{\alpha,\beta}[w](x) - I_\varepsilon^{\alpha,\beta}[w](z) \right| \lesssim \left(\frac{|x-z|}{\varepsilon^2} + \frac{\varpi(|x-z|)}{\varepsilon} \right) |w|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})},$$

where the hidden constant is independent of α, β, x, z, ε and w.

Proof. By definition of integral operator (2.3) and a change of variable $y := \varepsilon M^{\alpha,\beta}(x)\tilde{y}$, we obtain

$$I_\varepsilon^{\alpha,\beta}[w](x) = \frac{1}{\varepsilon^2} \int_{B_1} \mathfrak{d}w(x, \varepsilon M^{\alpha,\beta}(x) y) \varphi(y) \,\mathrm{d}y.$$

By the radial symmetry of φ

$$I_\varepsilon^{\alpha,\beta}[w](x) = \frac{2}{\varepsilon^2} \int_{B_1} \left(w(x + \varepsilon M^{\alpha,\beta}(x) y) - w(x) \right) \varphi(y) \,\mathrm{d}y.$$

Thus,

$$\left| I_\varepsilon^{\alpha,\beta}[w](x) - I_\varepsilon^{\alpha,\beta}[w](z) \right| \le \mathrm{I} + \mathrm{II},$$

with

$$
\begin{aligned}
\mathrm{I} &= \frac{2}{\varepsilon^2} \int_{B_1} \left| w(x+\varepsilon M^{\alpha,\beta}(x)y) - w(z+\varepsilon M^{\alpha,\beta}(z)y) \right| \varphi(y)\,\mathrm{d}y, \\
\mathrm{II} &= \frac{2}{\varepsilon^2} \int_{B_1} |w(x)-w(z)|\,\varphi(y)\,\mathrm{d}y.
\end{aligned}
$$

Evidently,

$$
\mathrm{II} \lesssim \frac{|x-z|}{\varepsilon^2} |w|_{\mathrm{C}^{0,1}(\bar\Omega)}.
$$

It remains then to estimate the first term. To do so we, again, use that $w \in \mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$ to obtain

$$
\mathrm{I} \lesssim \frac{|w|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})}}{\varepsilon^2} \sup_{y\in B_1} \left| x - z + \varepsilon \left(M^{\alpha,\beta}(x) - M^{\alpha,\beta}(z) \right) y \right|.
$$

an application of the triangle inequality, together with (2.4) imply

$$
\mathrm{I} \lesssim \left(\frac{|x-z|}{\varepsilon^2} + \frac{\varpi(|x-z|)}{\varepsilon} \right) |w|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})},
$$

where the hidden constant is uniform in α and β. Gathering the obtained bounds for I and II allows us to conclude. ☐

In [8] Caffarelli and Silvestre proposed an approximation of the Isaacs equation (1.1) by the following integro-differential problem

$$
\begin{cases}
\mathfrak{F}^\varepsilon[u^\varepsilon](x) := \dfrac{\lambda}{2}\Delta u^\varepsilon(x) + \inf\limits_{\alpha\in\mathcal{A}} \sup\limits_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u^\varepsilon](x) = \mathsf{f}(x) & \text{in } \Omega, \\
u^\varepsilon = 0 & \text{in } (\partial\Omega)_\varepsilon.
\end{cases}
\tag{2.5}
$$

Notice that, for $x \in \Omega$, the definition of $\mathfrak{F}^\varepsilon[u^\varepsilon](x)$ requires values of u^ε in $\overline{\Omega_\varepsilon}$. Thus in (2.5) the boundary condition has been replaced by a so-called volume constraint over $(\partial\Omega)_\varepsilon$. To define this volume constraint, the function that defines the boundary condition in (1.1) (in this case $g \equiv 0$) has been extended in a way that its extension belongs to $\mathrm{C}^{0,1}((\partial\Omega)_\varepsilon)$.

The approximation property of the integral operator $I_\varepsilon^{\alpha,\beta}[\cdot]$ given in Lemma 2.1, allows us to relate u^ε and u, which solve (2.5) and (1.1), respectively, as follows [8].

Proposition 2.3 (Approximation of the Isaacs equation)**.** *Let $\Omega \subset \mathbb{R}^d$ be convex and bounded, and $\mathsf{f} \in \mathrm{C}^{0,s}(\bar\Omega)$ for some $s \in [0,1)$. There exists a unique function $u^\varepsilon \in \mathrm{C}^{2,s}(\Omega)$ that solves (2.5). Moreover, if $\mathsf{f} \in \mathrm{C}^{0,1}(\bar\Omega)$, this solution satisfies*

$$
\|u^\varepsilon\|_{\mathrm{C}^{1,s}(U)} + \|u^\varepsilon\|_{\mathrm{C}^{0,1}(\bar\Omega)} \lesssim \|u^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} + \|\mathsf{f}\|_{\mathrm{L}^\infty(\Omega)},
$$

where U is any open set compactly contained in Ω and the hidden constant depends on the distance between U and $\partial\Omega$. In addition, there is a $\sigma > 0$ such that

$$
\|\mathsf{u} - u^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim \varepsilon^\sigma \|\mathsf{f}\|_{\mathrm{C}^{0,1}(\bar\Omega)},
\tag{2.6}
$$

where u is the (unique) viscosity solution to (1.1).

An important consequence of Proposition 2.3 is that, in the case when the coefficient matrices $A^{\alpha,\beta}$ are independent of x, we have the following regularity result for the function u, see Theorem 4.8 of [8] and the comments at the beginning of Section 4, for the $\mathrm{C}^{1,s}$ estimate and Theorem 5.2 and the comment at the beginning of Section 5 for the $\mathrm{C}^{0,1}$ estimate up to the boundary.

Corollary 2.4 (Regularity of u)**.** *Assume that, for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ $A^{\alpha,\beta}(x) = A^{\alpha,\beta}$, that Ω is convex and that $\mathsf{f} \in \mathrm{C}^{0,1}(\bar{\Omega})$. If u is the viscosity solution of (1.1) then there is a $s > 0$ such that $\mathsf{u} \in \mathrm{C}^{1,s}(\Omega) \cap \mathrm{C}^{0,1}(\bar{\Omega})$.*

The importance of this result lies in the fact that $\mathfrak{F}$ is not convex nor concave and, for that reason, the maximal regularity we can assert for u is $\mathrm{C}^{1,s}(\Omega)$, for some $s > 0$. We refer the reader, for instance, to the works by Nadirashvili and Vlăduţ [28–30], who have constructed viscosity solutions to nonconvex fully nonlinear elliptic equations whose Hessian is not bounded. This is in sharp contrast with, for instance, the uniformly elliptic Hamilton Jacobi Bellman equation where, under suitable assumptions on Ω and the “coefficients” of the equation, it can be shown that the solution belongs to $\mathrm{C}^{2,s}(\Omega)$ for some $s > 0$. Let us, finally, comment that the rate of convergence given in (2.6) cannot be improved; see the last paragraph of [8].

In this paper, based on the idea of an integro-differential approximation, we propose a finite element method for Isaacs equation (1.1).

2.2. Numerical quadrature to compute $I_\varepsilon^{\alpha,\beta}[\cdot]$

The approximation properties of the integral operator in Lemmas 2.1 and 2.2 are essential to ensure the consistency of the method. Let us briefly discuss here the use of numerical quadrature rules that preserve these properties. For each fixed $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, we aim to design a quadrature and choice of φ to approximate $I_\varepsilon^{\alpha,\beta}[w]$. The change of variable $y = \varepsilon M^{\alpha,\beta}(x)\zeta$ reveals that we must approximate

$$
\begin{aligned}
I_\varepsilon^{\alpha,\beta}[w](x) &= \frac{1}{\varepsilon^{d+2} \det M^{\alpha,\beta}(x)} \int_{\mathbb{R}^d} \mathfrak{d}w(x,y)\varphi_{x,\varepsilon}^{\alpha,\beta}(y)\,\mathrm{d}y \\
&= \frac{1}{\varepsilon^2} \int_{\mathbb{R}^d} \mathfrak{d}w(x, \varepsilon M(x)^{\alpha,\beta}\zeta)\varphi(\zeta)\,\mathrm{d}\zeta \approx \frac{1}{\varepsilon^2} \sum_{j=1}^{J} \omega_j \mathfrak{d}w(x, \varepsilon M^{\alpha,\beta}(x) q_j)\varphi(q_j).
\end{aligned}
$$

This shows that, to preserve the consistency and convergence properties of the integral operator, we must require the quadrature to satisfy the following three conditions:

- For all $j = 1, \ldots, J$, the quadrature points q_j must satisfy $q_j \in \operatorname{supp}(\varphi)$.
- The weights must be positive: $\omega_j > 0$ for all $j = 1, \ldots, J$.
- The quadrature must be exact for all quadratic polynomials, *i.e.* for a given kernel function φ with

$$
\int_{\mathbb{R}^d} |y|^2 \varphi(y)\,\mathrm{d}y = d
$$

we must have that, if $p \in \mathbb{P}_2$, then

$$
\sum_{j=1}^{J} \omega_j \mathfrak{d}p(x, \varepsilon M^{\alpha,\beta}(x) q_j)\varphi(q_j) = A^{\alpha,\beta}(x) : \mathrm{D}^2 p(x).
$$

The construction of quadrature formulas of this form which are of maximal degree of precision can be found in classical references like [27], see also [4,5,33]. For instance, for dimension $d = 2$, given $\varphi(y) = 4/\pi$ in the unit ball and $\varphi(y) = 0$ outside the unit ball, reference [36] shows that the following collection of weights and nodes has the requisite properties:

$$
\omega_j = \frac{\pi}{6}, \quad \rho_j = \frac{\sqrt{2}}{2}, \quad \theta_j = \frac{j\pi}{3}, \quad q_j = (\rho_j, \theta_j) \qquad j = 1, \ldots, 6,
$$

where the weights $q_j = (\rho_j, \theta_j)$ are given in polar coordinates. Indeed, since $\varphi(q_j) = 4/\pi$, this formula is exact for quadratic polynomials and it preserves the consistency in Lemmas 2.1 and 2.2.

3. Finite element discretization and convergence

Here we describe the scheme we use to approximate (1.1) and show that it converges to the viscosity solution. To keep technicalities to a minimum, in what follows we assume that Ω is a convex polytope.

3.1. Description of the scheme

Let $\{\mathscr{T}_h\}_{h>0}$ be a quasi-uniform, in the classical sense for finite elements [7, 11, 13], family of triangulations of size h of the domain Ω. We denote by $\mathscr{N}_h$ and $\mathscr{N}_h^\partial$ the collection of interior and boundary nodes of $\mathscr{T}_h$, respectively. We define the finite element space

$$\mathbb{V}_h := \left\{ v \in \mathrm{C}^0(\bar{\Omega}) : v_{|T} \in \mathbb{P}_1, \quad \forall T \in \mathscr{T}_h \right\},$$

where we denoted by $\mathbb{P}_1$ the space of polynomials of degree one.

We also need to introduce a triangulation of Ω_ε. Since this may not be a polytope we proceed in a standard way, namely, we introduce a triangulation $\mathscr{T}_h^\varepsilon$ of the polytopal domain $\Omega_\varepsilon^{\mathscr{T}} \subset \Omega_\varepsilon$ in such a way that:

- On Ω it coincides with $\mathscr{T}_h$.
- If we denote by $\mathscr{N}_h^\varepsilon$ the nodes of $\mathscr{T}_h^\varepsilon$ that are not in $\mathscr{N}_h$, then $\mathscr{N}_h^\varepsilon \cap \partial\Omega_\varepsilon^{\mathscr{T}} \subset \partial\Omega_\varepsilon$.
- $|\Omega_\varepsilon \setminus \Omega_\varepsilon^{\mathscr{T}}| \lesssim h^2$.

Notice also that $\mathscr{N}_h^\partial \subset \mathscr{N}_h^\varepsilon$. We then define

$$\mathbb{V}_h^\varepsilon := \left\{ v \in \mathrm{C}^0(\overline{\Omega_\varepsilon^{\mathscr{T}}}) : v_{|T} \in \mathbb{P}_1 \quad \forall T \in \mathscr{T}_h^\varepsilon \right\},$$
$$\mathbb{V}_h^0 := \left\{ v \in \mathbb{V}_h^\varepsilon : v(z) = 0, \quad \forall z \in \mathscr{N}_h^\varepsilon \right\},$$

and, as Ω and Ω_ε are convex, we have that $\Omega \subset \Omega_\varepsilon^{\mathscr{T}} \subset \Omega_\varepsilon$ and we can extend continuously functions in these spaces from $\Omega_\varepsilon^{\mathscr{T}}$ to Ω_ε by a constant in the direction normal to $\partial\Omega_\varepsilon^{\mathscr{T}}$.

We denote by $\{\phi_z\}_{z\in\mathscr{N}_h\cup\mathscr{N}_h^\varepsilon}$ the Lagrange nodal basis of $\mathbb{V}_h^\varepsilon$. Notice that, by restriction, we can consider $\mathbb{V}_h^\varepsilon$ to be continuously embedded in $\mathbb{V}_h$.

Below we will require $\mathscr{T}_h$ to be weakly acute, *i.e.* for all $z_1, z_2 \in \mathscr{N}_h$ with $z_1 \neq z_2$ we must have

$$\int_\Omega \nabla\phi_{z_1}(x) \cdot \nabla\phi_{z_2}(x)\,\mathrm{d}x \leq 0. \tag{3.1}$$

It is known, see Section III.20 of [10] and Lemma 3.42 of [31], that this is a sufficient condition for the discrete Laplacian, defined below, to be a monotone operator.

Let $z \in \mathscr{N}_h$, we set

$$\omega_z = \int_\Omega \phi_z(x)\,\mathrm{d}x.$$

For $w_h \in \mathbb{V}_h$ and $z \in \mathscr{N}_h$, we define the discrete Laplacian of w_h at z by

$$\Delta_h w_h(z) = -\frac{1}{\omega_z}\int_\Omega \nabla w_h(x)\cdot\nabla\phi_z(x)\,\mathrm{d}x. \tag{3.2}$$

Finally, for $z \in \mathscr{N}_h$, we set

$$f_z = \frac{1}{\omega_z}\int_\Omega \mathsf{f}(x)\phi_z(x)\,\mathrm{d}x.$$

With this notation at hand, we now define our scheme. We seek a function $u_h^\varepsilon \in \mathbb{V}_h^0$ such that, for every $z \in \mathscr{N}_h$, satisfies

$$\mathfrak{F}_h^\varepsilon[u_h^\varepsilon](z) := \frac{\lambda}{2}\Delta_h u_h^\varepsilon(z) + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z) = f_z. \tag{3.3}$$

3.2. Galerkin projection

In the analysis that follows we will make repeated use of the Galerkin projection and its properties. The Galerkin projection is the map

$$\mathcal{G}_h : \left\{ w \in \mathrm{L}^1(\Omega_\varepsilon) : w_{|\Omega} \in \mathrm{W}^1_1(\Omega),\ w_{|\overline{\Omega_\varepsilon}\setminus\Omega} \in \mathrm{C}^0(\overline{\Omega_\varepsilon}\setminus\Omega) \right\} \to \mathbb{V}^\varepsilon_h$$

that verifies $\mathcal{G}_h w(z) = w(z)$ for $z \in \mathscr{N}^\varepsilon_h$ and

$$\int_\Omega \nabla \mathcal{G}_h w(x) \cdot \nabla v_h(x)\,\mathrm{d}x = \int_\Omega \nabla w(x) \cdot \nabla v_h(x)\,\mathrm{d}x \quad \forall v_h \in \mathbb{V}^0_h. \tag{3.4}$$

Notice that, if $w_{|\Omega} \in \mathrm{C}^2(\Omega)$, setting $v_h = \phi_z$ with $z \in \mathscr{N}_h$ in (3.4) and integrating by parts we get

$$\Delta_h \mathcal{G}_h w(z) = \frac{1}{\omega_z} \int_\Omega \Delta w(x) \phi_z(x)\,\mathrm{d}x,$$

that is, the discrete Laplacian of the Galerkin projection at a node is a weighted average of the Laplacian of the original function around this node.

Combining the well known near optimal approximation properties [25,34], in the L^∞-norm, of the Galerkin projection with a standard interpolation estimate over $\overline{\Omega_\varepsilon}\setminus\Omega$ we conclude that if $w \in \mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$, then we have

$$\|w - \mathcal{G}_h w\|_{\mathrm{L}^\infty(\Omega_\varepsilon)} \lesssim h|\log h| \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})}. \tag{3.5}$$

Owing to (3.5), for every $z \in \mathscr{N}_h$ we obtain

$$|\mathfrak{d}\mathcal{G}_h w(z,y) - \mathfrak{d} w(z,y)| \lesssim h|\log h| \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})}.$$

By definition (2.3) of the integral operator $I^{\alpha,\beta}_\varepsilon[\cdot]$, we deduce that for all nodes $z \in \mathscr{N}_h$ we have

$$\left| I^{\alpha,\beta}_\varepsilon[\mathcal{G}_h w](z) - I^{\alpha,\beta}_\varepsilon[w](z) \right| \lesssim \frac{h}{\varepsilon^2} |\log h| \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})} \tag{3.6}$$

uniformly in $\mathcal{A}$ and $\mathcal{B}$.

3.3. Existence and uniqueness of the numerical solution

Based on the notions of *monotonicity* and *weak consistency* advanced in [2,19,32], we show existence and uniqueness of solutions to (3.3). Moreover, the family $\{u^\varepsilon_h\}_{h>0,\varepsilon>0}$ of solutions to (3.3) converges to u, the unique viscosity solution of (1.1), provided the mesh size h and ε satisfy a suitable relation.

We begin with two elementary properties, namely monotonicity and continuity, of the inf-sup operator.

Lemma 3.1 (Monotonicity and continuity of inf-sup)**.** *Let $\{X^{\alpha,\beta} : \alpha \in \mathcal{A},\ \beta \in \mathcal{B}\} \subset \mathbb{R}$ and $\{Y^{\alpha,\beta} : \alpha \in \mathcal{A},\ \beta \in \mathcal{B}\} \subset \mathbb{R}$ be two families parametrized by two index sets $\mathcal{A}$ and $\mathcal{B}$. If for every fixed $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ we have that $X^{\alpha,\beta} \le Y^{\alpha,\beta}$, then*

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} \le \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta}. \tag{3.7}$$

Moreover, if there is a constant C such that for all α and β we have $\left|X^{\alpha,\beta} - Y^{\alpha,\beta}\right| \le C$, then

$$\left| \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} - \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta} \right| \le C. \tag{3.8}$$

Proof. If $X^{\alpha,\beta} \leq Y^{\alpha,\beta}$ for all fixed α and β, taking supremum on $Y^{\alpha,\beta}$ with respect to β first, we have

$$X^{\alpha,\beta} \leq \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta} \quad \Longrightarrow \quad \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} \leq \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta}.$$

Taking infimum on $\sup_{\beta\in\mathcal{B}} X^{\alpha,\beta}$ with respect to α, we get

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} \leq \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta} \quad \Longrightarrow \quad \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} \leq \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta}.$$

This proves the first inequality (3.7).

To prove the second inequality, we only need to note that if $X^{\alpha,\beta} \leq Y^{\alpha,\beta} + C$ for all α and β, then by (3.7)

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} \leq \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta} + C.$$

Similarly, if $Y^{\alpha,\beta} \leq X^{\alpha,\beta} + C$ for all α and β, then

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} Y^{\alpha,\beta} \leq \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} X^{\alpha,\beta} + C.$$

This proves the second inequality (3.8). □

We now establish the monotonicity and weak consistency of the proposed scheme.

Lemma 3.2 (Monotonicity of the numerical scheme)**.** *Assume that for every discretization parameter $h > 0$ the mesh $\mathscr{T}_h$ is weakly acute. Then the family of operators $\mathfrak{F}_h^\varepsilon$ is monotone in the sense that if $v_h, w_h \in \mathbb{V}_h^\varepsilon$ with $v_h \leq w_h$ in $\mathscr{T}_h^\varepsilon$ and equality holds at some node $z \in \mathscr{N}_h$, then*

$$\mathfrak{F}_h^\varepsilon[v_h](z) \leq \mathfrak{F}_h^\varepsilon[w_h](z).$$

Proof. Let $v_h, w_h \in \mathbb{V}_h^\varepsilon$ be as indicated. Since $\mathscr{T}_h$ is weakly acute, the operator Δ_h is monotone, *i.e.*

$$\Delta_h v_h(z) \leq \Delta_h w_h(z).$$

It remains then to deal with the inf-sup over the approximating integral operators. Since, over $\mathscr{T}_h^\varepsilon$, we have that $v_h \leq w_h$ and equality holds at $z \in \mathscr{N}_h$, we can conclude that, whenever $y \in B_{Q\varepsilon}$, the second order differences verify $\mathfrak{d} v_h(z,y) \leq \mathfrak{d} w_h(z,y)$. Therefore, for fixed $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, we get $I_\varepsilon^{\alpha,\beta}[v_h](z) \leq I_\varepsilon^{\alpha,\beta}[w_h](z)$.

Using the monotonicity of the inf-sup operator (3.7), we obtain

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[v_h](z) \leq \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[w_h](z).$$

Combining both inequalities, we then deduce that

$$\mathfrak{F}_h^\varepsilon[v_h](z) \leq \mathfrak{F}_h^\varepsilon[w_h](z).$$

This completes the proof. □

As a consequence, we obtain a discrete maximum principle for our numerical scheme.

Corollary 3.3 (Discrete maximum principle)**.** *Assume that $\mathscr{T}_h$ is weakly acute. If $v_h, w_h \in \mathbb{V}_h^\varepsilon$ are such that $v_h \leq w_h$ in $\mathscr{N}_h^\varepsilon$ and, for all $z \in \mathscr{N}_h$, we have*

$$\mathfrak{F}_h^\varepsilon[v_h](z) \geq \mathfrak{F}_h^\varepsilon[w_h](z),$$

then $v_h \leq w_h$ in $\mathscr{T}_h^\varepsilon$.

Proof. We argue by contradiction by assuming that $v_h - w_h$ attains a positive value. We then necessarily have that there is $z_0 \in \mathscr{N}_h$ for which $v_h(z_0) - w_h(z_0) > 0$.

We now set $p(x) = |x|^2 - C$, where the constant C is so large that, for all $x \in \overline{\Omega_\varepsilon}$, we have $p(x) \le 0$. Let $p_h \in \mathbb{V}_h^\varepsilon$ be the Lagrange interpolant of p and, for $\delta > 0$, we define $e_h = v_h - w_h + \delta p_h \in \mathbb{V}_h^\varepsilon$. By construction we have that $e_h \le 0$ in $\mathscr{N}_h^\varepsilon$. In addition, for $\delta > 0$ sufficiently small, $e_h(z_0) > 0$. Consequently, there is $z_1 \in \mathscr{N}_h$ where e_h attains a, necessarily positive, maximum. At this node we must have, for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ that

$$I_\varepsilon^{\alpha,\beta}[e_h](z_1) \le 0.$$

The weak acuteness of the mesh $\mathscr{T}_h$ allows us to conclude also that

$$\Delta_h e_h(z_1) \le 0.$$

Adding these two inequalities yields that, for all $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ we have

$$\frac{\lambda}{2}\Delta_h(v_h + \delta p_h)(z_1) + I_\varepsilon^{\alpha,\beta}[v_h + \delta p_h](z_1) \le \frac{\lambda}{2}\Delta_h w_h(z_1) + I_\varepsilon^{\alpha,\beta}[w_h](z_1). \tag{3.9}$$

Notice that since p is convex and p_h is its piecewise linear interpolant we have that, for all $y \in B_{Q\varepsilon}$, $\mathfrak{d}p_h(z_1, y) \ge \mathfrak{d}p(z_1, y)$. This, in conjunction with Lemma 2.1, implies that

$$I_\varepsilon^{\alpha,\beta}[p_h](z_1) \ge I_\varepsilon^{\alpha,\beta}[p](z_1) = \left(A^{\alpha,\beta}(z_1) - \frac{\lambda}{2}I\right) : \mathrm{D}^2 p(z_1) \ge \lambda. \tag{3.10}$$

Observe now that the linear function $\psi(x) = -2x{\cdot}z_1 + |z_1|^2 \in \mathbb{V}_h^\varepsilon$ is such that $p(x) + \psi(x) = |x - z_1|^2 - C$ and $\Delta_h \psi(z_1) = 0$. Consequently, by definition (3.2), we have that

$$\begin{aligned} -\omega_{z_1}\Delta_h p_h(z_1) = -\omega_{z_1}\Delta_h(p_h + \psi)(z_1) &= \int_\Omega \nabla(p_h + \psi)(x){\cdot}\nabla\phi_{z_1}(x)\,\mathrm{d}x \\ &= \sum_{z \in \mathscr{N}_h \cup \mathscr{N}_h^\partial} (p + \psi)(z) \int_\Omega \nabla\phi_z(x){\cdot}\nabla\phi_{z_1}(x)\,\mathrm{d}x \\ &= \sum_{z \in \mathscr{N}_h \cup \mathscr{N}_h^\partial \setminus \{z_1\}} |z - z_1|^2 \int_\Omega \nabla\phi_z(x){\cdot}\nabla\phi_{z_1}(x)\,\mathrm{d}x, \end{aligned}$$

where, to get rid of the constant term $-C$, we used the partition of unity property of the Lagrange basis functions $\{\phi_z\}_{z \in \mathscr{N}_h \cup \mathscr{N}_h^\partial}$. Finally, we recall that the weak acuteness of $\mathscr{T}_h$ implies that, for $z \ne z_1$, $\int_\Omega \nabla\phi_z(x){\cdot}\nabla\phi_{z_1}(x)\,\mathrm{d}x \le 0$ to obtain

$$\Delta_h p_h(z_1) \ge 0. \tag{3.11}$$

Inserting (3.10) and (3.11) in (3.9) yields

$$\frac{\lambda}{2}\Delta_h v_h(z_1) + I_\varepsilon^{\alpha,\beta}[v_h](z_1) + \lambda\delta \le \frac{\lambda}{2}\Delta_h w_h(z_1) + I_\varepsilon^{\alpha,\beta}[w_h](z_1).$$

Applying Lemma 3.1 to this inequality we deduce that

$$\mathfrak{F}_h^\varepsilon[v_h](z_1) < \mathfrak{F}_h^\varepsilon[v_h](z_1) + \lambda\delta \le \mathfrak{F}_h^\varepsilon[w_h](z_1),$$

which is a contradiction. $\square$

The analysis of our scheme is based on a careful estimation of the consistency error, which we define as the difference between the result of applying the operator to the Galerkin projection of the solution to (2.5) and the approximate right hand side, *i.e.* $\mathfrak{F}_h^\varepsilon[\mathcal{G}_h u^\varepsilon](z) - f_z$. The critical step in this analysis is to understand what happens with the integral operators. With this in mind, and for future use, for a function $w \in \mathrm{C}^0(\overline{\Omega_\varepsilon})$ and $z \in \mathscr{N}_h$ we denote

$$\mathcal{R}_{h,\varepsilon}[w](z) = \frac{1}{\omega_z}\int_\Omega \left[\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h w](z) - \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[w](x) \right] \phi_z(x)\,\mathrm{d}x. \tag{3.12}$$

The following result bounds $\mathcal{R}_{h,\varepsilon}[w](z)$ uniformly in $z \in \mathscr{N}_h$ when $w \in \mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$.

Lemma 3.4 (Galerkin projection *vs.* integral operator)**.** *Let $w \in \mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})$. If the coefficient matrices are such that (2.4) holds then, for every $z \in \mathscr{N}_h$, we have*

$$|\mathcal{R}_{h,\varepsilon}[w](z)| \lesssim \left(\frac{h}{\varepsilon^2}|\log h| + \frac{\varpi(h)}{\varepsilon} \right) \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})}, \tag{3.13}$$

where the hidden constant is independent of h, ε and w.

Proof. Let $z \in \mathscr{N}_h$ and consider, for $x \in \operatorname{supp}\phi_z$,

$$\begin{aligned} r_{h,\varepsilon}^{\alpha,\beta}[w](z) &= I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h w](z) - I_\varepsilon^{\alpha,\beta}[w](x) \\ &= \left(I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h w](z) - I_\varepsilon^{\alpha,\beta}[w](z) \right) + \left(I_\varepsilon^{\alpha,\beta}[w](z) - I_\varepsilon^{\alpha,\beta}[w](x) \right) \\ &= r_{1,h,\varepsilon}^{\alpha,\beta}[w](z) + r_{2,h,\varepsilon}^{\alpha,\beta}[w](z). \end{aligned}$$

Estimate (3.6) immediately yields

$$|r_{1,h,\varepsilon}^{\alpha,\beta}[w](z)| \lesssim \frac{h}{\varepsilon^2}|\log h| \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})}.$$

On the other hand, Lemma 2.2 implies

$$|r_{2,h,\varepsilon}^{\alpha,\beta}[w](z)| \lesssim \left(\frac{h}{\varepsilon^2} + \frac{\varpi(h)}{\varepsilon} \right) \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})},$$

where we used that $|x - z| \lesssim h$.

We now invoke Lemma 3.1 with $X^{\alpha,\beta} = I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h w](z)$ and $Y^{\alpha,\beta} = I_\varepsilon^{\alpha,\beta}[w](x)$ to conclude, using (3.8), that

$$\left| \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h w](z) - \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[w](x) \right| \lesssim \left(\frac{h}{\varepsilon^2}|\log h| + \frac{\varpi(h)}{\varepsilon} \right) \|w\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})},$$

which immediately implies (3.13). □

The monotonicity property given in Lemma 3.2 ensures that, for every $\varepsilon > 0$ and $h > 0$, the scheme (3.3) has a unique solution. The main idea behind the the existence and uniqueness of the numerical solution is a discrete variant of Perron's method ([15], Sect. 6.1).

Theorem 3.5 (Existence and uniqueness)**.** *Let the family of meshes $\{\mathscr{T}_h\}_{h>0}$ be weakly acute. For every $h > 0$ and $\varepsilon > 0$ the finite element scheme (3.3) has a unique solution.*

Proof. We begin by proving existence in several steps. First, we define the set of discrete super-solutions

$$\mathbb{S}_h = \{v_h \in \mathbb{V}_h^\varepsilon : \mathfrak{F}_h^\varepsilon[v_h](z) \le f_z \ \forall z \in \mathscr{N}_h,\ v_h(z) \ge 0 \ \forall z \in \mathscr{N}_h^\varepsilon\}.$$

We claim that the set $\mathbb{S}_h$ is nonempty. Set

$$\delta_{h,\varepsilon} = C\left(\frac{h}{\varepsilon^2}|\log h| + \frac{\varpi(h)}{\varepsilon}\right),$$

where $C > 0$ is a constant to be chosen later. Let $F = \min\{0, \min_{\bar\Omega} \mathsf{f}\} - \delta_{h,\varepsilon}$ and let $R > 0$ be so large that, for some $\xi \in \Omega$ the ball $\xi + B_R$ contains $\overline{\Omega_\varepsilon}$. Define $p(x) = \frac{1}{2}\lambda^{-1}F(|x-\xi|^2 - R^2)$, which is a nonnegative quadratic polynomial in $\xi + B_R$, $p(x) \ge 0$ on $(\partial\Omega)_\varepsilon$, and $\mathrm{D}^2 p = \lambda^{-1}FI \le 0$. Owing to the uniform ellipticity condition (1.2), for each α and β, we obtain

$$A^{\alpha,\beta}(x) : \mathrm{D}^2 p(x) \le \lambda(\lambda^{-1}F) \le \mathsf{f}(x) - \delta_{h,\varepsilon} \quad \Longrightarrow \quad \mathfrak{F}[p](x) \le \mathsf{f}(x) - \delta_{h,\varepsilon}.$$

We now claim that $\mathcal{G}_h p \in \mathbb{S}_h$. To see this, first notice that for every $z \in \mathscr{N}_h^\varepsilon$ we have $\mathcal{G}_h p(z) \ge 0$. Now, for $z \in \mathscr{N}_h$, consider

$$\begin{aligned}\mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z) &= \frac{\lambda}{2}\Delta_h \mathcal{G}_h p(z) + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h p](z) \\ &= \frac{1}{\omega_z}\int_\Omega \left(\frac{\lambda}{2}\Delta p(x) + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[p](x)\right)\phi_z(x)\,\mathrm{d}x \\ &+ \mathcal{R}_{h,\varepsilon}[p](z).\end{aligned}$$

Using the consistency of $I_\varepsilon^{\alpha,\beta}[\cdot]$ for quadratics, see Lemma 2.1, we conclude that

$$\mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z) = \frac{\int_\Omega \mathfrak{F}[p](x)\phi_z(x)\,\mathrm{d}x}{\int_\Omega \phi_z(x)\,\mathrm{d}x} + \mathcal{R}_{h,\varepsilon}[p](z) \le f_z - \delta_{h,\varepsilon} + \mathcal{R}_{h,\varepsilon}[p](z). \tag{3.14}$$

Lemma 3.4 and the fact that $\|p\|_{\mathrm{C}^{0,1}(\overline{\Omega_\varepsilon})} \lesssim \|\mathsf{f}\|_{\mathrm{L}^\infty(\Omega)}$, show that we can choose $C > 0$ sufficiently large so that $\mathcal{R}_{h,\varepsilon}[p](z) - \delta_{h,\varepsilon} \le 0$ for all $z \in \mathscr{N}_h$. Therefore,

$$\mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z) \le f_z \implies \mathcal{G}_h p \in \mathbb{S}_h.$$

Thus, $\mathbb{S}_h$ is nonempty.

We now show that the minimum between two super-solutions is a super-solution. Given $v_h, w_h \in \mathbb{S}_h$, we define $(v \wedge w)_h \in \mathbb{V}_h^\varepsilon$ by

$$(v \wedge w)_h(z) = \min\{v_h(z), w_h(z)\} \quad \forall z \in \mathscr{N}_h \cup \mathscr{N}_h^\varepsilon.$$

Since, for every $z \in \mathscr{N}_h^\varepsilon$, we have $v_h(z) \ge 0$ and $w_h(z) \ge 0$, then

$$(v \wedge w)_h(z) \ge 0, \quad \forall z \in \mathscr{N}_h^\varepsilon.$$

Notice that, if for $z_0 \in \mathscr{N}_h$ we have $(v \wedge w)_h(z_0) = w_h(z_0)$ then

$$(v \wedge w)_h \le w_h \qquad \text{and} \qquad (v \wedge w)_h(z_0) = w_h(z_0).$$

The monotonicity result of Lemma 3.2 implies then that

$$\mathfrak{F}_h^\varepsilon[(v \wedge w)_h](z_0) \le \mathfrak{F}_h^\varepsilon[w_h](z_0) \le f_{z_0}.$$

Since for every $z \in \mathscr{N}_h$ we either have that $(v \wedge w)_h(z) = v_h(z)$ or $(v \wedge w)_h(z) = w_h(z)$, we conclude that $(v \wedge w)_h \in \mathbb{S}_h$.

Finally, we show that the smallest super-solution is a solution. Let $u_h^\star \in \mathbb{V}_h^\varepsilon$ be defined by

$$u_h^\star(z) = \inf_{v_h \in \mathbb{S}_h} v_h(z) \quad \forall z \in \mathscr{N}_h \cup \mathscr{N}_h^\varepsilon.$$

The reasoning given above shows that $u_h^\star \in \mathbb{S}_h$. We claim that $u_h^\star$ is a solution, for if that is not the case, then either:

Case I: There is a node $z_0 \in \mathscr{N}_h^\varepsilon$ such that $u_h^\star(z_0) > 0$. Define, for $\delta > 0$, $v_h^\star = u_h^\star - \delta\phi_{z_0}$ and notice that for δ sufficiently small, we have $v_h^\star(z_0) \geq 0$. Consequently, this function verifies

$$v_h^\star(z) \leq u_h^\star(z), \quad \forall z \in \mathscr{N}_h \cup \mathscr{N}_h^\varepsilon, \qquad 0 \leq v_h^\star(z), \quad \forall z \in \mathscr{N}_h^\varepsilon. \tag{3.15}$$

In addition, for all $z' \in \mathscr{N}_h$, we have $v_h^\star(z') = u_h^\star(z')$. The monotonicity of the scheme shown in Lemma 3.2 implies then that

$$\mathfrak{F}_h^\varepsilon[v_h^\star](z') \leq \mathfrak{F}_h^\varepsilon[u_h^\star](z'), \quad \forall z' \in \mathscr{N}_h,$$

so that $v_h^\star \in \mathbb{S}_h$. However, this contradicts the fact that $u_h^\star$ is the smallest super-solution of (3.3). Consequently, $u_h^\star \in \mathbb{V}_h^0$.

Case II: There exists an interior node $z_0 \in \mathscr{N}_h$ such that

$$\mathfrak{F}_h^\varepsilon[u_h^\star](z_0) < f_{z_0}.$$

We, again, define $v_h^\star = u_h^\star - \delta\phi_{z_0}$. Obviously $v_h^\star(z_0) \leq u_h^\star(z_0)$ and, $v_h(z) = u_h(z)$ for all $z \neq z_0$. Invoking Lemma 3.2 we see that

$$\mathfrak{F}_h^\varepsilon[v_h^\star](z) \leq \mathfrak{F}_h^\varepsilon[u_h^\star](z), \quad \forall z \in \mathscr{N}_h \setminus \{z_0\}.$$

In addition, since ϕ_{z_0} attains its absolute strict maximum at z_0 we must have

$$\frac{\lambda}{2}\Delta_h \phi_{z_0}(z_0) + I_\varepsilon^{\alpha,\beta}[\phi_{z_0}](z_0) \leq 0.$$

This, in conjunction with Lemma 3.1 yields that, for δ sufficiently small,

$$\mathfrak{F}_h^\varepsilon[v_h^\star](z_0) \leq f_{z_0}$$

so that $v_h^\star \in \mathbb{S}_h$. In this case, we also obtain a contradiction. This proves the existence of a solution.

Uniqueness immediately follows from Corollary 3.3. □

Remark 3.6 (Alternative proof of existence)**.** An alternative proof of existence of solutions will be given in Theorem 5.2 *via* the convergence of Howard's algorithm.

3.4. The discrete Alexandrov Bakelman Pucci estimate

The next step towards proving convergence of the discrete solutions to the viscosity solution is to show that solutions of (3.3) are bounded uniformly with respect to h and ε. To achieve this, we will employ the discrete ABP maximum principle of Theorem 5.1 from [32], which will also be our main tool to derive a rate of convergence. As Lemma 3.2 shows, weak acuteness of the mesh is sufficient to ensure the monotonicity of the numerical scheme. However, the discrete ABP maximum principle requires a slightly stronger assumption on the mesh, which we state below. We say that the mesh $\mathscr{T}_h$ is weakly acute with respect to faces if for every face F and $z_1, z_2 \in \mathscr{N}_h$ with $z_1 \neq z_2$ we have

$$\int_{\omega_F} \nabla\phi_{z_1}(x) \cdot \nabla\phi_{z_2}(x)\,\mathrm{d}x \leq 0, \tag{3.16}$$

where $\omega_F = \cup\{K^\pm \in \mathscr{T}_h : F \subset K^\pm\}$. This condition must be compared with weak acuteness as expressed in (3.1). In two dimensions ($d = 2$), condition (3.16) is equivalent to (3.1) and is valid if and only if the sum of two angles opposite to a face is not greater than π. However, condition (3.16) is stronger than weak acuteness for $d > 2$. We refer the reader to Section 3.3 of [32] for further details and references on this condition.

To state the discrete ABP estimate, we also introduce the following notation. Let B_R be a ball of radius R which contains the domain Ω. We define the convex envelope of a function $v_h \in \mathbb{V}_h$ such that $v_h \geq 0$ over $\mathscr{N}_h^\partial$ by

$$\Gamma(v_h)(x) := \sup\left\{L(x) : L \in \mathbb{P}_1,\ L \leq -v_h^- \text{ in } B_R\right\}, \tag{3.17}$$

where we denote by v_h^- the negative part of v_h in Ω and $v_h^- := 0$ in $B_R \setminus \Omega$. We also define the *(lower) nodal contact set*

$$\mathcal{C}_h^-(v_h) = \{z \in \mathscr{N}_h : \Gamma(v_h)(z) = v_h(z)\}. \tag{3.18}$$

Lemma 3.7 (Discrete ABP estimate)**.** *Let the family of meshes $\{\mathscr{T}_h\}_{h>0}$ be quasi-uniform and satisfy* (3.16)*. If $v_h \in \mathbb{V}_h$ is such that $v_h \geq 0$ over $\mathscr{N}_h^\partial$ and satisfies*

$$\Delta_h v_h(z) \leq f_z \quad \forall z \in \mathscr{N}_h,$$

then

$$\sup_\Omega v_h^- \lesssim \left(\sum_{z\in\mathcal{C}^-(v_h)} \omega_z |f_z|^d\right)^{1/d},$$

where the hidden constant is independent of h.

This inequality gives us an estimate on the negative part of v_h, while a bound for its positive part can be derived in the same fashion by considering a concave envelope and the corresponding (upper) contact set. Combining these bounds yields that the L^∞-norm of v_h is controlled by a discrete L^d-like norm of $\{f_z\}_{z\in\mathscr{N}_h}$. We make this idea rigorous below.

Proposition 3.8 (Uniform boundedness)**.** *Let the family of meshes $\{\mathscr{T}_h\}_{h>0}$ be quasi-uniform and satisfy* (3.16)*. Then, the family $\{u_h^\varepsilon \in \mathbb{V}_h^0\}_{h>0}$ of solutions of* (3.3) *is uniformly bounded with respect to h and ε,* i.e.

$$\|u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim \|\mathsf{f}\|_{\mathrm{L}^d(\Omega)},$$

where the hidden constant is independent of h and ε.

Proof. Notice that, by construction $u_h^\varepsilon = 0$ on $\mathscr{N}_h^\partial$ and, consequently $(u_h^\varepsilon)^- \neq 0$ only in the interior of $\bar{\Omega}$.

Now, by definition of the convex envelope (3.17) we have that

$$u_h^\varepsilon(x) \geq \Gamma(u_h^\varepsilon)(x), \quad \forall x \in \Omega,$$

and that, if $z \in \mathcal{C}_h^-(u_h^\varepsilon)$ we have $u_h^\varepsilon(z) = \Gamma(u_h^\varepsilon)(z)$. Since $\Gamma(u_h^\varepsilon)(x)$ is convex, we obtain that, for every $z \in \mathcal{C}_h^-(u_h^\varepsilon)$ and $y \in B_{Q\varepsilon}$ we have $\mathfrak{d}u^\varepsilon(z,y) \geq \mathfrak{d}\Gamma(u_h^\varepsilon)(z,y) \geq 0$. Consequently,

$$I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z) \geq I_\varepsilon^{\alpha,\beta}[\Gamma(u_h^\varepsilon)](z) \geq 0, \quad \forall z \in \mathcal{C}_h^-(u_h^\varepsilon)$$

and

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z) \geq \inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\Gamma(u_h^\varepsilon)](z) \geq 0.$$

In conclusion, for every $z \in \mathcal{C}_h^-(u_h^\varepsilon)$, the scheme (3.3) reduces to

$$\frac{\lambda}{2}\Delta_h u_h^\varepsilon(z) \leq \mathfrak{F}_h^\varepsilon[u_h^\varepsilon](z) = f_z.$$

The previous inequality, together with the discrete ABP estimate of Lemma 3.7 then yields

$$\sup_{\Omega}(u_h^\varepsilon)^- \lesssim \left(\sum_{z\in\mathcal{C}_h^-(u_h^\varepsilon)} \omega_z |f_z|^d \right)^{1/d} \lesssim \|\mathsf{f}\|_{\mathrm{L}^d(\Omega)},$$

where the last inequality follows from the shape regularity of $\mathscr{T}_h$.

We have obtained a lower bound for u_h^ε. By considering the positive part, and the corresponding concave envelope, we can obtain an upper bound on u_h^ε and this yields its uniform boundedness. □

3.5. Convergence to the viscosity solution

Combining the discrete maximum principle of Corollary 3.3 and the consistency estimate encoded in Lemma 3.4 it is now possible, following the guidelines established in [2, 24], to prove convergence of our scheme (3.3). We begin by recalling the essence of the convergence results established in [2, 24]:

> If (1.1) has a strong comparison principle, then uniformly bounded solutions of a consistent, monotone, stable scheme converge to the viscosity solution of (1.1).

To be able to assert consistency, we must guarantee that the right hand side of (3.13) tends to zero as we refine the mesh and let $\varepsilon \to 0$. This imposes a relation between the parameter ε and the mesh size h of the form

$$\varepsilon = \mathcal{O}\left(\max\left\{(h|\log h|)^{1/2}, \varpi(h)\right\}\right), \qquad \text{and } h \to 0 \implies \varepsilon \to 0. \tag{3.19}$$

which from now on we assume. At this point the reason why we call our scheme a two-scale one becomes evident. The fine scale is given by the mesh size h and provides a scale for the discretization of functions. On the other hand, the coarse scale ε can be understood as a stencil width that, as is well known [20, 26], must be large enough to be able to guarantee monotonicity of a scheme. In this regard, our method bears resemblance to semi-Lagrangian schemes [12] in that functions are approximated with piecewise linears, but their derivatives are approximated at a coarser scale.

To establish convergence, we begin by constructing boundary barrier functions, which will be essential to show that the boundary conditions are attained in a classical sense.

Lemma 3.9 (Barrier functions)**.** *Assume that Ω is convex and $x_0 \in \partial\Omega$. Let h and ε be small enough and satisfy (3.19). With these assumptions, for every positive constant $E > 0$ there exist finite element functions $p_{x_0,h}^\pm \in \mathbb{V}_h^\varepsilon$ such that:*

(1) *for all $x \in (\partial\Omega)_\varepsilon$ we have*
$$p_{x_0,h}^+(x) \le 0 \le p_{x_0,h}^-(x), \tag{3.20}$$
(2) *for all $z \in \mathscr{N}_h$ they satisfy*
$$\pm\mathfrak{F}_h^\varepsilon[p_{x_0,h}^\pm](z) \ge \frac{\lambda}{2}E, \tag{3.21}$$
and,

(3) *for all $x \in x_0 + B_{Q\varepsilon} \cap \overline{\Omega_\varepsilon}$*
$$|p_{x_0,h}^\pm(x)| \lesssim EQ\varepsilon, \tag{3.22}$$
where the hidden constant depends only on the domain Ω.

Proof. Let $x_0^\star \in \partial\Omega_\varepsilon$ be the closest point in $\partial\Omega_\varepsilon$ to x_0. Without loss of generality we may assume that $x_0^\star$ is the origin. Now, since Ω is convex, it is not difficult to see that Ω_ε is convex as well. Consequently, it has a supporting hyperplane at $x_0^\star$ which, again without losing generality, we may assume is given by $x_d = 0$. Finally, since Ω_ε is bounded, there is $L > 0$ such that

$$\Omega_\varepsilon \subset \{x \in \mathbb{R}^d : 0 \le x_d \le L\}.$$

With this notation at hand we define the functions

$$p_{x_0}^{\pm}(x) = \pm\frac{1}{2}Ex_d(x_d - L),$$

and the discrete barriers are

$$p_{x_0,h}^{\pm} = \mathcal{G}_h p_{x_0}^{\pm}.$$

Let us now show that these functions satisfy the claimed properties.

(1) To show this we recall that, if $z \in \mathscr{N}_h^{\varepsilon}$,

$$p_{x_0,h}^{+}(z) = p_{x_0}^{+}(z) = \frac{1}{2}Ez_d(z_d - L) \le 0.$$

Similarly, we have

$$p_{x_0,h}^{-}(z) = p_{x_0}^{-}(z) = -\frac{1}{2}Ez_d(z_d - L) \ge 0.$$

(2) To show (3.21) we begin by noticing that $\mathrm{D}^2 p_{x_0}^{\pm} = \pm E\mathbf{e}_d \otimes \mathbf{e}_d$, where $\mathbf{e}_d$ is the d-th canonical basis vector. Therefore, we have that $\mathfrak{F}[p_{x_0}^{+}] \ge \lambda E$. Moreover, since Lemma 2.1 shows that $I_{\varepsilon}^{\alpha,\beta}[\cdot]$ is exact for quadratics, proceeding as in (3.14) yields, for every $z \in \mathscr{N}_h$,

$$\mathfrak{F}_h^{\varepsilon}[p_{x_0,h}^{+}](z) \ge \lambda E + \mathcal{R}_{h,\varepsilon}[p_{x_0}^{+}](z).$$

For ε and h small enough, assumption (3.19) allows us to estimate $\mathcal{R}_{h,\varepsilon}[p_{x_0}^{+}](z) \ge -\frac{\lambda}{2}E$. Similarly, we have that $\mathfrak{F}[p_{x_0}^{-}] \le -\lambda E$ and we can choose ε and h so that $\mathcal{R}_{h,\varepsilon}[p_{x_0}^{-}](z) \le \frac{\lambda}{2}E$.

(3) It remains to establish the growth bound (3.22), but this is an easy consequence of the fact that $p_{x_0}^{\pm}(0) = 0$ and estimate (3.5) for the Galerkin projection.

The functions $p_{x_0,h}^{\pm}$ have thus been constructed. □

We are now in position to establish convergence of our scheme.

Theorem 3.10 (Convergence)**.** *Assume that* $\mathsf{f} \in \mathrm{C}^{0,1}(\bar\Omega)$, *that the matrices* $\{A^{\alpha,\beta} : \alpha \in \mathcal{A},\ \beta \in \mathcal{B}\}$ *are uniformly elliptic* (1.2) *and satisfy*

$$\sup_{x_1,x_2 \in \bar\Omega, |x_1 - x_2| \le t} \|M^{\alpha,\beta}(x_1) - M^{\alpha,\beta}(x_2)\| \le \varpi(t), \quad \forall \alpha \in \mathcal{A}, \beta \in \mathcal{B},$$

where $M^{\alpha,\beta} = (A^{\alpha,\beta} - \frac{\lambda}{2}I)^{1/2}$. *For* $h > 0$ *and* $\varepsilon > 0$ *let* $u_h^{\varepsilon} \in \mathbb{V}_h^0$ *be the solution to the numerical scheme* (3.3), *obtained over a mesh* $\mathscr{T}_h$ *that is weakly acute with respect to faces, and belongs to a quasi-uniform family. Assume also that* (3.19) *holds. For* $x \in \bar\Omega$ *define*

$$u^\star(x) = \limsup_{\varepsilon,h \to 0, z \to x} u_h^{\varepsilon}(z), \qquad u_\star(x) = \liminf_{\varepsilon,h \to 0, z \to x} u_h^{\varepsilon}(z). \tag{3.23}$$

Then the upper semi-continuous function $u^\star$ *is a viscosity subsolution of* (1.1) *with* $u^\star \le 0$ *on* $\partial\Omega$ *and the lower semi-continuous function* $u_\star$ *is a viscosity supersolution of* (1.1) *with* $u_\star \ge 0$ *on* $\partial\Omega$.

Proof. We begin by remarking that, in light of the uniform boundedness shown in Proposition 3.8, the functions $u^\star$ and $u_\star$ are bounded.

Let us now show that $u^\star$ is a subsolution. Let p be a quadratic polynomial such that $u^\star - p$ has a local maximum at $x_0 \in \Omega$. We need to show that

$$\mathfrak{F}[p](x_0) \ge \mathsf{f}(x_0).$$

Without loss of generality we can assume that this is a strict maximum.

Let us now show that there are $\{z_h \in \mathscr{N}_h\}_{h>0}$ such that $u_h^\varepsilon - \mathcal{G}_h p$ attains its maximum at z_h and, as $h \to 0$, we have $z_h \to x_0$. If that is not the case, for any $\{z_h\}_{h>0}$, there is a subsequence $\{z_{h_k}\} \subset \{z_h\}_{h>0}$ that converges to some $y_0 \neq x_0$. By definition (3.23) we have

$$(u^\star - p)(y_0) \geq \lim_{h_k \to 0} u_{h_k}^\varepsilon(z_{h_k}) - p(y_0) = \lim_{h_k \to 0} \left(u_{h_k}^\varepsilon - \mathcal{G}_{h_k} p\right)(z_{h_k}).$$

On the other hand, since $u_{h_k}^\varepsilon - \mathcal{G}_{h_k} p$ attains its maximum at z_{h_k}, $(u_{h_k}^\varepsilon - \mathcal{G}_{h_k} p)(x_0) \leq (u_{h_k}^\varepsilon - \mathcal{G}_{h_k} p)(z_{h_k})$. Passing to the limit, we have

$$(u^\star - p)(x_0) \leq \lim_{h_k \to 0} u_{h_k}^\varepsilon(z_{h_k}) - p(y_0) \leq (u^\star - p)(y_0),$$

which contradicts that $u^\star - p$ attains a strict maximum at x_0.

In what follows we consider this sequence of nodes and, for simplicity, we suppress the subindex h. As in Corollary 3.3, the fact that $u_h^\varepsilon - \mathcal{G}_h p$ attains its maximum at z yields

$$\frac{\lambda}{2} \Delta_h (u_h^\varepsilon - \mathcal{G}_h p)(z) + I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon - \mathcal{G}_h p](z) \leq 0,$$

which combined with Lemma 3.1 allows us to conclude that

$$f_z = \mathfrak{F}_h^\varepsilon[u_h^\varepsilon](z) \leq \mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z).$$

Since f is continuous $f_z \to \mathsf{f}(x_0)$ and so it remains to study the behavior of the right hand side in this inequality.

Repeating the computations in the proof of Theorem 3.5 that led to (3.14) allows us to obtain

$$\mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z) = \frac{\int_\Omega \mathfrak{F}[p](x)\phi_z(x)\,\mathrm{d}x}{\int_\Omega \phi_z(x)\,\mathrm{d}x} + \mathcal{R}_{h,\varepsilon}[p](z),$$

where $\mathcal{R}_{h,\varepsilon}[p](z)$ is defined in (3.12). Consequently, if we are able to show that, for every $z \in \mathscr{N}_h$, $\mathcal{R}_{h,\varepsilon}[p](z) \to 0$ as $h \to 0$, $\varepsilon \to 0$ we would get that

$$\mathsf{f}(x_0) \leq \limsup_{h \to 0} \mathfrak{F}_h^\varepsilon[\mathcal{G}_h p](z) \leq \mathfrak{F}[p](x_0),$$

which is what we need to prove. The bound on $\mathcal{R}_{h,\varepsilon}[p](z)$, obtained in Lemma 3.4, and the choice of ε allow us to conclude.

We now show that if $x_0 \in \partial\Omega$, then we must have $u^\star(x_0) \leq 0$. To do so, let ε and h be small enough and set $E = \frac{2}{\lambda}\|\mathsf{f}\|_{\mathrm{L}^\infty(\Omega)}$ in the barrier function $p_{x_0,h}^-$ of Lemma 3.9. Property (3.21) yields that, for every $z \in \mathscr{N}_h$

$$\mathfrak{F}_h^\varepsilon[p_{x_0,h}^-](z) \leq -\|\mathsf{f}\|_{\mathrm{L}^\infty(\Omega)} \leq f_z = \mathfrak{F}_h^\varepsilon[u_h^\varepsilon](z).$$

Moreover, by (3.20), we have $p_{x_0,h}^- \geq 0 = u_h^\varepsilon$ in $(\partial\Omega)_\varepsilon$. Consequently, invoking the discrete maximum principle of Corollary 3.3 we obtain

$$u_h^\varepsilon(x) \leq p_{x_0,h}^-(x) \leq C\varepsilon, \quad \forall x \in x_0 + B_{Q\varepsilon} \cap \Omega_\varepsilon,$$

where the upper bound follows from (3.22). Letting $\varepsilon \to 0$ and $x \to x_0$ we conclude that $u^\star(x_0) \leq 0$.

Finally, the fact that $u_\star$ is a super-solution can be shown in a similar fashion. For brevity, we skip the details. □

Corollary 3.11 (Convergence). *Let the family of triangulations $\{\mathscr{T}_h\}_{h>0}$ be quasi-uniform and assume that, for every $h > 0$, $\mathscr{T}_h$ is weakly acute with respect to faces (3.16), and that, as $\varepsilon, h \to 0$, we have (3.19). In this setting the family $\{u_h^\varepsilon\}_{h>0,\varepsilon>0}$ of solutions to (3.3) converges pointwise to u, the viscosity solution of the uniformly elliptic Isaacs equation (1.1).*

Proof. Using the notation of Theorem 3.10, we observe that $u_\star \geq 0 \geq u^\star$ on $\partial\Omega$ and that $u^\star$ and $u_\star$ are viscosity sub and super solutions, respectively. By a comparison principle for (1.1); see Theorem 2.52(b) of [31], we then must have

$$u^\star(x) \leq u_\star(x) \quad \forall x \in \Omega.$$

On the other hand, by definition (3.23), $u_\star \leq u^\star$. Both inequalities readily imply that $u^\star = u_\star = \mathsf{u}$, the viscosity solution of (1.1). This proves convergence of u_h^ε to u. ☐

Remark 3.12 (Nonhomogeneous boundary conditions)**.** All the considerations given above can be extended to the case of nonhomogeneous boundary conditions, *i.e.* when in (1.1) the boundary condition reads $\mathsf{u} = g$. For that we need to assume that $g = \mathrm{tr}_{\partial\Omega} G$ for some $G \in \mathrm{C}^2(\bar{\Omega})$, and a suitable extension of g to $(\partial\Omega)_\varepsilon$ must be provided in the second equation of (2.5). As described in [8], this can be achieved by using the so-called sup-convolution:

$$\tilde{g}(x) = \sup_{x^\star \in \partial\Omega} \left[g(x^\star) - |x - x^*|^2 \right], \quad \forall x \notin \Omega,$$

which given the smoothness of g and Ω satisfies $\tilde{g} \in \mathrm{C}^{1,1}(\overline{\Omega_\varepsilon})$. The discrete solution $u_h^\varepsilon \in \mathbb{V}_h^\varepsilon$ then, satisfies $u_h^\varepsilon(z) = \tilde{g}(z)$ for $z \in \mathscr{N}_h^\varepsilon$. Let us now briefly indicate the modifications that need to be made to our arguments:

(1) *Existence*: To show existence we need to redefine the set of super-solutions. We now require that $v_h(z) \geq \mathcal{G}_h \tilde{g}(z)$ for $z \in \mathscr{N}_h^\varepsilon$. We need to show that this set is nonempty, and for that it suffices to show $\mathcal{G}_h p \in \mathbb{S}_h$ with

$$p(x) = \frac{1}{2}\lambda^{-1} F(|x - \xi|^2 - R^2) + g^-, \qquad g^- = \max_{x \in (\partial\Omega)_\varepsilon} \tilde{g}(x).$$

(2) *Uniform boundedness*: We consider the function $w_h = u_h^\varepsilon + M$ where M is sufficiently large so that $w_h \geq 0$ on $\mathscr{N}_h^\partial$. This is possible because g is bounded. Notice that, for all $z \in \mathscr{N}_h$, $I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z) = I_\varepsilon^{\alpha,\beta}[w_h](z)$ and $\Delta_h u_h(z) = \Delta_h w_h(z)$. Therefore

$$\frac{\lambda}{2} \Delta_h w_h(z) \leq f_z, \quad \forall z \in \mathcal{C}_h^-(w_h).$$

and by ABP

$$\|u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim \|\mathsf{f}\|_{\mathrm{L}^d(\Omega)} + \|g\|_{\mathrm{L}^\infty(\partial\Omega)}.$$

(3) *Barrier functions*: When proving convergence, we must be able to assert that the boundary conditions are attained in a classical sense. For that it suffices to show that $u^\star(x) \geq g(x)$ for $x \in \partial\Omega$ and this is achieved by constructing suitable barrier functions (*cf.* Lemma 3.9). In this case, the barrier functions will be defined as

$$q_{x_0,h}^\pm = p_{x_0,h}^\pm + g^\pm, \qquad g^+ = \min_{x \in (\partial\Omega)_\varepsilon} \tilde{g}(x).$$

We leave the details of this program to the reader.

4. Rate of convergence

Let us now study the rate of convergence of u_h^ε to u, under the smoothness assumptions of Corollary 2.4. We will achieve this by comparing the solution of the discrete scheme (3.3) with the solution of the integrodifferential approximation (2.5). In light of the technique that we are adopting, we must immediately note that since the approximation results of Proposition 2.3 and, as a consequence, the regularity obtained in Corollary 2.4 only apply in the case of constant coefficients, $A^{\alpha,\beta}(x) = A^{\alpha,\beta}$, in what follows we must assume this. In this setting (2.4) is trivially satisfied with $\varpi \equiv 0$. It is possible that the results of [37] can be used to extend Proposition 2.3 to variable coefficients and if that is the case, our results will immediately follow as well.

The main technical tool we will employ to obtain rates of convergence in the L^∞-norm will be the discrete ABP estimate of Lemma 3.7. Thus, we must require that the mesh is weakly acute with respect to faces, as defined in (3.16). We finally remark that since the relation (3.19) is required for convergence of our method we will study the rate of convergence under this assumption.

4.1. The error equation

We now derive an equation to determine the error. Multiply the integro-differential approximation to the Isaacs equation (2.5) by a test function ϕ_z with $z \in \mathscr{N}_h$, integrate over Ω and divide by ω_z. In view of the definition of the discrete Laplacian (3.2) and of the Galerkin projection (3.4), we obtain

$$\frac{\lambda}{2}\Delta_h \mathcal{G}_h u^\varepsilon(z) + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h u^\varepsilon](z) = f_z + \mathcal{R}_{h,\varepsilon}[u^\varepsilon](z).$$

Subtract from this equation the scheme (3.3) to obtain, for all $z \in \mathscr{N}_h$,

$$\begin{aligned}\frac{\lambda}{2}\Delta_h \left(\mathcal{G}_h u^\varepsilon - u_h^\varepsilon\right)(z) + \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h u^\varepsilon](z)& \\ - \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z) &= \mathcal{R}_{h,\varepsilon}[u^\varepsilon](z).\end{aligned} \tag{4.1}$$

Notice, finally, that $\mathcal{G}_h u^\varepsilon - u_h^\varepsilon \equiv 0$ over $\mathscr{N}_h^\varepsilon$.

4.2. Rate of convergence

With the error equation (4.1) at hand, we now readily obtain an error estimate. This is the content of the next result.

Theorem 4.1 (Rate of convergence)**.** *Let* $\mathsf{u} \in C^{1,s}(\Omega) \cap C^{0,1}(\bar\Omega)$ *be the viscosity solution of the Isaacs equation* (1.1) *and* $u_h^\varepsilon \in \mathbb{V}_h^\varepsilon$ *be the solution of* (3.3)*. Choose* ε *so that* (3.19) *holds with* $\varpi \equiv 0$*. Then, there is* $\sigma > 0$ *such that*

$$\|\mathsf{u} - u_h^\varepsilon\|_{L^\infty(\Omega)} \lesssim \left(\varepsilon^\sigma + \frac{h}{\varepsilon^2}|\log h|\right)\|\mathsf{f}\|_{C^{0,1}(\bar\Omega)},$$

where the hidden constant depends on λ *and* Λ*, but is independent of* ε *and* h*.*

Proof. We write

$$\|\mathsf{u} - u_h^\varepsilon\|_{L^\infty(\Omega)} \le \|\mathsf{u} - u^\varepsilon\|_{L^\infty(\Omega)} + \|u^\varepsilon - \mathcal{G}_h u^\varepsilon\|_{L^\infty(\Omega)} + \|\mathcal{G}_h u^\varepsilon - u_h^\varepsilon\|_{L^\infty(\Omega)},$$

and we examine each term separately.

By Proposition 2.3, there is $\sigma > 0$ such that

$$\|\mathsf{u} - u^\varepsilon\|_{L^\infty(\Omega)} \lesssim \varepsilon^\sigma \|\mathsf{f}\|_{C^{0,1}(\bar\Omega)}.$$

Estimate (3.5) on the Galerkin projection immediately yields

$$\|u^\varepsilon - \mathcal{G}_h u^\varepsilon\|_{L^\infty(\Omega)} \lesssim h|\log h| \|\mathsf{f}\|_{C^{0,1}(\bar\Omega)},$$

where, using Proposition 2.3, we bounded $\|u^\varepsilon\|_{C^{0,1}(\bar\Omega)}$ by $\|\mathsf{f}\|_{C^{0,1}(\bar\Omega)}$.

Let us denote $e_h = \mathcal{G}_h u^\varepsilon - u_h^\varepsilon \in \mathbb{V}_h^0$.

Using the convex envelope $\Gamma(e_h)$ of the error we can, as in the proof of Proposition 3.8, conclude that for every $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$ we have

$$I_\varepsilon^{\alpha,\beta}[\mathcal{G}_h u^\varepsilon](z) \ge I_\varepsilon^{\alpha,\beta}[u_h^\varepsilon](z), \quad \forall z \in \mathcal{C}_h^-(e_h),$$

so that using the monotonicity property of the inf-sup operator (3.7), we have

$$\inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}\mathcal{G}_h u^\varepsilon(z) \ge \inf_{\alpha\in\mathcal{A}}\sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta} u_h^\varepsilon(z), \quad \forall z \in \mathcal{C}_h^-(e_h).$$

In conclusion, at the nodal contact set $\mathcal{C}^-(e_h)$ equation (4.1) reduces to

$$\frac{\lambda}{2}\Delta_h e_h(z) \le \mathcal{R}_{h,\varepsilon}[u^\varepsilon](z).$$

An application of the discrete ABP estimate of Lemma 3.7 then yields

$$\sup_\Omega e_h^- \lesssim \left(\sum_{z\in\mathcal{C}^-(e_h)} \omega_z |\mathcal{R}_{h,\varepsilon}[u^\varepsilon](z)|^d\right)^{1/d} \lesssim \max_{z\in\mathscr{N}_h} |\mathcal{R}_{h,\varepsilon}[u^\varepsilon](z)|.$$

This yields a lower bound for e_h. An upper bound can be derived in a similar fashion by considering $-e_h$. Hence, we infer

$$\|\mathcal{G}_h u^\varepsilon - u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim \max_{z\in\mathscr{N}_h} |\mathcal{R}_{h,\varepsilon}[u^\varepsilon](z)|.$$

Using the fact that $u^\varepsilon \in \mathrm{C}^{0,1}(\bar\Omega)$ uniformly in ε we can invoke the bounds on $\mathcal{R}_{h,\varepsilon}[u^\varepsilon](z)$ obtained in Lemma 3.4 which, recalling that $\varpi \equiv 0$, can be combined with the estimate of Proposition 2.3 to yield

$$\max_{z\in\mathscr{N}_h} |\mathcal{R}_{h,\varepsilon}[u^\varepsilon](z)| \lesssim \frac{h}{\varepsilon^2}|\log h| \|\mathsf{f}\|_{\mathrm{C}^{0,1}(\bar\Omega)}.$$

Notice now that (3.19) implies that $h|\log h|\varepsilon^{-2} \to 0$ as $h \to 0$ and $\varepsilon \to 0$. Therefore, we obtain

$$\|\mathcal{G}_h u^\varepsilon - u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim \frac{h}{\varepsilon^2}|\log h| \|\mathsf{f}\|_{\mathrm{C}^{0,1}(\bar\Omega)}.$$

Combining the estimates of these three steps yields the result. □

Remark 4.2 (Choice of ε)**.** If one knows the value of σ in Proposition 2.3, setting $\varepsilon^{\sigma+2} = h|\log h|$ yields an error estimate of the form

$$\|\mathsf{u} - u_h^\varepsilon\|_{\mathrm{L}^\infty(\Omega)} \lesssim h^{\frac{\sigma}{\sigma+2}} |\log h|^{\frac{\sigma}{\sigma+2}} \|\mathsf{f}\|_{\mathrm{C}^{0,1}(\bar\Omega)}.$$

Notice that in the best case scenario, that is $\sigma = 1$, we would obtain a rate of convergence of order $\mathcal{O}(h^{1/3}|\log h|^{1/3})$.

Remark 4.3 (Explicit rate of convergence)**.** The rate of convergence in Theorem 4.1 is given rather implicitly. It seems that this is a recurring feature in the literature; see for instance the main result in [23].

5. Implementation details

Let us discuss how to obtain a solution to the nonlinear problem that scheme (3.3) entails. The main difficulty in devising a convergent algorithm for the solution of (3.3) is the fact that, due to the inf-sup operations, the underlying operator is neither convex nor concave. This is in sharp contrast with, for instance, the interpretation of Howard's algorithm [16] as a semi-smooth Newton method described in [19, 35] for the Hamilton-Jacobi-Bellman equation since, as it is shown in Remark 5.3 of [6] such a method may not converge.

On the other hand, Section 5 of [6] presents a convergent generalization of Howard's algorithm for max-min problems, which we readily adapt here. We will present an algorithm which requires the solution, at every iteration step, of a Hamilton Jacobi Bellman equation, which can be realized via a semi-smooth Newton method. We will also comment on an algorithm with inexact solves.

For a given $w_h \in \mathbb{V}_h^0$ and $z \in \mathscr{N}_h$ define $\alpha(w_h, z) \in \mathcal{A}$ as the element that infimizes the supremum of the integral operators when applied to w_h at the point z, that is

$$\inf_{\alpha\in\mathcal{A}} \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha,\beta}[w_h](z) = \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha(w_h,z),\beta}[w_h](z).$$

Set $\boldsymbol{\alpha}(w_h) = \{\alpha(w_h, z) : z \in \mathscr{N}_h\}$. For $v_h \in \mathbb{V}_h^0$ and $z \in \mathscr{N}_h$ define

$$\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}(v_h)}[w_h](z) := \frac{\lambda}{2}\Delta_h w_h(z) + \sup_{\beta\in\mathcal{B}} I_\varepsilon^{\alpha(v_h,z),\beta}[w_h](z). \tag{5.1}$$

Our algorithm can then be described as follows:

- **Initialization**: Choose $w_h^{-1} \in \mathbb{V}_h^0$ and set $\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}(w_h^{-1})$.
- **Iteration**: For $k \geq 0$ find $w_h^k \in \mathbb{V}_h^0$ that solves

$$\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_k}[w_h^k](z) = f_z, \quad \forall z \in \mathscr{N}_h \tag{5.2}$$

  and set

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}(w_h^k). \tag{5.3}$$

- **Convergence test**: If $\mathfrak{F}_h^\varepsilon[w_h^k](z) = f_z$ for all $z \in \mathscr{N}_h$ stop.

Notice that (5.3) is equivalent to

$$\mathfrak{F}_h^\varepsilon[w_h^k](z) = \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_{k+1}}[w_h^k](z), \quad \forall z \in \mathscr{N}_h.$$

The analysis of the algorithm (5.2) and (5.3) relies on the following properties of the operators $\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}}$.

Lemma 5.1 (Monotonicity and comparison)**.** *For every $\boldsymbol{\alpha} \in \mathcal{A}^{\#\mathscr{N}_h}$ the operator $\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}}$ is monotone and satisfies a comparison principle,* i.e. *if $v_h, w_h \in \mathbb{V}_h^0$ are such that*

$$\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}}[v_h](z) \leq \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}}[w_h](z), \quad \forall z \in \mathscr{N}_h,$$

then $v_h \geq w_h$ over $\mathscr{N}_h$.

Proof. The proof repeats the arguments of Lemma 3.2 and Corollary 3.3. For brevity, we skip details. □

The comparison principle will allow us to obtain convergence.

Theorem 5.2 (Convergence)**.** *The sequence $\{w_h^k\}_{k\geq 0} \subset \mathbb{V}_h^0$ obtained by algorithm (5.2) and (5.3) converges in a finite number of steps to $u_h^\varepsilon \in \mathbb{V}_h^0$, solution of (3.3).*

Proof. We proceed in two steps:

First, we show that the sequence is monotone, *i.e.* for every $z \in \mathscr{N}_h$, $w_h^k(z) \geq w_h^{k+1}(z)$. By construction:

$$\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_{k+1}}[w_h^k](z) = \mathfrak{F}_h^\varepsilon[w_h^k](z) \leq \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_k}[w_h^k](z).$$

Subtracting f_z from this inequality we realize that

$$\mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_{k+1}}[w_h^k](z) - f_z \leq \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_k}[w_h^k](z) - f_z = 0 = \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_{k+1}}[w_h^{k+1}](z) - f_z,$$

which by the comparison principle established in Lemma 5.1 implies $w_h^k \geq w_h^{k+1}$ over $\mathscr{N}_h$.

Next, since the set $\mathcal{A}$ is finite, there are at most $(\#\mathcal{A})^{\#\mathscr{N}_h}$ different variables and there must be two indices $\kappa > \ell$ for which (5.3) yields $\boldsymbol{\alpha}_\kappa = \boldsymbol{\alpha}_\ell$. This implies that $w_h^\kappa = w_h^\ell$ and by monotonicity $w_h^k = w_h^\ell$ for all $k \geq l$. But then, by uniqueness $w_h^\ell = u_h^\varepsilon$.

This concludes the proof. □

Notice that (5.2) requires the exact solution of a discrete version of a Hamilton Jacobi Bellman problem, which can be achieved by Howard's algorithm. We also propose a scheme with inexact solves in this step:

- **Initialization**: Choose $w_h^{-1} \in \mathbb{V}_h^0$ and set $\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}(w_h^{-1})$.

- **Iteration**: For $k \geq 0$ find $w_h^k \in \mathbb{V}_h^0$ such that

$$\max_{z \in \mathscr{N}_h} \left| \mathfrak{F}_{h,\varepsilon}^{\boldsymbol{\alpha}_k}[w_h^k](z) - f_z \right| < \eta_k. \tag{5.4}$$

Set

$$\boldsymbol{\alpha}_{k+1} = \boldsymbol{\alpha}(w_h^k). \tag{5.5}$$

- **Convergence test**: If $\mathfrak{F}_h^\varepsilon[w_h^k](z) = f_z$ for all $z \in \mathscr{N}_h$ stop.

The convergence of this algorithm follows *mutatis mutandis* the proof of Theorem 5.4 of [6].

Theorem 5.3 (Convergence with inexact solves)**.** *Assume that the sequence of errors* $\{\eta_k\}_{k\in\mathbb{N}} \in \ell^1$*. Then the sequence* $\{w_h^k\}_{k\geq 0} \subset \mathbb{V}_h^0$*, obtained by algorithm* (5.4)–(5.5)*, converges to* $u_h^\varepsilon \in \mathbb{V}_h^0$*, solution of* (3.3)*.*

Acknowledgements. AJS has been partially supported by NSF grants DMS-1418784 and DMS-1720213.

References

[1] G. Barles and E.R. Jakobsen, Error bounds for monotone approximation schemes for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.* **43** (2005) 540–558.

[2] G. Barles and P.E. Souganidis, Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.* **4** (1991) 271–283.

[3] I.H. Biswas, I. Chowdhury and E.R. Jakobsen, On the rate of convergence for monotone numerical schemes for nonlocal Isaacs' equations. Preprint arXiv:1709.07743 (2017).

[4] B. Bojanov and G. Petrova, Numerical integration over a disc. A new Gaussian quadrature formula. *Numer. Math.* **80** (1998) 39–59.

[5] B. Bojanov and G. Petrova, Uniqueness of the Gaussian quadrature for a ball. *J. Approx. Theory* **104** (2000) 21–44.

[6] O. Bokanowski, S. Maroso and H. Zidani, Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.* **47** (2009) 3001–3026.

[7] S.C. Brenner and L.R. Scott, The mathematical theory of finite element methods, 3rd edn. In Vol. 15 of *Texts in Applied Mathematics.* Springer, New York (2008).

[8] L. Caffarelli and L. Silvestre, Smooth approximations of solutions to nonconvex fully nonlinear elliptic equations. In: Nonlinear Partial Differential Equations and Related Topics. Vol. 229 of American Mathematical Society Translations, Series 2. AMS, Providence, RI (2010) 67–85.

[9] L.A. Caffarelli and P.E. Souganidis, A rate of convergence for monotone finite difference approximations to fully nonlinear, uniformly elliptic PDEs. *Comm. Pure Appl. Math.* **61** (2008) 1–17.

[10] P.G. Ciarlet, Basic error estimates for elliptic problems. In: *Handbook of Numerical Analysis.* North-Holland, Amsterdam (1991) 17–351.

[11] P.G. Ciarlet, The finite element method for elliptic problems. In: *Classics in Applied Mathematics.* SIAM, Philadelphia, PA **40** (2002).

[12] K. Debrabant and E.R. Jakobsen, Semi-Lagrangian schemes for linear and fully non-linear diffusion equations. *Math. Comp.* **82** (2013) 1433–1462.

[13] A. Ern and J.-L. Guermond, Theory and practice of finite elements. In Vol. 159 *Applied Mathematical Sciences.* Springer-Verlag, New York, NY (2004).

[14] X. Feng, R. Glowinski and M. Neilan, Recent developments in numerical methods for fully nonlinear second order partial differential equations. *SIAM Rev.* **55** (2013) 205–267.

[15] Q. Han and F. Lin, Elliptic partial differential equations, 2nd edition. In Vol. I of *Courant Lecture Notes in Mathematics.* American Mathematical Society, Providence, RI (2011).

[16] R.A. Hovard, Dinamicheskoe programmirovanie i Markovskie protsessy. Izdat. "Sovetskoe Radio", Moscow (1964).

[17] E.R. Jakobsen, On error bounds for approximation schemes for non-convex degenerate elliptic equations. *BIT* **44** (2004) 269–285.

[18] E.R. Jakobsen, On error bounds for monotone approximation schemes for multi-dimensional Isaacs equations. *Asymptot. Anal.* **49** (2006) 249–273.

[19] M. Jensen and I. Smears, On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.* **51** (2013) 137–162.

[20] M. Kocan, Approximation of viscosity solutions of elliptic partial differential equations on minimal grids. *Numer. Math.* **72** (1995) 73–92.

[21] N.V. Krylov, On the rate of convergence of finite-difference approximations for Bellman's equations. *Algebra i Analiz* **9** (1997) 245–256.

[22] N.V. Krylov, The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients. *Appl. Math. Optim.* **52** (2005) 365–399.

[23] N.V. Krylov, On the rate of convergence of finite-difference approximations for elliptic Isaacs equations in smooth domains. *Comm. Partial Differ. Equ.* **40** (2015) 1393–1407.

[24] H.J. Kuo and N.S. Trudinger, Discrete methods for fully nonlinear elliptic equations. *SIAM J. Numer. Anal.* **29** (1992) 123–135.

[25] D. Leykekhman and B. Vexler, Finite element pointwise results on convex polyhedral domains. *SIAM J. Numer. Anal.* **54** (2016) 561–587.

[26] T.S. Motzkin and W. Wasow, On the approximation of linear elliptic differential equations by difference equations with positive coefficients. *J. Math. Phys.* **31** (1953) 253–259.

[27] I.P. Mysovskikh, Interpolyatsionnye kubaturnye formuly. "Nauka", Moscow (1981).

[28] N. Nadirashvili, V. Tkachev and S. Vlăduţ, A non-classical solution to a Hessian equation from Cartan isoparametric cubic. *Adv. Math.* **231** (2012) 1589–1597.

[29] N. Nadirashvili and S. Vlăduţ, Nonclassical solutions of fully nonlinear elliptic equations. *Geom. Funct. Anal.* **17** (2007) 1283–1296.

[30] N. Nadirashvili and S. Vlăduţ, Singular viscosity solutions to fully nonlinear elliptic equations. *J. Math. Pures Appl.* **89** (2008) 107–113.

[31] M. Neilan, A.J. Salgado, and W. Zhang, Numerical analysis of strongly nonlinear PDEs. *Acta Numer.* **26** (2017) 137–303.

[32] R.H. Nochetto and W. Zhang, Discrete ABP estimate and convergence rates for linear elliptic equations in non-divergence form. *Found. Comput. Math.* **18** (2018) 537–593.

[33] G. Petrova, Cubature formulae for spheres, simplices and balls. *J. Comput. Appl. Math.* **162** (2004) 483–496.

[34] A.H. Schatz and L.B. Wahlbin, On the quasi-optimality in L_∞ of the $\dot{H}^1$-projection into finite element spaces. *Math. Comput.* **38** (1982) 1–22.

[35] I. Smears and E. Süli, Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients. *SIAM J. Numer. Anal.* **52** (2014) 993–1016.

[36] A.H. Stroud, Approximate Calculation of Multiple Integrals. *Prentice-Hall Series in Automatic Computation.* Prentice-Hall, Inc., Englewood Cliffs, NJ (1971).

[37] O. Turanova, Error estimates for approximations of nonhomogeneous nonlinear uniformly elliptic equations. *Calc. Var. Partial Differ. Equ.* **54** (2015) 2939–2983.