MATHEMATICS OF COMPUTATION
Volume 88, Number 316, March 2019, Pages 637–664
http://dx.doi.org/10.1090/mcom/3353
Article electronically published on May 17, 2018

TWO-SCALE METHOD FOR THE MONGE-AMPÈRE EQUATION: CONVERGENCE TO THE VISCOSITY SOLUTION

R. H. NOCHETTO, D. NTOGKAS, AND W. ZHANG

Abstract. We propose a two-scale finite element method for the Monge-Ampère equation with Dirichlet boundary condition in dimension $d \geq 2$ and prove that it converges to the viscosity solution uniformly. The method is inspired by a finite difference method of Froese and Oberman, but is defined on unstructured grids and relies on two separate scales: the first one is the mesh size h and the second one is a larger scale that controls appropriate directions and substitutes the need of a wide-stencil. The main tools for the analysis are a discrete comparison principle and discrete barrier functions that control the behavior of the discrete solution, which is continuous piecewise linear, both close to the boundary and in the interior of the domain.

1. Introduction

We consider the Monge-Ampère equation with Dirichlet boundary condition:

$$\begin{cases} \det D^2 u = f & \text{in } \Omega \subset \mathbb{R}^d, \\ u = g & \text{on } \partial\Omega, \end{cases} \tag{1.1}$$

where $f \geq 0$ is uniformly continuous, Ω is a uniformly convex domain (not necessarily W^2_∞), and g is a continuous function. We seek a *convex* solution u of (1.1), which is critical for (1.1) to be elliptic and have a unique viscosity solution [26].

The Monge-Ampère equation has a wide spectrum of applications in optimal mass transport problems, geometry, nonlinear elasticity, and meteorology. These applications lead to an increasing interest in the investigation of efficient numerical methods. There exist several methods for the Monge-Ampère equation. These include the early work by Oliker and Prussner [35] for the space dimension $d = 2$, the vanishing moment methods by Feng and Neilan [21, 22], the penalty method of Brenner, Gudi, Neilan, and Sung [10], least squares and augmented Lagrangian methods by Dean and Glowinski [14, 15, 25], the finite difference method proposed recently by Benamou, Collino, and Mirebeau [5, 29], and a new semi-Lagrangian

Received by the editor June 19, 2017, and, in revised form, December 10, 2017, December 11, 2017, and December 19, 2017.

2010 *Mathematics Subject Classification.* Primary 65N30, 65N12, 65N06, 35J96.

Key words and phrases. Monge-Ampère equation, viscosity solution, two-scale method, monotone scheme, convergence, regularization.

The first author was partially supported by the NSF Grant DMS -1411808, the Institut Henri Poincaré (Paris), and the Hausdorff Institute (Bonn).

The second author was partially supported by the NSF Grant DMS -1411808 and the 2016-2017 Patrick and Marguerite Sung Fellowship of the University of Maryland.

The third author was partially supported by the NSF Grant DMS -1411808 and the Brin Postdoctoral Fellowship of the University of Maryland.

method by Feng and Jensen [20]. Our work is mostly motivated by the wide-stencil scheme proposed by Froese and Oberman, who proved convergence of the scheme [24]. Awanou [2] proved a linear rate of convergence for classical solutions for the wide-stencil method, when applied to a perturbed Monge-Ampère equation with an extra lower order term δu; the parameter $\delta > 0$ is independent of the mesh and appears in reciprocal form in the rate. In contrast, our analysis hinges on the discrete comparison principle and two discrete barrier functions, which are instrumental in proving convergence to the viscosity solution of (1.1). Moreover, our methodology further leads to pointwise error estimates, which we derive in [30].

1.1. **Our contribution.** Our method hinges on the following formula for the determinant of the semi-positive Hessian D^2w of a smooth function w as in [24]:

$$\det D^2 w(x) = \min_{(v_1,\dots,v_d)\in\mathbb{S}^\perp} \prod_{j=1}^{d} v_j^T D^2 w(x)\ v_j, \tag{1.2}$$

where $\mathbb{S}^\perp$ is the set of all $d-$orthonormal bases in $\mathbb{R}^d$. The minimum is achieved for the eigenvectors of D^2w and is equal to $\prod_{j=1}^d \lambda_j$, where $\lambda_j, j = 1,\dots,d$ are the corresponding eigenvalues. To discretize (1.2) we introduce two scales h and δ. We discretize the domain Ω by a shape regular and quasi-uniform mesh $\mathcal{T}_h$ with spacing h, and construct a space $\mathbb{V}_h$ of continuous piecewise linear functions associated with the mesh $\mathcal{T}_h$. The second scale δ is the length of directions we use to approximate second directional derivatives by central second order differences

$$\nabla^2_\delta w(x;v) := \frac{w(x+\delta v) - 2w(x) + w(x-\delta v)}{\delta^2} \quad \text{and} \quad |v| = 1$$

for any $w \in C^0(\overline{\Omega})$; this formula will be appropriately modified close to $\partial\Omega$. We denote by u_ε our discrete solution, where $\varepsilon = (h,\delta)$ represents the two scales, and define the discrete Monge-Ampère operator to be

$$T_\varepsilon[u_\varepsilon](x_i) := \min_{(v_1,\dots,v_d)\in\mathbb{S}^\perp} \prod_{j=1}^{d} \nabla^2_\delta u_\varepsilon(x_i;v_j),$$

where x_i is a generic node of $\mathcal{T}_h$. This leads to a clear separation of scales, which is a key theoretical advantage over the original wide-stencil method of [24]. This also yields continuous dependence of u_ε on data, which we further exploit in [30]. In fact, such continuous dependence result, along with the discrete comparison principle and the use of some discrete barrier functions give rise to rates of convergence in $L^\infty(\Omega)$ for viscosity solutions of (1.1) under some additional regularity requirements [30]. To make the two-scale method practical, we resort to fast search techniques within [38, 39] to locate points $x_i \pm \delta v_j$, which may not be nodes of $\mathcal{T}_h$ in general.

The main tool in the current work is the discrete comparison principle that enables us to control the behavior of u_ε and prove its uniform convergence to the unique viscosity solution u of (1.1) as $\delta \to 0$ and $h\delta^{-1} \to 0$. It is important to realize, as already observed in [20], that such a convergence is not an immediate consequence of the theory developed by Barles and Souganidis [4]. This theory assumes that the discrete operator is consistent up to the boundary and that the boundary conditions are treated in the viscosity sense; our operator T_ε is only consistent at distance δ from the boundary and our notion of Dirichlet condition is classical. Moreover, the theory of [4] also hinges on a comparison principle for the

underlying equation, which in the case of the Monge-Ampère equation (1.1) requires that the subsolution and supersolution constructed through the limit supremum and limit infimum of u_ε be convex.

We present two proofs of uniform convergence. The first one, discussed in Sections 5.2 and 5.3, relies on regularization of data f, g and Ω and the discrete comparison principle. This approach sets the stage for proving rates of convergence for the two-scale method, which we derive in [30]. Regularization is a natural device used already for Monge-Ampère by De Philippis and Figalli [16] as a PDE tool and Awanou for numerical purposes [3]. The second approach is along the lines of Barles and Souganidis [4], uses techniques similar to those developed by Feng and Jensen [20], and circumvents the two main issues mentioned above. Controlling the behavior of u_ε in a δ-neighborhood of the boundary $\partial\Omega$ is critical to both approaches. This is achieved via a discrete barrier function discussed in Section 5.1; similar constructions are discussed in [20, 31, 32].

To showcase the performance of our two-scale method, we present computational experiments for a classical and a degenerate viscosity solution solved with a semi-smooth Newton method. We obtain linear rates for both cases. We also present an example with unbounded f, which does not fall within our theory, and still observe convergence although with a reduced rate.

It is worth comparing the two-scale method with the Oliker-Prussner method [32, 35]. The former is easier to implement because it does not require the explicit computation of subdifferentials, and is formulated on shape regular meshes $\mathcal{T}_h$ instead of cartesian meshes. Although the coarse and fine scales δ and h must only satisfy $h\delta^{-1} \to 0$ for convergence, rates of convergence require knowledge of regularity of the exact solution u of (1.1) to choose $\delta = \delta(h)$ [30] in contrast to [32].

1.2. **Outline.** In Section 2 we introduce our method and the main tool of our analysis, the discrete comparison principle. In Section 3 we prove the existence and uniqueness of our discrete solution. In Section 4 we prove the consistency of the discrete operator and in Section 5 we prove the uniform convergence of the discrete solution to the viscosity solution of (1.1). Lastly, in Section 6 we document the performance of our method with numerical experiments.

2. Two-scale method

2.1. **Ideal two-scale method.** Let $\mathcal{T}_h$ be a shape regular and quasi-uniform triangulation with mesh size h. We denote by Ω_h the union of elements of $\mathcal{T}_h$ and we call it the computational domain. Let $\mathcal{N}_h$ denote the nodes of $\mathcal{T}_h$, $\mathcal{N}_h^b := \{x_i \in \mathcal{N}_h : x_i \in \partial\Omega_h\}$ be the boundary nodes, and $\mathcal{N}_h^0 := \mathcal{N}_h \setminus \mathcal{N}_h^b$ be the interior nodes. We require that $\mathcal{N}_h^b \subset \partial\Omega$, which in view of the convexity of Ω implies that Ω_h is also convex and $\Omega_h \subset \Omega$. We denote by $\mathbb{V}_h$ the space of continuous piecewise linear functions over $\mathcal{T}_h$. We recall the notation $\mathbb{S}^\perp$ for the collection of all d-tuples of orthonormal bases and $\mathbf{v} := (v_1, \dots, v_d) \in \mathbb{S}^\perp$ for a generic element, whence each component v_i belongs to the unit sphere $\mathbb{S}$ of $\mathbb{R}^d$. For $x_i \in \mathcal{N}_h^0$, we use the formula of centered second differences

$$\nabla_\delta^2 w(x_i; v_j) := \frac{w(x_i + \rho\delta v_j) - 2w(x_i) + w(x_i - \rho\delta v_j)}{\rho^2\delta^2}, \tag{2.1}$$

where $0 < \rho \leq 1$ is the largest number such that both $x_i \pm \rho\delta v_j \in \overline{\Omega}_h$ for all $v_j \in \mathbb{S}$; we stress that ρ need not be computed exactly. This is well defined for any $w \in C^0(\overline{\Omega})$, in particular, for $w \in \mathbb{V}_h$.

We seek $u_\varepsilon \in \mathbb{V}_h$ such that $u^\varepsilon(x_i) = g(x_i)$ for $x_i \in \mathcal{N}_h^b$ and for $x_i \in \mathcal{N}_h^0$

$$T_\varepsilon[u_\varepsilon](x_i) := \min_{\mathbf{v}\in\mathbb{S}^\perp} \left(\prod_{j=1}^{d} \nabla_\delta^{2,+} u_\varepsilon(x_i; v_j) - \sum_{j=1}^{d} \nabla_\delta^{2,-} u_\varepsilon(x_i; v_j) \right) = f(x_i), \tag{2.2}$$

where from now on we use the notation

$$\nabla_\delta^{2,+} u_\varepsilon(x_i; v_j) = \max\left(\nabla_\delta^2 u_\varepsilon(x_i; v_j), 0\right), \quad \nabla_\delta^{2,-} u_\varepsilon(x_i; v_j) = -\min\left(\nabla_\delta^2 u_\varepsilon(x_i; v_j), 0\right).$$

A similar definition was first proposed by Froese and Oberman in [23, 24] for a finite difference method. The key idea behind (2.2) is to enforce a suitable notion of discrete convexity. To build intuition we explore this concept next.

Definition 2.1 (Discrete convexity)**.** We say that $w_h \in \mathbb{V}_h$ is discretely convex if

$$\nabla_\delta^2 w_h(x_i; v_j) \geq 0 \qquad \forall x_i \in \mathcal{N}_h^0 \quad \forall v_j \in \mathbb{S}.$$

It is important to realize that this definition does not imply convexity in the usual sense, which is rather tricky to achieve with piecewise polynomials [1, 32, 37]. On the other hand, if $w \in C^0(\overline{\Omega}_h)$ is convex, then its Lagrange interpolant $\mathcal{I}_h w$ satisfies $\mathcal{I}_h w \geq w$, whence $\mathcal{I}_h w$ is discretely convex but not necessarily convex.

Lemma 2.2 (Discrete convexity)**.** *If $w_h \in \mathbb{V}_h$ satisfies*

$$T_\varepsilon[w_h](x_i) \geq 0 \quad \forall x_i \in \mathcal{N}_h^0,$$

then w_h is discretely convex and as a consequence

$$T_\varepsilon[w_h](x_i) = \min_{\mathbf{v}\in\mathbb{S}^\perp} \prod_{j=1}^{d} \nabla_\delta^2 w_h(x_i; v_j), \tag{2.3}$$

namely

$$\nabla_\delta^{2,+} w_h(x_i; v_j) = \nabla_\delta^2 w_h(x_i; v_j), \quad \nabla_\delta^{2,-} w_h(x_i; v_j) = 0 \quad \forall x_i \in \mathcal{N}_h^0 \quad \forall v_j \in \mathbb{S}.$$

Conversely, if w_h is discretely convex, then $T_\varepsilon[w_h](x_i) \geq 0$ for all $x_i \in \mathcal{N}_h^0$.

Proof. We distinguish two cases depending on whether $T_\varepsilon[w_h](x_i) > 0$ or not. Let $\mathbf{v} = (v_j)_{j=1}^d \in \mathbb{S}^\perp$ be a d-tuple that realizes the mininum in the definition of $T_\varepsilon[w_h](x_i)$ and note that

$$\prod_{j=1}^{d} \nabla_\delta^{2,+} w_h(x_i; v_j) \geq 0, \quad \sum_{j=1}^{d} \nabla_\delta^{2,-} w_h(x_i; v_j) \geq 0.$$

Case 1. $T_\varepsilon[w_h](x_i) > 0$. If the difference of these two quantities is positive, then so must be the first one. This implies that each factor $\nabla_\delta^{2,+} w_h(x_i; v_j) > 0$, whence the second term must vanish. This readily yields (2.3).

Case 2. $T_\varepsilon[w_h](x_i) = 0$. If instead the difference of the two quantities above vanishes, then there are two possible situations. If the first quantity is strictly positive, then the argument in Case 1 implies that the second quantity vanishes, which is a contradiction. Therefore, the alternative option is that both quantities vanish, whence

$$\nabla_\delta^{2,-} w_h(x_i; v_j) = 0 \quad \forall j \quad \Rightarrow \quad \nabla_\delta^2 w_h(x_i; v_j) \geq 0 \quad \forall j.$$

This again implies that w_h is discretely convex along with (2.3). Since the converse is trivial the proof is complete. □

2.2. **Practical two-scale method.** The ideal two-scale method of Section 2.1 leads to the notion of discrete convexity and Lemma 2.2 but cannot be implemented, because the minimum in (2.2) entails infinitely many options for $\mathbf{v} \in \mathbb{S}^\perp$. To render the two-scale method practical, we introduce a finite discretization $\mathbb{S}_\theta$ of the unit sphere $\mathbb{S}$ governed by the parameter θ: given $v \in \mathbb{S}$, there exists $v^\theta \in \mathbb{S}_\theta$ such that

$$|v - v^\theta| \le \theta.$$

Likewise, we approximate the set $\mathbb{S}^\perp$ of d-orthonormal bases by the finite set $\mathbb{S}_\theta^\perp$: for any $\mathbf{v}^\theta = (v_j^\theta)_{j=1}^d \in \mathbb{S}_\theta^\perp$, $v_j^\theta \in \mathbb{S}_\theta$ and there exists $\mathbf{v} = (v_j)_{j=1}^d \in \mathbb{S}^\perp$ such that $|v_j - v_j^\theta| \le \theta$ for all $1 \le j \le d$ and conversely. Note that the vectors $(v_j^\theta)_{j=1}^d$ are almost orthogonal but not necessarily orthogonal, which could provide flexibility in the implementation for dimensions $d > 2$.

If $\varepsilon := (h, \delta, \theta)$, the practical two-scale method now reads: seek $u_\varepsilon \in \mathbb{V}_h$ such that $u_\varepsilon(x_i) = g(x_i)$ for $x_i \in \mathcal{N}_h^b$ and for $x_i \in \mathcal{N}_h^0$

$$T_\varepsilon[u_\varepsilon](x_i) := \min_{\mathbf{v} \in \mathbb{S}_\theta^\perp} \left(\prod_{j=1}^d \nabla_\delta^{2,+} u_\varepsilon(x_i; v_j) - \sum_{j=1}^d \nabla_\delta^{2,-} u_\varepsilon(x_i; v_j) \right) = f(x_i). \tag{2.4}$$

We observe that if we relax Definition 2.1 (discrete convexity) to be

$$\nabla_\delta^2 w_h(x_i; v_j) \ge 0 \qquad \forall x_i \in \mathcal{N}_h^0 \quad \forall v_j \in \mathbb{S}_\theta,$$

then Lemma 2.2 (discrete convexity) is still valid and we can take

$$T_\varepsilon[w_h](x_i) = \min_{\mathbf{v} \in \mathbb{S}_\theta^\perp} \prod_{j=1}^d \nabla_\delta^2 w_h(x_i; v_j), \tag{2.5}$$

provided $T_\varepsilon[w_h](x_i) \ge 0$, as is the case of u_ε in (2.4).

We now show that (2.4) is monotone and prove a comparison principle for $f \ge 0$.

Lemma 2.3 (Monotonicity)**.** *Let $u_h, w_h \in \mathbb{V}_h$ be discretely convex. If $u_h - w_h$ attains a maximum at an interior node $z \in \mathcal{N}_h^0$, then*

$$T_\varepsilon[w_h](z) \ge T_\varepsilon[u_h](z).$$

Proof. If $u_h - w_h$ attains a maximum at $z \in \mathcal{N}_h^0$, then

$$u_h(z) - w_h(z) \ge u_h(x_i) - w_h(x_i) \quad \forall x_i \in \mathcal{N}_h.$$

For suitably chosen $0 < \rho \le 1$, the points $z \pm \delta\rho v_j \in \overline{\Omega}_h$ satisfy

$$u_h(z) - w_h(z) \ge u_h(z \pm \delta\rho v_j) - w_h(z \pm \delta\rho v_j) \quad \forall v_j \in \mathbb{S}_\theta,$$

because this relation holds at the vertices of the simplices where $z \pm \delta\rho v_j$ belong and both u_h and w_h are piecewise linear. Hence, (2.1) implies

$$\nabla_\delta^2 u_h(z; v_j) \le \nabla_\delta^2 w_h(z; v_j) \quad \forall v_j \in \mathbb{S}_\theta.$$

Since discrete convexity of u_h and w_h implies (2.5), the proof is complete. □

Lemma 2.4 (Discrete comparison principle)**.** *Let $u_h, w_h \in \mathbb{V}_h$ with $u_h \le w_h$ on the boundary $\partial\Omega_h$ be such that*

$$T_\varepsilon[u_h](x_i) \ge T_\varepsilon[w_h](x_i) \ge 0 \quad \forall x_i \in \mathcal{N}_h^0.$$

Then, $u_h \le w_h$ in Ω_h.

Proof. Since $u_h, w_h \in \mathbb{V}_h$, it suffices to prove $u_h(x_i) \le w_h(x_i)$ for all $x_i \in \mathcal{N}_h^0$. In view of Lemma 2.2 (discrete convexity) and (2.5), we realize that both u_h and w_h are discretely convex and we can rewrite the operator inequality as follows:

$$\min_{\mathbf{v}\in\mathbb{S}_\theta^\perp} \prod_{j=1}^{d} \nabla_\delta^2 u_h(x_i; v_j) \ge \min_{\mathbf{v}\in\mathbb{S}_\theta^\perp} \prod_{j=1}^{d} \nabla_\delta^2 w_h(x_i; v_j) \ge 0 \quad \forall x_i \in \mathcal{N}_h^0.$$

The proof splits into two steps according to whether this inequality is strict or not.

Step 1. We first consider the strict inequality

$$\min_{\mathbf{v}\in\mathbb{S}_\theta^\perp} \prod_{j=1}^{d} \nabla_\delta^2 u_h(x_i; v_j) > \min_{\mathbf{v}\in\mathbb{S}_\theta^\perp} \prod_{j=1}^{d} \nabla_\delta^2 w_h(x_i; v_j) \quad \forall x_i \in \mathcal{N}_h^0.$$

We assume by contradiction that there exists an interior node $x_k \in \mathcal{N}_h^0$ such that

$$u_h(x_k) - w_h(x_k) > 0$$

and

$$u_h(x_k) - w_h(x_k) \ge u_h(x_i) - w_h(x_i) \quad \forall x_i \in \mathcal{N}_h.$$

Reasoning as in Lemma 2.3 we obtain $\nabla_\delta^2 u_h(x_k; v_j) \le \nabla_\delta^2 w_h(x_k; v_j)$ for all $v_j \in \mathbb{S}_\theta$. On the other hand, the original strict inequality at $x_i = x_k$ yields

$$\min_{\mathbf{v}\in\mathbb{S}_\theta^\perp} \prod_{j=1}^{d} \nabla_\delta^2 w_h(x_k; v_j) < \prod_{j=1}^{d} \nabla_\delta^2 u_h(x_k; \overline{v}_j)$$

for all possible directions $\overline{\mathbf{v}} = (\overline{v}_j)_{j=1}^d \in \mathbb{S}_\theta^\perp$. Choosing $\overline{\mathbf{v}}$ to be a d-tuple that realizes the minimum of the left-hand side leads to a contradiction.

Step 2. We now deal with the nonstrict inequality. We introduce the quadratic strictly convex function $q(x) = \frac{1}{2}\big(|x|^2 - R^2\big)$, which satisfies $q \le 0$ on $\overline{\Omega}$ for $R > 0$ sufficiently large and in particular $q \le 0$ on $\partial\Omega_h$. Its Lagrange interpolant $q_h = \mathcal{I}_h q$ is discretely convex and

$$\nabla_\delta^2 q_h(x_i; v_j) \ge \nabla_\delta^2 q(x_i; v_j) = \partial_{v_j v_j}^2 q(x_i) = 1 \quad \forall x_i \in \mathcal{N}_h^0 \quad \forall v_j \in \mathbb{S}_\theta.$$

For arbitrary $\alpha > 0$, the function $u_h + \alpha q_h$ satisfies $u_h + \alpha q_h \le u_h \le w_h$ on $\partial\Omega_h$ and

$$\nabla_\delta^2 (u_h + \alpha q_h)(x_i; v_j) \ge \nabla_\delta^2 u_h(x_i; v_j) + \alpha > \nabla_\delta^2 w_h(x_i; v_j) \quad \forall x_i \in \mathcal{N}_h^0 \quad \forall v_j \in \mathbb{S}_\theta,$$

whence $T_\varepsilon[u_h + \alpha q_h](x_k) > T_\varepsilon[w_h](x_k)$. Applying Step 1 we deduce

$$u_h + \alpha q_h \le w_h \quad \forall \alpha > 0.$$

Taking the limit as $\alpha \to 0$ gives the asserted inequality.

□

3. Existence and uniqueness

We now prove existence and uniqueness of a discrete solution $u_\varepsilon \in \mathbb{V}_h$ of (2.4).

Lemma 3.1 (Existence, uniqueness, and stability)**.** *There exists a unique $u_\varepsilon \in \mathbb{V}_h$ that solves the discrete Monge-Ampère equation* (2.4)*. The solution u_ε is stable in the sense that $\|u_\varepsilon\|_{L^\infty(\Omega)}$ does not depend on the parameters $\epsilon = (h, \delta, \theta)$ of the method.*

Proof. Since uniqueness is a trivial consequence of Lemma 2.4 (discrete comparison principle), we just have to prove existence. To this end, we construct a monotone sequence of discrete convex functions $\{u_h^k\}_{k=0}^{\infty}$, starting with the initial iterate $u_h^0 \in \mathbb{V}_h$ that satisfies $u_h^0 = \mathcal{I}_h g$ on $\partial\Omega_h$ and

$$T_\varepsilon[u_h^0](x_i) \geq f(x_i) \quad \forall x_i \in \mathcal{N}_h^0.$$

Step 1 (Existence of u_h^0). We repeat the calculations of Step 2 in Lemma 2.4 (discrete comparison principle) for

$$q(x) = \frac{1}{2}\|f\|_{L^\infty(\Omega)}^{1/d}|x|^2$$

to obtain that for $q_h = \mathcal{I}_h q$ and for all $x_i \in \mathcal{N}_h^0$

$$T_\varepsilon[q_h](x_i) \geq \|f\|_{L^\infty(\Omega)} \geq f(x_i).$$

We utilize the stability of $\mathcal{I}_h q$ in $L^\infty(\Omega_h)$ to deduce

$$\|q_h\|_{L^\infty(\Omega_h)} \leq C_R \, \|f\|_{L^\infty(\Omega)},$$

where C_R is a geometric constant that depends on the domain Ω.

We next observe that the set of convex functions w satisfying a continuous Dirichlet boundary condition on a uniformly convex domain is nonempty. The solution $w \in C^0(\overline{\Omega})$ of the homogeneous Dirichlet problem (1.1) is one such function [26, Theorem 1.5.2]. Let w be convex and solve (1.1) with $f = 0$ and Dirichlet condition $w = g - q$, whence $w_h := \mathcal{I}_h w$ satisfies

$$T_\varepsilon[w_h](x_i) \geq 0 \quad \forall x_i \in \mathcal{N}_h^0$$

and $w_h = \mathcal{I}_h g - q_h$ on $\mathcal{N}_h^b$. We define the initial iterate to be

$$u_h^0 := w_h + q_h$$

and note that u_h^0 is discretely convex and satisfies the Dirichlet condition $u_h^0 = \mathcal{I}_h g$ on $\partial\Omega_h$. Since all the terms in $T_\varepsilon[u_h^0](x_i)$ are nonnegative, we also deduce

$$T_\varepsilon[u_h^0](x_i) = \min_{\mathbf{v}\in\mathbb{S}_\theta^+} \prod_{j=1}^{d} \Big(\nabla_\delta^2 w_h(x_i; v_j) + \nabla_\delta^2 q_h(x_i; v_j)\Big) \geq f(x_i) \quad \forall x_i \in \mathcal{N}_h^0.$$

Step 2 (Perron construction). We proceed by induction. Suppose that we have already a discretely convex function $u_h^k \in \mathbb{V}_h$ that satisfies $u_h^k = \mathcal{I}_h g$ on $\partial\Omega_h$ and

$$T_\varepsilon[u_h^k](x_i) \geq f(x_i) \quad \forall x_i \in \mathcal{N}_h^0. \tag{3.1}$$

We now construct $u_h^{k+1} \in \mathbb{V}_h$ such that $u_h^{k+1} \geq u_h^k$ in Ω_h, $u_h^{k+1} = \mathcal{I}_h g$ on $\partial\Omega_h$ and satisfies (3.1). We consider all interior nodes in order and construct auxiliary functions $u_h^{k,i-1} \in \mathbb{V}_h$ using the first $i-1$ nodes and starting from $u_h^{k,0} := u_h^k$ as follows. At $x_i \in \mathcal{N}_h^0$ we check whether or not $T^\varepsilon[u_h^{k,i-1}](x_i) > f(x_i)$. If so, we increase the value of $u_h^{k,i-1}(x_i)$ and denote the resulting function by $u_h^{k,i}$, until

$$T_\varepsilon[u_h^{k,i}](x_i) = f(x_i).$$

This is possible because the centered second differences (2.2) are strictly decreasing with increasing central value for all directions. Expression (2.2) also shows that this process potentially increases the centered second differences at other nodes $x_j \neq x_i$, whence

$$T_\varepsilon[u_h^{k,i}](x_j) \geq T_\varepsilon[u_h^{k,i-1}](x_j) \geq f(x_j) \quad \forall x_j \neq x_i.$$

We repeat this process with the remaining nodes x_j for $i < j \le N$, and set $u_h^{k+1} := u_h^{k,N}$ to be the last intermediate function. By construction, we clearly obtain

$$T_\varepsilon[u_h^{k+1}](x_i) \ge f(x_i), \quad u_h^{k+1}(x_i) \ge u_h^k(x_i) \quad \forall x_i \in \mathcal{N}_h^0.$$

Our construction preserves the boundary values $u_h^{k+1} = \mathcal{I}_h g$ on $\partial\Omega_h$ and enforces the relation $u_h^{k+1} \ge u_h^k$ in Ω_h because both u_h^{k+1}, u_h^k are piecewise linear functions.

Step 3 (Bounds). If $b_h := \max_{x_i \in \mathcal{N}_h^b} g(x_i)$, then we see that $b_h \in \mathbb{V}_h$ and

$$T_\varepsilon[b_h](x_i) = 0 \le f(x_i) \le T_\varepsilon[u_h^k](x_i) \quad \forall x_i \in \mathcal{N}_h^0 \quad \forall k \ge 0.$$

We apply Lemma 2.4 (discrete comparison principle) to infer that $u_h^k \le b_h$ for all $k \ge 0$. On the other hand, since $\|u_h^0\|_{L^\infty(\Omega_h)}$ is bounded uniformly in h and $u_h^0 \le u_h^k$, we deduce the uniform bound

$$\|u_h^k\|_{L^\infty(\Omega)} \le \Lambda$$

with $\Lambda > 0$ independent of the discretization parameters h, δ, and θ.

Step 4 (Convergence). The sequence $\{u_h^k(x_i)\}_{k=1}^\infty$ is monotone and bounded above for all $x_i \in \mathcal{N}_h^0$, and hence converges. The limit

$$u_\varepsilon(x_i) = \lim_{k\to\infty} u_h^k(x_i) \quad \forall x_i \in \mathcal{N}_h^0$$

defines $u_\varepsilon \in \mathbb{V}_h$ and satisfies $u_\varepsilon = \mathcal{I}_h g$ on $\partial\Omega_h$. It also satisfies the desired equality

$$T_\varepsilon[u_\varepsilon](x_i) = f(x_i) \quad \forall x_i \in \mathcal{N}_h^0,$$

since $T_\varepsilon[u_\varepsilon](x_i) = \lim_{k\to\infty} T_\varepsilon[u_h^k](x_i) \ge f(x_i)$ and if the last inequality were strict, then Step 2 could be applied to improve u_ε. This shows existence of a discrete solution u_ε of (2.2) as well as the uniform bound $\|u_\varepsilon\|_{L^\infty(\Omega)} \le \Lambda$.

□

4. Consistency

We now quantify the operator consistency error in terms of Hölder regularity of $D^2 u$. We start with the definitions of δ-interior region

$$\Omega_{h,\delta} = \{T \in \mathcal{T}_h \ : \ \operatorname{dist}(x, \partial\Omega_h) \ge \delta \ \forall x \in T\}, \tag{4.1}$$

and δ-boundary region

$$\omega_{h,\delta} = \Omega_h \setminus \Omega_{h,\delta}.$$

Moreover, given a node $x_i \in \mathcal{N}_h^0$ we denote by

$$B_i := \cup\{\overline{T} : T \in \mathcal{T}_h, \text{ dist } (x_i, T) \le \hat{\delta}\}, \tag{4.2}$$

where $\hat{\delta} := \rho\delta$ with $0 < \rho \le 1$ is the largest number so that $x_i \pm \hat{\delta} v_j \in \overline{\Omega}_h$ for all $v_j \in \mathbb{S}_\theta$.

Lemma 4.1 (Consistency of $\nabla_\delta^2 \mathcal{I}_h u$)**.** *Let $u \in W_\infty^2(B_i)$, $\mathcal{I}_h u$ be its Lagrange interpolant in Ω_h, and let B_i be defined in* (4.2)*. The following two estimates are then valid:*

(i) *For all $x_i \in \mathcal{N}_h^0$ and all $v_j \in \mathbb{S}_\theta$, we have*

$$\left|\nabla_\delta^2 \mathcal{I}_h u(x_i; v_j)\right| \le C |u|_{W_\infty^2(B_i)}.$$

(ii) *If in addition* $u \in C^{2+k,\alpha}(B_i)$ *for* $k = 0, 1$ *and* $\alpha \in (0, 1]$, *then for all* $x_i \in \mathcal{N}_h^0 \cap \Omega_{h,\delta}$ *and all* $v_j \in \mathbb{S}_\theta$, *we have*

$$\left|\nabla_\delta^2 \mathcal{I}_h u(x_i; v_j) - \frac{\partial^2 u}{\partial v_j^2}(x_i)\right| \leq C\left(|u|_{C^{2+k,\alpha}(B_i)}\delta^{k+\alpha} + |u|_{W^2_\infty(B_i)}\frac{h^2}{\delta^2}\right).$$

In both cases C *stands for a constant independent of the two scales* h *and* δ, *the parameter* θ *and* u.

Proof. We split the proof into three steps.

Step 1. Let $x_i \in \mathcal{N}_h^0$ and $v_j \in \mathbb{S}_\theta$. Since

$$u(x_i + \hat{\delta} v_j) - u(x_i) = \hat{\delta} \int_0^1 \nabla u(x_i + t\hat{\delta} v_j) \cdot v_j \, dt,$$

definition (2.1) yields

$$\nabla_\delta^2 u(x_i; v_j) = \hat{\delta}^{-1} \int_0^1 \Big(\nabla u(x_i + t\hat{\delta} v_j) - \nabla u(x_i - t\hat{\delta} v_j)\Big) \cdot v_j \, dt.$$

Adding and subtracting $\nabla u(x_i) \cdot v_j$ inside the integral, we similarly arrive at

$$\nabla_\delta^2 u(x_i; v_j) = \int_0^1 \int_0^1 t \, \Big(D^2 u(x_i + st\hat{\delta} v_j) + D^2 u(x_i - st\hat{\delta} v_j)\Big) : v_j \otimes v_j \, ds \, dt,$$

which implies

$$\left|\nabla_\delta^2 u(x_i; v_j)\right| \leq |u|_{W^2_\infty(B_i)}.$$

Step 2. Let $x_i \in \Omega_{h,\delta}$ and assume that $u \in C^{2,\alpha}(B_i)$. We prove the estimate

$$\left|\nabla_\delta^2 u(x_i; v_j) - \frac{\partial^2 u}{\partial v_j^2}(x_i)\right| \leq C|u|_{C^{2,\alpha}(B_i)} \, \delta^\alpha.$$

Write $\nabla_\delta^2 u(x_i; v_j) = I_1 + I_2$, where

$$I_1 = 2\int_0^1 \int_0^1 t D^2 u(x_i) : v_j \otimes v_j \, ds \, dt = \frac{\partial^2 u}{\partial v_j^2}(x_i)$$

and

$$I_2 = \int_0^1 \int_0^1 t \left(D^2 u(x_i + st \, \delta v_j) - 2D^2 u(x_i) + D^2 u(x_i - st \, \delta v_j)\right) : v_j \otimes v_j \, ds \, dt.$$

The fact that $u \in C^{2,\alpha}(B_i)$ gives

$$|D^2 u(x_i \pm st\delta v_j) - D^2 u(x_i)| \leq C|u|_{C^{2,\alpha}(B_i)}\delta^\alpha,$$

whence

$$I_2 \leq C|u|_{C^{2,\alpha}(B_i)}\delta^\alpha.$$

Combining I_1 and I_2, we deduce the asserted estimate for $u \in C^{2,\alpha}(B_i)$ and $k = 0$. For $u \in C^{3,\alpha}(B_i)$, we exploit the symmetry of I_2 to express the integrand in terms of differences of $D^3 u$ at points $x_i \pm stz \, \delta v_j$ for $0 < z < 1$ and thus deduce

$$I_2 \leq C|u|_{C^{3,\alpha}(B_i)}\delta^{1+\alpha}.$$

This implies the estimate for $k = 1$

$$\left|\nabla_\delta^2 u(x_i; v_j) - \frac{\partial^2 u}{\partial v_j^2}(x_i)\right| \leq C|u|_{C^{3,\alpha}(B_i)}\delta^{1+\alpha}.$$

Step 3. We now study the effect of interpolation, for which it is known that [9]

$$\|u - \mathcal{I}_h u\|_\infty \leq C\ |u|_{W^2_\infty(B_i)} h^2.$$

Therefore, applying definition (2.1), we deduce for $x_i \in \Omega_{h,\delta}$

$$|\nabla^2_\delta (u - \mathcal{I}_h u)(x_i; v_j)| \leq C |u|_{W^2_\infty(B_i)} \frac{h^2}{\delta^2}.$$

This completes the proof of (ii) for $k = 0, 1$. Otherwise, δ must be replaced by $\hat{\delta} = \rho\delta \geq Ch$ with $C > 0$ depending only on shape regularity. Therefore, we see that $h^2 \hat{\delta}^{-2} \leq C$, which combined with Step 1 yields the estimate in (i).

□

We now extend the consistency analysis to the practical two-scale operator T_ϵ.

Lemma 4.2 (Consistency of $T_\varepsilon[\mathcal{I}_h u]$)**.** *Let $x_i \in \mathcal{N}_h^0 \cap \Omega_{h,\delta}$ and B_i be defined as in* (4.2)*. If $u \in C^{2+k,\alpha}(B_i)$ is convex with $0 < \alpha \leq 1$ and $k = 0, 1$, and $\mathcal{I}_h u$ is its piecewise linear interpolant, then*

$$\left|\det D^2 u(x_i) - T_\varepsilon[\mathcal{I}_h u](x_i)\right| \leq C_1(d, \Omega, u)\delta^{k+\alpha} + C_2(d, \Omega, u)\left(\frac{h^2}{\delta^2} + \theta^2\right), \tag{4.3}$$

where

$$C_1(d, \Omega, u) = C|u|_{C^{2+k,\alpha}(B_i)} |u|^{d-1}_{W^2_\infty(B_i)}, \quad C_2(d, \Omega, u) = C|u|^d_{W^2_\infty(B_i)}.$$

If $x_i \in \mathcal{N}_h^0$ and $u \in W^2_\infty(B_i)$, then (4.3) *remains valid with $\alpha = k = 0$ and $C^{2+k,\alpha}(B_i)$ replaced by $W^2_\infty(B_i)$.*

Proof. We recall that $\mathcal{I}_h u$ is discretely convex, namely $\nabla^2_\delta \mathcal{I}_h u(x_i, v_j) \geq 0$ for all $x_i \in \mathcal{N}_h^0$ and $v_j \in \mathbb{S}_\theta$, because u is convex. Therefore, in view of Lemma 2.2 (discrete convexity), the definition of $T_\varepsilon[\mathcal{I}_h u]$ reduces to

$$T_\varepsilon[\mathcal{I}_h u](x_i) = \min_{\mathbf{v} \in \mathbb{S}^\perp_\theta} \prod_{j=1}^d \nabla^2_\delta \mathcal{I}_h u(x_i; v_j).$$

Step 1. Let $\mathbf{v} = (v_j)_{j=1}^d \in \mathbb{S}^\perp_\theta$ be the d-tuple that realizes the above minimum. Applying (1.2) to the determinant of the Hessian of u, we see that

$$\det D^2 u(x_i) - T_\varepsilon[\mathcal{I}_h u](x_i) \leq \prod_{j=1}^d \frac{\partial^2 u}{\partial v_j^2}(x_i) - \prod_{j=1}^d \nabla^2_\delta \mathcal{I}_h u(x_i; v_j).$$

We now invoke Lemma 4.1 (ii) (consistency of $\nabla^2_\delta \mathcal{I}_h u$) to write

$$\left| \frac{\partial^2 u}{\partial v_j^2}(x_i) - \nabla^2_\delta \mathcal{I}_h u(x_i; v_j) \right| \leq C|u|_{C^{2+k,\alpha}(B_i)} \delta^{k+\alpha} + C|u|_{W^2_\infty(B_i)} \frac{h^2}{\delta^2},$$

where $k = 0, 1$. Given the multiplicative structure above, utilizing Lemma 4.1 (i) we deduce

$$\det D^2 u(x_i) - T_\varepsilon[\mathcal{I}_h u](x_i) \leq C_1(d, \Omega, u)\delta^{k+\alpha} + C_2(d, \Omega, u)\frac{h^2}{\delta^2},$$

where C_1 and C_2 are defined above.

Step 2. We now choose $\mathbf{v} = (v_j)_{j=1}^d \in \mathbb{S}^\perp$ to be the d-tuple that realizes the minimum in (1.2) for $\det D^2u(x_i)$. We can then write

$$T_\varepsilon[\mathcal{I}_h u](x_i) - \det D^2u(x_i) \le I_1 + I_2,$$

where

$$I_1 = \prod_{j=1}^d \nabla_\delta^2 \mathcal{I}_h u(x_i;\hat{v}_j) - \prod_{j=1}^d \frac{\partial^2 u}{\partial \hat{v}_j^2}(x_i), \qquad I_2 = \prod_{j=1}^d \frac{\partial^2 u}{\partial \hat{v}_j^2}(x_i) - \prod_{j=1}^d \frac{\partial^2 u}{\partial v_j^2}(x_i),$$

and $\hat{\mathbf{v}} = (\hat{v}_j)_{j=1}^d \in \mathbb{S}_\theta^\perp$ is an approximation of $\mathbf{v}$ satisfying $|v_j - \hat{v}_j| \le \theta$ for all $1 \le j \le d$. The first term I_1 obeys a similar estimate to Step 1. For the second term I_2 we notice that $\hat{v}_j = v_j + w_j$ with $|w_j| \le \theta$, whence

$$\frac{\partial^2 u}{\partial \hat{v}_j^2}(x_i) = \hat{v}_j^T D^2u(x_i)\hat{v}_j = \frac{\partial^2 u}{\partial v_j^2}(x_i) + 2w_j^T D^2u(x_i)v_j + w_j^T D^2u(x_i)w_j.$$

Using that $\hat{v}_j = v_j + w_j$, we observe that

$$1 = |\hat{v}_j|^2 = |v_j|^2 + 2v_j \cdot w_j + |w_j|^2 \qquad \Rightarrow \qquad |v_j \cdot w_j| = \frac{1}{2}|w_j|^2 \le \frac{1}{2}\theta^2.$$

Since $D^2u(x_i)v_j = \lambda_j v_j$, we thus obtain

$$\left| \frac{\partial^2 u}{\partial \hat{v}_j^2}(x_i) - \frac{\partial^2 u}{\partial v_j^2}(x_i) \right| \le C\ \theta^2\ |u|_{W^2_\infty(B_i)}$$

as well as

$$I_2 \le C\ \theta^2\ |u|^d_{W^2_\infty(B_i)}.$$

This proves (4.3).

The remaining statement for $u \in W^2_\infty(B_i)$ is a simple consequence of Lemma 4.1 (i) and the above 2-step argument.

□

Remark 4.3 (Regularity). We give sufficient conditions for the regularity of u in Lemma 4.2: if $0 < f_0 \le f(x) \le f_1$ for all $x \in \Omega$ and $f \in C^\alpha(\overline{\Omega}), g \in C^3(\overline{\Omega})$, and $\partial\Omega \in C^3$, then $u \in C^{2,\alpha}(\overline{\Omega})$ [36, Theorem 1.1]. In such a case, there exist $0 < \lambda \le \Lambda < \infty$ depending on f, g, and Ω such that [17, Theorem 2.10]

$$\lambda I \le D^2u(x) \le \Lambda I \quad \forall x \in \Omega.$$

Since $\left|\frac{\partial^2 u}{\partial v_j^2}(x_i)\right| \le \Lambda$, the constants C_1 and C_2 in Lemma 4.2 could also be written

$$C_1(d,\Omega,u) = C\Lambda^{d-1}|u|_{C^{2+k,\alpha}(B_i)}, \quad C_2(d,\Omega,u) = C\Lambda^{d-1}|u|_{W^2_\infty(B_i)}.$$

5. Convergence

Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u]$) shows interior consistency at distance δ to $\partial\Omega_h$ for $u \in C^2(\overline{\Omega})$; hence the Barles-Souganidis theory [4] does not apply directly, as stated in [20]. We compensate with the fact that $\mathcal{I}_h u - u_\varepsilon$ vanishes on $\partial\Omega_h$ and cannot grow faster than $C\delta$ at distance δ to $\partial\Omega_h$. We make this statement rigorous via a barrier argument similar to those in [20, 31, 32]. To handle the behavior of $u - u_\varepsilon$ inside Ω_h we utilize Lemma 2.4 (discrete comparison principle) and Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u]$). In both cases we need the solution to be $C^2(\overline{\Omega})$, which may in general be false for the viscosity solution and thus requires

a regularization argument involving data (f, g, Ω). We discuss these topics in this section and give a variation of the Barles-Souganidis approach as well.

5.1. **Barrier functions.** We now introduce two discrete barrier functions, one to deal with the boundary behavior and the other one to handle the interior behavior.

Lemma 5.1 (Discrete boundary barrier)**.** *Let Ω be uniformly convex and $E > 0$ be arbitrary. For each node $z \in \mathcal{N}_h^0$ with $\operatorname{dist}(z, \partial\Omega_h) \le \delta$, there exists a function $p_h \in \mathbb{V}_h$ such that $T_\varepsilon[p_h](x_i) \ge E$ for all $x_i \in \mathcal{N}_h^0$, $p_h \le 0$ on $\partial\Omega_h$ and*

$$|p_h(z)| \le C E^{1/d} \delta$$

with C depending on Ω.

Proof. Take $z_1 \in \partial\Omega_h$ such that $|z - z_1| = \operatorname{dist}(z, \partial\Omega_h) \le \delta$. Upon extending the segment joining z and z_1, we find $z_2 \in \partial\Omega$ that satisfies the upper bound $|z_2 - z_1| \le C_1 h$ because Ω is uniformly convex and thus Lipschitz but not necessarily W^2_∞. This implies that for $z_0 \in \partial\Omega$ such that $|z - z_0| = \operatorname{dist}(z, \partial\Omega)$, we have $|z - z_0| \le |z - z_2| \le \delta + C_1 h \le C_2 \delta$. We now make a change of coordinates so that z_0 becomes the origin and $z = (0, \dots, 0, |z - z_0|)$. Since Ω is uniformly convex, it lies inside the ball

$$x_1^2 + x_2^2 + \ldots + x_{d-1}^2 + (x_d - R)^2 \le R^2,$$

where the radius R depends on $\partial\Omega$ which is not necessarily W^2_∞. Under this coordinate system, let $p(x)$ be the quadratic polynomial

$$p(x) = \frac{E^{1/d}}{2} \left(x_1^2 + x_2^2 + \ldots + x_{d-1}^2 + (x_d - R)^2 - R^2 \right)$$

and let $p_h = \mathcal{I}_h p$ be its piecewise linear Lagrange interpolant in Ω_h. We note that $p \le 0$ on $\overline{\Omega}$ yields $p_h \le 0$ on $\partial\Omega_h$. Since p is convex and $\mathcal{I}_h p \ge p$, we infer that

$$T_\varepsilon[p_h](x_i) \ge T_\varepsilon[p](x_i) = E \quad \forall\ x_i \in \mathcal{N}_h^0,$$

where the last equality is a consequence of p being quadratic and

$$\nabla^2_\delta p(x_i; v_j) = \partial^2_{v_j v_j} p(x_i) = E^{1/d} \quad \forall\ v_j \in \mathbb{S}_\theta.$$

Moreover, since $|z - z_0| \le C_2 \delta$, we deduce $|p_h(z)| \le C_\Omega E^{1/d} \delta$, as asserted. □

The following barrier function q_h and corresponding statement have already been used in the proof of Lemma 2.4 (discrete comparison principle).

Lemma 5.2 (Discrete interior barrier)**.** *Let Ω be contained in the ball $B(x_0, R)$ of center x_0 and radius R. If $q(x) := \frac{1}{2}\big(|x - x_0|^2 - R^2\big)$, then its interpolant $q_h := \mathcal{I}_h q \in \mathbb{V}_h$ satisfies*

$$T_\varepsilon[q_h](x_i) \ge 1 \quad \forall x_i \in \mathcal{N}_h^0, \qquad q_h(x_i) \le 0 \quad \forall x_i \in \mathcal{N}_h^b.$$

5.2. **Approximation by smooth problems.** For data f, g uniformly continuous in Ω, $f \ge 0$, and Ω uniformly convex, the regularity $u \in C^2(\overline{\Omega})$ which would yield small interior consistency error is not guaranteed. We thus embark on a regularization procedure similar to that used by DePhilippis-Figalli [16] and Awanou [3]. We start with a result about continuous dependence on data for viscosity solutions.

Lemma 5.3 (Continuous dependence on data)**.** *Given $f_1, f_2 \in C(\overline{\Omega})$, $f_1, f_2 \geq 0$, and $g_1, g_2 \in C(\partial\Omega)$, let $u_1, u_2 \in C(\overline{\Omega})$ be the corresponding convex viscosity solutions of* (1.1)*. Then there exists a constant C depending on Ω such that*

$$\|u_1 - u_2\|_{L^\infty(\Omega)} \leq C\|f_1 - f_2\|^{1/d}_{L^\infty(\Omega)} + \|g_1 - g_2\|_{L^\infty(\partial\Omega)}.$$

Proof. Let $q \leq 0$ be the barrier function of Lemma 5.2 (discrete interior barrier) and $F := \|f_1 - f_2\|^{1/d}_{L^\infty(\Omega)}$, $G := \|g_1 - g_2\|_{L^\infty(\partial\Omega)}$. We consider the auxiliary function

$$u_1^- := u_2 + Fq - G,$$

which is a convex viscosity subsolution of (1.1) with data (f_1, g_1). To prove this, let $\phi \in C^2(\Omega)$ and $x_0 \in \Omega$ be a point where $u_1^- - \phi$ attains a maximum. This implies that $u_2 - \big(\phi - Fq + G\big)$ attains also a maximum at x_0. Since u_2 is a viscosity subsolution of (1.1), and $D^2q(x_0) = I$ is the identity matrix, we deduce

$$\det\big(D^2\phi(x_0) - FI\big) \geq f_2(x_0) \geq 0.$$

Formula (1.2) for two symmetric positive semi-definite matrices A, B easily implies

$$\det(A + B) \geq \det(A) + \det(B).$$

Using this expression for $A = D^2\phi(x_0) - FI$ and $B = FI$ we obtain

$$\det(D^2\phi(x_0)) \geq f_2(x_0) + F^d = f_2(x_0) + \|f_1 - f_2\|_{L^\infty(\Omega)} \geq f_1(x_0).$$

In addition, since $q \leq 0$ in Ω, the function u_1^- satisfies on $\partial\Omega$

$$u_1^- \leq u_2 - G = g_2 - \|g_1 - g_2\|_{L^\infty(\partial\Omega)} \leq g_1.$$

These two properties of u_1^- imply that u_1^- is a viscosity subsolution of (1.1) with data (f_1, g_1). Since u_1^- is also convex, the comparison principle for (1.1) gives

$$u_1^- \leq u_1 \quad \Rightarrow \quad u_2 - u_1 \leq -Fq + G.$$

We similarly prove the reverse inequality, thus obtaining the desired estimate. □

We stress the monotonicity estimate

$$f_1 \geq f_2 \geq 0, \quad g_1 \leq g_2 \qquad \Rightarrow \qquad u_1 \leq u_2,$$

which is a consequence of u_1 being a convex subsolution of (1.1) with data (f_2, g_2).

Using the above result, we now show that we can approximate a viscosity solution u of (1.1) by regular (classical) solutions u_n.

Lemma 5.4 (Approximation of viscosity solutions by smooth solutions)**.** *Let Ω be uniformly convex, let f, g be uniformly continuous in Ω, $f \geq 0$, and let u be the viscosity solution of* (1.1) *with data (f, g, Ω). Then, there exist a decreasing sequence of uniformly convex and smooth domains Ω_n converging to Ω in the sense that the Hausdorff distance $\operatorname{dist}_H(\Omega_n, \Omega) \to 0$, a decreasing sequence of smooth functions $f_n > 0$ such that $f_n \to f$ uniformly in Ω, a sequence of smooth functions g_n such that $g_n \to g$ uniformly in Ω, and a sequence of smooth classical solutions u_n of* (1.1) *with data (f_n, g_n, Ω_n) such that $u_n \to u$ uniformly in Ω as $n \to \infty$.*

Proof. We prove the result in four steps.

Step 1 (Domain approximation). According to [8] there is a sequence of smooth and uniformly convex domains $\widetilde{\Omega}_n \subset \Omega$ that increase to Ω in the sense that the Hausdorff distance $\mathrm{dist}_H(\widetilde{\Omega}_n, \Omega) \to 0$. Since Ω is convex, it is star-shaped with respect to any of its points. Let's assume that the origin is contained in Ω and dilate the domains $\widetilde{\Omega}_n$ so that the ensuing domains Ω_n satisfy:

$$\Omega \subset \Omega_n \subset \Omega_m \quad m \le n; \qquad \mathrm{dist}_H(\Omega_n, \Omega) \to 0 \quad n \to \infty.$$

The domains Ω_n inherit the regularity of $\widetilde{\Omega}_n$ as well as their uniform convexity. Given $\delta_n \to 0$ as $n \to \infty$, to be chosen later in Step 4, we relabel Ω_n to be an approximate smooth domain so that $\mathrm{dist}_H(\Omega_n, \Omega) \le \delta_n$.

Step 2 (Data regularization). Let $\widetilde{\Omega}$ be an auxiliary domain such that $\Omega_n \subset \widetilde{\Omega}$ for all n. We now construct a sequence (f_n, g_n) of smooth functions defined in $\widetilde{\Omega}$ that converge uniformly in Ω to (f, g). We first extend (f, g) to $\widetilde{\Omega}$ and let $\sigma(t)$ be the modulus of continuity in $\widetilde{\Omega}$ for both (f, g) [18, Theorem 2.1.8.]:

$$|f(x) - f(y)|, |g(x) - g(y)| \le \sigma(|x - y|) \quad \forall x, y \in \widetilde{\Omega}.$$

Let $\rho < \mathrm{dist}_H(\Omega_n, \widetilde{\Omega})$ and let $\phi_\rho \ge 0$ be a standard smooth mollifier function with support in $B(0, \rho)$. We have for $f_\rho = f * \phi_\rho$ that

$$|f_\rho(x) - f(x)| = \left| \int_{\widetilde{\Omega}} (f(x - y) - f(x))\ \phi_\rho(y)\ dy \right| \le \sigma(\rho) \quad \forall x \in \Omega_n$$

because ϕ_ρ integrates to one. This implies that

$$\widetilde{f}_\rho(x) := f_\rho(x) + 2\sigma(\rho) \ge f(x) - \sigma(\rho) + 2\sigma(\rho) = f(x) + \sigma(\rho) > 0 \quad \forall x \in \Omega_n.$$

We now take $\rho_1 \le \rho_2$ and observe that for all $x \in \Omega_n$

$$\begin{aligned} \widetilde{f}_{\rho_1}(x) - \widetilde{f}_{\rho_2}(x) &= \big(f_{\rho_1}(x) + 2\sigma(\rho_1)\big) - \big(f_{\rho_2}(x) + 2\sigma(\rho_2)\big) \\ &\le f(x) + \sigma(\rho_1) + 2\sigma(\rho_1) - f(x) + \sigma(\rho_2) - 2\sigma(\rho_2) = 3\sigma(\rho_1) - \sigma(\rho_2) \le 0, \end{aligned}$$

if $\sigma(\rho_1) \le \frac{\sigma(\rho_2)}{3}$. We thus choose ρ_n such that $\sigma_n = \sigma(\rho_n) = 4^{-n}$ and define $f_n := \widetilde{f}_{\rho_n}$, which is a strictly positive and decreasing sequence of functions satisfying the error estimate

$$\sigma_n \le f_n(x) - f(x) \le 3\sigma_n \quad \forall\, x \in \Omega_n. \tag{5.1}$$

Similarly, we regularize g by convolution $g_\rho = g * \phi_\rho$ and define $g_n := g_{\rho_n}$ to obtain

$$\|g - g_n\|_{L^\infty(\Omega_n)} \le \sigma_n. \tag{5.2}$$

Step 3 (Boundary behavior). Let u_n be the smooth classical solution of (1.1) with data (f_n, g_n, Ω_n), which satisfies $u_n \in C^{2,\alpha}(\Omega_n)$ with norms depending on n but uniform α; this is possible because (f_n, g_n, Ω_n) are smooth, Ω_n is uniformly convex, and $f_n > 0$ [11] [36, Theorem 1.1].

We now compare g and u_n at $z \in \partial\Omega$ without invoking any regularity of u_n but rather using a barrier argument. We start with g: if $y \in \partial\Omega_n$ is the closest point to z, then $|z - y| \le \delta_n$ and

$$|g(z) - g(y)| \le \sigma(|z - y|) \le \sigma(\delta_n).$$

On the other hand, we know that

$$|g(y) - g_n(y)| \le \sigma(\rho_n).$$

Let p be the quadratic barrier function introduced in the proof of Lemma 5.1, but now associated with Ω_n and $y \in \partial\Omega_n$. We consider the (lower) barrier function

$$b_n^-(x) := p(x) + g_n(y) + \nabla g_n(y)(x-y),$$

which satisfies

$$\det D^2 b_n^- = \det D^2 p \geq f_n \quad \text{in } \Omega_n$$

for $E > \|f\|_{L^\infty(\widetilde{\Omega})}$ because b_n^- is a linear correction of p. We assert that $b_n^- \leq g_n$ on $\partial\Omega_n$ provided E also satisfies $E \geq C\|g_n\|_{W^2_\infty(\Omega_n)}$ where C depends on the uniform convexity of Ω. If this is true, then applying the comparison principle [26, Theorem 1.4.6] to the smooth functions b_n^- and u_n with data (f_n, g_n, Ω_n) yields

$$b_n^-(x) \leq u_n(x) \quad \forall\, x \in \Omega_n.$$

Taking $x = z$ and making use of the definition of b_n^- results in

$$-CE^{1/d}|z-y| + g_n(y) + \nabla g_n(y)(z-y) \leq u_n(z),$$

whence

$$u_n(z) - g_n(y) \geq -C_n|y-z| \geq -C_n\delta_n.$$

Similarly, upon letting $b_n^+(x) := -p(x) + g_n(y) + \nabla g_n(y)(x-y)$ be an upper barrier function, the preceding argument also shows

$$u_n(z) - g_n(y) \leq C_n|y-z| \leq C_n\delta_n,$$

whence the triangle inequality implies that for all $z \in \partial\Omega$

$$|g(z) - u_n(z)| \leq \sigma(\delta_n) + \sigma(\rho_n) + C_n\delta_n, \tag{5.3}$$

where the constant C_n depends on g_n but is independent of u_n. It remains to show

$$b_n^-(x) \leq g_n(x) \quad \forall\, x \in \partial\Omega_n.$$

We first observe that $b_n^-(y) = g_n(y)$ and the tangental gradients $\nabla_{\partial\Omega} b_n^-(y) = \nabla_{\partial\Omega} g_n(y)$ by construction, but g_n grows quadratically away from y on $\partial\Omega_n$ whereas p is just negative on $\partial\Omega_n$. To quantify the last statement, we let $y = 0$ for simplicity and resort to the uniform convexity of Ω (and thus to that of every Ω_n) to deduce the existence of two balls B_R and B_r tangent to Ω_n at $0 \in \partial\Omega_n$ and so that

$$\Omega_n \subset B_r \subset B_R;$$

hence $r < R$. Note that $0 \in \partial B_r, \partial B_R$ and the centers of these balls are $(0, \ldots, 0, r)$ and $(0, \ldots, 0, R)$, respectively. We denote $x' = (x_i)_{i=1}^{d-1}$ and note that $x = (x', x_d) \in \partial B_r$ satisfy $|x'|^2 + (x_d - r)^2 = r^2$, whence

$$x_d\Big(1 - \frac{x_d}{2r}\Big) = \frac{1}{2r}|x'|^2 \quad \Rightarrow \quad \frac{1}{2r}|x'|^2 \leq x_d \leq \frac{1}{r}|x'|^2,$$

provided $x_d \leq r$. This in turn implies for $1 < \xi < \frac{R}{r}$ fixed

$$p(x) \leq p\Big(x', \frac{1}{2r}|x'|^2\Big) = \frac{E^{1/d}}{2}\Big(1 - \frac{R}{r} + \frac{1}{4r^2}|x'|^2\Big)|x'|^2 \leq \frac{E^{1/d}}{2}(1-\xi)|x'|^2 < 0,$$

provided $|x'|^2 \leq C_1 := 4r^2\big(\frac{R}{r} - \xi\big)$ and R is used in the definition of p. Since $|x'|^2 \leq r^2$ and $x_d^2 \leq \frac{|x'|^4}{r^2} \leq |x'|^2$, we have that $|x|^2 = |x'|^2 + x_d^2 \leq 2|x'|^2$ and we deduce

$$|x'|^2 \leq C_1 \quad \Rightarrow \quad p(x) \leq -E^{1/d}\frac{(\xi-1)}{4}|x|^2 = -E^{1/d}C_2|x|^2.$$

On the other hand, for $x \in \partial B_r$ with $|x'| > C_1$ we infer that the distance from x to ∂B_R is strictly positive whence

$$p(x) \leq -C_3|x|^2.$$

Since both constants C_2, C_3 depend only on r, R, we see that p grows quadratically on ∂B_r with a constant independent of n, and thus on $\Omega_n \subset B_r$. To compare b_n^- with g_n, we recall that g_n is a smooth function for Taylor formula to give

$$\left| g_n(x) - g_n(0) - \nabla g_n(0)x \right| \leq \frac{1}{2}|g_n|_{W^2_\infty(\Omega_n)}|x|^2 \qquad \forall\, x \in \Omega_n.$$

We finally choose the factor E in b_n^- proportional to $|g_n|_{W^2_\infty(\Omega_n)}$ and realize that $b_n^-(x) \leq g_n(x)$ for all $x \in \partial\Omega_n$ as asserted.

Step 4 (Uniform convergence). We view both u and u_n as viscosity solutions of (1.1), the former with data (f, g, Ω) and the latter with data (f_n, u_n, Ω). Applying Lemma 5.3 (continuous dependence on data), along with (5.1) and (5.3), we obtain

$$\begin{aligned} \|u_n - u\|_{L^\infty(\Omega)} &\leq C\|f_n - f\|^{1/d}_{L^\infty(\Omega)} + \|u_n - g\|_{L^\infty(\partial\Omega)} \\ &\leq C\sigma(\rho_n)^{1/d} + \sigma(\rho_n) + \sigma(\delta_n) + C_n\delta_n. \end{aligned}$$

Given an arbitrary number β we first choose ρ_n so that $C\sigma(\rho_n)^{1/d} + \sigma_n(\rho_n) \leq \frac{\beta}{2}$. This choice determines the regularity of g_n, namely its W^2_∞ and $C^{2,\alpha}$ norms in $\widetilde{\Omega}$. Since C_n is proportional to $|g_n|_{W^2_\infty(\Omega_n)}$, we finally select δ_n so that $\sigma(\delta_n)+C_n\delta_n \leq \frac{\beta}{2}$. This shows the desired uniform convergence of u_n to u in Ω.

□

5.3. **Uniform convergence: Regularization approach.** In this section we combine Lemma 2.4 (discrete comparison principle), Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u]$), Lemma 5.1 (discrete boundary barrier), Lemma 5.2 (discrete interior barrier), and Lemma 5.4 (approximation of viscosity solutions by smooth solutions) to prove uniform convergence of u_ε to u in Ω.

Since u_ε is defined in the computational domain Ω_h, and $\Omega_h \subset \Omega$, we extend u_ε to Ω as follows. Given $x \in \Omega \setminus \Omega_h$ let $z \in \partial\Omega_h$ be the closest point to x, which is unique because Ω_h is convex, and let

$$u_\varepsilon(x) := u_\varepsilon(z) = \mathcal{I}_h g(z) \quad \forall\, x \in \Omega \setminus \Omega_h. \tag{5.4}$$

Theorem 5.5 (Uniform convergence)**.** *Let Ω be uniformly convex, $f, g \in C(\overline{\Omega})$ and $f \geq 0$ in Ω. The discrete solution u_ε of* (2.2) *and* (5.4) *converges uniformly to the unique viscosity solution $u \in C(\overline{\Omega})$ of* (1.1) *as $\varepsilon = (h, \delta, \theta) \to 0$ and $\frac{h}{\delta} \to 0$.*

Proof. We first split

$$\|u - u_\varepsilon\|_{L^\infty(\Omega)} \leq \|u - u_\varepsilon\|_{L^\infty(\Omega_h)} + \|u - u_\varepsilon\|_{L^\infty(\Omega\setminus\Omega_h)}$$

and then employ the triangle inequality to write

$$\|u - u_\varepsilon\|_{L^\infty(\Omega_h)} \leq \|u - u_n\|_{L^\infty(\Omega_h)} + \|u_n - \mathcal{I}_h u_n\|_{L^\infty(\Omega_h)} + \|\mathcal{I}_h u_n - u_\varepsilon\|_{L^\infty(\Omega_h)}.$$

Next, we recall that Lemma 5.4 yields $\|u - u_n\|_{L^\infty(\Omega_h)} \leq \|u - u_n\|_{L^\infty(\Omega)} \to 0$ as $n \to \infty$. In addition, polynomial interpolation theory gives

$$\|u_n - \mathcal{I}_h u_n\|_{L^\infty(\Omega_h)} \leq C|u_n|_{W^2_\infty(\Omega)}h^2 \to 0,$$

as $h \to 0$ for n fixed. On the other hand, (5.4) yields

$$|u(x) - u_\varepsilon(x)| = |u(x) - u_\varepsilon(z)| \le |u(x) - u(z)| + |u(z) - u_\varepsilon(z)| \quad \forall\, x \in \Omega \setminus \Omega_h$$

where $z \in \partial\Omega_h$. If τ is the modulus of continuity of $u \in C(\overline{\Omega})$, we have

$$\|u - u_\varepsilon\|_{L^\infty(\Omega\setminus\Omega_h)} \le \tau\big(\mathrm{dist}_H(\Omega, \Omega_h)\big) + \|u - u_\varepsilon\|_{L^\infty(\Omega_h)}.$$

Since $\mathrm{dist}_H(\Omega, \Omega_h) \to 0$, as $h \to 0$, the proof reduces to showing that $\|\mathcal{I}_h u_n - u_\varepsilon\|_{L^\infty(\Omega_h)}$ can be made arbitrarily small. We do this in three steps.

Step 1 (Boundary estimate). Let p_h be the function of Lemma 5.1 (discrete boundary barrier) with constant $E_{n,1} := C|u_n|^d_{W^2_\infty(\Omega)} + 3\sigma_n$, where $C|u_n|^d_{W^2_\infty(\Omega)}$ is the consistency error (4.3) from Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u]$) with u_n in place of u and $3\sigma_n$ is a bound (5.1) for $\|f - f_n\|_{L^\infty(\Omega)}$. Since both u_ε and p_h are discretely convex, we have

$$T_\varepsilon[u_\varepsilon + p_h](x_i) \ge T_\varepsilon[u_\varepsilon](x_i) + T_\varepsilon[p_h](x_i) \ge f(x_i) + E_{n,1} \ge T_\varepsilon[\mathcal{I}_h u_n](x_i)$$

for all $x_i \in \mathcal{N}_h^0$. Moreover, since (5.3) holds for all $z \in \partial\Omega$ and $\mathcal{N}_h^b \subset \partial\Omega$, linear interpolation implies that $\mathcal{I}_h u_n \ge \mathcal{I}_h g - \xi_n = u_\varepsilon - \xi_n$ on $\partial\Omega_h$ for all h, where $\xi_n := \sigma(\rho_n) + \sigma(\delta_n) + C_n\delta_n$ and $\delta_n \ge \mathrm{dist}_H(\Omega_n, \Omega)$, whence $u_\varepsilon + p_h - \xi_n \le \mathcal{I}_h u_n$ on $\partial\Omega_h$. Consequently, for all $z \in \mathcal{N}_h^0$ such that $\mathrm{dist}(z, \partial\Omega) \le 2\delta$, Lemma 2.4 (discrete comparison principle) yields

$$u_\varepsilon(z) - CE_{n,1}^{1/d}\delta - \xi_n \le \mathcal{I}_h u_n(z).$$

A similar argument with $u_\varepsilon - p_h + \xi_n$ gives rise to the reverse estimate.

Step 2 (Interior estimate). We resort to the function q_h of Lemma 5.2 (discrete interior barrier) to construct a discrete lower barrier b_ε^- as follows: let

$$E_{n,2} := C|u_n|_{C^{2+\alpha}(\overline{\Omega})}|u_n|^{d-1}_{W^2_\infty(\Omega)}\delta^\alpha + C|u_n|^d_{W^2_\infty(\Omega)}\Big(\frac{h^2}{\delta^2} + \theta^2\Big) + 3\sigma_n$$

and

$$b_\varepsilon^- := u_\varepsilon + E_{n,2}^{1/d} q_h - CE_{n,1}^{1/d}\delta - \xi_n.$$

Since $q_h \le 0$, Step 1 guarantees that $b_\varepsilon^- \le \mathcal{I}_h u_n$ on $\partial\Omega_{h,\delta}$, where $\Omega_{h,\delta}$ is defined in (4.1). Applying Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u]$) for u_n instead of u implies

$$T_\varepsilon[b_\varepsilon^-](x_i) \ge T_\varepsilon[u_\varepsilon](x_i) + E_{n,2} = f(x_i) + E_{n,2} \ge T_\varepsilon[\mathcal{I}_h u_n](x_i) \quad \forall\, x_i \in \mathcal{N}_h^0 \cap \Omega_{h,\delta}$$

where we have used that both u_ε and q_h are discretely convex as well as (5.1). Lemma 2.4 (discrete comparison principle) yields

$$b_\varepsilon^- = u_\varepsilon + E_{n,2}^{1/d} q_h - CE_{n,1}^{1/d}\delta - \xi_n \le \mathcal{I}_h u_n \quad \text{in} \quad \Omega_{h,\delta}.$$

A similar argument with $b_\varepsilon^+ := u_\varepsilon - E_{n,2}^{1/d} q_h + CE_{n,1}^{1/d}\delta + \xi_n$ results in $b_\varepsilon^+ \ge \mathcal{I}_h u_n$.

Combining these estimates with those of Step 1, we end up with

$$\|u_\varepsilon - \mathcal{I}_h u_n\|_{L^\infty(\Omega_h)} \le CE_{n,1}^{1/d}\delta + CE_{n,2}^{1/d} + \xi_n. \tag{5.5}$$

Step 3 (Uniform convergence in Ω). We finally proceed as in Step 4 of the proof of Lemma 5.4 (approximation of viscosity solutions by smooth solutions). Given an arbitrary number $\beta > 0$, we choose ρ_n so that $\sigma(\rho_n) \le \frac{\beta}{3}$. This dictates the regularity of g_n hidden in the constant C_n of ξ_n, as well as that of u_n, and allows us to select δ_n so that $\sigma(\delta_n) + C_n\delta_n \le \frac{\beta}{3}$; hence $\xi_n \le \frac{2\beta}{3}$. We next take $\delta, \frac{h}{\delta}$ and θ small enough, depending on u_n, so that the first two terms of (5.5) are $\le \frac{\beta}{3}$ and thus $\|u_\varepsilon - \mathcal{I}_h u_n\|_{L^\infty(\Omega_h)} \le \beta$. This completes the proof.

□

5.4. **Uniform convergence: Barles-Souganidis approach.** In this section we adapt the approach of [4] to our setting. Since (5.4) extends the definition of discrete solution u_ε to Ω, we let the limit supremum and limit infimum of u_ε be

$$u^*(x) = \limsup_{\varepsilon, \frac{h}{\delta} \to 0, z \to x} u_\varepsilon(z), \qquad u_*(x) = \liminf_{\varepsilon, \frac{h}{\delta} \to 0, z \to x} u_\varepsilon(z),$$

and observe that u^* is upper semi-continuous and u_* is lower semi-continuous. We show that they attain the Dirichlet boundary condition pointwise. Moreover, they are viscosity subsolution and supersolution of (1.1), respectively. An essential difficulty associated with (1.1), already mentioned in [20], is that viscosity sub and supersolutions of (1.1) must be convex for the comparison principle to be applicable. Since u_ε is only discretely convex, it is not obvious that u^* and u_* are convex.

To circumvent this issue we proceed as in [20]: we let $\partial^{2,+}_{v_j v_j} u := \max\left(\partial^2_{v_j v_j} u, 0\right)$, $\partial^{2,-}_{v_j v_j} u := -\min\left(\partial^2_{v_j v_j} u, 0\right)$, introduce the continuous version of our ideal discrete operator in (2.2)

$$T[u] := \min_{\mathbf{v}=(v_j)_{j=1}^d \in \mathbb{S}^\perp} \left(\prod_{j=1}^d \partial^{2,+}_{v_j v_j} u - \sum_{j=1}^d \partial^{2,-}_{v_j v_j} u \right),$$

and show that u is a convex viscosity solution of (1.1) if and only if u is a viscosity solution of the Dirichlet problem

$$T[u] = f \quad \text{in } \Omega, \qquad u = g \quad \text{on } \partial\Omega, \tag{5.6}$$

for which we do not require convexity because it is built in the operator.

Lemma 5.6 (Equivalence of viscosity solutions)**.** *If $f \in C(\Omega)$ satisfies $f \geq 0$, and $u \in C(\overline{\Omega})$, then u is a viscosity solution of* (5.6) *if and only if u is a convex viscosity solution of* (1.1)*.*

Proof. Since u is uniformly continuous in Ω the notion of Dirichlet condition is classical in both cases. We thus verify the equation in the viscosity sense.

Step 1 (Necessity). We rely on the notion of convexity of a function $v \in C(\Omega)$ in the viscosity sense: for test function $\phi \in C^2(\Omega)$ that touches v from above at a point $x \in \Omega$ the smallest eigenvalue $\lambda_1[D^2\phi](x)$ of $D^2\phi$ at x satisfies

$$\lambda_1[D^2\phi](x) \geq 0.$$

It is proven in [34] that a continuous function v is convex if and only if it is convex in the viscosity sense. We show that a viscosity solution u of (5.6) is convex in the viscosity sense and use this equivalence to deduce convexity of u.

We observe that u being a viscosity solution of (5.6) implies that for $\phi \in C^2(\Omega)$ touching u from above at $x \in \Omega$, we have

$$T[\phi](x) \geq f(x) \geq 0.$$

We argue as in Lemma 2.2: if there is a direction $v_j \in \mathbb{S}$ for which $\frac{\partial^2 \phi}{\partial v_j^2}(x) < 0$, then $T[\phi](x) < 0$ which contradicts the preceding statement. Therefore,

$$\frac{\partial^2 \phi}{\partial v_j^2}(x) \geq 0 \quad \forall\, v_j \in \mathbb{S} \qquad \Rightarrow \qquad \lambda_1[D^2\phi](x) \geq 0.$$

This proves that u is convex as well as

$$\det D^2\phi(x) = T[\phi](x) = \min_{\mathbf{v}=(v_j)_{j=1}^d \in \mathbb{S}^\perp} \prod_{j=1}^d \partial^2_{v_j v_j}\phi(x) \geq f(x)$$

according to (1.2). This implies that u is a convex subsolution of (1.1).

To prove that u is also a supersolution of (1.1), we recall that the definition of viscosity solutions for (1.1) uses convex test functions $\phi \in C^2(\Omega)$ [26]; hence $\det D^2\phi = T[\phi]$. Consequently, if $u - \phi$ attains a minimum at $x \in \Omega$, then

$$\det D^2\phi(x) = T[\phi](x) \leq f(x)$$

whence u is a supersolution of (1.1).

Step 2 (Sufficiency). Let's assume now that u is a convex viscosity solution of (1.1), and $\phi \in C^2(\Omega)$ is a test function that touches u at $x_0 \in \Omega$. Inspired by [26, Remark 1.3.3], we decompose $\phi = q + r$ into a quadratic q and a remainder r

$$q(x) = \phi(x_0) + D\phi(x_0)(x - x_0) + \frac{1}{2}(x - x_0)^T D^2\phi(x_0)(x - x_0), \quad r(x) = o(|x - x_0|^2);$$

hence $D^2\phi(x_0) = D^2q(x_0)$. If $q^\pm(x) := q(x) \pm \sigma|x - x_0|^2$, we then observe that $q^+ \geq \phi$ and $q^- \leq \phi$ in a suitable neighborhood of x_0 provided $\sigma > 0$. We take advantage of $q^\pm$ being quadratic to realize that $q^\pm$ is convex if and only if $D^2q^\pm(x_0) \geq 0$.

If $u - \phi$ attains a local max at x_0, so does $u - q^+$ and $D^2\phi(x_0) \geq 0$ because u is convex. Therefore, the quadratic q^+ is convex and must satisfy

$$\det D^2q^+(x_0) = \det(D^2q(x_0) + 2\sigma I) \geq f(x_0),$$

because u is a viscosity solution of (1.1). Take the limit $\sigma \downarrow 0$ to find out that $T[\phi](x_0) = \det D^2\phi(x_0) \geq f(x_0)$ whence u is a viscosity subsolution of (5.6).

On the other hand, if $u - \phi$ attains a local min at x_0, so does $u - q^-$. We now have two possible cases. If all the eigenvalues of $D^2\phi(x_0)$ are strictly positive, then q^- is a convex quadratic for σ sufficiently small. This in turn implies

$$\det D^2q^-(x_0) = \det(D^2q(x_0) - 2\sigma I) \leq f(x_0),$$

as u is a viscosity solution of (1.1); hence $T[\phi](x_0) = \det D^2\phi(x_0) \leq f(x_0)$ upon letting $\sigma \downarrow 0$. If any eigenvalue of $D^2\phi(x_0)$ is nonpositive, then $T[\phi](x_0) \leq 0$ by definition and $T[\phi](x_0) \leq f(x_0)$ because $f \geq 0$. We thus deduce that u is a viscosity supersolution of (5.6), whence a viscosity solution of (5.6), as asserted.

□

We are now ready to prove the convergence of our discrete solution u_ε to the viscosity solution u of (1.1).

Theorem 5.7 (Uniform convergence)**.** *Let Ω be uniformly convex, $f \in C(\Omega) \cap L^\infty(\Omega)$ satisfy $f \geq 0$, and $g \in C(\partial\Omega)$. The discrete solution u_ε of* (2.2) *converges uniformly to the unique viscosity solution $u \in C(\overline{\Omega})$ of* (1.1) *as $\varepsilon = (h, \delta, \theta) \to 0$ and $\frac{h}{\delta} \to 0$.*

Proof. In view of Lemma 5.6 (equivalence of viscosity solutions), we prove that u_ε converges to the viscosity solution of (5.6). To this end, we have to deal with a test function $\phi \in C^2(\Omega)$ and its Lagrange interpolant $\phi_h = \mathcal{I}_h\phi$. Without loss of generality we may assume $\phi \in C^{2,\alpha}(\Omega)$. We split the proof into five steps.

Step 1 (Consistency). We have the following alternative to (4.3):

$$\big|T[\phi](x_0) - T_\varepsilon[\phi_h](x_i)\big| \le C_1(\phi)\Big(\delta^\alpha + |x_0 - x_i|^\alpha\Big) + C_2(\phi)\Big(\frac{h^2}{\delta^2} + \theta^2\Big),$$

where the constants C_1, C_2 are defined in Lemma 4.2 (consistency of $T_\varepsilon[\mathcal{I}_h u](x_i)$) and depend on $|\phi|_{C^{2,\alpha}(B_i)}$ and $|\phi|_{W^2_\infty(B_i)}$ with B_i defined in (4.2), and $x_0 \in \Omega, x_i \in \mathcal{N}_h^0 \cap \Omega_{h,\delta}$. The proof of this inequality proceeds along the lines of those of Lemmas 4.1 and 4.2, except that now we need to deal with the functions $s \mapsto \max(s,0)$ and $s \mapsto \min(s,0)$ in the definitions of both T and T_ε because ϕ may not be convex. We exploit that these functions are Lipschitz with constant 1 to write

$$\big|\nabla^{2,+}_\delta \phi_h(x_i; v_j) - \partial^{2,+}_{v_j v_j}\phi(x_0)\big| \lesssim |\phi|_{C^{2,\alpha}(B_i)}\Big(\delta^\alpha + |x_0 - x_i|^\alpha\Big) + |u|_{W^2_\infty(B_i)}\frac{h^2}{\delta^2},$$

together with a similar bound for the operators $\nabla^{2,-}_\delta$ and $\partial^{2,-}_{v_j v_j}$.

Step 2 (Subsolutions). We show that u^* is a viscosity subsolution of (5.6); likewise u_* is a viscosity supersolution. This hinges on monotonicity and consistency [4]. We must show that if $u^* - \phi$ attains a local maximum at $x_0 \in \Omega$, we have

$$T[\phi](x_0) \ge f(x_0);$$

note that $u^* - \phi$ is upper semi-continuous and the local maximum is well defined. Without loss of generality, we may assume that $u^* - \phi$ attains a strict global maximum at $x_0 \in \Omega$ [28, Remark in p. 31], and $x_0 \in \Omega_h$ for h sufficiently small. Let u_ε and z_h be a sequence of functions and nodes such that

$$\lim_{\varepsilon, \frac{h}{\delta} \to 0, z_h \to x_0} u_\varepsilon(z_h) = u^*(x_0).$$

Let $x_h \in \mathcal{N}_h$ be a sequence of nodes so that $u_\varepsilon - \phi_h$ attains a maximum at x_h. We claim that $x_h \to x_0$ as $h \to 0$. If not, then there exists a subsequence $x_h \to y_0$ such that $y_0 \ne x_0$. Since $(u_\varepsilon - \phi_h)(x_h) \ge (u_\varepsilon - \phi_h)(z_h)$, passing to the limit we obtain

$$(u^* - \phi)(y_0) \ge \limsup_{\varepsilon, \frac{h}{\delta} \to 0} (u_\varepsilon - \phi_h)(x_h) \ge \lim_{\varepsilon, \frac{h}{\delta} \to 0, z_h \to x_0} (u_\varepsilon - \phi_h)(z_h) = (u^* - \phi)(x_0).$$

This contradicts the fact that $u^* - \phi$ attains a strict maximum at x_0. Exploiting the fact that $u_\varepsilon - \phi_h$ attains a maximum at x_h, Lemma 2.3 (monotonicity) yields

$$T_\varepsilon[\phi_h](x_h) \ge T_\varepsilon[u_\varepsilon](x_h) = f(x_h).$$

Since $f \in C(\Omega)$, to prove $T[\phi](x_0) \ge f(x_0)$ we only need to show that as $\varepsilon, \frac{h}{\delta} \to 0$

$$T_\varepsilon[\phi_h](x_h) \to T[\phi](x_0).$$

This is a consequence of Step 1 and the fact that $x_h \in \Omega_{h,\delta}$ for δ sufficiently small, because $x_0 \in \Omega$, $x_h \to x_0$ and the sequence of $\Omega_h \uparrow \Omega$ is nondecreasing.

Step 3 (Boundary behavior). We now prove that $u^* = u_* = g$ on $\partial\Omega$ via a barrier argument similar to those in [20, 31, 32]; we proceed as in [20]. This is essential to apply the comparison principle for operator T to relate u_*, u^* and u in Step 4.

Let p_k be the quadratic function in the proof of Lemma 5.1 (discrete boundary barrier) associated with an arbitrary boundary point $x \in \partial\Omega$ (the origin in the construction of p_k) and with constant $E = k$. We recall that $p_k(x) = 0$ and $p_k(z) \le 0$ for all $z \in \partial\Omega$ can be made arbitrarily large for $k \to \infty$ by virtue of the uniform convexity of Ω. A simple consequence is that the sequence of points

$x_k \in \partial\Omega$ where $g + p_k$ (resp., $g - p_k$) attains a maximum (resp., a minimum) over $\partial\Omega$ converges to x.

We now observe that taking $w_h \equiv 0$ in Lemma 2.3 (monotonicity) implies the following maximum principle: if a discretely convex function u_h satisfies $T_\varepsilon[u_h](x_i) > 0$ for all $x_i \in \mathcal{N}_h^0$, then u_h attains a maximum over $\overline{\Omega}_h$ on $\mathcal{N}_h^b \subset \partial\Omega$. Apply this to $T_\varepsilon[u_\varepsilon + \mathcal{I}_h p_k] > 0$ to deduce that $u_\varepsilon + \mathcal{I}_h p_k$ attains its maximum on $\mathcal{N}_h^b$. In view of (5.4), we may assume $z \in \Omega_h$ in $u^*(x) = \limsup_{\varepsilon, \frac{h}{\delta} \to 0, z \to x} u_\varepsilon(z)$. Consequently,

$$\begin{aligned} u^*(x) &\leq \limsup_{\varepsilon, \frac{h}{\delta} \to 0, z \to x} \big(u_\varepsilon(z) + \mathcal{I}_h p_k(z)\big) - \liminf_{\varepsilon, \frac{h}{\delta} \to 0, z \to x} \mathcal{I}_h p_k(z) \\ &\leq \limsup_{\varepsilon, \frac{h}{\delta} \to 0} \max_{z \in \mathcal{N}_h^b} \big(g + p_k\big)(z) - p_k(x) \leq g(x_k) + p_k(x_k) \leq g(x_k), \end{aligned}$$

because $\max_{\mathcal{N}_h^b} g + p_k \leq \max_{\partial\Omega} g + p_k$, whence taking $k \to \infty$ yields $u^*(x) \leq g(x)$.

On the other hand, since $T_\varepsilon[\mathcal{I}_h p_k](x_i) > T_\varepsilon[u_\varepsilon](x_i)$ for all $x_i \in \mathcal{N}_h^0$ and k large enough, Lemma 2.3 implies that $u_\varepsilon - \mathcal{I}_h p_k$ attains a minimum on $\mathcal{N}_h^b$. Therefore, arguing as before

$$u_*(x) \geq \liminf_{\varepsilon, \frac{h}{\delta} \to 0} \min_{z \in \mathcal{N}_h^b} \big(g - p_k\big)(z) + p_k(x) \geq g(x_k) - p_k(x_k) \geq g(x_k),$$

whence $u_*(x) \geq g(x)$. This in turn gives $u^* \leq g \leq u_* \leq u^*$ on $\partial\Omega$ as asserted.

Step 4 (Comparison). To prove that $u^* = u_*$ in $\overline{\Omega}$ we use the following comparison principle for (5.6): if v^- is a subsolution and is upper semi-continuous in $\overline{\Omega}$, v^+ is a supersolution and is lower semi-continuous in $\overline{\Omega}$, and $v^- \leq v^+$ on $\partial\Omega$, then $v^- \leq v^+$ on $\overline{\Omega}$. This result falls under the umbrella of [13, Theorem 3.3]. It hinges on an argument mentioned in [13, Section 5.C] that is briefly described for a more general form of the Monge-Ampère operator in [28, V.3]. Both operators in (1.1) and (5.6) satisfy the requirements posed in [28]. We apply this comparison principle to $v^- = u^*$ and $v^+ = u_*$, which satisfy the assumptions in view of Steps 2 and 3, to obtain $u^* \leq u_*$ in $\overline{\Omega}$. Since $u^* \geq u_*$ by definition, this results in $u^* = u_*$ in $\overline{\Omega}$.

Step 5 (Uniform convergence). Step 4 implies the pointwise limit

$$u(x) = \lim_{\varepsilon \to 0, z \to x} u_\varepsilon(z) \quad \forall\, x \in \overline{\Omega}.$$

To see that this gives rise to uniform convergence we argue by contradiction. We assume that for every ε there exist a point $x_\varepsilon \in \overline{\Omega}$ such that $|u(x_\varepsilon) - u_\varepsilon(x_\varepsilon)| \geq \sigma$, for some $\sigma > 0$. Since $\overline{\Omega}$ is compact, there exists a subsequence (not relabeled) $x_\varepsilon \to x_0 \in \overline{\Omega}$. Computing the limit $\varepsilon \to 0$ in the last inequality yields the contradiction $|u(x_0) - u(x_0)| \geq \sigma$. This concludes the proof.

□

6. Numerical experiments

We present three examples in the square domain $\Omega = \Omega_h = [0,1]^2$. The fact that Ω is not uniformly convex does not affect the existence of our discrete solution u_ε, as the Dirichlet datum g is the trace of a convex function; however this is beyond the assumptions of the convergence theory. We implement the 2-scale method within the MATLAB software FELICITY [38, 39]. We first consider two examples with smooth Hessian and with discontinuous Hessian, and observe linear experimental rates of convergence with respect to h; we further investigate rates theoretically

in [30]. The third example entails an unbounded right-hand side f and is not guaranteed to converge by theory. We still observe convergence experimentally.

6.1. **Semi-smooth Newton method.** We solve the nonlinear algebraic equation (2.2) via a damped semi-smooth Newton iteration. Let $\mathbf{z} := (z_h(x_i))_{i=1}^N \in \mathbb{R}^N$ stand for the vector of nodal values of a generic $z_h \in \mathbb{V}_h$; thus N is the cardinality of $\mathcal{N}_h$. If $\mathbf{u}_n = (u_\varepsilon^n(x_i))_{i=1}^N$, $\mathbf{DT}_\varepsilon[\mathbf{u}_n]$ is the Jacobian matrix of the nonlinear map $\mathbf{T}_\varepsilon : \mathbb{R}^N \to \mathbb{R}^N$ at $\mathbf{u}_n$, and $\mathbf{f} = (f(x_i))_{i=1}^N$, then a Newton increment is given by

$$\mathbf{DT}_\varepsilon[\mathbf{u}_n]\ \mathbf{w}_n = \mathbf{f} - \mathbf{T}_\varepsilon[\mathbf{u}_n]$$

and the n-th Newton step by $\mathbf{u}_{n+1} = \mathbf{u}_n + \tau \mathbf{w}_n$, where the damping parameter $\tau \in (0, 1]$, which might depend on n, satisfies

$$\|f - T_\varepsilon[u_\varepsilon^n + \tau w_n]\|_{L^2(\Omega)} < \|f - T_\varepsilon[u_\varepsilon^n]\|_{L^2(\Omega)}.$$

We now explain the construction of $\mathbf{DT}_\varepsilon[\mathbf{u}_n]$. Evaluating $\nabla_\delta^2 z_h(x_i; v_j)$, in view of (2.1), requires knowing z_h at $x_i^\pm = x_i \pm \rho\delta v_j$, which are not necessarily nodes of $\mathcal{N}_h$. Since $x_i^\pm$ belong to two simplices of $\mathcal{T}_h$, and $z_h \in \mathbb{V}_h$, the values $z_h(x_i^\pm)$ can be determined in terms of the barycentric coordinates of $x_i^\pm$. Therefore, if we define $\boldsymbol{\nabla}_\delta^2 \mathbf{z}(i; v_j) := \nabla_\delta^2 z_h(x_i; v_j)$, then we realize that this operator involves $2(d+1)+1$ components of $\mathbf{z}$ and is thus sparse. We likewise define $\boldsymbol{\nabla}_\delta^{2,+}$ and $\boldsymbol{\nabla}_\delta^{2,-}$ to be the component-wise versions of $\nabla_\delta^{2,+}$ and $\nabla_\delta^{2,-}$. The operator $\mathbf{T}_\varepsilon$ reads

$$\mathbf{T}_\varepsilon[\mathbf{z}](i) := \min_{\mathbf{v}=(v_j)_{j=1}^d \in \mathbb{S}_\theta^\perp} \left(\prod_{j=1}^d \boldsymbol{\nabla}_\delta^{2,+} \mathbf{z}(i; v_j) - \sum_{j=1}^d \boldsymbol{\nabla}_\delta^{2,-} \mathbf{z}(i; v_j) \right) = f(x_i), \tag{6.1}$$

according to (2.4).

Let now $\mathbf{v}^i = (v_j^i)_{j=1}^d \in \mathbb{S}_\theta^\perp$ be a set of directions that realize the minimum of $\mathbf{T}_\varepsilon[\mathbf{u}_n](i)$ and denote $\mathbf{V} := (\mathbf{v}^i)_{i=1}^N \in \mathbb{R}^{d\times d\times N}$, the collection of the minimizing d-tuples $\mathbf{v}^i$ for all $i = 1, \dots N$. Combining the above notation, we denote the matrix that contains the j-th minimizing directions for each node by $\mathbf{V}_j \in \mathbb{R}^{d\times N}$. This allows us to display our Jacobian in a vectorized form, using the notation $\boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(\mathbf{V}_j) := (\nabla_\delta^2 \mathbf{u}_n(i; v_j^i))_{i=1}^N$, since (6.1) gives for $\mathbf{z} = \mathbf{u}_n$:

$$\mathbf{T}_\varepsilon[\mathbf{u}_n] = \bigodot_{j=1}^d \boldsymbol{\nabla}_\delta^{2,+} \mathbf{u}_n(\mathbf{V}_j) - \sum_{j=1}^d \boldsymbol{\nabla}_\delta^{2,-} \mathbf{u}_n(\mathbf{V}_j),$$

where $\odot$ stands for the component-wise multiplication of vectors. Using Danskin's Theorem [7] and the product rule, we can then obtain $\mathbf{DT}_\varepsilon[\mathbf{u}_n]\mathbf{w}_n$. For that, we need to differentiate $\boldsymbol{\nabla}_\delta^{2,+} \mathbf{u}_n(\mathbf{V}_j)$. We observe that for each component, $\boldsymbol{\nabla}_\delta^{2,+} \mathbf{u}_n(i; v_j^i) = \max\{\boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(i; v_j^i), 0\}$ is not differentiable at $\boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(i; v_j^i) = 0$. As a result, we use the so-called slant derivative in the direction $\mathbf{w}_n$ [12, 27], in order to compute:

$$\mathbf{D}[\boldsymbol{\nabla}_\delta^{2,+} \mathbf{u}_n(\mathbf{V}_j)]\mathbf{w}_n = \mathbf{H}^+[\boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(\mathbf{V}_j)] \odot \boldsymbol{\nabla}_\delta^2 \mathbf{w}_n(\mathbf{V}_j),$$

each component of which is equal to

$$(\mathbf{D}[\boldsymbol{\nabla}_\delta^{2,+} \mathbf{u}_n(\mathbf{V}_j)]\mathbf{w}_n)_i = \begin{cases} \boldsymbol{\nabla}_\delta^2 \mathbf{w}_n(i; v_j^i) & \text{if } \boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(i; v_j^i) > 0, \\ 0 & \text{if } \boldsymbol{\nabla}_\delta^2 \mathbf{u}_n(i; v_j^i) \le 0. \end{cases}$$

Here $\mathbf{H}^+$ is the operator that assigns 1 to a strictly positive component and zero otherwise. Similarly, $\mathbf{D}[\boldsymbol{\nabla}_\delta^{2,-}\mathbf{u}_n(\mathbf{V}_j)]\mathbf{w}_n = \mathbf{H}^-[\boldsymbol{\nabla}_\delta^2\mathbf{u}_n(\mathbf{V}_j)] \odot \boldsymbol{\nabla}_\delta^2\mathbf{w}_n(\mathbf{V}_j)$ where $\mathbf{H}^-$ assigns -1 to a nonpositive component and 0 otherwise. We are now ready to employ Danskin's Theorem [7] and the product rule to obtain similarly to [24]:

$$\mathbf{DT}_\varepsilon[\mathbf{u}_n]\mathbf{w}_n = \sum_{j=1}^{d} \boldsymbol{\nabla}_\delta^2\mathbf{w}_n(\mathbf{V}_j) \bigodot \Big(\mathbf{H}^+[\boldsymbol{\nabla}_\delta^2\mathbf{u}_n(\mathbf{V}_j)] \bigodot_{k\neq j} \boldsymbol{\nabla}_\delta^{2,+}\mathbf{u}_n(\mathbf{V}_k) - \mathbf{H}^-[\boldsymbol{\nabla}_\delta^2\mathbf{u}_n(\mathbf{V}_j)]\Big),$$

The presence of both operators $\mathbf{H}^+$ and $\mathbf{H}^-$ enforces discrete convexity, and their definition at zero yields nonsingular Jacobians computationally. This flexibility in choosing $\mathbf{H}^+$ and $\mathbf{H}^-$ with vanishing argument is consistent with the definition of the slant derivative for the max and min functions [27].

We initialize the Newton iteration with $\mathbf{u}_0$ corresponding to the Galerkin solution in $\mathbb{V}_h$ to the auxiliary problem $\Delta u_0 = (d!f)^{1/d}$ in Ω and $u_0 = g$ on $\partial\Omega$, as proposed in [24], but only for the coarser mesh $h = 2^{-5}$. For all subsequent refinements we interpolate the discrete solution in the previous coarse mesh and use it as initial guess. This greatly improves the residual error and leads to minimal or no damping.

6.2. **Accuracy.** We examine the performance of our two-scale method mainly with two examples, with smooth and discontinuous Hessians; a third example entails an unbounded f. For the first two examples we choose $\delta = h^\alpha$ and $\theta = h^\beta$ for appropriate $\alpha, \beta > 0$ which yield provable rates of convergence according to theory [30]. We stress that smaller values of θ lead to similar convergence rates but affect the sparsity pattern of the matrix in the semi-smooth Newton iteration because the number of search directions within $\mathbb{S}_\theta$ increase. We thus choose θ consistent with theory [30]. The computation of ρ in (2.1) is exact, because Ω_h is a square, although need not be in general. We stop the Newton iterations when

$$\|f - T_\varepsilon[u_\varepsilon^{n+1}]\|_{L^2(\Omega)} < 10^{-8}\|f - T_\varepsilon[u_\varepsilon^0]\|_{L^2(\Omega)}.$$

Smooth Hessian: We choose the solution u and force f to be

$$u(x) = e^{|x|^2/2}, \quad f(x) = (1+|x|^2)e^{|x|^2} \quad \forall x \in \Omega.$$

We choose $\delta, \theta \approx h^{1/2}$ on the basis of [30, Theorem 5.3], and report the results in Table 1(a) and Figure 1(a). We observe linear experimental convergence rates with respect to h, thus better than predicted in [30]. The number $P = 4(D-1)$ stands for the number of points $x_i \pm \delta v_j$ used in the evaluation of the operator T_ε at each interior node $x_i \in \mathcal{N}_h^0$ and for D directions v_j in a quarter circle dictated by θ.

Discontinuous Hessian: We choose the solution u and forcing function f to be

$$u(x) = \frac{1}{2}\left(\max(|x-x_0| - 0.2, 0)\right)^2, \quad f(x) = \max\left(1 - \frac{0.2}{|x-x_0|}, 0\right) \quad \forall x \in \Omega,$$

where $x_0 = (0.5, 0.5)$. Since $f = 0$ in the ball centered at x_0 of radius 0.2, this example is degenerate elliptic. We choose $\delta = O(h^{4/5})$ and $\theta = O(h^{2/5})$ on the basis of [30, Theorem 5.6], and observe experimentally again a linear decay rate in h, which is better than predicted. This time [30] suggests a larger θ, but we choose a smaller θ without compromising the sparsity pattern of the Newton matrix. As illustrated on Table 1, despite its degeneracy and lack of global regularity, this example does not exhibit any problematic behavior compared to the smooth case.

We next explore the behavior of the operator T_ε in terms of the sign of the truncation error $E_\varepsilon[u_\varepsilon] := f - T_\varepsilon[u_\varepsilon]$. In Figure 2 (left) we split the interior nodes

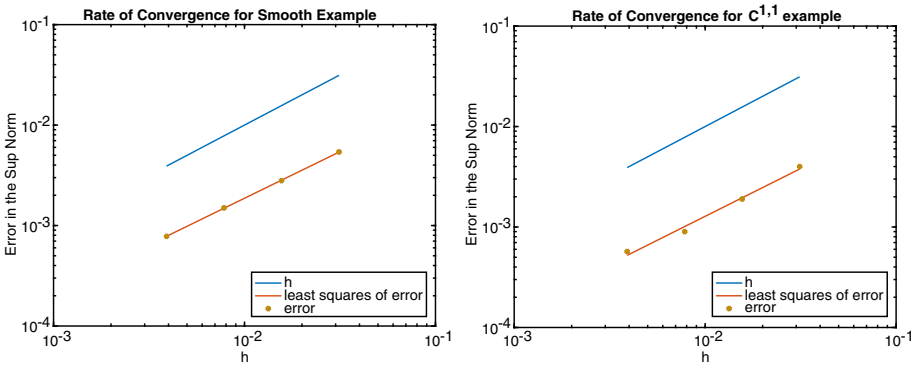

FIGURE 1. Experimental rates of convergence: the order is about 1 in terms of h for both the smooth Hessian with $\delta, \theta \approx h^{1/2}$ (left) and discontinuous Hessian $\delta \approx h^{4/5}, \theta \approx h^{2/5}$ (right).

TABLE 1. Smooth Hessian with $\delta, \theta \approx h^{1/2}$ (top), discontinuous Hessian with $\delta \approx h^{4/5}, \theta \approx h^{2/5}$ (bottom). The convergence rate is about linear in h for both cases (see Figure 1), whereas the number of Newton steps seem insensitive to the dimension N of the nonlinear system.

Degrees of freedom	P: number of points	$L_\infty-$error	Newton steps
N= 1089, $h = 2^{-5}$	16	$5.4\ 10^{-3}$	8
N=4225, $h = 2^{-6}$	24	$2.8\ 10^{-3}$	7
N=16641, $h = 2^{-7}$	36	$1.5\ 10^{-3}$	7
N= 66049, $h = 2^{-8}$	52	$7.8\ 10^{-4}$	8

Degrees of freedom	P: number of points	$L_\infty-$error	Newton steps
N= 1089, $h = 2^{-5}$	20	$4.0\ 10^{-3}$	10
N=4225, $h = 2^{-6}$	28	$1.9\ 10^{-3}$	9
N=16641, $h = 2^{-7}$	36	$9.0\ 10^{-4}$	9
N= 66049, $h = 2^{-8}$	48	$5.7\ 10^{-4}$	9

$\mathcal{N}_h^0$ into three sets, using the threshold $\texttt{eps} \approx 10^{-16}$ close to the machine precision of MATLAB: blue nodes x_i (34% of $\mathcal{N}_h^0$) correspond to $E_\varepsilon[u_\varepsilon](x_i) < -\texttt{eps}$; yellow nodes x_i (34% of $\mathcal{N}_h^0$) correspond to $E_\varepsilon[u_\varepsilon](x_i) > \texttt{eps}$; and magenta nodes x_i (32% of $\mathcal{N}_h^0$) correspond to $|E_\varepsilon[u_\varepsilon](x_i)| \leq \texttt{eps}$. Moreover, Figure 2 (left) displays with dashed lines the circle of discontinuity $|x - x_0| = 0.2$ and the two circles that are $\delta-$away from it. We point out that all points between the outer and inner circle are affected by the singularity, but they are mostly magenta nodes.

Lastly, we use the same example to provide some insight on the two-scale nature of our method. In Figure 2 (right) we display a node $x_i = (0.7656, 0.5391)$ within a zoomed mesh, and the thirty-six directions v_j in $\mathbb{S}_\theta$ scaled by δ which are used for the calculation of $T_\varepsilon[u_\varepsilon](x_i)$ for mesh size $h = 2^{-7}$. We see that for this specific instance, $\delta/h \approx 7$, and that most points $x_i \pm \delta v_j$ are not nodes. We employ a fast search routine within FELICITY to locate such points [38, 39].

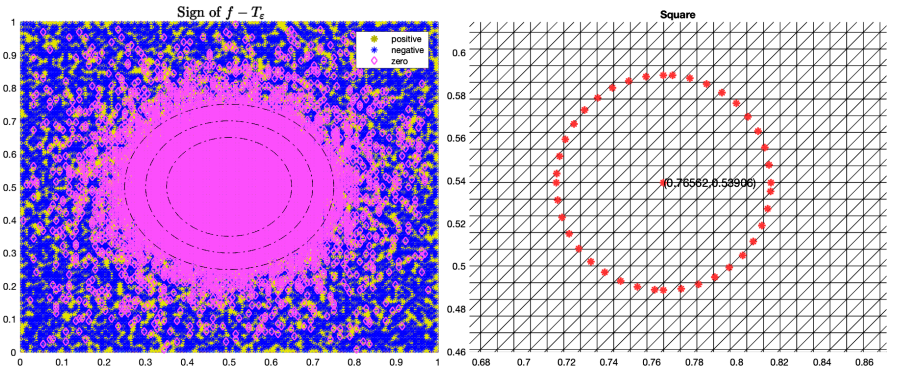

FIGURE 2. (left) Sign of the truncation error $E_\varepsilon[u_\varepsilon] = f - T_\varepsilon[u_\varepsilon]$ at nodes $x_i \in \mathcal{N}_h^0$ for the example with discontinuous Hessian and $h = 2^{-7}$: blue node x_i if $E_\varepsilon[u_\varepsilon](x_i) < -\texttt{eps}$, yellow node x_i if $E_\varepsilon[u_\varepsilon](x_i) > \texttt{eps}$, and magenta node x_i if $|E_\varepsilon[u_\varepsilon](x_i)| \leq \texttt{eps}$, where $\texttt{eps} \approx 10^{-6}$. We observe that the region $|x_i - x_0| < 0.2$, where $f \equiv 0$, is magenta. (right) Set of directions in $\mathbb{S}_\theta$ centered at a node $x_i \in \mathcal{N}_h^0$ and scaled by δ for the same example; note that $\delta/h \approx 7$.

Unbounded f: We finally present computational results for an example that does not fall within our theory because the right-hand side f is not uniformly bounded. More precisely, we consider the following f, which becomes unbounded near the corner $(1, 1)$ of Ω, and the corresponding exact solution u, which is twice differentiable in Ω but possesses an unbounded gradient near $(1, 1)$ [24]:

$$u(x) = -\sqrt{2 - |x|^2}, \quad f(x) = 2(2 - |x|^2)^{-2} \quad \forall x \in \Omega.$$

Table 2 shows that our method converges as the mesh size h decreases, but with a reduced rate and at the cost of an increased number of Newton iterations. We choose δ and θ similarly to the smooth Hessian case, but without any theoretical justification from [30]. We note that now we do not follow the approach of interpolating the coarse solution to the finer mesh, because $u \notin W^2_\infty(\Omega)$. Instead, we use the initial guess that corresponds to $\Delta u_0 = (d!f)^{1/d}$, which introduces more damping, say $\tau < 1$, in the Newton method.

TABLE 2. Unbounded f. We observe that the method converges, but with a rate slower than linear and at the cost of an increasing number of Newton iterations with each refinement.

Degrees of freedom	P: number of points	$L_\infty-$error	Newton steps
N= 1089, $h = 2^{-5}$	16	$8.3\ 10^{-3}$	8
N=4225, $h = 2^{-6}$	24	$5.0\ 10^{-3}$	15
N=16641, $h = 2^{-7}$	36	$3.3\ 10^{-3}$	18
N= 66049, $h = 2^{-8}$	52	$2.0\ 10^{-3}$	50

Computational performance: The process of locating the triangle of the mesh containing $x_i \pm \delta v_j$ and computing the barycentric coordinates is a rather small percentage of the total computing time. For instance, for $h = 2^{-7}$ and the smooth

Hessian, this represents just 3% (< 3 sec) of the total computation time (90 sec). Because of the reduced sparsity pattern of the Newton matrices, the most time demanding task of the method is solving the linear systems, for which we use MATLAB's backslash operator. This takes 42.7% of the total time. This computation is performed on an Intel 2.2 GHz i7 CPU, 16 GB RAM using MATLAB R2017b.

Acknowledgments

We are indebted to S. W. Walker for providing assistance and guidance with the software FELICITY and to H. Antil for numerous discussions about implementing the two-scale method and the semi-smooth Newton solver.

References

[1] N. E. Aguilera and P. Morin, *On convex functions and the finite element method*, SIAM J. Numer. Anal. **47** (2009), no. 4, 3139–3157. MR2551161

[2] G. Awanou, *Convergence rate of a stable, monotone and consistent scheme for the Monge-Ampère equation*, Symmetry **8** (2016), no. 4, Art. 18, 7. MR3488004

[3] G. Awanou, *Standard finite elements for the numerical resolution of the elliptic Monge-Ampère equation: Aleksandrov solutions*, ESAIM Math. Model. Numer. Anal. **51** (2017), no. 2, 707–725. MR3626416

[4] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Anal. **4** (1991), no. 3, 271–283. MR1115933

[5] J.-D. Benamou, F. Collino, and J.-M. Mirebeau, *Monotone and consistent discretization of the Monge-Ampère operator*, Math. Comp. **85** (2016), no. 302, 2743–2775. MR3522969

[6] J.-D. Benamou, B. D. Froese, and A. M. Oberman, *Two numerical methods for the elliptic Monge-Ampère equation*, M2AN Math. Model. Numer. Anal. **44** (2010), no. 4, 737–758. MR2683581

[7] D. P. Bertsekas, *Convex Analysis and Optimization*, Athena Scientific, Belmont, MA, 2003. With Angelia Nedić and Asuman E. Ozdaglar. MR2184037

[8] Z. Błocki, *Smooth exhaustion functions in convex domains*, Proc. Amer. Math. Soc. **125** (1997), no. 2, 477–484. MR1350934

[9] S. C. Brenner and L. R. Scott, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008. MR2373954

[10] S. C. Brenner, T. Gudi, M. Neilan, and L.-y. Sung, *C^0 penalty methods for the fully nonlinear Monge-Ampère equation*, Math. Comp. **80** (2011), no. 276, 1979–1995. MR2813346

[11] L. Caffarelli, L. Nirenberg, and J. Spruck, *The Dirichlet problem for nonlinear second-order elliptic equations. I. Monge-Ampère equation*, Comm. Pure Appl. Math. **37** (1984), no. 3, 369–402. MR739925

[12] X. Chen, Z. Nashed, and L. Qi, *Smoothing methods and semismooth methods for nondifferentiable operator equations*, SIAM J. Numer. Anal. **38** (2000), no. 4, 1200–1216. MR1786137

[13] M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 1–67. MR1118699

[14] E. J. Dean and R. Glowinski, *An augmented Lagrangian approach to the numerical solution of the Dirichlet problem for the elliptic Monge-Ampère equation in two dimensions*, Electron. Trans. Numer. Anal. **22** (2006), 71–96. MR2208483

[15] E. J. Dean and R. Glowinski, *On the Numerical Solution of the Elliptic Monge-Ampère Equation in Dimension Two: A Least-squares Approach*, Partial differential equations, Comput. Methods Appl. Sci., vol. 16, Springer, Dordrecht, 2008, pp. 43–63. MR2484684

[16] G. De Philippis and A. Figalli, *Second order stability for the Monge-Ampère equation and strong Sobolev convergence of optimal transport maps*, Anal. PDE **6** (2013), no. 4, 993–1000. MR3092736

[17] G. De Philippis and A. Figalli, *The Monge-Ampère equation and its link to optimal transportation*, Bull. Amer. Math. Soc. (N.S.) **51** (2014), no. 4, 527–580. MR3237759

[18] R. Engelking, *General Topology*, 2nd ed., Sigma Series in Pure Mathematics, vol. 6, Heldermann Verlag, Berlin, 1989. Translated from the Polish by the author. MR1039321

[19] X. Feng, R. Glowinski, and M. Neilan, *Recent developments in numerical methods for fully nonlinear second order partial differential equations*, SIAM Rev. **55** (2013), no. 2, 205–267. MR3049920

[20] X. Feng and M. Jensen, *Convergent semi-Lagrangian methods for the Monge-Ampère Equation on unstructured grids*, SIAM J. Numer. Anal. **55** (2017), no. 2, 691–712. MR3623696

[21] X. Feng and M. Neilan, *Mixed finite element methods for the fully nonlinear Monge-Ampère equation based on the vanishing moment method*, SIAM J. Numer. Anal. **47** (2009), no. 2, 1226–1250. MR2485451

[22] X. Feng and M. Neilan, *Vanishing moment method and moment solutions for fully nonlinear second order partial differential equations*, J. Sci. Comput. **38** (2009), no. 1, 74–98. MR2472219

[23] B. D. Froese, *A numerical method for the elliptic Monge-Ampère equation with transport boundary conditions*, SIAM J. Sci. Comput. **34** (2012), no. 3, A1432–A1459. MR2970259

[24] B. D. Froese and A. M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampère equation in dimensions two and higher*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1692–1714. MR2831067

[25] R. Glowinski, *Numerical methods for fully nonlinear elliptic equations*, ICIAM 07—6th International Congress on Industrial and Applied Mathematics, Eur. Math. Soc., Zürich, 2009, pp. 155–192. MR2588593

[26] C. E. Gutiérrez, *The Monge-Ampère Equation*, Progress in Nonlinear Differential Equations and their Applications, vol. 44, Birkhäuser Boston, Inc., Boston, MA, 2001. MR1829162

[27] M. Hintermüller, K. Ito, and K. Kunisch, *The primal-dual active set strategy as a semismooth Newton method*, SIAM J. Optim. **13** (2002), no. 3, 865–888 (2003). MR1972219

[28] H. Ishii and P.-L. Lions, *Viscosity solutions of fully nonlinear second-order elliptic partial differential equations*, J. Differential Equations **83** (1990), no. 1, 26–78. MR1031377

[29] J.-M. Mirebeau, *Discretization of the 3D Monge-Ampere operator, between wide stencils and power diagrams*, ESAIM Math. Model. Numer. Anal. **49** (2015), no. 5, 1511–1523. MR3423234

[30] R. H. Nochetto, D. Ntogkas, and W. Zhang, *Two-scale method for the Monge–Ampère equation: pointwise error estimates*, IMA J. Numer. Anal. DOI:10.1093/imanum/dry026

[31] R. H. Nochetto and W. Zhang, *Discrete ABP estimate and convergence rates for linear elliptic equations in non-divergence form*, Found. Comp. Math. (online), 2017.

[32] R. H. Nochetto and W. Zhang, *Pointwise rates of convergence for the Oliker-Prussner method for the Monge-Ampère equation*, arXiv:1611.02786.

[33] A. M. Oberman, *Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems*, SIAM J. Numer. Anal. **44** (2006), no. 2, 879–895. MR2218974

[34] A. M. Oberman, *The convex envelope is the solution of a nonlinear obstacle problem*, Proc. Amer. Math. Soc. **135** (2007), no. 6, 1689–1694. MR2286077

[35] V. I. Oliker and L. D. Prussner, *On the numerical solution of the equation* $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x \partial y))^2 = f$ *and its discretizations. I*, Numer. Math. **54** (1988), no. 3, 271–293. MR971703

[36] N. S. Trudinger and X.-J. Wang, *Boundary regularity for the Monge-Ampère and affine maximal surface equations*, Ann. of Math. (2) **167** (2008), no. 3, 993–1028. MR2415390

[37] G. Wachsmuth, *Conforming approximation of convex functions with the finite element method*, Numer. Math. **137** (2017), no. 3, 741–772. MR3712291

[38] S. W. Walker, *FELICITY: A Matlab/C++ Toolbox for Developing Finite Element Methods and Simulation Modeling*, submitted.

[39] S. W. Walker, *FELICITY: Finite ELement Implementation and Computational Interface Tool for You*. http://www.mathworks.com/matlabcentral/fileexchange/31141-felicity.

Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742
Email address: rhn@math.umd.edu

Department of Mathematics, University of Maryland, College Park, Maryland 20742
Email address: dimitnt@gmail.com

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08854
Email address: wujun@math.rutgers.edu