2021 UNC Greensboro PDE Conference,
Electronic Journal of Differential Equations, Conference 26 (2022), pp. 59–95.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu

A NARROW-STENCIL FRAMEWORK FOR CONVERGENT NUMERICAL APPROXIMATIONS OF FULLY NONLINEAR SECOND ORDER PDES

XIAOBING FENG, THOMAS LEWIS, KELLIE WARD

Abstract. This article develops a unified general framework for designing convergent finite difference and discontinuous Galerkin methods for approximating viscosity and regular solutions of fully nonlinear second order PDEs. Unlike the well-known monotone (finite difference) framework, the proposed new framework allows for the use of narrow stencils and unstructured grids which makes it possible to construct high order methods. The general framework is based on the concepts of consistency and g-monotonicity which are both defined in terms of various numerical derivative operators. Specific methods that satisfy the framework are constructed using numerical moments. Admissibility, stability, and convergence properties are proved, and numerical experiments are provided along with some computer implementation details.

1. Introduction

This article develops a unified general framework for designing convergent narrow-stencil finite difference (FD) and discontinuous Galerkin (DG) methods for approximating the viscosity (and regular) solution to the following fully nonlinear second order Dirichlet boundary value problem:

$$F[u](\mathbf{x}) \equiv F(D^2u, \nabla u, u, \mathbf{x}) = 0, \quad \forall \mathbf{x} \in \Omega, \tag{1.1a}$$

$$u(\mathbf{x}) = g(\mathbf{x}), \quad \forall \mathbf{x} \in \partial\Omega, \tag{1.1b}$$

where $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ is a bounded domain and $D^2u(\mathbf{x})$ denotes the Hessian matrix of u at $\mathbf{x}$. The partial differential equation (PDE) operator $F : \mathcal{S}^{d\times d} \times \mathbb{R}^d \times \mathbb{R} \times \overline{\Omega} \to \mathbb{R}$, where $\mathcal{S}^{d\times d} \subset \mathbb{R}^{d\times d}$ denotes the set of $d \times d$ symmetric real matrices, is a fully nonlinear second order differential operator in the sense that F is nonlinear in at least one component of the Hessian D^2u. The boundary data g is assumed to be continuous with F Lipschitz with respect to its first three arguments. Moreover, F is assumed to be *uniformly and proper elliptic* and satisfy a comparison principle (see Section 2 for the definitions). In this paper we focus our attention on two main classes of fully nonlinear second order PDEs, namely, the Monge-Ampère-type and Hamilton-Jacobi-Bellman-type equations (cf. [21, 9]).

2020 *Mathematics Subject Classification.* 65N06, 65N12.
Key words and phrases. Fully nonlinear PDEs; viscosity solutions; Monge-Ampère equation; Hamilton-Jacobi-Bellman equation; narrow-stencil; generalized monotonicity; g-monotonicity; numerical operators; numerical moment.

Published August 25, 2022.

Fully nonlinear second order PDEs arise from many scientific and engineering applications such as antenna design, astrophysics, economics, differential geometry, stochastic optimal control, and optimal mass transport; yet, they are a class of PDEs which are difficult to study analytically and even more challenging to approximate numerically. Due to the fully nonlinear structure, there is no general variational (or weak) formulation. As a result, its weak solution concept (called *viscosity solutions*, see Section 2 for the definition) is complicated and, in particular, very difficult to address numerically. Nevertheless, driven by the need for solving many emerging and intriguing application problems, numerical fully nonlinear PDEs has garnered a lot of attention and experienced rapid developments in recent years. See [9, 27] and the references therein for an overview of various numerical methods that have been proposed, analyzed, and tested.

To the best of our knowledge, there are only two main approaches in the literature which aim to approximate viscosity solutions. Both approaches have been used successfully to design and analyze (practical) numerical methods for approximating second order fully nonlinear PDEs. The first approach, which was adopted in the overwhelming majority of the existing works, is the *Barles-Souganidis' monotone (wide-stencil) finite difference framework* (cf. [1, 29, 7, 10, 22, 28]). We call the second approach the *numerical moment-enhanced g-monotone (narrow-stencil) finite difference and DG framework* (cf. [11, 12, 14, 24] and also the original vanishing moment method [17]). It is well-known that monotone (in the sense of Barles-Souganidis [1]) methods are difficult to construct; moreover, they are intrinsically low order, require the use of wide stencils, and yield strongly coupled nonlinear algebraic problems that need to be solved. The narrow-stencil approach aims to sidestep these limitations of the wide-stencil approach so high order methods can be constructed on both structured and non-structured grids. On the other hand, since the narrow-stencil approach abandons the standard monotonicity requirement, it prevents one to directly use the powerful Barles-Souganidis' framework for the convergence analysis. Consequently, new machineries and techniques must be developed for the analysis of the proposed narrow-stencil methods which, so far, has only been done on a case-by-case basis in [11, 12, 14, 24].

The primary goal of this paper is to re-examine (with a top-down view) and refine the *numerical moment-enhanced g-monotone (narrow-stencil) finite difference and DG approach*, which was initiated by us in [11, 12, 14, 24], and to formulate it into a unified framework which is parallel to the Barles-Souganidis' framework. A far-reaching goal is to provide a blueprint/framework for designing and analyzing practical convergent numerical methods for approximating viscosity (and regular) solutions of fully nonlinear second order PDEs.

The remainder of this paper is organized as follows. In Section 2, we first recall the basics of viscosity solution theory and elliptic operators as well as the necessary notations. We then introduce various FD and DG finite element numerical derivative operators (cf. [14, 16]) and unify the notations. Those discrete derivative operators are the building blocks for our narrow-stencil framework. In Section 3, we formulate our abstract framework and introduce the key concepts of *numerical operators, consistency, g-monotonicity,* and *numerical moments*. We then state a few structure conditions/assumptions for numerical operators. It should be noted that all of these concepts and conditions are motivated by and abstractions of similar ones in our earlier works [11, 12, 14, 24]. The *numerical moment* will play a critical

role in the specific examples of numerical operators that are presented. In Sections 4–6, we present a complete convergence, admissibility, and stability analysis for the narrow-stencil FD methods proposed in Section 3. Unlike the Barles-Souganidis' monotone (wide-stencil) framework where admissibility and the ℓ^∞-norm stability of the underlying numerical methods are almost free to obtain (thanks to the monotonicity), our results require entirely new techniques and their proofs given in Sections 4–6 are more technical and involved (as well as much longer). These technical issues are precisely the price to pay for using narrow stencils. Assuming the admissibility and ℓ^∞ stability, we first establish the convergence of the numerical solution to the viscosity solution of the underlying PDE problem in Section 4. The proof is adapted from the much more detailed version in [14]. We then prove the desired admissibility and ℓ^∞ stability in Sections 5 and 6. Our main idea of proving the admissibility is to use the Contractive Mapping Theorem in ℓ^2 instead of ℓ^∞. The ℓ^∞ stability is obtained by a novel numerical embedding technique first introduced in [14]. Finally, in Section 7, we present some numerical experiments to demonstrate the effectiveness of the proposed framework and to address some computer implementation issues.

2. Preliminaries and numerical derivatives

2.1. **Notation and definitions.** The narrow-stencil framework will rely upon two different partial orderings for matrices. Let $\mathcal{S}^{d\times d} \subset \mathbb{R}^{d\times d}$ denote the set of symmetric $d \times d$ matrices. We will utilize the convention that $A \geq B$ if $A - B$ is symmetric nonnegative definite for $A, B \in \mathcal{S}^{d\times d}$. We will also introduce the alternative convention that $A \succeq B$ if each component of $A - B$ is nonnegative. Note that the partial ordering induced by $\succeq$ does not require symmetric matrices. We also let $A : B$ denote the Frobenius inner product with $A : B \equiv \sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij}b_{ij}$ for all matrices $A, B \in \mathbb{R}^{d\times d}$.

For a bounded open domain $\Omega \subset \mathbb{R}^d$, let $B(\Omega)$, $USC(\Omega)$, and $LSC(\Omega)$ denote, respectively, the spaces of bounded, upper semi-continuous, and lower semicontinuous functions on Ω. For any $v \in B(\Omega)$, we define

$$v^*(\mathbf{x}) := \limsup_{\mathbf{y}\to\mathbf{x}} v(\mathbf{y}) \quad \text{and} \quad v_*(\mathbf{x}) := \liminf_{\mathbf{y}\to\mathbf{x}} v(\mathbf{y}).$$

Then, $v^* \in USC(\Omega)$ and $v_* \in LSC(\Omega)$, and they are called *the upper and lower semicontinuous envelopes* of v, respectively.

Below we use the convention of writing the boundary condition as a discontinuity of the PDE (cf. [1, p.274]). The Dirichlet boundary condition is assumed to hold in the viscosity sense. The following two definitions can be found in [21, 3, 1].

Definition 2.1. Equation (1.1) is said to be *proper elliptic* if for all $(\mathbf{q}, \mathbf{x}) \in \mathbb{R}^d \times \overline{\Omega}$, it holds

$$F(A, \mathbf{q}, v, \mathbf{x}) \leq F(B, \mathbf{q}, w, \mathbf{x}) \quad \forall A, B \in \mathcal{S}^{d\times d},\ A \geq B,\ v, w \in \mathbb{R},\ v \leq w.$$

Definition 2.2. Equation (1.1) is said to be *uniformly elliptic* if there exists $\Lambda \geq \lambda > 0$ such that, for all $(\mathbf{q}, v, \mathbf{x}) \in \mathbb{R}^d \times \mathbb{R} \times \overline{\Omega}$, it holds

$$0 \geq -\lambda \operatorname{tr}(A - B) \geq F(A, \mathbf{q}, v, \mathbf{x}) - F(B, \mathbf{q}, v, \mathbf{x}) \geq -\Lambda \operatorname{tr}(A - B)$$

for all $A, B \in \mathcal{S}^{d\times d}$ with $A \geq B$.

We note that when $F(A, \mathbf{q}, v, \mathbf{x})$ is differentiable with respect to the first parameter, then the proper ellipticity definition is equivalent to requiring that the matrix $\frac{\partial F}{\partial A}$ is negative semi-definite and the value $\frac{\partial F}{\partial v}$ is nonnegative (cf. [21, p. 441]). If F is also uniformly elliptic, then it holds $0 > -\lambda|\vec{\xi}|^2 \geq \vec{\xi} \cdot \frac{\partial F}{\partial A}\vec{\xi} \geq -\Lambda|\vec{\xi}|^2$ for all $\vec{\xi} \neq \vec{0}$. Thus, $\lambda I \leq -\frac{\partial F}{\partial A} \leq \Lambda I$.

Definition 2.3. A function $u \in B(\Omega)$ is called a viscosity subsolution (resp. supersolution) of (1.1) if, for all $\varphi \in C^2(\overline{\Omega})$, if $u^* - \varphi$ (resp. $u_* - \varphi$) has a local maximum (resp. minimum) at $\mathbf{x}_0 \in \overline{\Omega}$, then we have

$$F_*(D^2\varphi(\mathbf{x}_0), \nabla\varphi(\mathbf{x}_0), u^*(\mathbf{x}_0), \mathbf{x}_0) \leq 0$$

(resp. $F^*(D^2\varphi(\mathbf{x}_0), \nabla\varphi(\mathbf{x}_0), u_*(\mathbf{x}_0), \mathbf{x}_0) \geq 0$). The function u is said to be a viscosity solution of (1.1) if it is simultaneously a viscosity subsolution and a viscosity supersolution of (1.1).

Definition 2.4. Problem (1.1) is said to satisfy a *comparison principle* if the following statement holds. For any upper semicontinuous function u and lower semicontinuous function v on $\overline{\Omega}$, if u is a viscosity subsolution and v is a viscosity supersolution of (1.1), then $u \leq v$ on $\overline{\Omega}$.

Since we assume that F in (1.1) satisfies the comparison principle, we have that the underlying viscosity solution u must be continuous. Furthermore, if F is continuous with respect to $\mathbf{x}$, then by the Lipschitz continuity with respect to D^2u and u, we can drop the upper and lower $*$ indices in Definition 2.3.

2.2. **Finite difference derivative operators.** We introduce several difference operators for approximating first and second order partial derivatives. The narrow-stencil framework will use multiple difference operators to help resolve the underlying viscosity solution. The notation and difference operators used are the same as those in [14]. The section ends with a formal result for comparing various discrete second order operators.

2.2.1. *Finite difference grids.* Assume Ω is a d-rectangle, i.e., $\Omega = (a_1, b_1)\times(a_2, b_2)\times\cdots\times(a_d, b_d)$, and let $\{\mathbf{e}_i\}_{i=1}^d$ denote the canonical basis vectors for $\mathbb{R}^d$. We shall only consider grids that are uniform in each coordinate x_i, $i = 1, 2, \ldots, d$. Let $J_i(\geq 2)$ be an integer and $h_i = \frac{b_i - a_i}{J_i - 1}$ for $i = 1, 2, \ldots, d$. Define $\mathbf{h} = (h_1, h_2, \ldots, h_d) \in \mathbb{R}^d$, $h = \max_{i=1,2,\ldots,d} h_i$, $J = \prod_{i=1}^d J_i$, and $\mathbb{N}_\mathbb{J} = \{\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_d) \mid 1 \leq \alpha_i \leq J_i, i = 1, 2, \ldots, d\}$. Then, $|\mathbb{N}_\mathbb{J}| = J$. We partition Ω into $\prod_{i=1}^d (J_i - 1)$ sub-d-rectangles with grid points $\mathbf{x}_\alpha = \big(a_1+(\alpha_1-1)h_1, a_2+(\alpha_2-1)h_2, \ldots, a_d+(\alpha_d-1)h_d\big)$ for each multi-index $\alpha \in \mathbb{N}_\mathbb{J}$. We call $\mathcal{T}_\mathbf{h} = \{\mathbf{x}_\alpha\}_{\alpha\in\mathbb{N}_\mathbb{J}}$ a mesh (set of nodes) for $\overline{\Omega}$. We also introduce an extended mesh $\mathcal{T}'_\mathbf{h}$ which extends $\mathcal{T}_\mathbf{h}$ by a collection of ghost grid points that are at most one layer exterior to $\overline{\Omega}$ in each coordinate direction. In particular, we choose ghost grid points $\mathbf{x}$ such that $\mathbf{x} = \mathbf{y} \pm 2h_i\mathbf{e}_i$ for some $\mathbf{y} \in \mathcal{T}_\mathbf{h} \cap \Omega$, $i \in \{1, 2, \ldots, d\}$. We set $J'_i = J_i + 2$ and $\mathbb{N}_\mathbb{J}{}'$ is defined by replacing J_i by J'_i in the definition of $\mathbb{N}_\mathbb{J}$ and then removing the extra multi-indices that would correspond to ghost grid points that are not in the set $\mathcal{T}'_\mathbf{h}$ to ensure $|\mathbb{N}_\mathbb{J}{}'| = |\mathcal{T}'_\mathbf{h}|$.

2.2.2. *First order finite difference operators.* Define the (first order) forward and backward difference operators by

$$\delta^+_{x_i,h_i} v(\mathbf{x}) \equiv \frac{v(\mathbf{x} + h_i\mathbf{e}_i) - v(\mathbf{x})}{h_i}, \quad \delta^-_{x_i,h_i} v(\mathbf{x}) \equiv \frac{v(\mathbf{x}) - v(\mathbf{x} - h_i\mathbf{e}_i)}{h_i} \tag{2.1}$$

for a function v defined on $\mathbb{R}^d$. We also consider the central difference operator $\overline{\delta}_{x_i,h_i}$ defined by $\overline{\delta}_{x_i,h_i} \equiv \frac{1}{2}\delta^+_{x_i,h_i} + \frac{1}{2}\delta^-_{x_i,h_i}$ so that

$$\overline{\delta}_{x_i,h_i} v(\mathbf{x}) \equiv \frac{v(\mathbf{x}+h_i\mathbf{e}_i) - v(\mathbf{x}-h_i\mathbf{e}_i)}{2h_i}.$$

The corresponding forward, backward, and central discrete gradient operators are denoted by $\nabla^+_{\mathbf{h}}$, $\nabla^-_{\mathbf{h}}$, and $\overline{\nabla}_{\mathbf{h}}$, respectively.

2.2.3. *Second order finite difference operators.* Using the forward and backward difference operators introduced in the previous subsection, we have the following four possible approximations of the second order differential operator $\partial^2_{x_ix_j}$ given by $D^{\mu\nu}_{\mathbf{h},ij} \equiv \delta^\nu_{x_j,h_j}\delta^\mu_{x_i,h_i}$ for $\mu,\nu \in \{+,-\}$, which in turn leads to the definition of the following four approximations of the Hessian operator $D^2 \equiv [\partial^2_{x_ix_j}]$:

$$D^{\mu\nu}_{\mathbf{h}} \equiv \big[D^{\mu\nu}_{\mathbf{h},ij}\big]^d_{i,j=1} \quad \text{for } \mu,\nu \in \{+,-\}.$$

To analyze our narrow-stencil framework in the next section, we also need to introduce the following two sets of averaged second order difference operators:

$$\begin{aligned}
\widehat{\delta}^2_{x_i,x_j;h_i,h_j} &\equiv \frac{1}{2}\big(D^{+-}_{\mathbf{h},ij} + D^{-+}_{\mathbf{h},ij}\big) = \frac{1}{2}\big(\delta^-_{x_j,h_j}\delta^+_{x_i,h_i} + \delta^+_{x_j,h_j}\delta^-_{x_i,h_i}\big),\\
\widetilde{\delta}^2_{x_i,x_j;h_i,h_j} &\equiv \frac{1}{2}\big(D^{--}_{\mathbf{h},ij} + D^{++}_{\mathbf{h},ij}\big) = \frac{1}{2}\big(\delta^+_{x_j,h_j}\delta^+_{x_i,h_i} + \delta^-_{x_j,h_j}\delta^-_{x_i,h_i}\big),
\end{aligned} \tag{2.2}$$

for all $i,j = 1,2,\dots,d$. Note that, while the various components of $D^{\mu\nu}_{\mathbf{h}}$ may not be self-adjoint operators, the discrete second order partial derivative operators $\widehat{\delta}^2_{x_i,x_j;h_i,h_j}$ and $\widetilde{\delta}^2_{x_i,x_j;h_i,h_j}$ defined by (2.2) are self-adjoint as can be verified by the component forms in [14].

Using the above difference operators, we define the following two "centered" approximations of the Hessian operator $D^2 \equiv [\partial^2_{x_ix_j}]$:

$$\widehat{D}^2_{\mathbf{h}} \equiv \big[\widehat{\delta}^2_{x_i,x_j;h_i,h_j}\big]^d_{i,j=1}, \quad \widetilde{D}^2_{\mathbf{h}} \equiv \big[\widetilde{\delta}^2_{x_i,x_j;h_i,h_j}\big]^d_{i,j=1}. \tag{2.3}$$

We will also consider the average central approximation of the Hessian operator $\overline{D}^2_{\mathbf{h}}$ defined by

$$\overline{D}^2_{\mathbf{h}} \equiv \frac{1}{2}\widehat{D}^2_{\mathbf{h}} + \frac{1}{2}\widetilde{D}^2_{\mathbf{h}}. \tag{2.4}$$

A visual representation of the local stencils for the various discrete Hessians $\widetilde{D}^2_{\mathbf{h}}$, $\widehat{D}^2_{\mathbf{h}}$, and $\overline{D}^2_{\mathbf{h}}$ can be found in Figure 1.

For notation brevity, we set $\delta^2_{x_i,h_i} \equiv \widehat{\delta}^2_{x_i,h_i} \equiv \widehat{\delta}^2_{x_i,x_i;h_i,h_i}$ and $\overline{\delta}^2_{x_i,h_i} \equiv \overline{\delta}^2_{x_i,x_j;h_i,h_j}$. Then

$$\delta^2_{x_i,h_i} v(\mathbf{x}) = \frac{v(\mathbf{x}-h_i\mathbf{e}_i) - 2v(\mathbf{x}) + v(\mathbf{x}+h_i\mathbf{e}_i)}{h_i^2},$$

$$\overline{\delta}^2_{x_i,h_i} = \delta_{x_i,h_i}\delta_{x_i,h_i} = \delta^2_{x_i,2h_i}$$

for all $i = 1,2,\dots,d$. Lastly, we denote the discrete Laplacian operator by $\Delta_{\mathbf{h}} \equiv \sum_{i=1}^d \delta^2_{x_i,h_i}$.

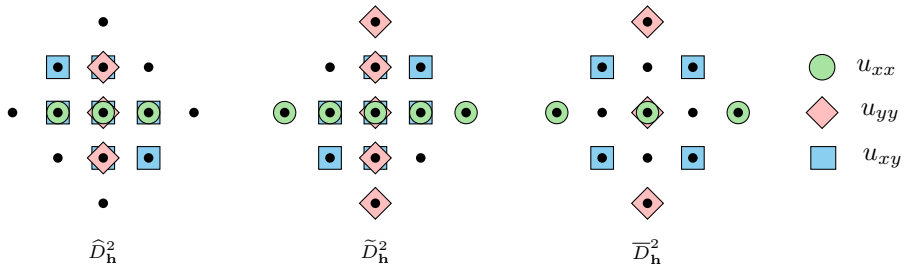

FIGURE 1. Illustration of the local stencils for the three discrete Hessian operators $\widehat{D}^2_{\mathbf{h}}$, $\widetilde{D}^2_{\mathbf{h}}$, and $\overline{D}^2_{\mathbf{h}}$.

2.2.4. *Properties of second order finite difference operators.* We now derive a relationship between $\widehat{D}^2_{\mathbf{h}}$ and $\widetilde{D}^2_{\mathbf{h}}$ that will be essential to the analysis of our narrow-stencil methods. The following expands upon observations in [14].

Let V be a grid function defined over $\mathcal{T}'_{\mathbf{h}}$. Suppose $V = 0$ over $\mathcal{T}_{\mathbf{h}} \cap \partial\Omega$ and V satisfies $\Delta_{\mathbf{h}} V_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}}$, where $\mathcal{S}_{\mathbf{h}} \subset \mathcal{T}_{\mathbf{h}} \cap \partial\Omega$ is defined by

$$\mathcal{S}_{\mathbf{h}} \equiv \big\{ \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega : \mathbf{x}_\alpha + h_i \mathbf{e}_i \in \mathcal{T}_{\mathbf{h}} \cap \Omega \text{ or } \mathbf{x}_\alpha - h_i \mathbf{e}_i \in \mathcal{T}_{\mathbf{h}} \cap \Omega \tag{2.5}$$

for some $i \in \{1, 2, \dots, d\}\big\}$. Since $V = 0$ over $\mathcal{T}_{\mathbf{h}} \cap \partial\Omega$, it also holds $\delta^2_{x_i,h_i} V_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}}$ and $i \in \{1, 2, \dots, d\}$ such that either $\mathbf{x}_{\alpha - \mathbf{e}_i} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$ or $\mathbf{x}_{\alpha + \mathbf{e}_i} \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Note that the additional boundary information is needed to incorporate the ghost points associated with the extended grid $\mathcal{T}'_{\mathbf{h}}$.

Let $\widehat{D}_{ij,0}$ and $\widetilde{D}_{ij,0}$ denote the matrix representations of $[\widehat{D}^2_{\mathbf{h}}]_{ij}$ and $[\widetilde{D}^2_{\mathbf{h}}]_{ij}$, respectively, restricted to grid functions defined over $\mathcal{T}_{\mathbf{h}} \cap \Omega$ with the boundary assumptions for V built in. Notationally, the zero subscript is used to denote the boundary conditions. Then $\widehat{D}_{ii,0}$ is symmetric positive definite for all $i = 1, 2, \dots, d$. Lastly, let $\delta^2_{x_i,h_i;0}$ denote the central difference operator $\delta^2_{x_i,h_i}$ with the zero Dirichlet boundary data assumption.

It is easy to check that (cf. [14]) it holds

$$\widetilde{\delta}^2_{x_i,x_j;h_i,h_j} V_\alpha - \widehat{\delta}^2_{x_i,x_j;h_i,h_j} V_\alpha = \frac{h_i h_j}{2} \delta^2_{x_i,h_i} \delta^2_{x_j,h_j} V_\alpha$$

for all $i, j = 1, 2, \dots, d$ and $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Suppose $i \neq j$. If $\mathbf{x}_{\alpha \pm \mathbf{e}_j} \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega$, then $\delta^2_{x_i,h_i} V_{\alpha \pm \mathbf{e}_j} = 0$ using the fact that x_i is orthogonal to x_j and $V = 0$ over $\mathcal{T}_{\mathbf{h}} \cap \partial\Omega$. Then

$$\delta^2_{x_i,h_i;0} \delta^2_{x_j,h_j;0} V_\alpha = \delta^2_{x_i,h_i;0} \delta^2_{x_j,h_j} V_\alpha = \delta^2_{x_i,h_i} \delta^2_{x_j,h_j} V_\alpha,$$

and, by a simple computation,

$$\delta^2_{x_i,h_i} \delta^2_{x_j,h_j} V_\alpha = \delta^2_{x_j,h_j} \delta^2_{x_i,h_i} V_\alpha = \delta^2_{x_j,h_j;0} \delta^2_{x_i,h_i;0} V_\alpha$$

for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Thus,

$$\widetilde{D}_{ij,0} - \widehat{D}_{ij,0} = \frac{h_i h_j}{2} \widehat{D}_{ii,0} \widehat{D}_{jj,0} = \frac{h_i h_j}{2} \widehat{D}_{jj,0} \widehat{D}_{ii,0}$$

for all $i \neq j$, and it follows that the matrix is symmetric positive definite.

Choose $i \in \{1, 2, \dots, d\}$. Observe that, since $\delta^2_{x_i,h_i} V_\alpha = 0$ over $\mathcal{S}_{\mathbf{h}}$, it holds $\delta^2_{x_i,h_i} \delta^2_{x_i,h_i} V_\alpha = \delta^2_{x_i,h_i;0} \delta^2_{x_i,h_i} V_\alpha$. Thus, by the Dirichlet boundary condition, we have

$$\delta^2_{x_i,h_i} \delta^2_{x_i,h_i} V_\alpha = \delta^2_{x_i,h_i;0} \delta^2_{x_i,h_i;0} V_\alpha,$$

and it follows that

$$\widetilde{D}_{ii,0} - \widehat{D}_{ii,0} = \frac{h_i^2}{2}\widehat{D}_{ii,0}\widehat{D}_{ii,0}.$$

Therefore, the matrix $\widetilde{D}_{ii,0} - \widehat{D}_{ii,0}$ is symmetric positive definite.

We have proved the following lemma, where the result for $\overline{D}_{ij,0} \equiv \frac{1}{2}\widehat{D}_{ij,0} + \frac{1}{2}\widetilde{D}_{ij,0}$ is an immediate consequence:

Lemma 2.5. *The matrix $\widetilde{D}_{ij,0} - \widehat{D}_{ij,0}$ is symmetric positive definite for all $i, j \in \{1, 2, \ldots, d\}$. Furthermore, it holds $-\widehat{D}_{ij,0} > -\overline{D}_{ij,0} > -\widetilde{D}_{ij,0}$.*

2.3. **Discontinuous Galerkin finite element derivative operators.** We note that the above FD numerical derivative operators are only defined on uniform Cartesian grids. In order to extend them to arbitrary grids, and, in particular, to triangular/tetrahedral grids, we utilize finite element DG numerical derivatives which were first introduced in [16] (also see [12]). Below we recall their definitions and some useful properties, but we shall use new notations which are consistent with the above FD discrete derivative operators.

2.3.1. *DG mesh and space notations.* Let Ω be a polygonal domain, and let $\mathcal{T}_h$ denote a locally quasi-uniform shape-regular partition of the domain Ω with $h \equiv \max_{K\in\mathcal{T}_h}(\mathrm{diam}K)$. We introduce the broken L^2-space, broken H^1-space, and broken C^0-space defined by

$$L^2(\mathcal{T}_h) \equiv \prod_{K\in\mathcal{T}_h} L^2(K), \quad H^1(\mathcal{T}_h) \equiv \prod_{K\in\mathcal{T}_h} H^1(K), \quad C^0(\mathcal{T}_h) \equiv \prod_{K\in\mathcal{T}_h} C^0(\overline{K})$$

and the broken L^2-inner product $(\cdot,\cdot)_{\mathcal{T}_h}$ defined by

$$(v, w)_{\mathcal{T}_h} \equiv \sum_{K\in\mathcal{T}_h} \int_K vw\, dx \quad \forall v, w \in L^2(\mathcal{T}_h).$$

Let $\mathcal{E}_h^I$ denote the set of all interior faces/edges of $\mathcal{T}_h$, $\mathcal{E}_h^B$ denote the set of all boundary faces/edges of $\mathcal{T}_h$, and $\mathcal{E}_h \equiv \mathcal{E}_h^I \cup \mathcal{E}_h^B$. Then, for a set $\mathcal{S}_h \subset \mathcal{E}_h$, we define the broken L^2-space by $L^2(\mathcal{S}_h) \equiv \prod_{e\in\mathcal{S}_h} L^2(e)$ and the broken L^2-inner product over $\mathcal{S}_h$ by

$$\langle v, w\rangle_{\mathcal{S}_h} \equiv \sum_{e\in\mathcal{S}_h} \int_e v\, w\, ds \quad \forall v, w \in L^2(\mathcal{S}_h).$$

For a fixed integer $r \geq 0$, we define the standard DG finite element space $V^h \subset H^1(\mathcal{T}_h) \subset L^2(\Omega)$ by

$$V^h \equiv \prod_{K\in\mathcal{T}_h} \mathbb{P}_r(K),$$

where $\mathbb{P}_r(K)$ denotes the set of all polynomials on K with degree not exceeding r.

For $K, K' \in \mathcal{T}_h$, let $e = \partial K \cap \partial K' \in \mathcal{E}_h^I$, and let $v_K, v_{K'}$ denote the restrictions of v to K, K', respectively. Without a loss of generality, we assume that the global labeling number of K is smaller than that of K' and define the following (standard) jump and average notations:

$$[v]\big|_e \equiv v_K\big|_e - v_{K'}\big|_e, \quad \{v\}\big|_e \equiv \frac{v_K\big|_e + v_{K'}\big|_e}{2} \tag{2.6}$$

for any $v \in H^m(\mathcal{T}_h)$. We also define $\mathbf{n}_e \equiv \mathbf{n}_K = -\mathbf{n}_{K'}$ as the normal vector to e.

Based on the formulation in [12], we will extend the jump and average operators to the boundary of the domain in a nonstandard way. Let $K \in \mathcal{T}_h$ such that $e \subset$

$\partial K \in \mathcal{E}_h^B$, and define $\mathbf{n}_e$ as the unit outward normal for the underlying boundary simplex. To unify notation, we impose the convention that the set exterior to the domain Ω has a global labeling number of 0 with the indexing starting at 1 for the "first" label for the simplices in $\mathcal{T}_h$. We also assume $v \in L^2(\mathcal{T}_h) \cap C^0(\mathcal{T}_h) \subset L^2(\Omega)$ can be extended to be in $\big(L^2(\mathcal{T}_h) \times L^2(\Omega^c)\big) \cap \big(C^0(\mathcal{T}_h) \times C^0(\Omega^c)\big)$, where $\Omega^c = \mathbb{R}^d \backslash \Omega$. Then, we can use the same convention as for interior edges and define

$$[v]\big|_e \equiv v_{\Omega^c}\big|_e - v_K\big|_e, \quad \{v\}\big|_e \equiv \frac{v_{\Omega^c}\big|_e + v_K\big|_e}{2}, \tag{2.7}$$

where v_{Ω^c}, v_K denote the restrictions of v to Ω^c, K, respectively. Below we will specify how to choose values for v_{Ω^c} and how to interpret v_K in order to naturally impose a boundary condition.

Using the jump and average operators for $\mathcal{E}_h$, we define the labelling-dependent trace operators $T_i^{\pm}(v) : \mathcal{E}_h \to \mathbb{R}$ for each $i = 1, 2, \ldots, d$ for a given function $v \in H^m(\mathcal{T}_h)$ by

$$T_i^{\pm}(v) \equiv \big\{v\big\} \mp \frac{1}{2}\operatorname{sgn}(n_e^{(i)})\big[v\big] \quad \text{where} \quad \operatorname{sgn}(y) = \begin{cases} 1 & \text{if } y > 0, \\ -1 & \text{if } y < 0, \\ 0 & \text{if } y = 0 \end{cases} \tag{2.8}$$

for all $y \in \mathbb{R}$ and $n_e^{(i)}$ denoting the i-th component of $\mathbf{n}_e$. Note that the exact trace values for v along the boundary still need to be specified for the jump and average operators.

Let $K \in \mathcal{T}_h$ such that $e \subset \partial K \in \mathcal{E}_h^B$. Suppose we have Dirichlet boundary data for the given function v, denoted by g. Then, for $r \geq 1$, we use the convention that $v_{\Omega^c}\big|_e = v_K\big|_e = g\big|_e$ so that $T_i^{\pm}(v) = g$. The convention yields the standard interpretation as introduced in [16]. If $r = 0$, we use the convention that $v_{\Omega^c}\big|_e = g\big|_e$ and $v_K\big|_e$ is given by the interior limit for v. Thus, $T_i^{\pm}(v)$ is given by either g or the interior limit for v depending on the choice for $\pm$ and the sign of the i-th component of the unit normal vector. Such a nonstandard approach allows for weighting degrees of freedom associated with the Dirichlet data against degrees of freedom associated with the value on the interior of K. Since $r = 0$ implies only one degree of freedom is available on K, such a weighting is essential to not overly emphasize the boundary condition. Notationally, we write $T_i^{\pm,g}$ to denote the natural enforcement of the Dirichlet boundary data g. When no boundary data is explicitly given and $r \geq 1$, we use the convention that $v_{\Omega^c}\big|_e = v_K\big|_e$ for v_K given by the interior limit for v. If $r = 0$, we always let $v_K\big|_e$ be given by the interior limit for v and have to assign values for $v_{\Omega^c}\big|_e$ appropriately. Typically we use the convention that $v_{\Omega^c}\big|_e = v_K\big|_e$ or $T_i^+(v) = T_i^-(v)$ with the understanding that $v_K\big|_e$ is given by the interior limit for v. More explicit values for $v_{\Omega^c}\big|_e$ can be assigned based on context as in [12]. Notationally, the use of $T_i^{\pm}$ denotes the lack of given Dirichlet data.

Remark 2.6.

(a) The trace operators $T_i^{\pm}$ and $T_i^{\pm,g}$ are nonstandard in that their values depend on the individual components of the edge normal $\mathbf{n}_e$. The standard definition used for LDG assigns a single-value (called a numerical flux) based on the edge normal vector as a whole.

(b) A labelling-independent definition can also be used so that $T_i^{\pm}$ can be associated with the upwind or downwind direction with respect to the x_i axis. The conventions are equivalent on a uniform Cartesian mesh using the natural ordering. See [25] for more details.

2.3.2. *First order DG derivative operators.* The main idea in [16] for defining DG derivative operators is to use the following local integration by parts formula for a given function $v \in H^1(\mathcal{T}_{\mathbf{h}}) \cap C^0(\overline{\Omega})$:

$$\int_K v_{x_i}\varphi\,dx = \int_{\partial K} v\varphi(x^I)\,n_i\,ds - \int_K v\varphi_{x_i}\,dx, \quad i = 1, 2, \dots, d,\ K \in \mathcal{T}_h, \tag{2.9}$$

with test functions φ chosen from the DG space V^h, $\varphi(x^I)$ denoting the limit from the interior of K, and n_i denoting the i-th component of the unit outward normal vector for K. Thus, the DG (partial in x_i) derivative intends to approximate the weak partial derivative v_{x_i} for all $v \in H^1(\mathcal{T}_h)$. To this end, the trace value $v|_{\partial K}$ must be appropriately chosen/defined when v is not continuous.

We define DG first order partial derivative operators $\partial^{\pm}_{x_i,h}$ for $i = 1, 2, \dots, d$ as follows: for $u \in H^1(\mathcal{T}_h)$,

$$\int_K \partial^{\pm}_{x_i,h} u\phi\,dx \equiv \int_{\partial K} T_i^{\pm}(u)\phi(x^I)\,n_K^{(i)}\,ds - \int_K u\phi_{x_i}\,dx \quad \forall \phi \in V^h \tag{2.10}$$

for all $K \in \mathcal{T}_h$. Notice that the "forward/backward" DG first order derivative operators $\partial^{\pm}_{x_i,h}$ are different if the values of $T_i^{\pm}(u)$ are different due to a discontinuity in u. It is easy to check that $\partial^{\pm}_{x_i,h}$ coincides with the FD operators $\delta^{\pm}_{x_i,h_i}$ on Cartesian grids when using the natural ordering (cf. [16]). Hence, the forward/backward DG derivative operators are indeed generalizations of the forward/backward difference operators to unstructured grids. When boundary trace data g is known, we define the DG first order partial derivative operators $\partial^{\pm,g}_{x_i,h}$ that naturally enforce the boundary data by

$$\int_K \partial^{\pm,g}_{x_i,h} u\phi\,dx = \int_{\partial K} T_i^{\pm,g}(u)\phi(x^I)\,n_K^{(i)}\,ds - \int_K u\phi_{x_i}\,dx \quad \forall \phi \in V^h \tag{2.11}$$

for all $K \in \mathcal{T}_h$ and $i = 1, 2, \dots, d$.

Using the DG first order partial derivative operators as building blocks, we can define various central DG first order derivative operators and corresponding DG finite element gradient operators. Let $i \in \{1, 2, \dots, d\}$. Then, we define

$$\partial_{x_i,h} \equiv \frac{1}{2}\Big(\partial^{+}_{x_i,h} + \partial^{-}_{x_i,h}\Big)$$

as a generalization of the central difference operator $\overline{\delta}_{x_i,h_i}$. If boundary data g is given, then we define the following two central DG first derivative operators that naturally enforce the boundary condition:

$$\partial^{g}_{x_i,h} \equiv \frac{1}{2}\Big(\partial^{+,g}_{x_i,h} + \partial^{-,g}_{x_i,h}\Big), \quad \overline{\partial}^{g}_{x_i,h} \equiv \frac{1}{2}\Big(\partial^{g}_{x_i,h} + \partial_{x_i,h}\Big).$$

The first operator $\partial^{g}_{x_i,h}$ naturally generalizes the form of the central difference operator δ_{x_i,h_i} when acting on a grid function with known boundary values, and the operator $\overline{\partial}^{g}_{x_i,h}$ generalizes the central difference operator δ_{x_i,h_i} in the sense that both correspond to antisymmetric matrices when vectorized with $g = 0$ (see [15] for the motivation for $\overline{\partial}^{g}_{x_i,h}$). In general, we use $\partial^{g}_{x_i,h}$ when $r = 0$ and $\overline{\partial}^{g}_{x_i,h}$ when

$r \geq 1$ to naturally enforce boundary conditions while appropriately weighting interior degrees of freedom versus fixed boundary data. Corresponding finite element gradient operators $\nabla_h^{\pm}$, ∇_h, $\nabla_h^{\pm,g}$, ∇_h^g, and $\overline{\nabla}_h^g$ are naturally defined by letting all components be given by the appropriate DG first order partial derivative operator. For example, $\nabla_h^{\pm} \equiv \left(\partial_{x_1,h}^{\pm}, \partial_{x_2,h}^{\pm}, \ldots, \partial_{x_d,h}^{\pm}\right)^T$.

2.3.3. *Second order DG derivative operators.* Similar to the finite difference (and to the classical calculus) construction, using first order DG derivative operators as the building blocks, we can easily define their high order extensions. Below we only define the second order operators. We also only define the operators that will be used directly in our framework for approximating fully nonlinear elliptic equations. We will only consider the case when boundary conditions correspond to Dirichlet boundary data. More information about Neumann boundary data can be found in [8, 16].

We first define the following one-sided second order DG partial derivatives:

$$\partial_{x_i x_j,h}^{\mu\nu} \equiv \partial_{x_j,h}^{\nu} \partial_{x_i,h}^{\mu}, \quad \partial_{x_i x_j,h}^{\mu\nu,g} \equiv \partial_{x_j,h}^{\nu} \partial_{x_i,h}^{\mu,g} \quad \mu,\nu \in \{+,-\}, \tag{2.12}$$

where g corresponds to given Dirichlet boundary data. Then we define the eight "sided" $d \times d$ matrix-valued DG Hessian operators

$$D_h^{\mu\nu} \equiv \left[\partial_{x_i x_j,h}^{\mu\nu}\right]_{i,j=1}^{d}, \quad D_h^{\mu\nu,g} \equiv \left[\partial_{x_i x_j,h}^{\mu\nu,g}\right]_{i,j=1}^{d}, \quad \mu,\nu \in \{+,-\}, \tag{2.13}$$

and the six central $d \times d$ matrix-valued DG Hessian operators

$$\widehat{D}_h^2 \equiv \frac{1}{2}\left(D_h^{+-} + D_h^{-+}\right), \quad \widehat{D}_h^{2,g} \equiv \frac{1}{2}\left(D_h^{+-,g} + D_h^{-+,g}\right), \tag{2.14a}$$

$$\widetilde{D}_h^2 \equiv \frac{1}{2}\left(D_h^{++} + D_h^{--}\right), \quad \widetilde{D}_h^{2,g} \equiv \frac{1}{2}\left(D_h^{++,g} + D_h^{--,g}\right), \tag{2.14b}$$

$$\overline{D}_h^2 \equiv \frac{1}{2}\left(\widehat{D}_h^2 + \widetilde{D}_h^2\right), \quad \overline{D}_h^{2,g} \equiv \frac{1}{2}\left(\widehat{D}_h^{2,g} + \widetilde{D}_h^{2,g}\right) \tag{2.14c}$$

that can be used when assuming the underlying method has reduced form as introduced below.

Remark 2.7. It can be shown ([16]) that the above second order DG operators coincide with their corresponding FD operators on Cartesian grids. Moreover, it is easy to see that all of the DG operators defined above can be applied to any piecewise "nice" functions on $\mathcal{T}_h$ including those in V^h.

3. A narrow-stencil and g-monotone numerical framework

In this section we formulate a general framework for both FD and DG methods that can be used to approximate fully nonlinear elliptic boundary value problems using narrow-stencil methods. We first introduce the ideas using FD methods. We then provide examples and extend the ideas to DG methods.

3.1. **A narrow-stencil FD framework.** The narrow-stencil FD schemes that we consider will all correspond to seeking a grid function $U_\alpha : \mathbb{N}_{\mathbb{J}}{}' \to \mathbb{R}$ such that

$$\widehat{F}[U_\alpha, \mathbf{x}_\alpha] = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{3.1a}$$

$$U_\alpha = g(\mathbf{x}_\alpha) \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega, \tag{3.1b}$$

$$\Delta_{\mathbf{h}} U_\alpha = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}} \subset \mathcal{T}_{\mathbf{h}} \cap \partial\Omega \tag{3.1c}$$

for all $\alpha \in \mathbb{N}_{\mathbb{J}}$, where

$$\widehat{F}[U_\alpha, \mathbf{x}_\alpha] \equiv \widehat{F}\big(D_{\mathbf{h}}^{--}U_\alpha, D_{\mathbf{h}}^{-+}U_\alpha, D_{\mathbf{h}}^{+-}U_\alpha, D_{\mathbf{h}}^{++}U_\alpha, \overline{\nabla}_{\mathbf{h}}U_\alpha, U_\alpha, \mathbf{x}_\alpha\big). \tag{3.2}$$

Since the schemes only depend upon the discrete Hessian operators $D_{\mathbf{h}}^{\mu\nu}$ and the discrete gradient operator $\overline{\nabla}_{\mathbf{h}}$, they are inherently narrow-stencil. The multiple Hessian operators are used to avoid the directional resolution approach used for monotone schemes that often lead to the use of wide-stencils. We also note that the auxiliary boundary condition (3.1c) needed to define the ghost points that arise when calculating $D_{\mathbf{h}}^{\pm\pm}U_\alpha$ for nodes $\mathbf{x}_\alpha$ near the boundary could be generalized to setting $\Delta_{\mathbf{h}}U_\alpha = h(\mathbf{x}_\alpha)$ for some bounded function h. A well chosen h can increase the accuracy of the underlying scheme by removing any boundary layer error associated with the auxiliary boundary condition. Such an h can be chosen using a refining process by solving various iterations of (3.1) with increasingly better chosen functions h based on the previous iteration.

The main goal for this paper is to define sufficient conditions that $\widehat{F}$ can satisfy in order to guarantee the scheme (3.1) is admissible and convergent. The following definitions are adapted from the 1D definitions presented in [11].

Definition 3.1.

(i) A function $\widehat{F} : \left(\mathbb{R}^{d\times d}\right)^4 \times \mathbb{R}^d \times \mathbb{R} \times \Omega \to \mathbb{R}$ is called a *numerical operator*.

(ii) A numerical operator $\widehat{F}$ is said to be *consistent* (with the differential operator F) if $\widehat{F}$ satisfies

$$\liminf_{\substack{P^{\mu\nu}\to P;\mu,\nu=-,+\\ \mathbf{q}\to\mathbf{v},\lambda\to v,\xi\to\mathbf{x}}} \widehat{F}(P^{--}, P^{-+}, P^{+-}, P^{++}, \mathbf{q}, \lambda, \xi) \geq F_*(P, \mathbf{v}, v, \mathbf{x}),$$

$$\limsup_{\substack{P^{\mu\nu}\to P;\mu,\nu=-,+\\ \mathbf{q}\to\mathbf{v},\lambda\to v,\xi\to\mathbf{x}}} \widehat{F}(P^{--}, P^{-+}, P^{+-}, P^{++}, \mathbf{q}, \lambda, \xi) \leq F_*(P, \mathbf{v}, v, \mathbf{x}),$$

where F_* and F^* denote, respectively, the lower and upper semi-continuous envelopes of F.

(iii) A numerical operator $\widehat{F}$ is said to be *generalized-monotone or g-monotone* if

$$\widehat{F}(A^{++}, B^{+-}, B^{-+}, A^{--}, \mathbf{q}, v, \mathbf{x}) \leq \widehat{F}(B^{++}, A^{+-}, A^{-+}, B^{--}, \mathbf{q}, w, \mathbf{x})$$

for all $A^{\mu\nu}, B^{\mu\nu} \in \mathbb{R}^{d\times d}$; $\mathbf{q} \in \mathbb{R}^d$; $v, w \in \mathbb{R}$; $\mathbf{x} \in \Omega$ such that $B^{\mu\nu} \succeq A^{\mu\nu}$ and $w \geq v$ for all $\mu, \nu \in \{+, -\}$. A numerical operator $\widehat{F}$ is said to be *uniformly g-monotone* if there exists a constant $\kappa_* > 0$ such that $\widehat{F}(P^{++}, P^{+-}, P^{-+}, P^{--}, \mathbf{q}, v, \mathbf{x})$ is increasing in P^{++}, P^{--}, and v at a rate bounded below by κ_* and decreasing in P^{+-} and P^{-+} at a rate bounded above by $-\kappa_*$ using the partial ordering imposed by $\succeq$.

(iv) A numerical operator $\widehat{F}$ can be written in *reduced form* if there exists a function $\widehat{G} : \left(\mathbb{R}^{d\times d}\right)^2 \times \mathbb{R} \times \Omega \to \mathbb{R}$ such that

$$\widehat{F}(P^{--}, P^{-+}, P^{+-}, P^{++}, \mathbf{v}, v, \mathbf{x}) = \widehat{G}(\widetilde{P}, \widehat{P}, \mathbf{v}, v, \mathbf{x})$$

for $\widetilde{P} \equiv \frac{1}{2}(P^{--} + P^{++})$ and $\widehat{P} \equiv \frac{1}{2}(P^{-+} + P^{+-})$ for all P^{--}, P^{-+}, P^{+-}, $P^{++} \in \mathbb{R}^{d\times d}$; $\mathbf{v} \in \mathbb{R}^d$; $v \in \mathbb{R}$; and $\mathbf{x} \in \Omega$.

Remark 3.2.

(a) When F and $\widehat{F}$ are continuous, the definition of consistency can be simplified to $\widehat{F}(P,P,P,P,\mathbf{v},v,\mathbf{x}) = F(P,\mathbf{v},v,\mathbf{x})$ for all $P \in \mathbb{R}^{d\times d}$, $\mathbf{v} \in \mathbb{R}^d$, $v \in \mathbb{R}$, and $\mathbf{x} \in \Omega$.

(b) When $\widehat{F}$ is differentiable, g-monotonicity can be defined by requiring that the matrices $\frac{\partial \widehat{F}}{\partial P^{--}}$ and $\frac{\partial \widehat{F}}{\partial P^{++}}$ have all nonnegative entries, the matrices $\frac{\partial \widehat{F}}{\partial P^{-+}}$ and $\frac{\partial F}{\partial P^{+-}}$ have all nonpositive entries, and $\frac{\partial \widehat{F}}{\partial v}$ is nonnegative. In other words, $\widehat{F}(\uparrow,\downarrow,\downarrow,\uparrow,\cdot,\uparrow,\cdot)$. For a uniformly g-monotone numerical operator, $\frac{\partial \widehat{F}}{\partial P^{--}} \succeq \kappa_* \mathbf{1}_{d\times d}$, $\frac{\partial \widehat{F}}{\partial P^{++}} \succeq \kappa_* \mathbf{1}_{d\times d}$, and $\frac{\partial \widehat{F}}{\partial v} \geq \kappa_*$ while $-\frac{\partial \widehat{F}}{\partial P^{+-}} \succeq \kappa_* \mathbf{1}_{d\times d}$ and $-\frac{\partial \widehat{F}}{\partial P^{-+}} \succeq \kappa_* \mathbf{1}_{d\times d}$, where $\mathbf{1}_{d\times d}$ denotes the matrix with all components equal to 1.

(c) The g-monotonicity approach for narrow-stencil methods uses a component partial ordering instead of the SPD partial ordering for symmetric matrices. The approach also directly compares high order differences instead of looking directly at function values as in the standard monotonicity approach of Barles and Souganidis making it easier to design g-monotone schemes for a wide class of problems. The consistency of $\widehat{F}$ with F will allow the scheme to also take advantage of the SPD partial ordering associated with a proper elliptic operator.

(d) To simplify notation, we will assume $\widehat{F}$ can be written in reduced form and write $\widehat{F}(\widetilde{P},\widehat{P},\mathbf{v},v,\mathbf{x})$ instead of introducing the new function $\widehat{G}$.

A key tool for designing the g-monotone numerical operators in Section 3.2 is the introduction of a numerical moment as defined in [14]:

Definition 3.3. Let $A : \mathbb{R}^J \times \mathcal{T}_\mathbf{h} \to \mathbb{R}^{d\times d}$ and V be a given grid function. The discrete operator $M : \mathbb{R}^J \to \mathbb{R}$ defined by

$$M[V,\mathbf{x}_\alpha] \equiv A\big(V_\alpha,\mathbf{x}_\alpha\big) : \big(\widetilde{D}^2_\mathbf{h} V_\alpha - \widehat{D}^2_\mathbf{h} V_\alpha\big)$$

for all $\mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h} \cap \Omega$ is called a *numerical moment operator*.

3.2. **Examples of g-monotone FD methods.** We now introduce particular examples of g-monotone FD methods that fulfill the structure assumptions of the narrow-stencil framework. The first method is the Lax-Friedrichs-like method proposed in [14] that uses both a numerical moment and a numerical viscosity (where the g-monotone definition could be extended for multiple discrete gradient arguments). The (general) method is defined by

$$\begin{aligned}\widehat{F}[U_\alpha,\mathbf{x}_\alpha] \equiv F\Big(\overline{D}^2_\mathbf{h} U_\alpha, \overline{\nabla}_\mathbf{h} U_\alpha, U_\alpha, \mathbf{x}_\alpha\Big) + A(U_\alpha,\mathbf{x}_\alpha) : \big(\widetilde{D}^2_\mathbf{h} U_\alpha - \widehat{D}^2_\mathbf{h} U_\alpha\big)\\ - \vec{\beta}(U_\alpha,\mathbf{x}_\alpha)\cdot\big(\nabla^+_\mathbf{h} U_\alpha - \nabla^-_\mathbf{h} U_\alpha\big),\end{aligned}$$

where $\vec{\beta} : \mathbb{R}^J \times \mathcal{T}_\mathbf{h} \to \mathbb{R}^d$ is a vector-valued function and $-\vec{\beta}(U_\alpha,\mathbf{x}_\alpha)\cdot\big(\nabla^+_\mathbf{h} U_\alpha - \nabla^-_\mathbf{h} U_\alpha\big)$ is called a numerical viscosity. Note that the method is (globally) g-monotone and consistent using the framework above for the particular choices $\vec{\beta} = \vec{0}$ and $A = \sigma \mathbf{1}_{d\times d}$ for the constant $\sigma > K/2$, where K denotes the global Lipschitz constant of F with respect to the Hessian argument and $\mathbf{1}_{d\times d}$ denotes the matrix with all entries equal to one. The choices $\vec{\beta} = \vec{0}$ and $A = \sigma \mathbf{1}_{d\times d}$ for $\sigma > K/2$ were the focus in the admissibility, stability, and convergence analysis in [14].

In Section 7 we test the performance of the consistent FD method corresponding to the choice

$$\begin{aligned}\widehat{F}_{\gamma,\sigma}[U_\alpha,\mathbf{x}_\alpha] = F\left(\overline{D}^2_{\mathbf{h}}U_\alpha, \overline{\nabla}_{\mathbf{h}}U_\alpha, U_\alpha, \mathbf{x}_\alpha\right)\\ + (M_\alpha + \gamma I_{d\times d} + \sigma \mathbf{1}_{d\times d}) : \left(\widetilde{D}^2_{\mathbf{h}}U_\alpha - \widehat{D}^2_{\mathbf{h}}U_\alpha\right)\end{aligned} \tag{3.3}$$

for $\sigma \geq 0$ and $\gamma + \sigma \geq 0$, where

$$[M_\alpha]_{ij} \equiv \frac{1}{2}\Big|\frac{\partial F}{\partial P_{ij}}\Big|_{(\overline{D}^2_{\mathbf{h}}U_\alpha, \overline{\nabla}_{\mathbf{h}}U_\alpha, U_\alpha, \mathbf{x}_\alpha)}$$

for all $i,j = 1,2,\dots,d$ using the convention that $F = F(P,\mathbf{v},v,\mathbf{x})$ for $P \in \mathbb{R}^{d\times d}$, $\mathbf{v}\in\mathbb{R}^d$, $v\in\mathbb{R}$, and $\mathbf{x}\in\Omega$. By construction, the method is only locally g-monotone in the sense that the linearization of the method is g-monotone. Furthermore, for $\sigma > 0$, the method is only locally uniformly g-monotone. The admissibility proof in Section 5 when applied to $\widehat{F}_{\gamma,\sigma}$ holds for $\gamma \geq 0$ and $\sigma = 0$ and $\gamma \geq -\sigma$ with $\sigma \geq 0$ if the problem is uniformly elliptic. If the operator F is only degenerate elliptic but globally Lipschitz, the proof would hold for either $\gamma > 0$ or $\sigma > 0$.

Consider the linear problem $A : D^2u = f$. Then, the method $\widehat{F}_{0,0}$ is equivalent to

$$\widehat{F}_{0,0}[U_\alpha,\mathbf{x}_\alpha] = F\left(D^2_{\mathbf{h}}U_\alpha,\mathbf{x}_\alpha\right) = A : D^2_{\mathbf{h}}U_\alpha - f(\mathbf{x}_\alpha)$$

for the discrete Hessian $D^2_{\mathbf{h}}$ defined by

$$\left[D^2_{\mathbf{h}}U_\alpha\right]_{ij} = \begin{cases}\widetilde{\delta}^2_{x_i,x_j;h_i,h_j}U_\alpha & \text{if } a_{ij}(\mathbf{x}_\alpha) > 0,\\ \widehat{\delta}^2_{x_i,x_j;h_i,h_j}U_\alpha & \text{if } a_{ij}(\mathbf{x}_\alpha) \leq 0\end{cases} \tag{3.4}$$

for all $i,j \in \{1,2,\dots,d\}$ and $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega$. Thus, the method takes the form of an upwinding-type method where, instead of matching the choice of the discrete partial derivative approximation to the advection field, we match the choice of the discrete second-order partial derivative approximation to the sign of the corresponding diffusion coefficient in A. For $-A$ symmetric nonnegative definite, we would have $a_{ii} \leq 0$ for all $i = 1,2,\dots,d$. Consequently, when $\gamma = \sigma = 0$, the method would only have a nine-point stencil instead of a 13-point stencil in two-dimensions, and the auxiliary boundary condition $\Delta_{\mathbf{h}}U_\alpha = 0$ would not be required. Similarly, the choice $\gamma = -\sigma$ would also only have a nine-point stencil and would not require the auxiliary boundary condition. We would further reduce the stencil to only seven points if the diffusion coefficient a_{12} has a fixed sign. Thus, the methods based on choosing $\gamma = \sigma = 0$ or $\gamma = -\sigma$ are of particular interest since the choices $\gamma = -\sigma$ or $\gamma \geq 0$ with $\sigma = 0$ represent limiting choices for enforcing the g-monotonicity of a numerical operator $\widehat{F}$ while remaining consistent with the PDE operator F.

3.3. **A narrow-stencil DG framework.** We can also naturally formulate narrow-stencil and g-monotone DG methods by seeking a piecewise polynomial function $u_h \in V^h$ such that

$$\left(\widehat{F}[u_h],\varphi_h\right)_{\mathcal{T}_h} + \gamma^B \sum_{e\in\mathcal{E}^B}\frac{1}{h_e}\langle u_h - g,\varphi_h\rangle_e + \gamma^I\sum_{e\in\mathcal{E}^I}\frac{1}{h_e}\langle[u_h],[\varphi_h]\rangle_e = 0 \tag{3.5}$$

for all $\varphi_h \in V^h$, where $\gamma^B,\gamma^I \geq 0$ and

$$\widehat{F}[u_h] = \widehat{F}\big(D_h^{--,g}u_h, D_h^{-+,g}u_h, D_h^{+-,g}u_h, D_h^{++,g}u_h, \overline{\nabla}_h^g u_h, u_h, \cdot\big)$$

is the same as the numerical operator used for FD methods but is now evaluated using DG derivatives. Notationally, h_e denotes the diameter of e and u_h, φ_h are evaluated using interior limits over $e \in \mathcal{E}^B$ when $\gamma^B > 0$. When $r = 0$, we use ∇_h^g instead of $\overline{\nabla}_h^g$ to approximate the gradient operator since the corresponding trace operators naturally weight exterior limits versus interior limits when assuming only the exterior limit corresponds to g. We also set $\gamma^B = \gamma^I = 0$ when $r = 0$ to ensure consistency with the DG method and the underlying FD method on uniform Cartesian grids.

Observe that the term $\left(\widehat{F}[u_h], \varphi_h\right)_{\mathcal{T}_h}$ corresponds to projecting the numerical operator $\widehat{F}[u_h]$ into the discrete space V^h using an L^2 projection. For a quasi-uniform mesh, the penalty terms can be controlled using the uniform ellipticity assumption for F and the properties of the DWDG method for approximating Poisson's equation derived in [25]. Consequently, we can choose $\gamma^B = \gamma^I = 0$ even when $r \geq 1$. As such, the formulation for DG methods requires projecting the FD formulation into the discrete space and optionally adding penalization. We note that the auxiliary boundary condition is not required for $r \geq 1$ based on the definitions of the DG derivative operators and, in particular, the way in which the boundary trace operators are defined. For $r = 0$, explicit rules for defining the exterior values for the boundary trace operators are provided in [12], and they are consistent with the Dirichlet boundary data and the auxiliary boundary condition (3.1c). Letting $q_i^{\pm} = \partial_{x_i,h}^{\pm,g} u_h$, the difficulty addressed in [12] is how to define $q_i^{\pm}\big|_{\Omega^c}$, which can be thought of as defining ghost points for the partial derivative with respect to x_i when $n_i \neq 0$. We refer the reader to [12] for the complete formulation when $r = 0$.

Remark 3.4.

(a) By construction, the proposed DG methods can be considered "narrow-stencil."

(b) When utilizing ∇_h^g instead of $\overline{\nabla}_h^g$, the DG method (3.5) is equivalent to the nonstandard LDG methods in [12] written in a compact form using the DG finite element calculus. For the unified framework we utilize $\overline{\nabla}_h^g$ to more closely mimic the antisymmetric property of the FD operator δ_{x_i,h_i} as inspired by [15] where DG methods were formulated for approximating stationary Hamilton-Jacobi equations.

(c) The DG method (3.5) is equivalent to the FD method (3.1) when $\mathcal{T}_h$ is a uniform Cartesian mesh and the natural ordering is used. As such, all of the analytical results in Sections 4, 5, and 6 can be extended to (3.5) in this special case while, in general, the DG approach formally allows for higher degree bases and more general meshes.

4. Convergence analysis

In this section we prove that consistent, stable, and g-monotone methods converge to the underlying viscosity solution of (1.1). Similar to the proof in [11], the result will assume the numerical operator can be written in reduced form. The theorem requires the methods defined by (3.1) are *admissible* in the sense that a solution U_α exists. The admissibility of the schemes will be verified in Section 5. We also use the definition in [14] that defines a piecewise constant extension $u_{\mathbf{h}}$ for

a given grid function $U \in S(\mathcal{T}'_{\mathbf{h}})$ by

$$u_{\mathbf{h}}(\mathbf{x}) \equiv U_\alpha, \quad \mathbf{x} \in B_\alpha \tag{4.1}$$

for all $\alpha \in \mathbb{N}'_J$, where $B_\alpha \equiv \prod_{i=1,2,\dots,d} \left(\mathbf{x}_\alpha - \frac{h_i}{2}\mathbf{e}_i, \mathbf{x}_\alpha + \frac{h_i}{2}\mathbf{e}_i\right]$ for all $\mathbf{x} \in \Omega' \equiv \cup_{\alpha\in\mathbb{N}'_J} B_\alpha \supset \overline{\Omega}$.

For transparency, we will only explicitly consider operators F that have the form

$$F[u](\mathbf{x}) = F\left(D^2u, u, \mathbf{x}\right)$$

in (1.1a). We note that the proof can be readily extended to the more general case $F[u](\mathbf{x}) = F\left(D^2u, \nabla u, u, \mathbf{x}\right)$ using the techniques in [14]. Indeed, in the proof below, we would have $\overline{\nabla}_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to \nabla\varphi(\mathbf{x}_0)$ in Case (i) which exploits the consistency of the scheme and (4.6) could be rewritten as

$$\begin{aligned}
&F_*\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \overline{\nabla}_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) - F_*\big(D^2\varphi(\mathbf{x}_0), \nabla\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x_k}\big)\\
&\geq -\lambda\delta^2_{x_\ell, h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{x_k}) - \lambda\left|\varphi_{x_\ell,x_\ell}(\mathbf{x}_0)\right| - K\left(|u_{\mathbf{h_k}}(\mathbf{x_k})| + |\varphi(\mathbf{x}_0)|\right)\\
&\quad - K\sum_{i=1}^{d}\Big(|[\nabla^+_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_i| + |[\nabla^-_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_i| + |\varphi_{x_i}(\mathbf{x}_0)|\Big)\\
&\quad - K\sum_{i=1}^{d}\sum_{j=1,\,(i,j)\neq(\ell,\ell)}^{d}\Big(|[\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |\varphi_{x_ix_j}(\mathbf{x}_0)|\Big)
\end{aligned}$$

so that the fact

$$f_\ell\Big(h_\ell^{([\mathbf{k}_0]_\ell)}\Big)\sum_{i=1}^{d}\Big(|[\nabla^+_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_i| + |[\nabla^-_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_i| + |\varphi_{x_i}(\mathbf{x}_0)|\Big) \to 0$$

for $\min \mathbf{k}$ sufficiently large and $k_\ell \to \infty$ can be exploited in Case (ii).

Theorem 4.1. *Suppose the operator F in* (1.1) *is proper and uniformly elliptic with $\lambda > 0$, g is continuous on $\partial\Omega$, F is Lipschitz continuous with respect to its first two arguments, and* (1.1) *satisfies the comparison principle. Suppose $\widehat{F}$ is consistent, is uniformly g-monotone, and can be written in reduced form, and suppose that $\widehat{F}$ is Lipschitz continuous with respect to its first three arguments when written in reduced form. Let $U \in S(\mathcal{T}'_{\mathbf{h}})$ be the solution to the scheme* (3.1)*, and let $u_{\mathbf{h}}$ denote the piecewise constant extension of U defined by* (4.1)*. If* (3.1) *is admissible and ℓ^∞-norm stable, then $u_{\mathbf{h}}$ converges to u locally uniformly as $\mathbf{h} \to \mathbf{0}^+$.*

Proof. The following is a sketch of the proof that highlights the differences from the complete convergence proof for the Lax-Friedrich's-like method given in [14].

Step 1. Since the underlying FD scheme is assumed to be ℓ^∞-norm stable, there exists a constant $C > 0$ such that $\|u_{\mathbf{h}}\|_{L^\infty(\Omega)} \leq C$ independent of $\mathbf{h}$. Define the upper and lower semicontinuous functions $\overline{u}$ and $\underline{u}$ by

$$\overline{u}(\mathbf{x}) \equiv \limsup_{\mathbf{h}\to\mathbf{0}^+,\,\xi\to\mathbf{x}} u_{\mathbf{h}}(\xi), \quad \underline{u}(\mathbf{x}) \equiv \liminf_{\mathbf{h}\to\mathbf{0}^+,\,\xi\to\mathbf{x}} u_{\mathbf{h}}(\xi),$$

where the limits are understood as multi-limits. We show $\overline{u}$ is a viscosity subsolution of (1.1). The proof that $\underline{u}$ is a viscosity supersolution of (1.1) is analogous. By the comparison principle, we have $\overline{u} = \underline{u}$, and it follows that $u = \overline{u} = \underline{u}$ is the viscosity solution of (1.1).

Let $\varphi \in C^2(\overline{\Omega})$ be a quadratic polynomial such that $\overline{u} - \varphi$ takes a strict local maximum at $\mathbf{x}_0 \in \overline{\Omega}$ with $\overline{u}(\mathbf{x}_0) = \varphi(\mathbf{x}_0)$. Then there exists a ball, $B_{r_0}(\mathbf{x}_0) \subset \mathbb{R}^d$, centered at $\mathbf{x}_0$ with radius $r_0 > 0$ (in the ℓ^∞ metric) such that

$$\overline{u}(\mathbf{x}) - \varphi(\mathbf{x}) < \overline{u}(\mathbf{x}_0) - \varphi(\mathbf{x}_0) = 0 \quad \forall \mathbf{x} \in \big(B_{r_0}(\mathbf{x}_0) \cap \overline{\Omega}\big) \setminus \{\mathbf{x}_0\}. \tag{4.2}$$

Suppose $\mathbf{x}_0 \in \Omega$. We show that

$$F_*\big(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x}_0\big) \le 0 \tag{4.3}$$

based on various cases determined by the regularity of $\overline{u}$ at $\mathbf{x}_0$. Note that if $\mathbf{x}_0 \in \partial\Omega$, then, by the argument in [14], $\overline{u}$ can be shown to satisfy the boundary condition (1.1b) in the viscosity sense.

By the definition of $\overline{u}$ and (4.2), there exists (maximizing) sequences $\{\mathbf{h_k}\}$, $\{\mathbf{x_k}\}$, and $\{\mathbf{z_k}\}$ and a constant $K_0 > 0$ such that

$$\mathbf{h_k} \to \mathbf{0}^+, \tag{4.4a}$$

$$\mathbf{x_k} \to \mathbf{x}_0 \text{ with } \mathbf{x_k} \in \mathcal{T}_{\mathbf{h_k}}, \tag{4.4b}$$

$$u_{\mathbf{h_k}}(\mathbf{x_k}) \to \overline{u}(\mathbf{x}_0), \tag{4.4c}$$

$$\mathbf{z_k} \to \mathbf{x}_0 \quad \text{with } |x_i^{(k_i)} - z_i^{(k_i)}| \le \frac{1}{2} h_i^{(k_i)} \text{ and } u_{\mathbf{h_k}}(\mathbf{z_k}) = u_{\mathbf{h_k}}(\mathbf{x_k}), \tag{4.4d}$$

$$u_{\mathbf{h_k}}(\mathbf{z}) - \varphi(\mathbf{z}) \text{ is locally maximized at } \mathbf{z} = \mathbf{z_k} \text{ for all } \min \mathbf{k} \ge K_0. \tag{4.4e}$$

Let $H^{(\mathbf{k})} \in \mathbb{R}^{d\times d}$ be defined by $H^{(\mathbf{k})} = D^2 u^{(\mathbf{k})}(\mathbf{z_k})$ using the convention in [14] to define the local interpolation functions $u^{(\mathbf{k})}$.

Case (i): $\{H^{(\mathbf{k})}\}$ has a uniformly bounded subsequence. In this case, there exists a symmetric matrix $H \in \mathbb{R}^{d\times d}$ and a subsequence (not relabeled) such that $H^{(\mathbf{k})} \to H$, $\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to H$, and $\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to H$ with $D^2\varphi - H$ symmetric positive semidefinite (see [14] for details). Thus,

$$\begin{aligned}
0 &= \lim_{\min \mathbf{k}\to\infty} \widehat{F}\left[u_{\mathbf{h_k}}, \mathbf{x_k}\right] \\
&= \lim_{\min \mathbf{k}\to\infty} \widehat{F}\left[u_{\mathbf{h_k}}, \mathbf{z_k}\right] \\
&= \lim_{\min \mathbf{k}\to\infty} \widehat{F}\big(\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}), u_{\mathbf{h_k}}(\mathbf{z_k}), \mathbf{z_k}\big) \\
&\ge F_*\left(H, \varphi(\mathbf{x}_0), \mathbf{x}_0\right) \\
&\ge F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x}_0\right)
\end{aligned}$$

by the consistency of the scheme and the ellipticity of F.

Case (ii): $\{H^{(\mathbf{k})}\}$ does not have a uniformly bounded subsequence and there is no set of local interpolation functions $\widetilde{u}_{\mathbf{h_k}}$ such that the sequence $\widetilde{H}^{(\mathbf{k})}$ has a bounded subsequence (see [14] for the definition of $\widetilde{u}_{\mathbf{h_k}}$). If such functions $\widetilde{u}_{\mathbf{h_k}}$ exist, then the argument in Case (i) can be easily updated to show $F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x}_0\right) \le 0$.

There exists a pair of indices (i, j) such that the sequence $[\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k})]_{ij}$ or $[\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k})]_{ij}$ does not have a bounded subsequence. Thus, by (4.4e), there exists an index $\ell \in \{i, j\}$ and a subsequence such that $\delta^2_{\xi_i^j, \mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, $\delta^2_{\eta_i^j, \mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, $\delta^2_{x_\ell, 2h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, or $\delta^2_{x_\ell, h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$ using the notation in [14] to rewrite the components of $[\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k})]_{ij}$ and $[\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k})]_{ij}$ in terms of central difference operators. Since $\widehat{F}$ is uniformly g-monotone, $\widehat{F}$ must be

uniformly increasing with respect to $\delta^2_{\xi_i^j,\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, $\delta^2_{\eta_i^j,\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, or $\delta^2_{x_\ell,2h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$. By sending $k_i \to \infty$ and $k_j \to \infty$ while ensuring $\min \mathbf{k}$ is sufficiently large, if it holds $\delta^2_{\xi_i^j,\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, $\delta^2_{\eta_i^j,\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, or $\delta^2_{x_\ell,2h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$, then there must hold $\delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) \to -\infty$ to ensure $\widehat{F}[u_{\mathbf{h_k}}, \mathbf{x_k}] = 0$ for all $\mathbf{k}$. Therefore, there exists an index ℓ such that the sequence $\delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{z_k}) = \delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{x_k})$ does not have a bounded subsequence as $\min \mathbf{k} \to \infty$.

Choose sequences $\{\mathbf{h_k}\}$, $\{\mathbf{x_k}\}$ that maximize the rate at which $\delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{x_k}) \to -\infty$. By the definition of the scheme, we have

$$\begin{aligned}
0 &= \widehat{F}[u_{\mathbf{h_k}}, \mathbf{x_k}] \\
&= \widehat{F}\big(\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \\
&= F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x_k}\right) \\
&\quad + F_*\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) - F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x_k}\right) \\
&\quad + \widehat{F}\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \\
&\quad - F_*\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \\
&\quad + \widehat{F}\big(\widetilde{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \\
&\quad - \widehat{F}\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big).
\end{aligned} \tag{4.5}$$

Then, by the mean value theorem, the Lipschitz continuity of F, and the uniform and proper ellipticity of F, there exists a constant $K \geq 0$ such that

$$\begin{aligned}
&F_*\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) - F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x_k}\right) \\
&\geq -\lambda\Big(\delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{x_k}) - \varphi_{x_\ell,x_\ell}(\mathbf{x}_0)\Big) - K\left(|u_{\mathbf{h_k}}(\mathbf{x_k})| + |\varphi(\mathbf{x}_0)|\right) \\
&\quad - K\sum_{i=1}^{d} \sum_{j=1,\,(i,j)\neq(\ell,\ell)}^{d} \Big(|[\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |\varphi_{x_i x_j}(\mathbf{x}_0)|\Big) \\
&\geq -\lambda\delta^2_{x_\ell,h_\ell^{(k_\ell)}} u_{\mathbf{h_k}}(\mathbf{x_k}) - \lambda\left|\varphi_{x_\ell,x_\ell}(\mathbf{x}_0)\right| - K\left(|u_{\mathbf{h_k}}(\mathbf{x_k})| + |\varphi(\mathbf{x}_0)|\right) \\
&\quad - K\sum_{i=1}^{d} \sum_{\substack{j=1 \\ (i,j)\neq(\ell,\ell)}}^{d} \Big(|[\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |\varphi_{x_i x_j}(\mathbf{x}_0)|\Big).
\end{aligned} \tag{4.6}$$

Using the consistency of the scheme,

$$\begin{aligned}
&\widehat{F}\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \\
&- F_*\big(\widehat{D}^2_{\mathbf{h_k}} u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big) \geq 0.
\end{aligned} \tag{4.7}$$

Lastly, by the mean value theorem, the Lipschitz continuity of $\widehat{F}$, and the g-monotonicity of $\widehat{F}$, there exists a constant $\widehat{K} \geq 0$ and a sequence $0 \leq a_{\mathbf{k}} \leq K$

such that

$$\begin{aligned}
&\widehat{F}\big(\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big)\\
&- \widehat{F}\big(\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}), \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}), u_{\mathbf{h_k}}(\mathbf{x_k}), \mathbf{x_k}\big)\\
&\geq a_{\mathbf{k}}\left[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}) - \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})\right]_{\ell\ell}\\
&\quad - \widehat{K}\sum_{i=1}^{d}\sum_{j=1,\,(i,j)\neq(\ell,\ell)}^{d}\Big(|[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}|\Big).
\end{aligned} \tag{4.8}$$

Plugging (4.6), (4.7), and (4.8) into (4.5), it follows that

$$\begin{aligned}
0 \geq\ & F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x_k}\right)\\
&- \lambda\delta^2_{x_\ell, h_\ell^{(k_\ell)}}u_{\mathbf{h_k}}(\mathbf{x_k}) + a_{\mathbf{k}}\left[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}) - \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})\right]_{\ell\ell}\\
&- \lambda|\varphi_{x_\ell,x_\ell}(\mathbf{x}_0)| - K\left(|u_{\mathbf{h_k}}(\mathbf{x_k})| + |\varphi(\mathbf{x}_0)|\right)\\
&- K\sum_{i=1}^{d}\sum_{j=1,\,(i,j)\neq(\ell,\ell)}^{d}\Big(|[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |\varphi_{x_ix_j}(\mathbf{x}_0)|\Big)\\
&- \widehat{K}\sum_{i=1}^{d}\sum_{j=1,\,(i,j)\neq(\ell,\ell)}^{d}\Big(|[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}|\Big).
\end{aligned} \tag{4.9}$$

Choose the corresponding optimal function f_ℓ (defined in [14]) and subsequences such that

$$\lim_{\min\mathbf{k}\to\infty} f_\ell\big(h_\ell^{(k_\ell)}\big)\delta^2_{x_{k_\ell}, h_\ell^{(k_\ell)}}u_{\mathbf{h_k}}(\mathbf{x_k}) = -C_\ell \tag{4.10}$$

for some constant $C_\ell > 0$. Then, by [14], it holds

$$\liminf_{\min\mathbf{k}\to\infty} f_\ell\big(h_\ell^{(k_\ell)}\big)\big[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}) - \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})\big]_{\ell\ell} \geq 0$$

implying

$$f_\ell\big(h_\ell^{(k_\ell)}\big)a_{\mathbf{k}}[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k}) - \widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{\ell\ell} \geq -\frac{\lambda}{4}C_\ell \tag{4.11}$$

for all $\min\mathbf{k}$ sufficiently large. Furthermore, by combining the observations in Subcases iia, iib, and iic in the proof of Theorem 6.1 in [14], there exists subsequences such that, for $\min\mathbf{k}$ sufficiently large and $k_\ell >> \max k_j$ for all $j \neq \ell$,

$$\begin{aligned}
\frac{\lambda}{4}C_\ell \geq\ & f_\ell\big(h_\ell^{(k_\ell)}\big)\Big(\lambda|\varphi_{x_\ell,x_\ell}(\mathbf{x}_0)| + K\left(|u_{\mathbf{h_k}}(\mathbf{x_k})| + |\varphi(\mathbf{x}_0)|\right)\Big)\\
&+ Kf_\ell\big(h_\ell^{(k_\ell)}\big)\sum_{i=1}^{d}\sum_{\substack{j=1\\(i,j)\neq(\ell,\ell)}}^{d}\Big(|[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |\varphi_{x_ix_j}(\mathbf{x}_0)|\Big)\\
&+ \widehat{K}f_\ell\big(h_\ell^{(k_\ell)}\big)\sum_{i=1}^{d}\sum_{\substack{j=1\\(i,j)\neq(\ell,\ell)}}^{d}\Big(|[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}| + |[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{ij}|\Big).
\end{aligned} \tag{4.12}$$

(Note that the primary difficulty in showing (4.12) is controlling the contributions of $[\widetilde{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{\ell,j}$ and $[\widehat{D}^2_{\mathbf{h_k}}u_{\mathbf{h_k}}(\mathbf{x_k})]_{\ell,j}$ for $j \neq \ell$ due to the $\frac{1}{h_\ell^{(k_\ell)}}$ factor when approximating mixed derivatives.) Thus, scaling (4.9) by $f_\ell\big(h_\ell^{(k_\ell)}\big)$ and plugging

in (4.10), (4.11), and (4.12), there exists an index $\mathbf{k}_0$ with $\min \mathbf{k}_0$ sufficiently large such that

$$\begin{aligned} 0 &\geq f_\ell\big(h_\ell^{([\mathbf{k}_0]_\ell)}\big) F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x}_{\mathbf{k}_0}\right) + \frac{3\lambda}{4} C_\ell - \frac{\lambda}{4} C_\ell - \frac{\lambda}{4} C_\ell \\ &> f_\ell\big(h_\ell^{([\mathbf{k}_0]_\ell)}\big) F_*\left(D^2\varphi(\mathbf{x}_0), \varphi(\mathbf{x}_0), \mathbf{x}_{\mathbf{k}_0}\right). \end{aligned} \tag{4.13}$$

The bound $0 \geq F_*[\varphi](\mathbf{x}_0)$ follows since $f_\ell\big(h_\ell^{([\mathbf{k}_0]_\ell)}\big) > 0$ and $\mathbf{x}_{\mathbf{k}} \to \mathbf{x}_0$. Hence, (4.3) has been verified.

The remainder of the proof is identical to Steps 4-6 in [14], and the result follows. □

Remark 4.2.

(a) g-monotonicity allowed us to identify a sufficiently positive term when $\lambda > 0$. Consequently, we could strongly exploit the uniformly elliptic structure of the PDE operator F.

(b) Theorem 4.1 is proved under the assumption that the numerical scheme is admissible and ℓ^∞-norm stable. The remainder of the paper verifies sufficient conditions under which the assumptions hold.

5. Admissibility analysis

The goal of this section is to show that the proposed narrow stencil scheme (3.1) has a unique solution whenever the numerical operator $\widehat{F}$ is consistent, is g-monotone, and can be written in reduced form. For transparency, we will only consider operators F that have the form $F[u](\mathbf{x}) = F\left(D^2u, u, \mathbf{x}\right)$ in (1.1a). The proofs can be adapted for the more general case using the techniques in [15] by exploiting the fact that the matrix representation of $\overline{\nabla}_{\mathbf{h}}$ is anti-symmetric.

The idea for proving the well-posedness is to equivalently reformulate the proposed scheme as a fixed point problem and to prove the mapping is contractive in the ℓ^2-norm. To this end, let $S(\mathcal{T}'_{\mathbf{h}})$ denote the space of all grid functions on $\mathcal{T}'_{\mathbf{h}}$, and introduce the mapping $\mathcal{M}_\rho : S(\mathcal{T}'_{\mathbf{h}}) \to S(\mathcal{T}'_{\mathbf{h}})$ defined by

$$\widehat{U} \equiv \mathcal{M}_\rho U, \tag{5.1}$$

where the grid function $\widehat{U} \in S(\mathcal{T}'_{\mathbf{h}})$ is defined by

$$\widehat{U}_\alpha = U_\alpha - \rho \widehat{F}[U_\alpha, \mathbf{x}_\alpha], \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{5.2a}$$

$$\widehat{U}_\alpha = g(\mathbf{x}_\alpha), \quad \text{if } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega, \tag{5.2b}$$

$$\Delta_{\mathbf{h}} \widehat{U}_\alpha = 0, \quad \text{if } \mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}} \tag{5.2c}$$

for $\rho > 0$ an undetermined constant. Clearly, the iteration defined in (5.2) is the standard forward Euler method with pseudo time-step ρ complemented with a boundary condition consistent with (3.1). To show $\mathcal{M}_\rho$ is a contraction, we will linearize the operator via the mean value theorem. As a preliminary result in Section 5.1, we will first consider a simple case when F is linear with constant-valued coefficients and a simple scheme based on $\overline{D}^2_{\mathbf{h}}$ is used to discretize the Hessian. The general case will be considered in Section 5.2.

We do require one additional structure assumption on the numerical operator $\widehat{F}$ to assist in the admissibility and stability proofs. The condition will ensure that the method based on using multiple Hessian operators is compatible with the

uniform ellipticity property of the corresponding PDE problem. We first motivate the property before defining it. Note that the Lax-Friedrich's-like method in [14] and the examples in Section 3.2 satisfy the additional structure assumption.

Suppose that $F = F(P, v, x)$ is uniformly elliptic and differentiable with respect to its first two arguments and $\widehat{F} = \widehat{F}(\widetilde{P}, \widehat{P}, v, x)$ is differentiable with respect to its first three arguments. Then, if $\widehat{F}$ is consistent with F, it holds

$$\frac{\partial F}{\partial P} = \frac{\partial}{\partial P} F(P, v, x) = \frac{\partial}{\partial P} \widehat{F}(P, P, v, x) = \frac{\partial \widehat{F}}{\partial \widetilde{P}} + \frac{\partial \widehat{F}}{\partial \widehat{P}}, \tag{5.3a}$$

$$\frac{\partial F}{\partial v} = \frac{\partial}{\partial v} F(P, v, x) = \frac{\partial}{\partial v} \widehat{F}(P, P, v, x) = \frac{\partial \widehat{F}}{\partial v} \tag{5.3b}$$

for all $P \in \mathbb{R}^{d\times d}$, $v \in \mathbb{R}$, and $x \in \Omega$. Let $A, B \in \mathbb{R}^{d\times d}$, and suppose $P = \frac{1}{2}A + \frac{1}{2}B$. Then, for the Lax-Friedrich's-like method in [14] where

$$\widehat{F}(A, B, v, x) = F\Big(\frac{1}{2}A + \frac{1}{2}B, v, x\Big) + \gamma \mathbf{1}_{d\times d} : (A - B)$$

for γ sufficiently large, it holds

$$\begin{aligned}
&\frac{\partial}{\partial \widetilde{P}} \widehat{F}(A, B, v, x) + \frac{\partial}{\partial \widehat{P}} \widehat{F}(A, B, v, x) \\
&= \frac{1}{2}\frac{\partial}{\partial P} F(P, v, x) + \gamma \mathbf{1}_{d\times d} + \frac{1}{2}\frac{\partial}{\partial P} F(P, v, x) - \gamma \mathbf{1}_{d\times d} \\
&= \frac{\partial}{\partial P} F(P, v, x) \le -\lambda I,
\end{aligned}$$

and for the g-monotone method

$$\widehat{F}(A, B, v, x) = F\left(B, v, x\right) + \gamma \mathbf{1}_{d\times d} : (A - B),$$

it holds

$$\begin{aligned}
\frac{\partial}{\partial \widetilde{P}} \widehat{F}(A, B, v, x) + \frac{\partial}{\partial \widehat{P}} \widehat{F}(A, B, v, x) &= \gamma \mathbf{1}_{d\times d} + \frac{\partial}{\partial P} F(B, v, x) - \gamma \mathbf{1}_{d\times d} \\
&= \frac{\partial}{\partial P} F(B, v, x) \le -\lambda I.
\end{aligned}$$

Note that $\frac{\partial}{\partial P} F(B, v, x)$ may not equal $\frac{\partial}{\partial P} F(P, v, x)$ for $P \neq B$. However, the same uniform ellipticity bound holds. Similarly, for the g-monotone method

$$\widehat{F}(A, B, v, x) = F\left(A, v, x\right) + \gamma \mathbf{1}_{d\times d} : (A - B),$$

it holds

$$\begin{aligned}
\frac{\partial}{\partial \widetilde{P}} \widehat{F}(A, B, v, x) + \frac{\partial}{\partial \widehat{P}} \widehat{F}(A, B, v, x) &= \frac{\partial}{\partial P} F(A, v, x) + \gamma \mathbf{1}_{d\times d} - \gamma \mathbf{1}_{d\times d} \\
&= \frac{\partial}{\partial P} F(A, v, x) \le -\lambda I,
\end{aligned}$$

and again the same uniform ellipticity bound holds. Thus, we assume the following compatibility condition when proving the admissibility and stability of our proposed narrow-stencil schemes.

Definition 5.1. Suppose $F = F(P, v, \mathbf{x})$ is proper elliptic and differentiable with respect to its first two arguments with $\frac{\partial}{\partial P} F(A, w, \mathbf{x}) \le -\lambda I$ and $\frac{\partial}{\partial v} F(A, w, \mathbf{x}) \ge \kappa_0$ for all $A \in \mathcal{S}^{d\times d}$, $w \in \mathbb{R}$, and $\mathbf{x} \in \Omega$. Suppose the numerical operator $\widehat{F}$ is consistent with F and differentiable with respect to its first three arguments. The numerical

operator $\widehat{F} = \widehat{F}(\widetilde{P}, \widehat{P}, v, \mathbf{x})$ is *elliptic compatible* if there exists a constant $c > 0$ independent of $\mathbf{h}$ such that

$$\frac{\partial}{\partial \widetilde{P}} \widehat{F}(A, B, w, \mathbf{x}) + \frac{\partial}{\partial \widehat{P}} \widehat{F}(A, B, w, \mathbf{x}) \leq -c\lambda I$$

and $\frac{\partial}{\partial v} \widehat{F}(A, B, w, \mathbf{x}) \geq c\kappa_0$ for all $A, B \in \mathcal{S}^{d\times d}$, $w \in \mathbb{R}$, and $\mathbf{x} \in \Omega$.

5.1. **Admissibility of a simple method for linear, constant coefficient PDEs.** Consider the linear elliptic boundary value problem

$$\mathcal{L}[u] \equiv -A : D^2 u = -\sum_{i=1}^{d}\sum_{j=1}^{d} a_{ij} u_{x_i x_j} = f \quad \text{in } \Omega, \tag{5.4a}$$

$$u = g \quad \text{on } \partial\Omega, \tag{5.4b}$$

where A is constant-valued and symmetric positive definite, and consider the simple FD scheme corresponding to finding a grid function $U_\alpha : \mathbb{N}_{\mathbb{J}}{}' \to \mathbb{R}$ such that

$$\mathcal{L}_{\mathbf{h}} U_\alpha \equiv -A : \overline{D}^2_{\mathbf{h}} U_\alpha = f(\mathbf{x}_\alpha) \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{5.5a}$$

$$U_\alpha = g(\mathbf{x}_\alpha) \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega, \tag{5.5b}$$

$$\Delta_{\mathbf{h}} U_\alpha = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}} \subset \mathcal{T}_{\mathbf{h}} \cap \partial\Omega \tag{5.5c}$$

for all $\alpha \in \mathbb{N}_{\mathbb{J}}$. We show that (5.5) is equivalent to solving a linear system $L\vec{U} = \mathbf{b}$ with the matrix L symmetric positive definite.

Let $\lambda_0 > 0$ denote the smallest eigenvalue of A. Define $A_0 \equiv A - \lambda_0 I$. Then, A_0 is symmetric nonnegative definite. Thus, there exists an eigenvalue decomposition $A_0 = Q\Lambda Q^T = \sum_{k=1}^{d} \lambda_k \mathbf{q}_k \mathbf{q}_k^T$, where $\lambda_k \geq 0$ and $\{\mathbf{q}_k\}$ forms an orthonormal basis for $\mathbb{R}^d$. Define constants $b_i^{(k)} \in \mathbb{R}$ by $\mathbf{q}_k = \sum_{i=1}^{d} b_i^{(k)} \mathbf{e}_i$, and observe that

$$\mathbf{q}_k \mathbf{q}_k^T = \sum_{i=1}^{d} \big(b_i^{(k)}\big)^2 \mathbf{e}_i \mathbf{e}_i^T + \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d} b_i^{(k)} b_j^{(k)} \mathbf{e}_i \mathbf{e}_j^T.$$

Thus,

$$\begin{aligned} A_0 &= \sum_{k=1}^{d} \lambda_k \Big[\sum_{i=1}^{d} \big(b_i^{(k)}\big)^2 \mathbf{e}_i \mathbf{e}_i^T + \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d} b_i^{(k)} b_j^{(i)} \mathbf{e}_i \mathbf{e}_j^T \Big] \\ &= \sum_{i=1}^{d} \Big[\sum_{k=1}^{d} \lambda_k \big(b_i^{(k)}\big)^2 \Big] \mathbf{e}_i \mathbf{e}_i^T + \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d} \Big[\sum_{k=1}^{d} \lambda_k b_i^{(k)} b_j^{(k)} \Big] \mathbf{e}_i \mathbf{e}_j^T, \end{aligned}$$

and it follows that

$$[A_0]_{ii} = a_{ii} - \lambda_0 = \sum_{k=1}^{d} \lambda_k \big(b_i^{(k)}\big)^2, \quad [A_0]_{ij} = a_{ij} = \sum_{k=1}^{d} \lambda_k b_i^{(k)} b_j^{(k)}$$

for all $i, j = 1, 2, \ldots, d$ with $j \neq i$.

Let $J_0 = |\mathcal{T}_{\mathbf{h}} \cap \Omega|$, $D_i \in \mathbb{R}^{J_0 \times J_0}$ denote the matrix representation of $\overline{\delta}_{x_i, h_i}$ with the boundary condition (5.5b) and $M \in \mathbb{R}^{J_0 \times J_0}$ denote the matrix representation of $-\Delta_{2\mathbf{h}}$ and the boundary conditions (5.5b) and (5.5c). Then $(D_i)^T = -D_i$ and

there exists diagonal matrices $B_i \in \mathbb{R}^{J_0\times J_0}$ for $i = 1, 2, \dots, d$ with nonnegative components such that

$$M = -\sum_{i=1}^{d} (D_iD_i - B_i) \geq -\sum_{i=1}^{d} D_iD_i = \sum_{i=1}^{d} (D_i)^T D_i.$$

The positive components of B_i correspond to nodes $\mathbf{x}_\alpha$ near the boundary and increase the coefficients for U_α. Such corrections are needed in the matrix form to account for the values of $\delta_{x_i,h_i}U_{\alpha'}$ when $\mathbf{x}_{\alpha'} \in \mathcal{T}_\mathbf{h} \cap \partial\Omega$ and for the implementation of the auxiliary boundary condition (5.5c). Indeed, suppose $\mathbf{x}_\alpha - h_i\mathbf{e}_i \in \partial\Omega$. Then

$$\begin{aligned}
\delta^2_{x_i,2h_i}U_\alpha &= \delta_{x_i,h_i}\delta_{x_i,h_i}U_\alpha = \frac{1}{2h_i}\delta_{x_i,h_i}\left(U_{\alpha+\mathbf{e}_i} - U_{\alpha-\mathbf{e}_i}\right)\\
&= \frac{1}{2h_i}\delta_{x_i,h_i}U_{\alpha+\mathbf{e}_i} - \frac{1}{2h_i}\delta_{x_i,h_i}U_{\alpha-\mathbf{e}_i}\\
&= \frac{1}{2h_i}\delta_{x_i,h_i}U_{\alpha+\mathbf{e}_i} - \frac{1}{4h_i^2}U_\alpha + \frac{1}{4h_i^2}g(\mathbf{x}_\alpha - h_i\mathbf{e}_i).
\end{aligned}$$

However, when computing using the matrix representation, D_i treats the boundary value $U_{\alpha-\mathbf{e}_i}$ as a known value in its representation and removes it when calculating D_iU. Consequently, the second application of D_i does not act on the boundary node leading to a smaller coefficient for the adjacent interior node involved in the calculation of $\delta_{x_i,h_i}U_{\alpha-\mathbf{e}_i}$. Thus, M would contain the contribution $\frac{1}{4h_i^2}$ to the coefficient for U_α while $-D_iD_i$ would not.

We can see that (5.5c) ensures that the ghost value $U_{\alpha\pm2\mathbf{e}_i}$ satisfies

$$\frac{1}{2h_i^2}U_{\alpha\pm2\mathbf{e}_i} = -\frac{1}{2h_i^2}U_\alpha + \frac{1}{h_i^2}g(\mathbf{x}_{\alpha\pm\mathbf{e}_i}) - \sum_{j=1,\, j\neq i}^{d} \delta^2_{x_j,h_j}g(\mathbf{x}_{\alpha\pm\mathbf{e}_i})$$

for $\mathbf{x}_{\alpha\pm\mathbf{e}_i} \in \mathcal{T}_\mathbf{h} \cap \partial\Omega$, where $\frac{1}{2h_i^2}U_{\alpha\pm2\mathbf{e}_i}$ is directly involved in the computation of $\overline{\delta}^2_{x_i,h_i}U_\alpha$. Lastly, note that correction terms are not needed when considering the relationship of $\delta_{x_i,h_i}\delta_{x_j,h_j}$ to D_iD_j for $i \neq j$ since the computation of $\delta_{x_i,h_i}U_{\alpha'}$ would only include boundary nodes whenever $\mathbf{x}_\alpha \pm h_j\mathbf{e}_j \in \mathcal{T}_\mathbf{h} \cap \partial\Omega$.

Observe that

$$\begin{aligned}
-A_0 : \overline{D}^2_\mathbf{h} &= -\sum_{i=1}^{d}\sum_{j=1}^{d}[A_0]_{ij}\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_j,h_j}\\
&= -\sum_{i=1}^{d}(a_{ii} - \lambda_0)\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_i,h_i} - \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d} a_{ij}\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_j,h_j}\\
&= -\sum_{i=1}^{d}\sum_{k=1}^{d}\lambda_k\big(b_i^{(k)}\big)^2\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_i,h_i} - \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d}\Big[\sum_{k=1}^{d}\lambda_k b_i^{(k)}b_j^{(k)}\Big]\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_j,h_j}\\
&= -\sum_{k=1}^{d}\lambda_k\Big[\sum_{i=1}^{d}\big(b_i^{(k)}\big)^2\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_i,h_i} + \sum_{i=1}^{d}\sum_{j=1,\, j\neq i}^{d} b_i^{(k)}b_j^{(k)}\overline{\delta}_{x_i,h_i}\overline{\delta}_{x_j,h_j}\Big]\\
&= -\sum_{k=1}^{d}\lambda_k\Big(\Big[\sum_{i=1}^{d} b_i^{(k)}\overline{\delta}_{x_i,h_i}\Big]\Big[\sum_{i=1}^{d} b_i^{(k)}\overline{\delta}_{x_i,h_i}\Big]\Big),
\end{aligned}$$

and it follows that

$$\begin{aligned}
0_{J_0\times J_0} &\le \sum_{k=1}^{d}\lambda_k\Big(\Big[\sum_{i=1}^{d}b_i^{(k)}D_i\Big]^T\Big[\sum_{i=1}^{d}b_i^{(k)}D_i\Big]\Big)\\
&= \sum_{i=1}^{d}\sum_{k=1}^{d}\lambda_k\big(b_i^{(k)}\big)^2D_i^TD_i + \sum_{i=1}^{d}\sum_{j=1,\,j\neq i}^{d}\Big[\sum_{k=1}^{d}\lambda_k b_i^{(k)}b_j^{(k)}\Big]D_i^TD_j\\
&= \sum_{i=1}^{d}(a_{ii}-\lambda_0)D_i^TD_i + \sum_{i=1}^{d}\sum_{j=1,\,j\neq i}^{d}a_{ij}D_i^TD_j\\
&= \sum_{i=1}^{d}\sum_{j=1}^{d}[A_0]_{ij}D_i^TD_j.
\end{aligned}$$

Using the fact that $D_i^T=-D_i$,

$$L \equiv -\sum_{i=1}^{d}\sum_{j=1}^{d}a_{ij}D_iD_j + \sum_{i=1}^{d}a_{ii}B_i \tag{5.6}$$

$$= -\lambda_0\sum_{i=1}^{d}\big(D_i^2-B_i\big) - \sum_{i=1}^{d}\sum_{j=1}^{d}[A_0]_{ij}D_iD_j + \sum_{i=1}^{d}(a_{ii}-\lambda_0)B_i \tag{5.7}$$

$$= \lambda_0 M + \sum_{i=1}^{d}\sum_{j=1}^{d}[A_0]_{ij}D_i^TD_j + \sum_{i=1}^{d}(a_{ii}-\lambda_0)B_i \tag{5.8}$$

$$\ge \lambda_0 M > 0_{J_0\times J_0}. \tag{5.9}$$

Therefore, (5.5) has a unique solution and the matrix representation $L\vec{U}=\mathbf{b}$ yields a symmetric positive definite matrix L.

5.2. **Admissibility for fully nonlinear PDEs.** To show that the mapping $\mathcal{M}_\rho$ has a unique fixed point in $S(\mathcal{T}_{\mathbf{h}}')$, we first establish a lemma that specifies conditions under which $\mathcal{M}_\rho$ is a contraction in ℓ^2. The proof will assume F is differentiable; however, the assumption is for ease of notation and the proof can be extended for F Lipschitz but not differentiable. The result will utilize the following result found in [14]:

Lemma 5.2. *Let $B,F\in\mathbb{R}^{J\times J}$ such that B is symmetric nonnegative definite and F is symmetric positive definite. Define $R\in\mathbb{R}^{J\times J}$ such that R is upper triangular and $F=R^*R$. Then*

$$\|\sigma I - FB\|_2 \le \sigma$$

for all positive constants σ such that $\sigma I > RBR^*$.

Lemma 5.3. *Suppose the operator F in* (1.1) *is proper and uniformly elliptic, differentiable, and Lipschitz continuous with respect to its first two arguments. Suppose $\widehat{F}$ is consistent, g-monotone, can be written in reduced form, and is differentiable with respect to its first three arguments. Choose $U,V\in S(\mathcal{T}_{\mathbf{h}}')$ that satisfy the boundary conditions* (5.5b) *and* (5.5c)*, and let $\widehat{U}=\mathcal{M}_\rho U$ and $\widehat{V}=\mathcal{M}_\rho V$ for $\mathcal{M}_\rho$ defined by* (5.1) *and* (5.2)*. Then, for $\widehat{F}$ elliptic compatible, it holds*

$$\|\widehat{U}-\widehat{V}\|_{\ell^2(\mathcal{T}_{\mathbf{h}})} \le \Big(1-\rho\frac{c\lambda\kappa}{4}-\rho\frac{ck_0}{4}\Big)\|U-V\|_{\ell^2(\mathcal{T}_{\mathbf{h}})}$$

for all $\rho > 0$ sufficiently small, where c is from the definition of elliptic compatibility, $4 > \rho c\lambda\kappa + \rho c k_0$, $\frac{\partial F}{\partial D^2 u} \geq \lambda I$, $\frac{\partial F}{\partial u} \geq k_0$, and $M \geq \kappa I$ for M the matrix representation of $-\Delta_{2\mathbf{h}}$.

Proof. Let $W \equiv V - U$ and $\widehat{W} \equiv \widehat{V} - \widehat{U}$. Then, by the boundary conditions, $\widehat{W}_\alpha = W_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h} \cap \partial\Omega$ and $\Delta_\mathbf{h}\widehat{W}_\alpha = \Delta_\mathbf{h} W_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{S}_\mathbf{h}$. Thus, by the mean value theorem for $\widehat{F} = \widehat{F}(\widetilde{P}, \widehat{P}, v, \mathbf{x})$, it holds

$$
\begin{aligned}
\widehat{W}_\alpha &= W_\alpha - \rho\Big(\widehat{F}[V_\alpha, \mathbf{x}_\alpha] - \widehat{F}[U_\alpha, \mathbf{x}_\alpha]\Big) \\
&= W_\alpha - \rho\Big(\widehat{F}(\widetilde{D}^2_\mathbf{h} V_\alpha, \widehat{D}^2_\mathbf{h} V_\alpha, V_\alpha, \mathbf{x}_\alpha) - \widehat{F}(\widetilde{D}^2_\mathbf{h} U_\alpha, \widehat{D}^2_\mathbf{h} U_\alpha, U_\alpha, \mathbf{x}_\alpha)\Big) \\
&= W_\alpha - \rho\widehat{F}\Big(\overline{D}^2_\mathbf{h} V_\alpha + \frac{1}{2}\widetilde{D}^2_\mathbf{h} V_\alpha - \frac{1}{2}\widehat{D}^2_\mathbf{h} V_\alpha, \overline{D}^2_\mathbf{h} V_\alpha - \frac{1}{2}\widetilde{D}^2_\mathbf{h} V_\alpha + \frac{1}{2}\widehat{D}^2_\mathbf{h} V_\alpha, V_\alpha, \mathbf{x}_\alpha\Big) \\
&\quad + \rho\widehat{F}\Big(\overline{D}^2_\mathbf{h} U_\alpha + \frac{1}{2}\widetilde{D}^2_\mathbf{h} U_\alpha - \frac{1}{2}\widehat{D}^2_\mathbf{h} U_\alpha, \overline{D}^2_\mathbf{h} U_\alpha - \frac{1}{2}\widetilde{D}^2_\mathbf{h} U_\alpha + \frac{1}{2}\widehat{D}^2_\mathbf{h} U_\alpha, U_\alpha, \mathbf{x}_\alpha\Big) \\
&= \Big(1 - \rho\frac{\partial \widehat{F}}{\partial v}\Big) W_\alpha - \rho\sum_{i=1}^d\sum_{j=1}^d \Big(\frac{\partial \widehat{F}}{\partial \widetilde{P}_{ij}} + \frac{\partial \widehat{F}}{\partial \widehat{P}_{ij}}\Big)\overline{\delta}^2_{x_i,x_j;h_i,h_j} W_\alpha \\
&\quad - \rho\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \Big(\frac{\partial \widehat{F}}{\partial \widetilde{P}_{ij}} - \frac{\partial \widehat{F}}{\partial \widehat{P}_{ij}}\Big)\widetilde{\delta}^2_{x_i,x_j;h_i,h_j} W_\alpha \\
&\quad + \rho\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \Big(\frac{\partial \widehat{F}}{\partial \widetilde{P}_{ij}} - \frac{\partial \widehat{F}}{\partial \widehat{P}_{ij}}\Big)\widehat{\delta}^2_{x_i,x_j;h_i,h_j} W_\alpha
\end{aligned}
\tag{5.10}
$$

for all $\mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h} \cap \Omega$.

Let $J_0 = |\mathcal{T}_\mathbf{h} \cap \Omega|$ and $\widehat{\mathbf{W}}, \mathbf{W} \in \mathbb{R}^{J_0}$ denote the vectorization of the grid functions $\widehat{W}$ and W restricted to $\mathcal{T}_\mathbf{h} \cap \Omega$, respectively. We introduce several matrix operators in $\mathbb{R}^{J_0 \times J_0}$ that act on $\mathbf{W}$ and correspond to the FD operators with the boundary data naturally incorporated directly into the definition of the matrix. Then, using the notation of Section 2.2.4, $\widetilde{D}_{ij,0}$ and $\widehat{D}_{ij,0}$ denote the matrix representations of $\widetilde{\delta}^2_{x_i,x_j;h_i,h_j}$ and $\widehat{\delta}^2_{x_i,x_j;h_i,h_j}$, respectively. Similarly, $\overline{D}_{ij,0}$ denotes the matrix representations of $\overline{\delta}^2_{x_i,x_j;h_i,h_j}$. We also let $\widehat{F}_0$ denote the diagonal matrix corresponding to the nodal values of $\frac{\partial \widehat{F}}{\partial v}$, $\widetilde{F}_{ij}$ denote the diagonal matrix corresponding to $\frac{\partial \widehat{F}}{\partial \widetilde{P}_{ij}}$, and $\widehat{F}_{ij}$ denote the diagonal matrix corresponding to $\frac{\partial \widehat{F}}{\partial \widehat{P}_{ij}}$. Then $\widehat{F}_0$, $\widetilde{F}_{ij}$, and $-\widehat{F}_{ij}$ are all nonnegative definite, and we have (5.10) becomes

$$
\begin{aligned}
\widehat{\mathbf{W}} &= (I - \rho\widehat{F}_0)\mathbf{W} - \rho\sum_{i=1}^d\sum_{j=1}^d \Big(\widetilde{F}_{ij} + \widehat{F}_{ij}\Big)\overline{D}_{ij,0}\mathbf{W} \\
&\quad - \rho\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \Big(\widetilde{F}_{ij} - \widehat{F}_{ij}\Big)\widetilde{D}_{ij,0}\mathbf{W} + \rho\frac{1}{2}\sum_{i=1}^d\sum_{j=1}^d \Big(\widetilde{F}_{ij} - \widehat{F}_{ij}\Big)\widehat{D}_{ij,0}\mathbf{W}.
\end{aligned}
$$

Letting M denote the matrix corresponding to $-\Delta_{2\mathbf{h}}$, it follows that

$$\begin{aligned}\widehat{\mathbf{W}} &= (I-\rho\widehat{F}_0)\mathbf{W} - \rho\frac{1}{2}c\lambda M\mathbf{W} - \rho\Big[\sum_{i=1}^{d}\sum_{j=1}^{d}\big(\widetilde{F}_{ij}+\widehat{F}_{ij}\big)\overline{D}_{ij,0} - \frac{1}{2}c\lambda M\Big]\mathbf{W}\\ &\quad - \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\big(\widetilde{F}_{ij}-\widehat{F}_{ij}\big)\big(\widetilde{D}_{ij,0}-\overline{D}_{ij,0}\big)\mathbf{W} \qquad (5.11)\\ &\quad + \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\big(\widetilde{F}_{ij}-\widehat{F}_{ij}\big)\big(\widehat{D}_{ij,0}-\overline{D}_{ij,0}\big)\mathbf{W}.\end{aligned}$$

Choose $\mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h}\cap\Omega$. Let $A(\mathbf{x}_\alpha)\in\mathbb{R}^{d\times d}$ be defined by $A_{ij}(\mathbf{x}_\alpha) = \frac{\partial\widehat{F}}{\partial\widetilde{P}_{ij}} + \frac{\partial\widehat{F}}{\partial\widehat{P}_{ij}}$. Then, $-A(\mathbf{x}_\alpha)\geq c\lambda I$ by the elliptic compatible condition. Observe that

$$\begin{aligned}&\Big(A(\mathbf{x}_\alpha)-\frac{1}{2}c\lambda I\Big):\overline{D}^2_\mathbf{h}W_\alpha\\ &= \sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\frac{\partial\widehat{F}}{\partial\widetilde{P}_{ij}}+\frac{\partial\widehat{F}}{\partial\widehat{P}_{ij}}\Big)\overline{\delta}^2_{x_i,x_j;h_i,h_j}W_\alpha - \frac{1}{2}c\lambda\sum_{i=1}^{d}\overline{\delta}^2_{x_i,h_i}W_\alpha\\ &= \sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\frac{\partial\widehat{F}}{\partial\widetilde{P}_{ij}}+\frac{\partial\widehat{F}}{\partial\widehat{P}_{ij}}\Big)\overline{\delta}^2_{x_i,x_j;h_i,h_j}W_\alpha - \frac{1}{2}c\lambda\Delta_{2\mathbf{h}}W_\alpha.\end{aligned}$$

Then, the term $[\sum_{i=1}^{d}\sum_{j=1}^{d}(\widetilde{F}_{ij}+\widehat{F}_{ij})\overline{D}_{ij,0} - \frac{1}{2}c\lambda M]\mathbf{W}$ in (5.11) is the matrix representation of $\big(A(\mathbf{x}_\alpha)-\frac{1}{2}c\lambda I\big):\overline{D}^2_\mathbf{h}W_\alpha$.

We next utilize the frozen coefficient technique. Define the matrices $L_\alpha\in\mathbb{R}^{J_0\times J_0}$ as the matrix representations of $\big(A(\mathbf{x}_\alpha)-\frac{1}{2}c\lambda I\big):\overline{D}^2_\mathbf{h}$ for all $\mathbf{x}_\alpha\in\mathcal{T}_\mathbf{h}\cap\Omega$. Then, by Section 5.1, we have L_α is symmetric positive definite for all multi-indices α since the coefficient matrix is constant-valued for each fixed value of α. Define $E_k\in\mathbb{R}^{J_0\times J_0}$ by $[E_k]_{ij}=1$ only if $i=j=k$ and 0 otherwise. Notationally, we let $\alpha(k)$ be the multi-index corresponding to the single-index k. Then, (5.11) can be rewritten as

$$\begin{aligned}\widehat{\mathbf{W}} &= (I-\rho\widehat{F}_0)\mathbf{W} - \rho\frac{1}{2}c\lambda M\mathbf{W} - \rho\sum_{k=1}^{J_0}E_kL_{\alpha(k)}\mathbf{W}\\ &\quad - \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\widetilde{F}_{ij}-\widehat{F}_{ij}\Big)\Big(\widetilde{D}_{ij,0}-\overline{D}_{ij,0}\Big)\mathbf{W}\\ &\quad - \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\widetilde{F}_{ij}-\widehat{F}_{ij}\Big)\Big(\overline{D}_{ij,0}-\widehat{D}_{ij,0}\Big)\mathbf{W}\\ &\equiv G\mathbf{W}\end{aligned}$$

for the iteration matrix $G\in\mathbb{R}^{J_0\times J_0}$ defined by

$$\begin{aligned}G &\equiv (I-\rho\widehat{F}_0) - \rho\frac{1}{2}c\lambda M - \rho\sum_{k=1}^{J_0}E_kL_{\alpha(k)}\\ &\quad - \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\widetilde{F}_{ij}-\widehat{F}_{ij}\Big)\Big(\widetilde{D}_{ij,0}-\overline{D}_{ij,0}\Big)\end{aligned}$$

$$- \rho\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\left(\widetilde{F}_{ij} - \widehat{F}_{ij}\right)\left(\overline{D}_{ij,0} - \widehat{D}_{ij,0}\right).$$

Choose $\epsilon > 0$. Let $N = 2 + J_0 + 2d^2$. Observe that $E_k + \epsilon I$ and $\widetilde{F}_{ij} - \widehat{F}_{ij} + \epsilon I$ are symmetric positive definite for all k and all i, j. Choose K such that $0_{J_0\times J_0} \leq \sum_{k=1}^{J_0} L_{\alpha(k)} \leq KI$, $0_{J_0\times J_0} \leq \widetilde{D}_{ij,0} - \overline{D}_{ij,0} \leq KI$, and $0_{J_0\times J_0} \leq \overline{D}_{ij,0} - \widehat{D}_{ij,0} \leq KI$ for all i, j, where the first bound uses the fact that $L_{\alpha(k)}$ is symmetric positive definite for all indices k and the other bounds follow from Lemma 2.5. Then, by Lemma 5.2, it holds

$$\begin{aligned}
\|G\|_2 &\leq \frac{1}{2} + \|\frac{1}{2N}I - \rho\widehat{F}_0\|_2 + \|\frac{1}{2N}I - \rho\frac{1}{2}c\lambda M\|_2 \\
&\quad + \sum_{k=1}^{J_0}\|\frac{1}{2N}I - \rho(E_k + \epsilon I)L_{\alpha(k)}\|_2 + \rho\epsilon\|\sum_{k=1}^{J_0}L_{\alpha(k)}\|_2 \\
&\quad + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\|\frac{1}{N}I - \rho\left(\widetilde{F}_{ij} - \widehat{F}_{ij} + \epsilon I\right)\left(\widetilde{D}_{ij,0} - \overline{D}_{ij,0}\right)\|_2 \\
&\quad + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\|\frac{1}{N}I - \rho\left(\widetilde{F}_{ij} - \widehat{F}_{ij} + \epsilon I\right)\left(\overline{D}_{ij,0} - \widehat{D}_{ij,0}\right)\|_2 \\
&\quad + \rho\epsilon\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\|\widetilde{D}_{ij,0} - \overline{D}_{ij,0}\|_2 + \rho\epsilon\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\|\overline{D}_{ij,0} - \widehat{D}_{ij,0}\|_2 \\
&\leq \frac{1}{2} + \frac{1}{2N} - \rho ck_0 + \frac{1}{2N} - \rho\frac{1}{2}c\lambda\kappa + \sum_{k=1}^{J_0}\frac{1}{2N} + d^2\frac{1}{N} + \rho\epsilon(d^2+1)K \\
&< 1 - \rho\frac{3}{4}ck_0 - \rho\frac{1}{4}c\lambda\kappa
\end{aligned}$$

for all $\rho > 0$ sufficiently small and $\epsilon < c\frac{k_0+\lambda\kappa}{4(d^2+1)K}$ chosen independently of ρ. The bound

$$\|\widehat{W}\|_{\ell^2(\mathcal{T}_{\mathbf{h}}\cap\Omega)} \leq \left(1 - \rho\frac{c\lambda\kappa}{4} - \rho\frac{ck_0}{4}\right)\|W\|_{\ell^2(\mathcal{T}_{\mathbf{h}}\cap\Omega)}$$

follows since $\|\widehat{\mathbf{W}}\|_2 \leq \|G\|_2\|\mathbf{W}\|_2$, and the bound over $\mathcal{T}_{\mathbf{h}}$ follows since $\widehat{W}_\alpha = W_\alpha = 0$ over $\mathcal{T}_{\mathbf{h}} \cap \partial\Omega$. The proof is complete. □

As an immediate corollary to Lemma 5.3, we have the following well-posedness result by the contractive mapping theorem.

Theorem 5.4. *Suppose the operator F in* (1.1) *is proper and uniformly elliptic, differentiable, and Lipschitz continuous with respect to its first two arguments. Suppose $\widehat{F}$ is consistent, is g-monotone, can be written in reduced form, is differentiable with respect to its first three arguments, and is elliptic compatible. The scheme* (3.1) *for approximating problem* (1.1) *has a unique solution.*

Remark 5.5. We emphasize that the properties of F and $\widehat{F}$ guarantee a unique solution whenever $k_0 > 0$ or $\lambda > 0$.

6. Stability analysis

For transparency and consistency with the results in Section 5, assume the operator F in (1.1a) has the form $F[u](\mathbf{x}) = F\left(D^2u, u, \mathbf{x}\right)$.

Theorem 6.1. *Suppose the operator F in (1.1) is proper and uniformly elliptic and is Lipschitz continuous and differentiable with respect to its first two arguments with $\lambda > 0$ or $k_0 > 0$. Suppose $\widehat{F}$ is consistent, is g-monotone, can be written in reduced form, is Lipschitz continuous and differentiable with respect to its first three arguments, and is elliptic compatible. Then the solution U to the scheme (3.1) for approximating problem (1.1) is ℓ^2-norm stable in the sense that*

$$\Big(\prod_{i=1,2,\dots,d} h_i 1/2\Big)\|U\|_{\ell^2(\mathcal{T}_{\mathbf{h}}\cap\Omega)} \le C,$$

where C is a positive $\mathbf{h}$-independent constant which depends on Ω, the lower (proper) ellipticity constants k_0 and λ, F, and g.

Proof. Define the function $v \in C^0(\overline{\Omega}) \cap H^2(\Omega)$ to be the solution to

$$-\Delta v = 0 \quad \text{in } \Omega, \tag{6.1a}$$

$$v = g \quad \text{on } \partial\Omega, \tag{6.1b}$$

and define $V : \mathcal{T}_{\mathbf{h}}' \to \mathbb{R}$ by $V_\alpha = v(\mathbf{x}_\alpha)$ for all $\mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \overline{\Omega}$ and introduce the ghost points so that $\Delta_{\mathbf{h}} V_\alpha = 0$ for all $\mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}}$. Then, by the mean value theorem, there exists a linear operator $\mathcal{L}_{\mathbf{h}}$ such that

$$\widehat{F}[U_\alpha, \mathbf{x}_\alpha] - \widehat{F}[V_\alpha, \mathbf{x}_\alpha] = \mathcal{L}_{\mathbf{h}}[U_\alpha - V_\alpha].$$

Furthermore, since U is a solution to (3.1),

$$\widehat{F}[U_\alpha, \mathbf{x}_\alpha] - \widehat{F}[V_\alpha, \mathbf{x}_\alpha] = 0 - \widehat{F}[V_\alpha, \mathbf{x}_\alpha] = -\widehat{F}[V_\alpha, \mathbf{x}_\alpha].$$

Thus, $U_\alpha - V_\alpha$ is a solution to

$$\mathcal{L}_{\mathbf{h}}[U_\alpha - V_\alpha, \mathbf{x}_\alpha] + \widehat{F}[V_\alpha, \mathbf{x}_\alpha] = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \Omega, \tag{6.2a}$$

$$U_\alpha - V_\alpha = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{T}_{\mathbf{h}} \cap \partial\Omega, \tag{6.2b}$$

$$\Delta_{\mathbf{h}}(U_\alpha - V_\alpha) = 0 \quad \text{for } \mathbf{x}_\alpha \in \mathcal{S}_{\mathbf{h}} \subset \mathcal{T}_{\mathbf{h}} \cap \partial\Omega, \tag{6.2c}$$

where $\mathcal{L}_{\mathbf{h}}$ has the form

$$\mathcal{L}_{\mathbf{h}} \equiv A_{\mathbf{h}} : \widetilde{D}^2_{\mathbf{h}} - B_{\mathbf{h}} : \widehat{D}^2_{\mathbf{h}} + c_{\mathbf{h}}$$

for $A_{\mathbf{h}}, B_{\mathbf{h}} \in \mathcal{S}^{d\times d}$ with all nonnegative components and $c_{\mathbf{h}} \in \mathbb{R}$ with $c_{\mathbf{h}} \ge 0$ by the g-monotonicity of $\widehat{F}$. Furthermore, $\mathcal{L}_{\mathbf{h}}$ is consistent with the linear elliptic boundary value problem

$$\mathcal{L}w \equiv -A : D^2 w + cw = -F[v], \quad \text{in } \Omega, \tag{6.3a}$$

$$u = g, \quad \text{on } \partial\Omega \tag{6.3b}$$

with the matrix A symmetric positive definite with $\lambda I \le A \le \Lambda I$ and $c \ge k_0$, where $A_{\mathbf{h}} - B_{\mathbf{h}} \le -c_0\lambda I$ and $c_{\mathbf{h}} \ge c_0 k_0$ for c_0 the constant based on the elliptic compatibility of $\widehat{F}$.

Applying Theorem 5.4, $U_\alpha - V_\alpha$ is the unique solution to (6.2). Furthermore, by Lemma 5.3 and using the technique in [14, Theorem 4.3], it holds

$$\|U - V\|_{\ell^2(\mathcal{T}_{\mathbf{h}}\cap\Omega)} \le \frac{4}{c_0\lambda\kappa + c_0 k_0}\|\mathcal{L}_{\mathbf{h}}[0] + \widehat{F}[V_\alpha, \mathbf{x}_\alpha]\|_{\ell^2(\mathcal{T}_{\mathbf{h}}\cap\Omega)}$$

$$= \frac{4}{c_0\lambda\kappa + c_0 k_0}\|\widehat{F}[V_\alpha, \mathbf{x}_\alpha]\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)}.$$

Therefore,

$$\|U\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)} \le \|V\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)} + \frac{4}{c_0\lambda\kappa + c_0 k_0}\|\widehat{F}[V_\alpha, \mathbf{x}_\alpha]\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)}$$

and the result follows by the properties of v. The proof is complete. □

Extending the techniques above and following the proofs in Section 5 of [14], we can also prove the following results.

Lemma 6.2. *Let $i \in \{1, 2, \dots, d\}$. Suppose the operator F in* (1.1) *is proper and uniformly elliptic, differentiable, and Lipschitz continuous with respect to its first two arguments. Suppose $\widehat{F}$ is consistent, g-monotone, can be written in reduced form, is differentiable with respect to its first three arguments, and is elliptic compatible. Choose $U, V \in S(\mathcal{T}'_\mathbf{h})$ that satisfy the boundary conditions* (5.5b) *and* (5.5c)*, and let $\widehat{U} = \mathcal{M}_{\rho,2}U$ and $\widehat{V} = \mathcal{M}_{\rho,2}V$, where $\widehat{W} \in S(\mathcal{T}'_\mathbf{h})$ is defined by $\widehat{W} = \mathcal{M}_{\rho,2}W$ for some $W \in S(\mathcal{T}'_\mathbf{h})$ if*

$$\begin{gathered}
-\delta^2_{x_i,2h_i}\widehat{W}_\alpha = -\delta^2_{x_i,2h_i}W_\alpha - \rho\widehat{F}[W_\alpha, \mathbf{x}_\alpha] \quad \forall \mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h}\cap\Omega,\\
\widehat{W}_\alpha = g(\mathbf{x}_\alpha) \quad \forall \mathbf{x}_\alpha \in \mathcal{T}_\mathbf{h}\cap\partial\Omega,\\
\Delta_\mathbf{h}\widehat{W}_\alpha = 0 \quad \forall \mathbf{x}_\alpha \in \mathcal{S}_\mathbf{h}.
\end{gathered}$$

Then it holds

$$\|\delta^2_{x_i,2h_i}(\widehat{U} - \widehat{V})\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)} \le \Big(1 - \rho\frac{c\lambda\kappa}{4}\Big)\|\delta^2_{x_i,2h_i}(U - V)\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)}$$

for all $\rho > 0$ sufficiently small, where $4 > \rho c\lambda\kappa$, $\frac{\partial F}{\partial D^2 u} \ge \lambda I$, $\frac{\partial F}{\partial u} \ge k_0$, and $M \ge \kappa I$ for M the matrix representation of $-\Delta_{2\mathbf{h}}$.

Theorem 6.3. *Suppose the operator F in* (1.1) *is proper and uniformly elliptic, and is Lipschitz continuous and differentiable with respect to its first two arguments with $\lambda > 0$. Suppose $\widehat{F}$ is consistent, is g-monotone, can be written in reduced form, is Lipschitz continuous and differentiable with respect to its first three arguments, and is elliptic compatible. Then the solution U to the scheme* (3.1) *for approximating problem* (1.1) *satisfies*

$$\Big(\prod_{i=1,2,\dots,d} h_i 1/2\Big)\|\delta^2_{x_i,2h_i}U\|_{\ell^2(\mathcal{T}_\mathbf{h}\cap\Omega)} \le C$$

for all $i = 1, 2, \dots, d$, where C is a positive $\mathbf{h}$-independent constant which depends on Ω, the lower (proper) ellipticity constant λ, F, and g.

Lastly, by [14], combining Theorems 6.1 and 6.3 using a novel numerical embedding technique yields the following ℓ^∞ stability result.

Theorem 6.4. *Under the assumptions of Lemma 6.2, the numerical solution U is stable in the ℓ^∞-norm for $d \le 3$; that is, U satisfies $\|U\|_{\ell^\infty(\mathcal{T}_\mathbf{h})} \le C$ for $d \le 3$, where C is a positive constant independent of $\mathbf{h}$.*

7. Numerical experiments

In this section we test the performance of the FD method based on the numerical operator $\widehat{F}_{\gamma,\sigma}$ defined by (3.3) for $\sigma \geq 0$ and $\gamma + \sigma \geq 0$. We consider test problems based on choosing F to be linear and uniformly elliptic with non-divergence form, the Hamilton-Jacobi-Bellman operator, the Monge-Ampère operator, and the operator coming from the equation of prescribed Gauss curvature. We will see that for the linear problem, $\widehat{F}_{0,0}$ yields a nonsingular sparse matrix. However, the nonlinear PDE problems considered are degenerate, and, consequently, do not conform to all of the assumptions used in the convergence, admissibility, and stability analysis. Observationally, the nonlinear solver 'fsolve' in MATLAB has trouble in the degenerate cases finding a zero when using the zero function as the initial guess and σ and γ are small. For the Hamilton-Jacobi-Bellman problem, we can successfully solve the limiting cases by forming a sequence of approximations for decreasing values of γ starting with an initial large value for γ when initializing fsolve with the zero function. For the Monge-Ampère operator and in the equation of prescribed Gauss curvature, we need to form the sequence of approximations with σ large to more strictly impose g-monotonicity with respect to all of the components of the discrete Hessian operators. Once fsolve is in a neighborhood of the solution we can successfully find a zero for all $\sigma \geq 0$ and $\gamma \geq -\sigma$. For the Monge-Ampére problem and the prescribed Gauss curvature problem, we choose an exact solution for which the Monge-Ampère operator is locally uniformly elliptic. It is worth noting that fsolve does not find a false solution but instead always reported no solution found when staring with a poor initial guess. This is in direct contrast to the experiment in [9] that found false solutions to the Monge-Ampère problem when using the standard nine-point finite difference formula for approximating the discrete Hessian.

Additional numerical tests for g-monotone FD methods can be found in [24, 11, 13] and tests for the corresponding DG scheme can be found in [24, 12].

7.1. **Test 1: Linear F with non-aligned grids.** In this test, we consider the linear uniformly elliptic problem $F[u] = -A : D^2u - f$ for a discontinuous coefficient matrix A and uniform grids chosen such that no monotone finite difference method exists (see [26]). The matrix $A(\mathbf{x})$ will be uniformly symmetric positive definite, and f will be a uniformly bounded function. We form a non-singular linear system $LU = F$ that is solved using MATLAB's backslash command. The matrix L is formed using sparse storage.

Let $\Omega = (-1, 1)^2$. We form a sequence of uniform grids with $N_x = N_y = 10, 40, 80, 120, 180, 240, 300$. Let $h_x = h_y = \frac{2}{301}$ based on the finest mesh, define the directions $\mathbf{v}^{(i)} \in \mathbb{R}^2$ by

$$\mathbf{v}^{(1)} = \begin{bmatrix} 1 \\ \frac{h_y}{2} \end{bmatrix}, \quad \mathbf{v}^{(2)} = \begin{bmatrix} h_x \\ 5h_y \end{bmatrix}, \quad \mathbf{v}^{(3)} = \begin{bmatrix} 10h_x \\ h_y \end{bmatrix}, \quad \mathbf{v}^{(4)} = \begin{bmatrix} \frac{h_x}{2} \\ 2 \end{bmatrix},$$

and define the unit length vectors $\mathbf{q}_1^{(i)} = \frac{1}{\|\mathbf{v}^{(i)}\|_2}\mathbf{v}^{(i)}$ for all $i = 1, 2, 3, 4$. Note that, by construction, the directions $\mathbf{v}^{(1)}$ and $\mathbf{v}^{(4)}$ do not align with grid points in any of the meshes while $\mathbf{v}^{(2)}$ aligns with nodes at least 5 layers away and $\mathbf{v}^{(3)}$ aligns with nodes at least 10 layers away. For each $i = 1, 2, 3, 4$, choose unit length vectors $\mathbf{q}_2^{(i)}$ that are orthogonal to $\mathbf{q}_1^{(i)}$. Define the orthogonal matrices $Q_i = [\mathbf{q}_1^{(i)}, \mathbf{q}_2^{(i)}]$ for each

$i = 1, 2, 3, 4$, and define the diagonal matrix $\Lambda(x, y) \in \mathbb{R}^{2\times 2}$ by

$$\Lambda(x,y) = \begin{bmatrix} 2 - \sin(e^{5x})\cos(e^{-3y}) & 0 \\ 0 & 2 - \operatorname{sign}\big(\cos(6\pi x)\sin(6\pi y)\big) \end{bmatrix}$$

for all $(x, y) \in \Omega$. Then, $\Lambda_{11} \in [1, 3]$ oscillates several times and $\Lambda_{22} \in \{2, 3, 4\}$ has several discontinuities over Ω. Finally, we define

$$A(x,y) = \begin{cases} Q_1\,\Lambda(x,y)\,Q_1^T, & \text{for } x \geq 0,\ \ y \geq 0, \\ Q_2\,\Lambda(x,y)\,Q_2^T, & \text{for } x < 0,\ \ y \geq 0, \\ Q_3\,\Lambda(x,y)\,Q_3^T, & \text{for } x < 0,\ \ y < 0, \\ Q_4\,\Lambda(x,y)\,Q_4^T, & \text{otherwise} \end{cases}$$

so that A is strictly symmetric positive definite and discontinuous. No monotone method exists for this problem on the specified grids due to the choices for Q_1 and Q_4, and Q_2 and Q_3 would lead to wider-stencils.

We consider the problem $-A : D^2u = f$ for two different solutions

$$u^{(1)}(x,y) = \sin\Big(\frac{\pi}{2}(x+y)^2\Big),$$

$$u^{(2)}(x,y) = \frac{x^3}{18}\big(3\log(x^2) - 11\big) + \big(y - \frac{1}{2}\big)^{8/3}\sqrt{|x + \frac{1}{5}|^5}$$

so that $u^{(1)} \in C^\infty(\Omega)$ and $u^{(2)} \in C^2(\Omega) \setminus C^3(\Omega)$. The source function f and boundary data g are chosen so that the solution is given by either $u^{(1)}$ or $u^{(2)}$. Using $\widehat{F}_{0,0}$, we observe optimal second order rates when approximating $u^{(1)}$ and expected deteriorated rates when approximating $u^{(2)}$ in Table 1.

TABLE 1. Approximation results for Test 1 using $\widehat{F}_{0,0}$.

	$u^{(1)}$		$u^{(2)}$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order
1.82e-01	5.77e-02		5.21e-03	
4.88e-02	3.86e-03	2.06	8.29e-04	1.40
2.47e-02	9.79e-04	2.01	3.26e-04	1.37
1.65e-02	4.40e-04	1.99	1.86e-04	1.40
1.10e-02	1.96e-04	2.00	1.05e-04	1.41
8.30e-03	1.11e-04	2.00	7.02e-05	1.42
6.64e-03	7.10e-05	2.00	5.11e-05	1.43

7.2. **Test 2: Hamilton-Jacobi-Bellman equations.** This example is adapted from [32], and it was considered in [14] using the Lax-Friedrich's-like method. Let $\Lambda = [0, \pi/3] \times SO(2)$, where $SO(2)$ is the set of 2×2 rotation matrices and define σ^θ by

$$\sigma^\theta \equiv R^T \begin{bmatrix} 1 & \sin(\phi) \\ 0 & \cos(\phi) \end{bmatrix}, \quad \theta = (\phi, R) \in \Lambda.$$

Consider the Hamilton-Jacobi-Bellman equation

$$F[u] \equiv \inf_{\theta\in\Lambda}\big(L_\theta u - f_\theta\big) = 0,$$

$$L_\theta u = -A^\theta(\mathbf{x}) : D^2 u + b^\theta(\mathbf{x}) \cdot \nabla u + c^\theta(\mathbf{x})\, u$$

with $\Omega = (0,1)^2$, $A^\theta = \frac{1}{2}\sigma^\theta \left(\sigma^\theta\right)^T$, $\beta^\theta = \vec{0}$, $c^\theta = \pi^2$, $f_\theta = \sqrt{3}\sin^2(\phi/\pi^2) + g$ for g chosen independent of θ, and Dirichlet boundary data chosen such that the exact solution is given by $u(x,y) = e^{xy}\sin(\pi x)\sin(\pi y)$. The optimal controls vary significantly throughout the domain and the corresponding diffusion coefficient is not diagonally dominant in parts of Ω.

We approximate u using $\widehat{F}_{\gamma,0}$ for $\gamma = 1000, 10, 1, 0$ in Table 2. The case $\gamma = 1000$ appears to be in a pre-asymptotic regime with the error dominated by the numerical moment term. Otherwise, as expected, we see that the methods are around second order accuracy with the approximations increasing in accuracy as γ decreases. In the implementation of $\widehat{F}_{\gamma,\sigma}$, we do not calculate derivatives of F to define M_α. Instead, we define $\widehat{F}^\theta_{\gamma,\sigma}$ by

$$\begin{aligned}\widehat{F}^\theta_{\gamma,\sigma}[U_\alpha] \equiv A^\theta(\mathbf{x}_\alpha) : D^{2,\theta}_{\mathbf{h}} U_\alpha + c^\theta(\mathbf{x}_\alpha) U_\alpha - f_\theta(\mathbf{x}_\alpha) \\ + (\gamma I_{d\times d} + \sigma 1_{d\times d}) : \left(\widetilde{D}^2_{\mathbf{h}} U_\alpha - \widehat{D}^2_{\mathbf{h}} U_\alpha\right)\end{aligned}$$

for $D^{2,\theta}_{\mathbf{h}} U_\alpha$ defined analogously to (3.4) for each value of α and θ. We then solve the optimization problem $\inf_{\theta\in\Lambda} \widehat{F}^\theta_{\gamma,\sigma}[U_\alpha] = 0$.

When choosing $\gamma = 0$, the solver does not successfully find a solution even for coarse meshes. The solver also takes several iterations for $\gamma = 1$ sometimes struggling to find a solution for the finer meshes. Thus, we first find the solution corresponding to $\gamma = 1000$, and then we sequentially find the solutions corresponding to $\gamma = 100, 10, 1, 0$ by using the solution for the next largest value of γ as an initial guess. This iterative technique allows fsolve to successfully find a solution for each value of γ and often allows fsolve to significantly decrease the number of iterations needed to converge. We also can form a better initial guess by first solving the problem corresponding to a small finite number of controls θ before optimizing over Λ.

7.3. **Test 3: Monge-Ampère equation.** Consider the Monge-Ampère problem

$$F[u] = -\det(D^2 u) + f = -u_{xx}u_{yy} + |u_{xy}|^2 + f = 0$$

over $\Omega = (0,1)^2$. The problem has a unique convex viscosity solution whenever $f \geq 0$. We choose the source term f and boundary function g such that the exact solution is $u(x,y) = \exp\left(\frac{x^2+y^2}{2}\right)$. The matrix M_α is easily found since $|\frac{\partial F}{\partial P_{ij}}| = |P_{k\ell}|$ for $(k,\ell) = (2,2)$ if $(i,j) = (1,1)$, $(k,\ell) = (1,1)$ if $(i,j) = (2,2)$, and $(k,\ell) = (1,2)$ if $(i,j) = (1,2)$ or $(i,j) = (2,1)$.

This problem is degenerate and the uniformity of the g-monotonicity of $\widehat{F}_{\gamma,\sigma}[U]$ strongly depends upon U similarly to how the ellipticity of $F[u]$ strongly depends upon the convexity of u. Consequently, we now form initial guesses for fsolve by solving a sequence of problems based on a decreasing sequence of values for σ starting with σ large. As such, we are using a stronger form of the numerical moment to overcome the conditional ellipticity and potential degeneracy of the problem.

Rates of convergence for the various tests can be found in Table 3 where we consider the method $\widehat{F}_{\gamma,0}$ for $\gamma = 1000, 100, 10, 1, 0$ and Table 4 where we consider the method $\widehat{F}_{-\sigma,\sigma}$ for $\sigma = 1000, 100, 10, 1$. We observe optimal / near optimal rates

TABLE 2. Approximation results for Test 2 using $\widehat{F}_{\gamma,0}$ with $\gamma = 1000, 100, 10, 1, 0$.

	$\gamma = 1000$		$\gamma = 100$		$\gamma = 10$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order
1.57e-01	1.30e+00		1.17e+00		6.08e-01	
9.43e-02	1.27e+00	0.05	9.83e-01	0.35	3.34e-01	1.18
6.15e-02	1.22e+00	0.09	7.51e-01	0.63	1.73e-01	1.54
4.56e-02	1.15e+00	0.20	5.67e-01	0.94	1.01e-01	1.80
3.63e-02	1.08e+00	0.29	4.34e-01	1.16	6.47e-02	1.95
2.89e-02	9.78e-01	0.43	3.19e-01	1.35	4.05e-02	2.05
2.21e-02	8.32e-01	0.61	2.12e-01	1.53	2.30e-02	2.12

	$\gamma = 1$		$\gamma = 0$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order
1.57e-01	1.32e-01		2.35e-02	
9.43e-02	5.08e-02	1.87	1.03e-02	1.61
6.15e-02	2.15e-02	2.01	4.96e-03	1.71
4.56e-02	1.17e-02	2.03	2.92e-03	1.77
3.63e-02	7.43e-03	1.99	1.94e-03	1.79
2.89e-02	4.74e-03	1.97	1.28e-03	1.81
2.21e-02	2.81e-03	1.96	7.92e-04	1.80

of convergence as γ decreases in Table 3 and as σ decreases in Table 4. For γ and σ large the rates appear to be suboptimal but improving towards a rate of 2 as h decreases. We also note that the method $\widehat{F}_{-\sigma,\sigma}$ appears more accurate than the analogous method $\widehat{F}_{\gamma,0}$ when $\gamma = \sigma$. Such a relationship is expected based on the sensitivity of the Monge-Ampère problem to the auxiliary boundary condition. Indeed, setting $\Delta_{\mathbf{h}} U_\alpha = 0$ along the boundary and observing that g must be convex along the tangential direction implies U must be concave along the normal direction which brings a qualitative error into the interior of the domain. We can see this directly in Figure 2 where we plot the approximation for varying h, γ, and σ values. For $\gamma > 0$ on a coarse mesh we can see that the convexity of the approximation is incorrect near the boundary and that for γ large this forces the curvature to be incorrect throughout the interior. Consequently, for γ large, the method does not appear to enforce the convexity of the underlying viscosity solution. Instead, based on the solver's performance and based on the tests in [24], the numerical moment for γ large appears to minimize the number of times the function changes convexity over the domain by penalizing discontinuities in the second derivative and steering the approximation towards the correct viscosity solution as $h \to 0$. We also note that, in contrast, choosing σ large and minimizing the effect of the auxiliary boundary condition by setting $\gamma = -\sigma$ appears to completely eliminate the convexity issue. Another way to decrease the convexity issue when $\gamma > 0$ is to set only $\delta^2_{x_i,h_i} U_\alpha = 0$ along the normal direction instead of setting $\Delta_{\mathbf{h}} U_\alpha = 0$ or to choose a more appropriate positive value for the auxiliary boundary condition consistent with the convex nature of the viscosity solution.

TABLE 3. Approximation results for Test 3 using $\widehat{F}_{\gamma,0}$ with $\gamma = 1000, 100, 10, 1, 0$.

	$\gamma = 1000$		$\gamma = 100$		$\gamma = 10$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	5.59e-01		5.49e-01		4.02e-01	
1.29e-01	5.86e-01	-0.06	5.14e-01	0.08	9.65e-02	1.81
6.15e-02	5.64e-01	0.05	2.22e-01	1.14	2.09e-02	2.08
3.01e-02	4.22e-01	0.40	5.07e-02	2.06	5.23e-03	1.94
1.99e-02	2.32e-01	1.46	2.18e-02	2.04	2.39e-03	1.90
1.49e-02	1.31e-01	1.97	1.22e-02	2.00	1.38e-03	1.89
1.19e-02	8.13e-02	2.11	7.95e-03	1.90	8.98e-04	1.91

	$\gamma = 1$		$\gamma = 0$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	4.19e-02		1.57e-02	
1.29e-01	9.30e-03	1.91	3.41e-03	1.94
6.15e-02	2.31e-03	1.89	7.87e-04	1.99
3.01e-02	5.87e-04	1.91	1.88e-04	2.01
1.99e-02	2.64e-04	1.94	8.21e-05	2.01
1.49e-02	1.50e-04	1.93	4.58e-05	2.00
1.19e-02	9.70e-05	1.94	2.91e-05	2.00

TABLE 4. Approximation results for Test 3 using $\widehat{F}_{\gamma,\sigma}$ for $\sigma =$ 1000, 100, 10, 1 and $\gamma = -\sigma$.

	$\sigma = 1000$		$\sigma = 100$		$\sigma = 10$		$\sigma = 1$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	3.65e-02		3.36e-02		1.95e-02		1.47e-02	
1.29e-01	3.56e-02	0.03	2.49e-02	0.38	6.22e-03	1.45	3.05e-03	1.99
6.15e-02	3.07e-02	0.20	1.19e-02	1.00	1.63e-03	1.81	7.00e-04	2.00
3.01e-02	1.97e-02	0.62	3.83e-03	1.59	4.07e-04	1.95	1.67e-04	2.00
1.99e-02	1.24e-02	1.13	1.79e-03	1.84	1.79e-04	1.98	7.31e-05	2.00
1.49e-02	8.14e-03	1.44	1.02e-03	1.92	1.00e-04	1.99	4.08e-05	2.00
1.19e-02	5.66e-03	1.62	6.59e-04	1.95	6.41e-05	1.99	2.60e-05	2.00

7.4. **Test 4: Prescribed Gauss curvature equations.** This example is adapted from [19]. Let $K > 0$. The equation of prescribed Gauss curvature corresponds to the choice

$$F[u] = -\frac{\det(D^2u)}{(1+|\nabla u|^2)^{\frac{d+2}{2}}} + Kf = 0.$$

The problem is based on the Monge-Ampère operator, and it is known that the problem when $f = 1$ has a unique convex viscosity solution for each $K \in [0, K^*)$ for some positive constant K^*. We let $K = 0.1$ and $\Omega = (0,1)^2$, and we choose f

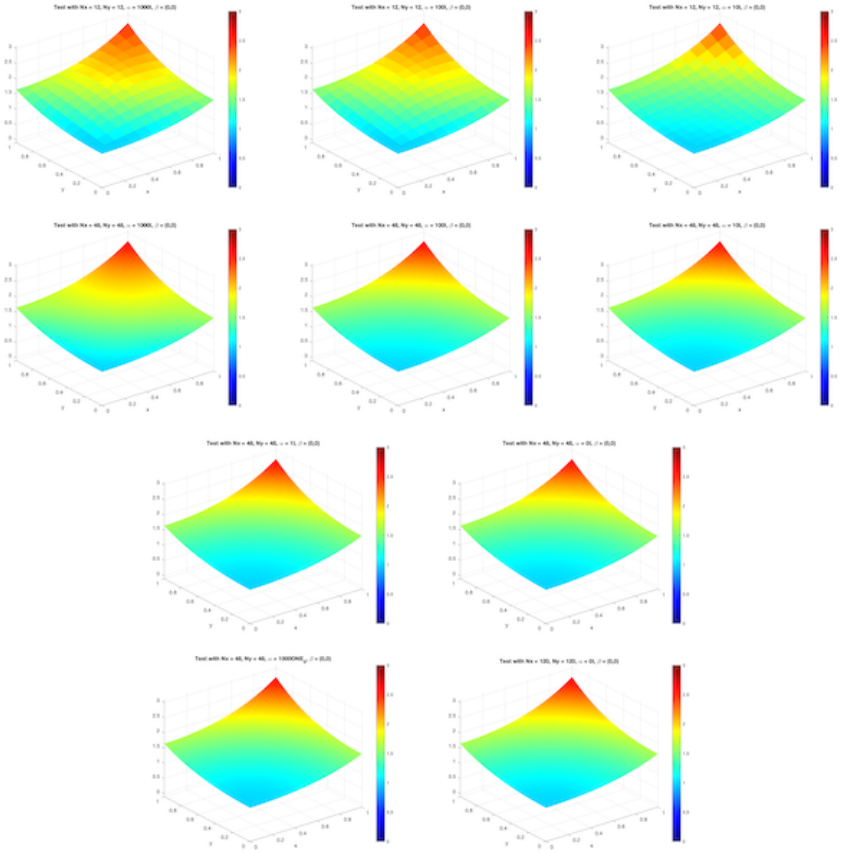

FIGURE 2. Plots of various approximations for Test 3 using $\widehat{F}_{\gamma,\sigma}$. The first row corresponds to $h = 1.29\text{e-}01$ with $\sigma = 0$ and $\gamma = 1000, 100, 10$ from left to right. The second row corresponds to $h = 3.01\text{e-}02$ with $\sigma = 0$ and $\gamma = 1000, 100, 10$ from left to right. The third row corresponds to $h = 3.01\text{e-}02$ with $\sigma = 0$ and $\gamma = 1, 0$ from left to right. The first plot on the fourth row corresponds to $h = 3.01\text{e-}02$ with $\sigma = 1000$ and $\gamma = -1000$. The second plot on the fourth row corresponds to $h = 1.19\text{e-}02$ with $\gamma = \sigma = 0$.

and g such that the exact solution is given by $u(x, y) = e^{\frac{x^2+y^2}{2}}$. The matrix M_α is easily found since $|\frac{\partial F}{\partial P_{ij}}| = |P_{k\ell}|/(1 + |\mathbf{v}|^2)^2$ for $(k, \ell) = (2, 2)$ if $(i, j) = (1, 1)$, $(k, \ell) = (1, 1)$ if $(i, j) = (2, 2)$, and $(k, \ell) = (1, 2)$ if $(i, j) = (1, 2)$ or $(i, j) = (2, 1)$.

Rates of convergence for the various tests can be found in Table 5 where we consider the method $\widehat{F}_{\gamma,0}$ for $\gamma = 1000, 100, 10, 1, 0$ and Table 6 where we consider the method $\widehat{F}_{-\sigma,\sigma}$ for $\sigma = 1000, 100, 10, 1$. Overall we observe similar behavior as Test 3 but less accuracy and lower rates when $\gamma > 0$ or $\sigma > 0$. The method $\widehat{F}_{0,0}$ does exhibit an optimal convergence rate of 2 and is the most accurate of all of the methods tested. Since this problem involves the gradient operator, any monotone method that directly approximates the gradient would in general be limited to only

first order accuracy until the mesh is fine enough to use the local uniform ellipticity to enforce the monotonicity with respect to $\nabla^{\pm}_{\mathbf{h}} U_\alpha$.

TABLE 5. Approximation results for Test 4 using $\widehat{F}_{\gamma,0}$ with $\gamma = 1000, 100, 10, 1, 0$.

	$\gamma = 1000$		$\gamma = 100$		$\gamma = 10$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	1.30e+00		5.57e-01		5.33e-01	
1.29e-01	1.27e+00	0.05	5.80e-01	-0.05	4.05e-01	0.35
6.15e-02	1.22e+00	0.09	5.24e-01	0.14	1.67e-01	1.20
3.01e-02	1.15e+00	0.20	2.79e-01	0.88	8.64e-02	0.93
1.99e-02	1.08e+00	0.29	1.72e-01	1.17	5.76e-02	0.98
1.49e-02	9.78e-01	0.43	1.31e-01	0.96	4.24e-02	1.06
1.19e-02	8.32e-01	0.61	1.06e-01	0.91	3.29e-02	1.12

	$\gamma = 1$		$\gamma = 0$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	2.56e-01		2.19e-02	
1.29e-01	1.08e-01	1.10	3.89e-03	2.19
6.15e-02	5.52e-02	0.91	8.20e-04	2.11
3.01e-02	2.51e-02	1.10	1.90e-04	2.05
1.99e-02	1.49e-02	1.26	8.25e-05	2.02
1.49e-02	1.01e-02	1.35	4.59e-05	2.01
1.19e-02	7.30e-03	1.43	2.92e-05	2.01

TABLE 6. Approximation results for Test 4 using $\widehat{F}_{\gamma,\sigma}$ for $\sigma =$ 1000, 100, 10, 1 and $\gamma = -\sigma$.

	$\sigma = 1000$		$\sigma = 100$		$\sigma = 10$		$\sigma = 1$	
h	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order	ℓ^∞ Error	Order
2.83e-01	3.68e-02		3.63e-02		3.20e-02		1.81e-02	
1.29e-01	3.70e-02	-0.01	3.40e-02	0.08	2.11e-02	0.53	6.83e-03	1.24
6.15e-02	3.59e-02	0.04	2.70e-02	0.31	1.02e-02	0.98	2.09e-03	1.60
3.01e-02	3.19e-02	0.17	1.61e-02	0.72	3.84e-03	1.38	5.75e-04	1.81
1.99e-02	2.73e-02	0.38	1.04e-02	1.06	1.96e-03	1.62	2.61e-04	1.92
1.49e-02	2.31e-02	0.57	7.22e-03	1.26	1.18e-03	1.74	1.47e-04	1.96
1.19e-02	1.96e-02	0.73	5.28e-03	1.39	7.88e-04	1.81	9.45e-05	1.97

Acknowledgements. X. Feng was partially supported by the NSF grants DMS-1620168 and DMS-2012414. T. Lewis and K. Ward were partially supported by the NSF grant DMS-2111059.

References

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

[2] F. Bonnans, H. Zidani; *Consistency of generalized finite difference schemes for the stochastic HJB equation*, SIAM J. Numer. Anal., **41** (2003), 1008–1021.

[3] L. A. Caffarelli, X. Cabré; *Fully nonlinear elliptic equations*, Vol. 43 of *American Mathematical Society Colloquium Publications, AMS*, Providence, RI, 1995.

[4] M. G. Crandall, H. Ishii, P.-L. Lions; *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc., **27** (1992), 1–67.

[5] M. G. Crandall, P.-L. Lions; *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **277** (1983), 1–42.

[6] M. G. Crandall, L. C. Evans, P.-L. Lions; *Some properties of viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., **282** (1984), 487–502.

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

[8] W. Feng, T. Lewis, ,S. Wise; *Discontinuous Galerkin derivative operators with applications to second order elliptic problems and stability*, Mathematical Meth. in App. Sciences, **38**(18) (2015), 5160–5182.

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

[10] X. Feng, M. Jensen; *Convergent semi-Lagrangian methods for the Monge-Ampère equation on unstructured grids*, SIAM J. Numer. Anal., **55** (2017), 691–712.

[11] X. Feng, C. Kao, T. Lewis; *Convergent FD methods for one-dimensional fully nonlinear second order partial differential equations*, J. Comput. Appl. Math., **254** (2013), 81–98.

[12] X. Feng, T. Lewis; *Nonstandard local discontinuous Galerkin methods for fully nonlinear second order elliptic and parabolic equations in high dimensions*, J. Scient. Comput., **77** (2018), 1534–1565.

[13] X. Feng, T. Lewis; *A Narrow-stencil finite difference method for approximating viscosity solutions of fully nonlinear elliptic partial differential equations with applications to Hamilton-Jacobi-Bellman equations*, `http://arxiv.org/abs/1907.10204` (2019).

[14] X. Feng, T. Lewis; *A Narrow-stencil finite difference method for approximating viscosity solutions of Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal., **59** (2021), 886–924.

[15] X. Feng, T. Lewis, A. Rapp; *Dual-Wind Discontinuous Galerkin Methods for Stationary Hamilton-Jacobi Equations and Regularized Hamilton-Jacobi Equations*, Commun. Appl. Math. Comput., **4** (2022), 563–596.

[16] X. Feng, T. Lewis, M. Neilan; *Discontinuous Galerkin finite element differential calculus and applications to numerical solutions of linear and nonlinear partial differential equations*, J. Comput. Appl. Math., **299** (2016), 68–91.

[17] X. Feng, M. Neilan; *Vanishing moment method and moment solutions for second order fully nonlinear partial differential equations*, J. Scient. Comput., **38** (2008), 74–98.

[18] X. Feng, M. Neilan; *Mixed finite element methods for the fully nonlinear Monge-Ampére equation based on the vanishing moment method*, SIAM J. Numer. Anal., **47** (2009), 1226–1250.

[19] X. Feng, M. Neilan; *Finite element approximations of general fully nonlinear second order elliptic partial differential equations based on the vanishing moment method*, Computers & Mathematics with Applications, **68** (2014), 2182–2204.

[20] W. H. Fleming, H. M. Soner; *Controlled Markov Process and Viscosity Solutions*, *Springer*, New York, 2006.

[21] D. Gilbarg, N. S. Trudinger; *Elliptic Partial Differential Equations of Second Order*, *Classics in Mathematics, Springer-Verlag*, Berlin, 2001.

[22] M. Jensen, I. Smears; *On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal., **51** (2013), 37–162.

[23] H. Kushner, P. G. Dupuis; *Numerical methods for stochastic optimal control problems in continuous time*, Volume 24 of *Applications of Mathematics, Springer*, New York, 1992.

[24] T. Lewis, *Finite Difference and Discontinuous Galerkin Finite Element Methods for Fully Nonlinear Second Order Partial Differential Equations*, Ph.D. Thesis, University of Tennessee, 2013.
[25] T. Lewis, M. Neilan; *Convergence analysis of a symmetric dual-wind discontinuous Galerkin method*, J. Sci. Comput., **59** (2014), 602-625.
[26] T. S. Motzkin, W. Wasow; *On the approximation of linear elliptic elliptic differential equations by difference equations with positive coefficients*, J. Math. Phys., **31** (1953), 253–259.
[27] M. Neilan, A.J. Salgado, W. Zhang; *Numerical analysis of strongly nonlinear PDEs*, Acta Numerica, **26** (2020), 137–303.
[28] R. H. Nochetto, D. Ntogakas, W. Zhang; *Two-scale method for the Monge-Ampére equation: convergence rates*, IMA J. Numer. Anal., **39**(3) (2019), 1085–1109.
[29] A. Oberman, *Wide stencil finite difference schemes for the elliptic Monge-Ampère equation and functions of the eigenvalues of the Hessian*, Discrt. Cont. Dynam. Syst. series B, **10** (2008), 221–238.
[30] M. Safonov, *Nonuniqueness for second-order elliptic equations with measurable coefficients*, SIAM J. Numeri. Anal., **30**(4) (1999), 379–395.
[31] A. J. Salgado, W. Zhang; *Finite element approximation of the Isaacs equation*, ESAIM: Math. Model. Numer. Anal., **53**(2) (2019), 351–374.
[32] I. Smears, E. Süli; *Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients*, SIAM J. Numer. Anal., **52** (2014), 993–1016.

Xiaobing Feng
Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, USA
Email address: xfeng@utk.edu

Thomas Lewis
Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27412, USA
Email address: tllewis3@uncg.edu

Kellie Ward
Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27412, USA
Email address: kmward7@uncg.edu