Published in ESAIM : control, optimisation and calculus of variations, 2013, vol. 19, no. 3, p.780-810 which should be cited to refer to this work

ESAIM: Control, Optimisation and Calculus of Variations Will be set by the publisher
URL: http://www.emath.fr/cocv/

ON THE LEAST-SQUARES SOLUTION OF THE DIRICHLET PROBLEM FOR THE ELLIPTIC MONGE-AMPÈRE EQUATION IN DIMENSION TWO *

Alexandre Caboussat[1], Roland Glowinski[1] and Danny C. Sorensen[2]

Abstract. We address the numerical solution of the Dirichlet problem for the real elliptic Monge-Ampère equation for arbitrary domains in two dimensions. The numerical method we discuss combines a least-squares formulation with a relaxation method. This approach leads to a sequence of Poisson-Dirichlet problems and another sequence of low dimensional algebraic eigenvalue problems of a new type. Mixed finite element approximations with a smoothing procedure are used for the computer implementation of our least-squares/relaxation methodology. Domains with curved boundaries are easily accommodated. Numerical experiments show the convergence of the computed solutions to their exact counterparts when such solutions exist. On the other hand, when smooth solutions do not exist, our least-squares based methodology produces generalized solutions which can be viewed as viscosity solutions, but in a sense different from Ishii & Lions'.

Résumé. Nous étudions dans cet article la résolution numérique de l'équation de Monge-Ampère elliptique dans des domaines de forme arbitraire en deux dimensions. Une méthode de moindres carrés est couplée à un algorithme de relaxation, conduisant à la résolution d'une suite de problèmes variationnels linéaires, et d'une suite de problèmes de valeurs propres en deux dimensions. Une approximation par éléments finis mixtes couplée à une méthode de régularisation est utilisée, de sorte que les domaines avec frontière courbe sont traités facilement. Des expériences numériques montrent l'efficacité de la méthode, ainsi que des bonnes propriétés de convergence.

1991 Mathematics Subject Classification. 65N30, 65K10, 65F30, 49M15, 49K20,

January 26, 2011.

1. Introduction

If f is *positive*, the canonical *Monge-Ampère equation*

$$\det \mathbf{D}^2\psi = f,$$

Keywords and phrases: Monge-Ampère equation, Least-squares method, Biharmonic problem, conjugate gradient method, Quadratic constraint minimization, Mixed finite element method.

* *This work was partially supported by the National Science Foundation (Grants NSF DMS-0913982 and DMS-0412267).*

[1] University of Houston, Department of Mathematics, 4800 Calhoun Rd, Houston, Texas 77204 - 3008, USA
e-mail: caboussat@math.uh.edu, roland@math.uh.edu

[2] Rice University, Department of Computational and Applied Mathematics, MS 134, Houston, TX 77251-1892, USA
e-mail: sorensen@rice.edu

is considered by many mathematicians as the prototypical *fully nonlinear elliptic equation*. As such, it has recently received considerable attention from both the analytical and computational standpoints as shown by, e.g., [1, 3, 7, 16, 28–30], with applications in geometry, mechanics and physics.

In particular, *augmented Lagrangian algorithms* and *least-squares techniques* have been used for the numerical solution of the Dirichlet problem for the Monge-Ampère equation in dimension two. These methods are discussed in [6, 8–13, 21, 22]; actually, [21] contains a review of several methods for the solution of the Monge-Ampère equation and related fully nonlinear elliptic equations such as Pucci's.

For a bounded, convex, two-dimensional domain Ω, let us assume that $f \in L^1(\Omega)$ and $g \in H^{3/2}(\partial\Omega)$. With this kind of data, it makes sense to look in $H^2(\Omega)$ for the solutions of the following Dirichlet problem for the Monge-Ampère equation

$$\det \mathbf{D}^2\psi = f \ \text{ in } \Omega, \quad \psi = g \ \text{ on } \partial\Omega. \tag{1}$$

Using the augmented Lagrangian and least-squares methods discussed in [6, 8–13, 21, 22], it has become possible to solve problem (1) with $\Omega = (0,1)^2$ when it has solutions with $H^2(\Omega)$ regularity. Numerical experiments reported in the above references indicate the L^2 approximation error is $\mathcal{O}(h^2)$, which is optimal for second order elliptic problems using such approximations. With these methods, we have also been able to compute *generalized solutions* of (1) when this problem has no classical solutions, as is the case for example when $\Omega = (0,1)^2$, $f = 1$, and $g = 0$. Hence, this approach provides an alternative to the *viscosity solution methods* discussed in, e.g., [5, 29]. As shown in Section 11, this limit can be viewed as a viscosity solution but in a sense different from *Ishii and P.L. Lions'* [27].

The least-squares methodology discussed in this article was introduced in [9] and further discussed in [13, 21, 22]. Actually, the most detailed account-published so far-of our least-squares approach can be found in [13] (for a detailed description of the augmented Lagrangian based methodology see [11]). The methodology discussed in [9, 13, 21, 22] relies on the following ingredients:

(i) A well-suited least-squares formulation in appropriate Hilbert spaces.
(ii) Associating with the optimality conditions of the above least-squares problem an initial value problem (flow in the dynamical system terminology).
(iii) The time-discretization of the above initial value problem by an operator-splitting scheme decoupling nonlinearity and differential operators.
(iv) The solution of the nonlinear (resp., linear) problems resulting from the splitting by a Newton's type algorithm (resp., by a preconditioned conjugate gradient algorithm).
(v) A mixed finite element approximation of the Monge-Ampère problem (1) based on piecewise linear continuous approximations of ψ and of its three second order derivatives.

In [13] and related publications, all the test problems considered were posed in $\Omega = (0,1)^2$ and the finite element spaces were associated with uniform triangulations like the one on the left in Figure 2 (see Section 10). When applied to problems where Ω has a curved boundary requiring unstructured meshes, or when using uniform meshes like the one on the right in Figure 2, we observed a deterioration of the convergence properties when $h \to 0$, and even divergence for some test problems. In this paper, we address this issue. An obvious way to overcome this difficulty is to proceed as in, e,g., [16, 17], that is, use mixed finite element approximations of the solutions of problem (1), and of their second order derivatives, based on continuous, piecewise polynomial functions of degree ≥ 2. This approach has several drawbacks, the main ones being that: (i) Unlike piecewise linear approximations, the higher order ones do not preserve the maximum principle when this principle holds. (ii) Compared to piecewise linear approximations, the higher order ones are not easy to implement for domains Ω with a curved boundary. Instead, in order to "rescue" the piecewise linear approximations, we advocate a *Tychonoff-like regularization method* when defining the discrete analogues of the second order derivatives. With this approach we recover convergence of optimal (or nearly optimal) order, as $h \to 0$, even for unstructured meshes, or for pathological structured ones like the triangulation on the right in Figure 2.

To summarize, in this article, we advocate a *relaxation* algorithm for the solution of a well-chosen *least-squares* variant of problem (1). With such an algorithm we are able to *decouple* the treatment of the differential

operators from the treatment of the nonlinearities. Indeed, the treatment of the differential operators leads to the solution of a sequence of elliptic linear biharmonic problems. The nonlinearity requires the solution of an infinite family of low dimensional constrained minimization problems, one for almost every point of Ω. In practice, there is such a minimization problem for each vertex of the finite element triangulation, after an appropriate spatial discretization.

To solve the above linear biharmonic problems we advocate a conjugate gradient algorithm operating in well-chosen sub-spaces of $H^2(\Omega)$; on the other hand, two quite different methods are considered for the solution of the low dimensional constrained minimization problems: the first one based on the Newton's method combined with an appropriate parametrization of the set $\left\{\mathbf{z} = \{z_i\}_{i=1}^3\,,\ z_1 > 0,, z_2 > 0,\ z_1 z_2 - z_3^2 = 1\right\}$. The second method is based on a novel algorithm for quadratically constrained minimization problems (denoted by $\mathbf{Q_{min}}$ and introduced in [33]). Following [8–14, 24], mixed finite element approximations are used for the discretization of (1). A regularization procedure for the approximation of second derivatives on arbitrary meshes allows obtaining optimal (or nearly optimal) convergence properties.

The structure of the article is as follows: In Section 2 we introduce some fundamental function spaces and sets and use them to provide a least-squares formulation of problem (1). The relaxation algorithm is described in Section 3. In Sections 4 and 5, we discuss the solutions of the local low dimensional constrained minimization problems and of the linear variational bi-harmonic problems. The mixed finite element approximation of problem (1) is discussed in Section 6, while Sections 7, 8 and 9 are dedicated to the discrete analogues of the problems discussed in Sections 3, 4 and 5. In Section 10, the methodology discussed in the preceding sections is applied to the solution of test problems, some of them borrowed from [6, 8–13, 21, 22]; these numerical experiments include test cases where Ω has a curved boundary and/or when problem (1) has no solution in $H^2(\Omega)$.

The methodology described in this article owes much to *Calculus of Variations* and *Optimal Control*. Indeed the least-squares criterion that we use is nothing but a multi-dimensional integral defined on the subset of a functional space *à la* Sobolev. Moreover *adjoint equation techniques* are used to compute some of the derivatives of the discrete cost functional, resulting in substantial memory and computational time savings.

2. Formulation of the Dirichlet Problem For the Elliptic Monge-Ampère Equation in Two Dimensions

Let Ω be a bounded convex domain of $\mathbb{R}^2$; we denote by Γ the boundary of Ω. The Dirichlet problem for the canonical Monge-Ampère equation reads as follows:

$$\det \mathbf{D}^2\psi = f\,(>0) \ \text{ in } \Omega, \quad \psi = g \ \text{ on } \Gamma, \tag{2}$$

where $\mathbf{D}^2\psi$ is the *Hessian* of the unknown function ψ, that is $\mathbf{D}^2\psi = \left(\dfrac{\partial^2\psi}{\partial x_i \partial x_j}\right)_{1\le i,j\le 2}$. Among the various methods available for the solution of (2), we advocate the following one of *nonlinear least-squares* type:

$$\left\{\begin{array}{l} \text{Find } (\psi, \mathbf{p}) \in V_g \times \mathbf{Q}_f \text{ such that} \\ \\ J(\psi, \mathbf{p}) \le J(\varphi, \mathbf{q}), \quad \forall (\varphi, \mathbf{q}) \in V_g \times \mathbf{Q}_f, \end{array}\right. \tag{3}$$

where:

$$J(\varphi, \mathbf{q}) = \frac{1}{2}\int_\Omega \left|\mathbf{D}^2\varphi - \mathbf{q}\right|^2 d\mathbf{x}, \tag{4}$$

using the Fröbenius norm and scalar product defined by $|\mathbf{T}| = \sqrt{\mathbf{T} : \mathbf{T}}$, with $\mathbf{S} : \mathbf{T} = \sum_{i,j=1}^2 s_{ij}t_{ij}$, for all $\mathbf{S} = (s_{ij})$, $\mathbf{T} = (t_{ij}) \in \mathbb{R}^{2\times 2}$. The functional spaces in (3) are defined by:

$$V_g = \left\{\varphi \in H^2(\Omega)\,,\ \varphi = g \ \text{ on } \Gamma\right\}, \tag{5}$$

$$\mathbf{Q}_f = \left\{\mathbf{q} \in \mathbf{Q}\,,\ \det \mathbf{q} = f,\ q_{11} > 0,\ q_{22} > 0\right\}, \quad \mathbf{Q} = \left\{\mathbf{q} \in L^2(\Omega)^{2\times 2},\ \mathbf{q} = \mathbf{q}^t\right\}. \tag{6}$$

The space $\mathbf{Q}$ is a Hilbert space for the scalar product $(\mathbf{q}, \mathbf{q}') \to \int_\Omega \mathbf{q} : \mathbf{q}' d\mathbf{x}$, and the associated norm. In order to have V_g and $\mathbf{Q}_f$ both non-empty, we assume from now on that $f \in L^1(\Omega)$ and $g \in H^{3/2}(\Gamma)$.

Remark 2.1. As shown in, e.g., [11–13], problem (3) may have smooth solutions, even if (2) has no such solutions as it is the case if $\Omega = (0,1)^2$, $f = 1$ and $g = 0$. Generally speaking, (2) admits a smooth solution when $\mathbf{D}^2 V_g \cap \mathbf{Q}_f \neq \emptyset$, as illustrated in Figure 1 (left). On the other hand, when $\mathbf{D}^2 V_g \cap \mathbf{Q}_f = \emptyset$, it makes perfect sense to search for a *least-squares solution*, in the sense of (3) (see Figure 1 (right)).

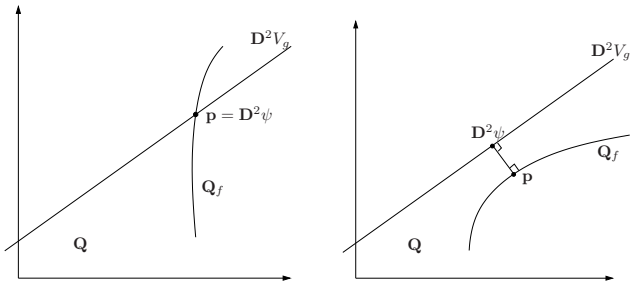

FIGURE 1. The Monge-Ampère problem (2) has a solution in V_g (left), or no solution in V_g (right).

3. A Relaxation Algorithm for the Solution of Problem (3)

In order to compute a *convex* solution (or at least to force the convexity of the solution) to problem (3) we suggest the following relaxation algorithm: Solve

$$-\Delta\psi^0 = -2\sqrt{f} \ \text{ in } \Omega, \quad \psi^0 = g \ \text{ on } \Gamma. \tag{7}$$

Then, for $n \geq 0$, assuming that ψ^n is known, compute $\mathbf{p}^n, \psi^{n+1/2}$ and ψ^{n+1} as follows:

$$\mathbf{p}^n = \arg\min_{\mathbf{q}\in\mathbf{Q}_f} J(\psi^n, \mathbf{q}), \tag{8}$$

$$\psi^{n+1/2} = \arg\min_{\varphi\in V_g} J(\varphi, \mathbf{p}^n), \tag{9}$$

$$\psi^{n+1} = \psi^n + \omega(\psi^{n+1/2} - \psi^n), \tag{10}$$

with ω, $0 < \omega < \omega_{\max} \leq 2$, a relaxation parameter.

Remark 3.1 (Initialization strategy). The rationale behind (7) is as follows: Denote by λ_1 and λ_2 the eigenvalues of $\mathbf{D}^2\psi$ so that $\lambda_1\lambda_2 = f$. If λ_1 and λ_2 are close to each other, then $\Delta\psi = \lambda_1 + \lambda_2 \simeq 2\sqrt{\lambda_1\lambda_2} = 2\sqrt{f}$, justifying the initialization (7).

The relaxation algorithm (7)-(10) looks simple but the solution of the problems (8) and (9) leads to technical issues that we will address in the following sections.

4. Numerical Solution of the Sub-problems (8)

4.1. Explicit Formulation of Problem (8)

An explicit formulation of problem (8) is given by

$$\mathbf{p}^n = \arg\min_{\mathbf{q}\in\mathbf{Q}_f} \left[\frac{1}{2}\int_\Omega |\mathbf{q}|^2 \, d\mathbf{x} - \int_\Omega \mathbf{D}^2\psi^n : \mathbf{q} d\mathbf{x} \right]. \tag{11}$$

Since neither integrands in (11) contains derivatives of $\mathbf{q}$, the minimization problem (11) can be solved *point-wise* (in practice at the vertices of a finite element or finite difference grid). This leads us, a.e. in Ω, to the solution of the following finite dimensional minimization problem

$$\mathbf{p}^n(\mathbf{x}) = \arg \min_{\mathbf{q}\in\mathbf{E}_f(\mathbf{x})} \left[\frac{1}{2}\left|\mathbf{q}\right|^2 - \mathbf{D}^n(\mathbf{x}) : \mathbf{q}\right], \tag{12}$$

where $\mathbf{D}^n(\mathbf{x}) = \mathbf{D}^2\psi^n(\mathbf{x})$ is a symmetric matrix and $\mathbf{E}_f(\mathbf{x}) = \left\{\mathbf{q}\in\mathbb{R}^{2\times 2}\,,\, \mathbf{q}=\mathbf{q}^t,\, \det\mathbf{q} = f(\mathbf{x}),\, q_{11}>0,\, q_{22}>0\right\}$.

4.2. **A Newton-Type Method for the Numerical Solution of Problem (12)**

Taking advantage of the symmetry of $\mathbf{q}$ and $\mathbf{D}^n(\mathbf{x})$, and using the notation $z_1 = q_{11}, z_2 = q_{22}$, $z_3 = q_{12} = q_{21}$ and $\mathbf{D}^n(\mathbf{x})_{ij} = d^n_{ij}(\mathbf{x})$, the minimization problem in (12) can be rewritten as

$$\min_{\mathbf{z}\in\mathbf{Z}_f(\mathbf{x})} \left[\frac{1}{2}(z_1^2+z_2^2+2z_3^2) - d^n_{11}(\mathbf{x})z_1 - d^n_{22}(\mathbf{x})z_2 - 2d^n_{12}(\mathbf{x})z_3\right], \tag{13}$$

with $\mathbf{Z}_f(\mathbf{x}) = \left\{\mathbf{z}\in\mathbb{R}^3,\, z_1>0,\, z_2>0,\, z_1z_2 - z_3^2 = f(\mathbf{x})\right\}$. To transform (13) into an unconstrained minimization problem in $\mathbb{R}^2$, we perform the change of variables $z_1 = \sqrt{f(\mathbf{x})}e^{\rho}\cosh\theta$, $z_2 = \sqrt{f(\mathbf{x})}e^{-\rho}\cosh\theta$, $z_3 = \sqrt{f(\mathbf{x})}\sinh\theta$, for $(\rho,\theta)\in\mathbb{R}^2$, so that (13) becomes

$$\min_{(\rho,\theta)\in\mathbb{R}^2} j(\rho,\theta),$$

with $j(\rho,\theta) = \frac{\sqrt{f(\mathbf{x})}}{2}\left(\cosh 2\rho\cosh 2\theta + \cosh 2\rho + \cosh 2\theta - 1\right) - \left(d^n_{11}(\mathbf{x})e^{\rho} + d^n_{22}(\mathbf{x})e^{-\rho}\right)\cosh\theta - 2d^n_{12}(\mathbf{x})\sinh\theta$. This leads us in turn to the solution of $Dj(\rho,\theta) = \mathbf{0}$, where $Dj(\cdot)$ is the differential of the functional $j(\cdot)$. This 2×2 nonlinear system actually reads as follows:

$$\begin{aligned}
Dj(\rho,\theta)_1 &= \sqrt{f(\mathbf{x})}(1+\cosh 2\theta)\sinh 2\rho - \left(d^n_{11}(\mathbf{x})e^{\rho} - d^n_{22}(\mathbf{x})e^{-\rho}\right)\cosh\theta &&= 0,\\
Dj(\rho,\theta)_2 &= \sqrt{f(\mathbf{x})}(1+\cosh 2\rho)\sinh 2\theta - \left(d^n_{11}(\mathbf{x})e^{\rho} + d^n_{22}(\mathbf{x})e^{-\rho}\right)\sinh\theta - 2d^n_{12}(\mathbf{x})\cosh\theta &&= 0.
\end{aligned}$$

This system can be solved by using a *Newton method*. Let $(\rho^0,\theta^0)\in\mathbb{R}^2$ be given. For $k\geq 0$, we compute $(\rho^{k+1},\theta^{k+1})$ from (ρ^k,θ^k) via the solution of

$$D^2j(\rho^k,\theta^k)\begin{pmatrix}\rho^{k+1}-\rho^k\\ \theta^{k+1}-\theta^k\end{pmatrix} = -Dj(\rho^k,\theta^k),$$

where $D^2j(\rho,\theta) = (D^2j(\rho,\theta)_{ij})_{1\leq i,j\leq 2}$ is given by:

$$\begin{aligned}
D^2j(\rho,\theta)_{11} &= 2\sqrt{f(\mathbf{x})}\cosh 2\rho(1+\cosh 2\theta) - \left(d^n_{11}(\mathbf{x})e^{\rho} + d^n_{22}(\mathbf{x})e^{-\rho}\right)\cosh\theta,\\
D^2j(\rho,\theta)_{12} &= D^2j(\rho,\theta)_{21} = 2\sqrt{f(\mathbf{x})}\sinh 2\rho\sinh 2\theta - \left(d^n_{11}(\mathbf{x})e^{\rho} - d^n_{22}(\mathbf{x})e^{-\rho}\right)\sinh\theta,\\
D^2j(\rho,\theta)_{22} &= 2\sqrt{f(\mathbf{x})}\cosh 2\theta(1+\cosh 2\rho) - \left(d^n_{11}(\mathbf{x})e^{\rho} + d^n_{22}(\mathbf{x})e^{-\rho}\right)\cosh\theta - 2d_{12}(\mathbf{x})\sinh\theta.
\end{aligned}$$

Remark 4.1 (**Choice of the scalar product).** Since we are dealing with symmetric matrices, we can equip $\mathbf{Q}$ with the following scalar product $(\mathbf{q},\mathbf{q}')\rightarrow\int_\Omega(q_{11}q'_{11}+q_{22}q'_{22}+q_{12}q'_{12})d\mathbf{x}$. As shown in [6,10,21], this new scalar product has given better results than the one defined by $(\mathbf{q},\mathbf{q}')\rightarrow\int_\Omega\mathbf{q}:\mathbf{q}'d\mathbf{x}$ when applied to the numerical solution of the two-dimensional Dirichlet problem for the *Pucci's equation*, that is $\alpha\lambda^+ + \lambda^- = f$ in Ω, together with $\psi = g$ on Γ, where λ^+ (resp., λ^-) denotes the largest (resp., the smallest) eigenvalue of the Hessian $\mathbf{D}^2\psi$ of the unknown function ψ, and where $\alpha\geq 1$. Using this new scalar product, (13) would be replaced by

$$\min_{\mathbf{z}\in\mathbf{Z}_f(\mathbf{x})}\left[\frac{1}{2}(z_1^2+z_2^2+z_3^2)-d_{11}^n(\mathbf{x})z_1-d_{22}^n(\mathbf{x})z_2-d_{12}^n(\mathbf{x})z_3\right],$$

with $\mathbf{Z}_f(\mathbf{x})$ defined similarly. The same change of variables and Newton method can be applied to this problem.

4.3. The Quadratically Constrained Minimization Method for the Numerical Solution of Problem (13)

In [33], a class of quadratically constrained minimization problems has been addressed with a new algorithm denoted by $\mathbf{Q_{min}}$. This algorithm allows the solution of some specific eigenvalue-constrained matrix optimization problems of dimension N ($N \geq 2$), and is of complexity $\mathcal{O}(N^3)$. The particular case $N = 2$ corresponds to (12). This method relies on the equivalence between (12) and the following formulation:

$$\mathbf{p}^n(\mathbf{x}) = \mathbf{S}^n(\mathbf{x})\mathbf{\Lambda}(\mathbf{x})\mathbf{S}^n(\mathbf{x})^T, \quad (\mathbf{\Lambda}(\mathbf{x}), \mathbf{S}^n(\mathbf{x})) = \arg\min_{(\mathbf{\Lambda},\mathbf{S})\in\mathcal{E}_f}\left[\frac{1}{2}(\mu_1^2+\mu_2^2)-\text{trace}\left(\mathbf{D}^n(\mathbf{x})\mathbf{S}\mathbf{\Lambda}\mathbf{S}^T\right)\right], \tag{14}$$

where $\mathcal{E}_f(\mathbf{x}) = \left\{(\mathbf{\Lambda},\mathbf{S}),\ \mathbf{\Lambda} = \text{diag}(\mu_1,\mu_2),\ \mu_1\mu_2 = f(\mathbf{x}),\ \mathbf{S}^T\mathbf{S} = \mathbf{I}\right\}$. The algorithm developed in [33] applies beautifully to the solution of (14). After normalization of $\mathbf{D}^n(\mathbf{x})$ by $\sqrt{f(\mathbf{x})}$, (14) is equivalent to

$$\arg\min_{\mathbf{A}\in\mathcal{A}_1}\text{trace}\left[\mathbf{A}\mathbf{A}-2\mathbf{D}^n\mathbf{A}\right], \tag{15}$$

where $\mathcal{A}_1 = \left\{\mathbf{A}\in\mathbf{R}^{2\times 2},\ \ell^t\mathbf{M}\ell = 2,\ \mathbf{M}\ell \geq 2\right\}$, $\mathbf{M} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, and $\ell = (\mu_1,\mu_2)^T$, with $\{\mu_1,\mu_2\}$ being the spectrum of $\mathbf{A}$. Ultimately, for $N = 2$ the solution is found by solving a simple rational equation of the form

$$\frac{\beta_1^2}{(1+\mu)^2} = 2+\frac{\beta_2^2}{(1-\mu)^2},$$

where $\beta_1 = (\lambda_1+\lambda_2)/\sqrt{2}$ and $\beta_2^2 = (\lambda_1^2+\lambda_2^2)/2-\lambda_1\lambda_2$ with $\{\lambda_1,\lambda_2\}$ being the spectrum of $\mathbf{D}^n(\mathbf{x})/\sqrt{f(\mathbf{x})}$. Remarkably, essentially the same rational equation results for arbitrary $N \geq 2$. This equation is efficiently solved numerically by first taking reciprocals and then square roots on both sides and applying Newton's method. With a starting guess $\mu_0 = -1$, this typically converges in 3 to 5 iterations. This happens because the reciprocal square root transformation yields a problem that is essentially the intersection of straight lines. See [33] for full detail where this algorithm is developed for arbitrary $N \geq 2$.

5. Conjugate Gradient Solution of the Sub-problems (9)

Written in variational form, the Euler-Lagrange equation of the sub-problem (9) reads as follows:

$$\text{Find}\ \ \psi^{n+1/2}\in V_g\ \ \text{such that}\ \ \int_\Omega \mathbf{D}^2\psi^{n+1/2}:\mathbf{D}^2\varphi d\mathbf{x} = \int_\Omega \mathbf{p}^n:\mathbf{D}^2\varphi d\mathbf{x}, \quad \forall\varphi\in V_0, \tag{16}$$

where $V_0 = H^2(\Omega)\cap H_0^1(\Omega)$. The linear variational problem (16) is well-posed and belongs to the following family of linear variational problems:

$$u\in V_g\ :\ \int_\Omega \mathbf{D}^2u:\mathbf{D}^2v d\mathbf{x} = L(v), \quad \forall v\in V_0, \tag{17}$$

with the functional $L(\cdot)$ linear and continuous over $H^2(\Omega)$; problem (17) is clearly of the biharmonic type. The *conjugate gradient solution* of linear variational problems in Hilbert spaces, such as (17), has been addressed in, e.g., [19, Chapter 3]. Following the above reference, we are going to solve (17) by a conjugate gradient algorithm operating in the spaces V_0 and V_g, both spaces being equipped with the scalar product defined by $(v,w)\rightarrow\int_\Omega \Delta v\Delta w d\mathbf{x}$, and the corresponding norm. This conjugate gradient algorithm reads as follows:

Step 1

$$u^0 \in V_g \quad \text{given.} \tag{18}$$

Step 2 Solve:

$$\text{Find} \quad g^0 \in V_0 \quad \text{such that} \int_\Omega \Delta g^0 \Delta v d\mathbf{x} = \int_\Omega \mathbf{D}^2 u^0 : \mathbf{D}^2 v d\mathbf{x} - L(v), \quad \forall v \in V_0, \tag{19}$$

and set the first descent direction:

$$w^0 = g^0. \tag{20}$$

Then, for $k \geq 0$, u^k, g^k, and w^k being known, the last two different from zero, we compute u^{k+1}, g^{k+1} and, if necessary, w^{k+1} as follows.

Step 3 Solve:

$$\text{Find} \quad \bar{g}^k \in V_0 \quad \text{such that} \int_\Omega \Delta \bar{g}^k \Delta v d\mathbf{x} = \int_\Omega \mathbf{D}^2 w^k : \mathbf{D}^2 v d\mathbf{x}, \quad \forall v \in V_0, \tag{21}$$

and compute the new iterates as follows:

$$\rho_k = \frac{\int_\Omega \left|\Delta g^k\right|^2 d\mathbf{x}}{\int_\Omega \Delta \bar{g}^k \Delta w^k d\mathbf{x}}, \tag{22}$$

$$u^{k+1} = u^k - \rho_k w^k, \tag{23}$$

$$g^{k+1} = g^k - \rho_k \bar{g}^k. \tag{24}$$

Step 4 Compute

$$\delta_k = \frac{\int_\Omega \left|\Delta g^{k+1}\right|^2 d\mathbf{x}}{\int_\Omega \left|\Delta g^0\right|^2 d\mathbf{x}}. \tag{25}$$

If $\delta_k < \varepsilon$ (meaning that the residual is small enough), take $u = u^{k+1}$; otherwise, compute:

$$\gamma_k = \frac{\int_\Omega \left|\Delta g^{k+1}\right|^2 d\mathbf{x}}{\int_\Omega \left|\Delta g^k\right|^2 d\mathbf{x}}, \tag{26}$$

and update the descent direction via

$$w^{k+1} = g^{k+1} + \gamma_k w^k. \tag{27}$$

Step 5 Do $k+1 \to k$ and return to **Step 3**.

Numerical experiments indicate that the conjugate gradient algorithm (18)-(27) has excellent convergence properties. Combined with an appropriate mixed finite element approximation of (2) it requires the solution of two discrete Poisson problems at each iteration.

6. On a Mixed Finite Element Approximation

6.1. Generalities

Considering the highly variational flavor of the methodology discussed in the preceding sections, it makes sense to look for finite element based methods for the approximation of (2). In order to avoid the complications associated with the construction of finite element sub-spaces of $H^2(\Omega)$ (see, however, [3,17] for such an approach), we employ a mixed finite element approximation (closely related to those discussed in, e.g., [14,15,20,25,32] for the solution of linear and nonlinear bi-harmonic problems). Following this approach, it is possible to solve (2) employing approximations commonly used for the solution of second order elliptic problems (piecewise linear and globally continuous over a triangulation of Ω for example).

6.2. Mixed Finite Element Approximation

For simplicity, we assume that Ω is a bounded polygonal domain of $\mathbb{R}^2$. Let us denote by $\mathcal{T}_h$ a finite element triangulation of Ω as discussed in, e.g., [20, Appendix 1]. From $\mathcal{T}_h$, we approximate the spaces $L^2(\Omega)$, $H^1(\Omega)$ and $H^2(\Omega)$ (respectively, $H^1_0(\Omega)$ and $H^2(\Omega)\cap H^1_0(\Omega)$) by the finite dimensional space V_h (respectively, V_{0h}) defined by:

$$V_h = \left\{ v \in C^0\left(\overline{\Omega}\right)\,,\ v|_T \in \mathbb{P}_1,\ \forall T \in \mathcal{T}_h \right\}, \quad V_{0h} = V_h \cap H^1_0(\Omega) = \left\{ v \in V_h\,,\ v = 0 \text{ on } \Gamma \right\}, \tag{28}$$

with $\mathbb{P}_1$ the space of the two-variables polynomials of degree ≤ 1. For a function φ being given in $H^2(\Omega)$, we denote $\partial^2\varphi/\partial x_i\partial x_j$ by $D^2_{ij}(\varphi)$. It follows from *Green's formula* that

$$\int_\Omega \frac{\partial^2\varphi}{\partial x_i\partial x_j} v d\mathbf{x} = -\frac{1}{2}\int_\Omega \left[\frac{\partial\varphi}{\partial x_i}\frac{\partial v}{\partial x_j} + \frac{\partial\varphi}{\partial x_j}\frac{\partial v}{\partial x_i}\right] d\mathbf{x},\ \forall v \in H^1_0(\Omega),\ \forall i,j = 1,2. \tag{29}$$

Consider now $\varphi \in V_h$. Taking advantage of the relations (29), we define the discrete analogues of the differential operators D^2_{ij} by

$$D^2_{hij}(\varphi) \in V_{0h}, \quad \int_\Omega D^2_{hij}(\varphi) v d\mathbf{x} = -\frac{1}{2}\int_\Omega \left[\frac{\partial\varphi}{\partial x_i}\frac{\partial v}{\partial x_j} + \frac{\partial\varphi}{\partial x_j}\frac{\partial v}{\partial x_i}\right] d\mathbf{x},\ \forall v \in V_{0h},\ \forall i,j = 1,2. \tag{30}$$

The functions $D^2_{hij}(\varphi)$ are uniquely defined by the relations (30). However, in order to simplify the computation of the above discrete second order partial derivatives, it is tempting to consider using the trapezoidal rule to evaluate the integrals in the left hand sides of (30). Owing to their practical importance, let us detail these calculations:

(i) First, we introduce the set Σ_h of the vertices of $\mathcal{T}_h$ and then $\Sigma_{0h} = \{P \in \Sigma_h\,,\ P \notin \Gamma\}$. Next, we define the integers N_h and N_{0h} by $N_h = \mathrm{Card}(\Sigma_h)$ and $N_{0h} = \mathrm{Card}(\Sigma_{0h})$. We have then $\dim V_h = N_h$ and $\dim V_{0h} = N_{0h}$. We suppose that $\Sigma_{0h} = \{P_j\}_{j=1}^{N_{0h}}$ and $\Sigma_h = \Sigma_{0h} \cup \{P_j\}_{j=N_{0h}+1}^{N_h}$.

(ii) To each $P_k \in \Sigma_h$, we associate the function w^k uniquely defined by

$$w^k \in V_h,\ w^k(P_k) = 1,\ w^k(P_l) = 0,\ \forall l = 1,\ldots,N_h,\ l \neq k.$$

It is well-known (see, e.g., [20, Appendix 1]) that the sets $\mathcal{B}_h = \{w^k\}_{k=1}^{N_h}$ and $\mathcal{B}_{0h} = \{w^k\}_{k=1}^{N_{0h}}$ are *vector bases* for V_h and V_{0h}, respectively.

(iii) Let us denote by A_k the area of the polygonal domain which is the union of those triangles of $\mathcal{T}_h$ which have P_k as a common vertex. Applying the trapezoidal rule to the integrals in the left-hand side of the relations (30), we obtain

$$D^2_{hij}(\varphi) \in V_{0h}, \quad D^2_{hij}(\varphi)(P_k) = -\frac{3}{2A_k}\int_\Omega \left[\frac{\partial\varphi}{\partial x_i}\frac{\partial w^k}{\partial x_j} + \frac{\partial\varphi}{\partial x_j}\frac{\partial w^k}{\partial x_i}\right] d\mathbf{x},\ \forall k = 1,\ldots,N_{0h},\ \forall i,j = 1,2. \tag{31}$$

Computing the integrals in the right hand side of (31) is quite simple since the first order derivatives of φ and w^k are *piecewise constant*. Finally, with $\varphi \in V_h$, we associate $\Delta_h\varphi \in V_{0h}$ uniquely defined by $\Delta_h\varphi(P_k) = D^2_{h11}(\varphi)(P_k) + D^2_{h22}(\varphi)(P_k)$, for $k = 1, \ldots, N_{0h}$.

Taking the above relations into account, approximating problem (2) is now fairly straightforward. Assuming that the boundary function g is continuous over Γ (which is definitely the case if $g \in H^{3/2}(\Gamma)$), let us denote by g_h the interpolant of g associated with the triangulation $\mathcal{T}_h$. We approximate the affine space V_g by $V_{gh} = \{\varphi \in V_h\,,\, \varphi(P) = g(P), \forall P \in \Sigma_h \cap \Gamma\}$ and then problem (2) by:

$$\text{Find} \quad \psi_h \in V_{gh} \quad \text{such that} \quad D^2_{h11}(\psi_h)(P_k)D^2_{h22}(\psi_h)(P_k) - \left|D^2_{h12}(\psi_h)(P_k)\right|^2 = f_h(P_k), \quad k = 1, \ldots, N_{0h}, \quad (32)$$

where f_h is a continuous approximation of f (we can always assume that $f_h \in V_h$). In addition, we define the discrete equivalent of $\mathbf{Q}_f$ as follows:

$$\mathbf{Q}_{fh} = \{\mathbf{q} \in \mathbf{Q}_h\,,\, \det \mathbf{q}(P_k) = f_h(P_k),\ q_{11}(P_k) > 0,\ q_{22}(P_k) > 0,\ k = 1, \ldots, N_{0h}\}\,,$$

with $\mathbf{Q}_h = \left\{\mathbf{q} \in (V_{0h})^{2\times 2},\ \mathbf{q}(P_k) = \mathbf{q}^t(P_k),\ k = 1, \ldots, N_{0h}\right\}$. We associate with V_{0h} and $\mathbf{Q}_h$ the following discrete scalar products and corresponding Euclidean norms:

$$\begin{aligned}
(v, w)_{0h} &= \frac{1}{3}\sum_{k=1}^{N_h} A_k v(P_k)w(P_k),\ \forall v, w \in V_{0h}, \qquad ||v||^2_{0h} = (v, v)_{0h},\ \forall v \in V_{0h},\\
((\mathbf{S}, \mathbf{T}))_{0h} &= \frac{1}{3}\sum_{k=1}^{N_{0h}} A_k \mathbf{S}(P_k) : \mathbf{T}(P_k),\ \forall \mathbf{S}, \mathbf{T} \in \mathbf{Q}_h, \quad |||\mathbf{S}|||^2_{0h} = ((\mathbf{S}, \mathbf{S}))_{0h},\ \forall \mathbf{S} \in \mathbf{Q}_h.
\end{aligned}$$

The solution of problem (32) will be discussed in the sequel.

Remark 6.1. Suppose that $\Omega = (0, 1)^2$ and that the triangulation $\mathcal{T}_h$ is uniform like the one shown in Figure 2 (left). Suppose that $h = 1/(I+1)$, I being a positive integer greater than one. In this particular case, the sets Σ_h and Σ_{0h} are given by $\Sigma_h = \{P_{ij} = (ih, jh)\,,\, 0 \le i, j \le I+1\}$, and $\Sigma_{0h} = \{P_{ij} = (ih, jh)\,,\, 1 \le i, j \le I\}$. implying that $N_h = (I+2)^2$ and $N_{0h} = I^2$. It follows then from the relations (31) that (with obvious notation):

$$\begin{aligned}
D^2_{h11}(\varphi)(P_{ij}) &= \frac{\varphi_{i+1,j} + \varphi_{i-1,j} - 2\varphi_{ij}}{h^2}, \quad 1 \le i, j \le I,\\
D^2_{h22}(\varphi)(P_{ij}) &= \frac{\varphi_{i,j+1} + \varphi_{i,j-1} - 2\varphi_{ij}}{h^2}, \quad 1 \le i, j \le I,\\
D^2_{h12}(\varphi)(P_{ij}) &= \frac{\varphi_{i+1,j+1} + \varphi_{i-1,j-1} + 2\varphi_{ij} - (\varphi_{i+1,j} + \varphi_{i-1,j} + \varphi_{i,j+1} + \varphi_{i,j-1})}{2h^2}, \quad 1 \le i, j \le I.
\end{aligned}$$

The above discrete second order derivatives of finite difference type have the easily verified yet remarkable properties that they are *exact for polynomial functions of degree* ≤ 2.

6.3. **Smoothing Procedure for the Approximation of the Second Derivatives**

As emphasized in [31], when using piecewise linear mixed finite elements, the a priori estimates for the error on the second derivatives of the solution ψ are, in general, $\mathcal{O}(1)$ in the L^2-norm. Therefore the convergence properties of the global solution method strongly depends on the type of triangulation. Indeed, assuming that the discrete second order derivatives have been computed via (30) and (31), numerical experiments performed by the authors showed the triangulation dependence of the convergence; non-convergence cases (in the L^2-norm) were also observed. Unfortunately, the approximations of $\dfrac{\partial^2\varphi}{\partial x_i \partial x_j}$ provided by (30) and (31) converge to the above second derivative, no better than in $H^{-1}(\Omega)$ in general. This allows oscillations and explains the growth

of the approximation error in $L^2(\Omega)$ and $H^1(\Omega)$ as $h \to 0$. Such pathological behavior can be observed in the results presented in Section 10. From that point of view a dramatic confirmation of these non-convergence properties is provided by the numerical results associated with the structured symmetric mesh shown on the right of Figure 2 (also called 'British flag' mesh or 'crisscross' pattern). To cure the non-convergence properties associated with the approximations (30) and (31) of the second derivatives, we see two options:

(i) Use, as in, e.g., [16, 17], mixed finite elements methods based on piecewise polynomial approximations of degree ≥ 2. This approach has several drawbacks, among them: (a) it is more complicated to implement than the mixed methods described in Section 6.2, particularly if Ω has a curved boundary. (b) These higher order polynomial approximations do not preserve the maximum principle, if this principle takes place for the continuous problem.

(ii) Use a *regularization procedure à la Tychonoff*, while keeping a piecewise linear approximation based mixed finite element approach.

Focusing on the second approach, a simple and novel (in this context) way to obtain better convergence properties of the discrete second order derivatives is to use the following regularization procedure: with $C > 0$ and $|K| = \text{meas}(K)$, when computing the discrete second derivatives $D^2_{hij}(\varphi)$ replace (30) by:

$$\begin{array}{l} \text{Find } D^2_{hij}(\varphi) \in V_{0h} \text{ such that, } \forall v \in V_{0h},\ i,j = 1,2, \\ \displaystyle\int_\Omega D^2_{hij}(\varphi) v d\mathbf{x} + C \sum_{K \in \mathcal{T}_h} |K| \int_K \nabla D^2_{hij}(\varphi) \cdot \nabla v d\mathbf{x} = -\frac{1}{2} \int_\Omega \left[\frac{\partial \varphi}{\partial x_i} \frac{\partial v}{\partial x_j} + \frac{\partial \varphi}{\partial x_j} \frac{\partial v}{\partial x_i} \right] d\mathbf{x}. \end{array} \tag{33}$$

and (31) by

$$\begin{array}{l} \text{Find } D^2_{hij}(\varphi) \in V_{0h} \text{ such that, } \forall v \in V_{0h},\ i,j = 1,2, \\ \displaystyle (D^2_{hij}(\varphi), v)_{0h} + C \sum_{K \in \mathcal{T}_h} |K| \int_K \nabla D^2_{hij}(\varphi) \cdot \nabla v d\mathbf{x} = -\frac{1}{2} \int_\Omega \left[\frac{\partial \varphi}{\partial x_i} \frac{\partial v}{\partial x_j} + \frac{\partial \varphi}{\partial x_j} \frac{\partial v}{\partial x_i} \right] d\mathbf{x}. \end{array} \tag{34}$$

The above linear systems can be solved by a sparse Cholesky solver (with the Cholesky factorization made once and for all at the beginning of the algorithm). The overhead in computational time appears to be non significant. Numerical results in Section 10 show that the above regularization procedure generally provides a significant improvement to the orders of convergence of the approximations of the solution ψ of problem (2). On the other hand, in the particular case of triangulations like the one on the left of Figure 2, the regularization associated with (33) or (34), deteriorates significantly the $L^2(\Omega)$-approximation error, while preserving optimal orders of convergence.

7. Discrete Least-Squares Formulation and Discrete Relaxation Algorithm

We advocate the following *nonlinear least-squares* method for the solution of problem (32):

$$\text{Find } \ (\psi_h, \mathbf{p}_h) \in V_{gh} \times \mathbf{Q}_{fh} \ \text{ such that } \ J_h(\psi_h, \mathbf{p}_h) \leq J_h(\varphi, \mathbf{q}), \quad \forall (\varphi, \mathbf{q}) \in V_{gh} \times \mathbf{Q}_{fh}, \tag{35}$$

where

$$J_h(\varphi, \mathbf{q}) = \frac{1}{2} \left| \left| \left| \mathbf{D}^2_h(\varphi) - \mathbf{q} \right| \right| \right|^2_{0h}.$$

In order to solve the nonlinear least-squares problem (35), we suggest the following *relaxation* algorithm:

$$\text{Find } \ \psi^0_h \in V_{gh} \ \text{ such that } \ \int_\Omega \nabla \psi^0_h \cdot \nabla \varphi d\mathbf{x} = -2(\sqrt{f_h}, \varphi)_{0h}, \quad \forall \varphi \in V_{0h}. \tag{36}$$

For $n \geq 0$, assuming that ψ^n_h is known, compute $\mathbf{p}^n_h, \psi^{n+1/2}_h$ and ψ^{n+1}_h as follows:

$$\mathbf{p}^n_h = \arg \min_{\mathbf{q} \in \mathbf{Q}_{fh}} J_h(\psi^n_h, \mathbf{q}), \tag{37}$$

$$\psi_h^{n+1/2} = \arg\min_{\varphi \in V_{gh}} J_h(\varphi, \mathbf{p}_h^n), \tag{38}$$

$$\psi_h^{n+1} = \psi_h^n + \omega(\psi_h^{n+1/2} - \psi_h^n), \tag{39}$$

with $0 < \omega < \omega_{\max} \leq 2$. The solution of the finite dimensional problems (37) and (38) will be addressed in the following sections.

8. Numerical Solution of the Discrete Sub-problems (37)

An explicit formulation of problem (37) is given by

$$\mathbf{p}_h^n = \arg\min_{\mathbf{q} \in \mathbf{Q}_{fh}} \left[\frac{1}{2} |||\mathbf{q}|||_{0h}^2 - ((\mathbf{D}_h^2(\psi_h^n), \mathbf{q}))_{0h} \right].$$

This minimization problem can be solved *point-wise*, at each vertex of $\mathcal{T}_h$ belonging to Σ_{0h}, that is:

$$\mathbf{p}_h^n(P_k) = \arg\min_{\mathbf{q} \in \mathbf{E}_{fh}(P_k)} \left[\frac{1}{2} |\mathbf{q}|^2 - \mathbf{D}_h^n(P_k) : \mathbf{q} \right], \ k = 1, \ldots, N_{0h},$$

where $\mathbf{D}_h^n(P_k) = \mathbf{D}_h^2(\psi_h^n)(P_k)$ and $\mathbf{E}_{fh}(P_k) = \left\{ \mathbf{q} \in \mathbb{R}^{2\times 2}, \mathbf{q} = \mathbf{q}^t, \det \mathbf{q} = f_h(P_k), q_{11} > 0, q_{22} > 0 \right\}$. Both the Newton's and the $\mathbf{Q_{min}}$ methods presented in Section 4 apply here, after replacing $\mathbf{x}$ by P_k, $k = 1, \ldots, N_{0h}$.

9. Conjugate Gradient Solution of the Discrete Sub-problems (38)

9.1. Formulation of (38) as a Discrete Linear Variational Problem

The *Euler-Lagrange equation* associated with problem (38) reads as follows:

$$\text{Find } \psi_h^{n+1/2} \in V_{gh} \text{ such that } ((\mathbf{D}^2(\psi_h^{n+1/2}), \mathbf{D}^2(\varphi)))_{0h} = ((\mathbf{p}_h^n, \mathbf{D}^2(\varphi)))_{0h}, \quad \forall \varphi \in V_{0h}. \tag{40}$$

Problem (40) is a well-posed linear variational problem in the affine space V_{gh}. Following [19, Chapter 3], the solution of problem (40) will be discussed in Section 9.3. However, as written, the linear problem (40) leads to excessive computer resource requirements. This is easy to understand: to derive the linear system equivalent to (40), we need to compute-via the solution of (33) or (34)-the matrix-valued functions $\mathbf{D}_h^2(w^j)$, where the functions w^j form a basis of V_{0h}. To avoid this difficulty, we are going to employ an *adjoint equation* approach to derive an equivalent formulation of (40), well-suited to solution by a conjugate gradient algorithm.

9.2. An adjoint equation based equivalent formulation of problem (40)

Problem (40) is equivalent to:

$$\text{Find } \psi_h^{n+1/2} \in V_{gh} \text{ such that } \left\langle \frac{\partial J_h}{\partial \varphi}(\psi_h^{n+1/2}, \mathbf{p}_h^n), \theta \right\rangle = 0, \quad \forall \theta \in V_{0h}, \tag{41}$$

where, more generally, $\left\langle \frac{\partial J_h}{\partial \varphi}(\varphi, \mathbf{q}), \theta \right\rangle$ denotes the action of the partial derivative $\frac{\partial J_h}{\partial \varphi}(\varphi, \mathbf{q})$ on the test function θ. Suppose that $\mathbf{D}_h^2(\varphi)$ is obtained from φ via relations (34); proceeding as in, e,g., [23] one can easily show that, for all $(\varphi, \mathbf{p}) \in V_{gh} \times \mathbf{Q}_h$:

$$\left\langle \frac{\partial J_h}{\partial \varphi}(\varphi, \mathbf{q}), \theta \right\rangle = \int_\Omega \left[\frac{\partial \lambda_{11}}{\partial x_1} \frac{\partial \theta}{\partial x_1} + \frac{\partial \lambda_{22}}{\partial x_2} \frac{\partial \theta}{\partial x_2} + \frac{\partial \lambda_{12}}{\partial x_1} \frac{\partial \theta}{\partial x_2} + \frac{\partial \lambda_{12}}{\partial x_2} \frac{\partial \theta}{\partial x_1} \right] d\mathbf{x}, \quad \forall \theta \in V_{0h}, \tag{42}$$

where $(\lambda_{11}, \lambda_{12}, \lambda_{22})$ is obtained from φ via the solution of the following (adjoint) system, for $1 \leq i \leq j \leq 2$:

$$\lambda_{ij} \in V_{0h}, \quad (\lambda_{ij}, \theta)_{0h} + C \sum_{K \in \mathcal{T}_h} |K| \int_K \nabla \lambda_{ij} \cdot \nabla \theta d\mathbf{x} = (q_{ij} - D_{hij}^2(\varphi), \theta)_{0h}, \quad \forall \theta \in V_{0h}. \tag{43}$$

Modifying the adjoint system (43), in order to handle (33) instead of (34) is straightforward. The solvers used to compute the $D^2_{hij}(\varphi)$ (via (33) or (34)) still apply to the solution of the linear problems in (43).

9.3. **Conjugate gradient solution of problem (40)**

Assume that $D^2_{hij}(\varphi)$ is obtained from φ via (34). Then, for the solution of problem (40), we can use a conjugate gradient algorithm operating in the spaces V_{0h} and V_{gh} equipped with the scalar product $(v,w) \to (\Delta_h v, \Delta_h w)_{0h}$ and the associated norm. Taking advantage of the results of Section 9.2, this algorithm reads as follows:

Step 1

$$\psi_h^{n+1/2,0} \in V_{gh} \ \text{ given } (\psi_h^{n+1/2,0} = \psi_h^n \text{ for example}). \tag{44}$$

Compute $D^2_{hij}(\psi_h^{n+1/2,0})$ via the solution of:

Find $D^2_{hij}(\psi_h^{n+1/2,0}) \in V_{0h}$ such that, for $1 \le i,j \le 2$:

$$\begin{aligned}(D^2_{hij}(\psi_h^{n+1/2,0}),\theta)_{0h} &+ C\sum_{K\in\mathcal{T}_h}|K|\int_\Omega \nabla D^2_{hij}(\psi_h^{n+1/2,0})\cdot\nabla\theta d\mathbf{x}\\ &= -\frac{1}{2}\int_\Omega\left[\frac{\partial\psi_h^{n+1/2,0}}{\partial x_i}\frac{\partial\theta}{\partial x_j}+\frac{\partial\psi_h^{n+1/2,0}}{\partial x_j}\frac{\partial\theta}{\partial x_i}\right]d\mathbf{x},\ \forall\theta\in V_{0h}.\end{aligned} \tag{45}$$

and then $(\lambda_{11}^{n+1/2,0},\lambda_{12}^{n+1/2,0},\lambda_{22}^{n+1/2,0}) \in (V_{0h})^3$ via the solution of the adjoint system:

Find $\lambda_{ij}^{n+1/2,0} \in V_{0h}$, for $1 \le i \le j \le 2$, such that:

$$(\lambda_{ij}^{n+1/2,0},\theta)_{0h} + C\sum_{K\in\mathcal{T}_h}|K|\int_K \nabla\lambda_{ij}^{n+1/2,0}\cdot\nabla\theta d\mathbf{x} = (p_{ij}^n - D^2_{hij}(\psi_h^{n+1/2,0}),\theta)_{0h},\quad \forall\theta\in V_{0h}. \tag{46}$$

Step 2 Solve:

Find $g^{n+1/2,0} \in V_{0h}$ such that

$$(\Delta_h g^{n+1/2,0},\Delta_h\varphi)_{0h} = \int_\Omega\left[\frac{\partial\lambda_{11}^{n+1/2,0}}{\partial x_1}\frac{\partial\varphi}{\partial x_1}+\frac{\partial\lambda_{22}^{n+1/2,0}}{\partial x_2}\frac{\partial\varphi}{\partial x_2}+\frac{\partial\lambda_{12}^{n+1/2,0}}{\partial x_1}\frac{\partial\varphi}{\partial x_2}+\frac{\partial\lambda_{12}^{n+1/2,0}}{\partial x_2}\frac{\partial\varphi}{\partial x_1}\right]d\mathbf{x},\quad \forall\varphi\in V_{0h}, \tag{47}$$

and set

$$w^{n+1/2,0} = g^{n+1/2,0}. \tag{48}$$

Then, for $k \ge 0$, assuming that $\psi_h^{n+1/2,k}, g^{n+1/2,k}$ and $w^{n+1/2,k}$ are known, the last two different from zero, we compute $\psi_h^{n+1/2,k+1}, g^{n+1/2,k+1}$ and, if necessary, $w^{n+1/2,k+1}$ as follows.

Step 3 Compute $D^2_{hij}(w^{n+1/2,k})$ via the solution of:

Find $D^2_{hij}(w^{n+1/2,k}) \in V_{0h}$ such that, for $1 \le i,j \le 2$:

$$\begin{aligned}(D^2_{hij}(w^{n+1/2,k}),\theta)_{0h} &+ C\sum_{K\in\mathcal{T}_h}|K|\int_\Omega \nabla D^2_{hij}(w^{n+1/2,k})\cdot\nabla\theta d\mathbf{x}\\ &= -\frac{1}{2}\int_\Omega\left[\frac{\partial w^{n+1/2,k}}{\partial x_i}\frac{\partial\theta}{\partial x_j}+\frac{\partial w^{n+1/2,k}}{\partial x_j}\frac{\partial\theta}{\partial x_i}\right]d\mathbf{x},\ \forall\theta\in V_{0h},\end{aligned} \tag{49}$$

and then $(\bar{\lambda}_{11}^{n+1/2,k}, \bar{\lambda}_{12}^{n+1/2,k}, \bar{\lambda}_{22}^{n+1/2,k}) \in (V_{0h})^3$ via the solution of the adjoint system:

Find $\bar{\lambda}_{ij}^{n+1/2,k} \in V_{0h}$, for $1 \leq i \leq j \leq 2$, such that:

$$(\bar{\lambda}_{ij}^{n+1/2,k}, \theta)_{0h} + C \sum_{K \in \mathcal{T}_h} |K| \int_K \nabla \bar{\lambda}_{ij}^{n+1/2,k} \cdot \nabla \theta d\mathbf{x} = -(D_{hij}^2(w^{n+1/2,k}), \theta)_{0h}, \quad \forall \theta \in V_{0h}. \tag{50}$$

Solve:

Find $\bar{g}^{n+1/2,k} \in V_{0h}$ such that

$$(\Delta_h \bar{g}^{n+1/2,k}, \Delta_h \varphi)_{0h} = \int_\Omega \left[\frac{\partial \bar{\lambda}_{11}^{n+1/2,k}}{\partial x_1} \frac{\partial \varphi}{\partial x_1} + \frac{\partial \bar{\lambda}_{22}^{n+1/2,k}}{\partial x_2} \frac{\partial \varphi}{\partial x_2} + \frac{\partial \bar{\lambda}_{12}^{n+1/2,k}}{\partial x_1} \frac{\partial \varphi}{\partial x_2} + \frac{\partial \bar{\lambda}_{12}^{n+1/2,k}}{\partial x_2} \frac{\partial \varphi}{\partial x_1} \right] d\mathbf{x}, \quad \forall \varphi \in V_{0h}, \tag{51}$$

and compute the new iterate and residual as follows:

$$\rho_k^{n+1/2} = \frac{\left|\left|\Delta_h g^{n+1/2,k}\right|\right|_{0h}^2}{(\Delta_h \bar{g}^{n+1/2,k}, \Delta_h w^{n+1/2,k})_{0h}}, \tag{52}$$

$$\psi_h^{n+1/2,k+1} = \psi_h^{n+1/2,k} - \rho_k^{n+1/2} w^{n+1/2,k}, \tag{53}$$

$$g^{n+1/2,k+1} = g^{n+1/2,k} - \rho_k^{n+1/2} \bar{g}^{n+1/2,k}. \tag{54}$$

Step 4 Compute

$$\delta_k^{n+1/2} = \frac{\left|\left|\Delta_h g^{n+1/2,k+1}\right|\right|_{0h}^2}{\left|\left|\Delta_h g^{n+1/2,0}\right|\right|_{0h}^2}. \tag{55}$$

If $\delta_k^{n+1/2} < \varepsilon$ (meaning that the residual is small enough), take $\psi_h^{n+1/2} = \psi_h^{n+1/2,k+1}$; otherwise, compute:

$$\gamma_k^{n+1/2} = \frac{\left|\left|\Delta_h g^{n+1/2,k+1}\right|\right|_{0h}^2}{\left|\left|\Delta_h g^{n+1/2,k}\right|\right|_{0h}^2}; \tag{56}$$

and update the descent direction via

$$w^{n+1/2,k+1} = g^{n+1/2,k+1} + \gamma_k^{n+1/2} w^{n+1/2,k}. \tag{57}$$

Step 5 Do $k+1 \to k$ and return to **Step 3**.

Remark 9.1. Modifying algorithm (44)-(57) in order to accommodate the construction of the discrete second order derivatives associated with (33) is straightforward. Since the results of numerical experiments (not reported in this article) have shown that the method based on (33) is no more accurate than the one based on (34) we will focus on the latter, which has also the advantage of being less computer time consuming, everything else being the same.

Remark 9.2. The choice of ε in the stopping criterion of algorithm (44)-(57) is a delicate issue which has been briefly discussed in [19, Chapter 3] (see also the references therein). As expected other stopping criteria are possible, a rather natural one being

$$\frac{\left(\Delta_h g^{n+1/2,k+1}, \Delta_h g^{n+1/2,k+1}\right)_{0h}}{\max\left\{\left(\Delta_h g^{n+1/2,0}, \Delta_h g^{n+1/2,0}\right)_{0h}, \left(\Delta_h \psi_h^{n+1/2,k+1}, \Delta_h \psi_h^{n+1/2,k+1}\right)_{0h}\right\}} < \varepsilon.$$

Remark 9.3 (**Solution of the biharmonic problems).** Concerning the solution of the discrete bi-harmonic problems in (47) and (51), let us observe that both problems are of the following type:

$$\text{Find} \quad r_h \in V_{0h} \quad \text{such that} \quad (\Delta_h r_h, \Delta_h v)_{0h} = \Lambda_h(v), \quad \forall v \in V_{0h}, \tag{58}$$

the functional $\Lambda_h(\cdot)$ being linear over V_h. Let us denote $-\Delta_h r_h$, by ω_h. It follows then from (31) that problem (58) is equivalent to the following system of two coupled discrete Poisson-Dirichlet problems

$$\begin{aligned} &\omega_h \in V_{0h}, \quad \int_\Omega \nabla \omega_h \cdot \nabla v d\mathbf{x} = \Lambda_h(v), \quad \forall v \in V_{0h}, \\ &r_h \in V_{0h}, \quad \int_\Omega \nabla r_h \cdot \nabla v d\mathbf{x} = (\omega_h, v)_{0h}, \quad \forall v \in V_{0h}. \end{aligned} \tag{59}$$

Both problems are well-posed. Actually, the solution (by direct or iterative methods) of discrete Poisson problems, such as (59) has motivated an important literature; some related references can be found in [19, Chapter 5].

10. Numerical Experiments

10.1. **Generalities**

In this section, we shall validate the methodology discussed in Sections 2 to 9. The validation will be achieved via the solution of a variety of test problems associated with domains Ω of different shapes, including some with curved boundaries. We will investigate, in particular, the mesh dependence of the computed solutions. The results of our numerical experiments suggest that the methodology based on the regularization procedure associated with relations (33) and (34) is the only one, so far, able to solve the Monge-Ampère problem (2) accurately on domains of arbitrary convex shapes using piecewise linear continuous approximations on unstructured finite element meshes.

The first test problems to be considered concern (not surprisingly) the case where Ω is the unit square $(0,1)^2$. In order to study the mesh dependence of the computed solution, the three types of triangulations visualized in Figure 2 have been used. The structured triangulations (resp., the un-structured one) have been built using Modulef [2] (resp., Gmsh [18]). The uniform triangulation on the left of Figure 2 is called asymmetric despite the fact that it has some (but not many) symmetry properties; this terminology has been used to distinguish it from triangulations, like the one on the right of Figure 2, which have many symmetry properties. Recall (see Remark 6.1) that on uniform asymmetric triangulations, the discrete second order derivatives provided by relation (31) are exact for polynomial functions of degree ≤ 2.

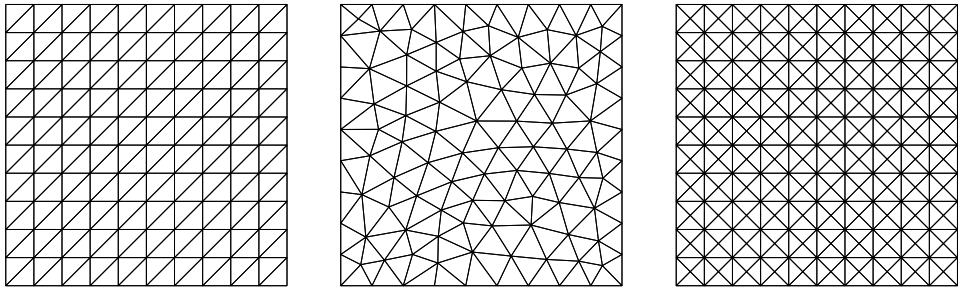

Figure 2. Typical triangulations of the unit square $\Omega = (0,1)^2$. Left: structured (asymmetric) mesh; middle: unstructured (isotropic) mesh; right: structured (symmetric) mesh.

10.2. **First Test Problem**

In this section, and below, we have denoted by $||\cdot||_{0h}$ the discrete variants of the L^2-errors (obtained by numerical integration). The *first test problem* that we consider is defined by

$$\det \mathbf{D}^2\psi(x_1,x_2)=1,\ \forall (x_1,x_2)\in\Omega=(0,1)^2,\quad \psi(x_1,x_2)=\frac{5}{2}x_1^2+2x_1x_2+\frac{1}{2}x_2^2,\ \forall (x_1,x_2)\in\Gamma. \tag{60}$$

The convex solution of the Monge-Ampère-Dirichlet problem (60) is the function ψ given by

$$\psi(x_1,x_2)=\frac{5}{2}x_1^2+2x_1x_2+\frac{1}{2}x_2^2,\ \forall (x_1,x_2)\in\Omega. \tag{61}$$

Its solution being a convex polynomial of degree 2, problem (60) looks rather simple. The condition number of $\mathbf{D}^2\psi$ ($\mathbf{D}^2\psi=\begin{pmatrix}5&2\\2&1\end{pmatrix}$ here) is $\frac{3+2\sqrt{2}}{3-2\sqrt{2}}\simeq 34$, making ψ fairly *anisotropic*. In general, this implies a strong mesh dependence of the approximate solution, particularly if one uses the non-smoothed discrete second order derivatives associated with either (30) or (31).

In Figure 3 and Table 1, we have reported on the three types of meshes shown in Figure 2: (i) Convergence results for the errors $||\psi_h-\psi||_{0h}$ and $||\nabla\psi_h-\nabla\psi||_{0h}$, as functions of the mesh size h, for both the non-regularized (relations (30) or (31)) and regularized (relations (33) and (34), with $C=2$) discrete second order derivatives; (ii) The number of relaxation iterations necessary to achieve convergence, with $\left|\left|\mathbf{D}^2(\psi_h^n)-\mathbf{p}_h^n\right|\right|_{0h}<10^{-4}$ as stopping criterion. Both the Newton's method and the $\mathbf{Q_{min}}$ algorithm have been used to solve the local nonlinear problems (see Sections 4.2 and 4.3). These results deserve several comments:

(i) When both algorithms *relaxation/Newton* and *relaxation/*$\mathbf{Q_{min}}$ converge, they lead essentially to the same solution. However, *relaxation/*$\mathbf{Q_{min}}$ requires significantly fewer iterations to achieve convergence, and there are situations where it converges while *relaxation/Newton* does not. Also, $\mathbf{Q_{min}}$ requires fewer iterations than the Newton algorithm. Also, it seems far less sensitive to initialization than Newton's. Actually, the (well-known) sensitivity to initialization of the standard Newton's method has forced us to take $\omega=0.5$ in some cases (identified with a $\star$ in Table 1), slowing down significantly the convergence of the relaxation method. On the contrary, the greater robustness of $\mathbf{Q_{min}}$ allowed us to work with $\omega=1.5$, making the overall algorithm about 20% faster. On the basis of the superior performances of *relaxation/*$\mathbf{Q}_{\min}$, this method has been retained for the solution of the test problems discussed in the following sections.

(ii) To illustrate how the various iterative methods embedded in the relaxation algorithm perform, let us assume that the stopping criterion for the relaxation iterations is the one mentioned above (that is, $\left|\left|\mathbf{D}_h^2(\psi_h^n)-\mathbf{p}_h^n\right|\right|_{0h}<10^{-4}$); if one takes a 10^{-5} tolerance to stop the conjugate gradient algorithm (44)-(57) (resp., the Newton's method or the $\mathbf{Q_{min}}$ algorithm), we observe the following behavior: The Newton's method (resp., $\mathbf{Q_{min}}$) requires on the average $5-10$ (resp., $2-5$) iterations to converge, while the number of conjugate gradient iterations varies between 9 and 25 and increases as h decreases (as does the number of relaxation iterations).

(iii) The best convergence results as $h\to 0$, and the fastest convergence of the relaxation method, are obtained by combining the uniform asymmetric triangulations (like the one on the left of Figure 2) with the non-regularized approximations of the second derivatives (given by relations (31)) and $\mathbf{Q_{min}}$. As expected, in this particular case, the (approximated) L^2-norm of the approximation errors is quite small (of the order of 10^{-7}), since (cf. Remark 6.1) for this type of triangulations, the discrete second order derivatives associated with (31) are exact for polynomial functions of degree ≤ 2, as is the convex solution (given by (61)) of problem (60). Considering the various errors associated with, among others, the solvers involved in our methodology and the mesh generator, we never expected results exact up to machine precision. On the other hand, the uniform asymmetric meshes associated with the non-regularized discrete second order derivatives defined by (31) lead to $||\nabla(\psi_h-\psi)||_{0h}=\mathcal{O}(h)$, which is generically optimal when approximating the solution of second-order elliptic equations, using piecewise linear continuous finite element approximations.

(iv) Unlike the uniform asymmetric triangulations, the other types of meshes lead to approximation results ranging from poor (for the unstructured isotropic meshes) to terrible (for the structured symmetric meshes) if one uses the non-regularized discrete second order derivatives defined by (31). We observe however that for the unstructured isotropic meshes, although $||\psi_h - \psi||_{0h}$ shows no tendency to converge to 0, we have $||\nabla(\psi_h - \psi)||_{0h} = \mathcal{O}(h)$ for the range of values of h which has been considered. However, there is no contradiction with the Poincaré inequality since, according to [31], we should expect, ultimately, a reduction of the order of convergence for $||\nabla(\psi_h - \psi)||_{0h}$ as $h \to 0$.

(v) For the three types of meshes the regularization of the discrete second order derivatives lead to approximation errors of optimal orders in the range of mesh sizes which has been considered.

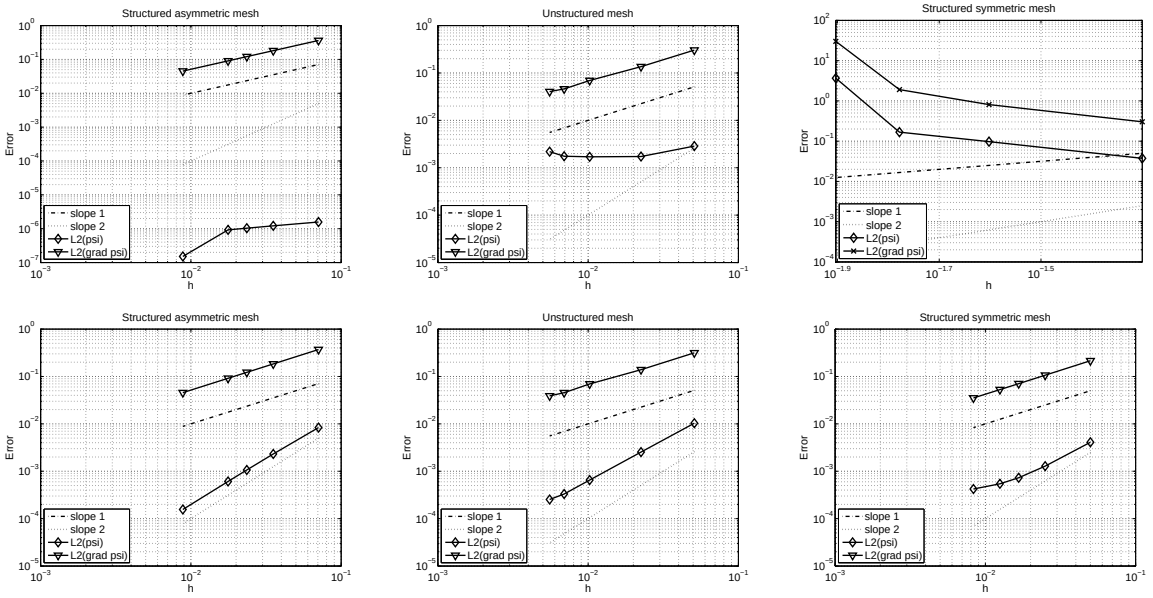

FIGURE 3. First test problem. Convergence (log-log scale) of the errors $||\psi_h - \psi||_{0h}$, $||\nabla(\psi_h - \psi)||_{0h}$; first row: when using non-smoothed approximation of the second derivatives (31). second row: when using smoothed approximation of the second derivatives (34). Left: structured asymmetric meshes; middle: unstructured meshes; right: structured symmetric meshes. All results obtained with $\mathbf{Q_{min}}$.

10.3. **Second Test problem**

Numerical results for test cases on the unit square $\Omega = (0,1)^2$ introduced, e.g., in [12] are presented. Let us consider the test problem defined by $f(x_1, x_2) = \left(1 + (x_1^2 + x_2^2)\right) e^{(x_1^2+x_2^2)}$, and $g(x_1, x_2) = e^{\frac{1}{2}(x_1^2+x_2^2)}$, whose exact solution is the radial function $\psi(x_1, x_2) = e^{\frac{1}{2}(x_1^2+x_2^2)}$, $(x_1, x_2) \in \Omega$. Figure 4 illustrates the solution ψ_h obtained with various types of triangulations. The method for solving the algebraic problems (11) is the $\mathbf{Q_{min}}$ algorithm. The CG algorithm for the solution of the biharmonic problem is stopped when $\delta_k < 10^{-5}$, and the tolerance for the $\mathbf{Q_{min}}$ algorithm is 10^{-5} on successive iterates. The relaxation parameter is $\omega = 1.0$.

Remark 10.1. The stopping criterion for the iterative solution method can be any one of the following three: (i) $\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|_{0h} < 10^{-4}$; (ii) $\left|\left|\psi^{n+1} - \psi^n\right|\right|_{0h} < 10^{-9}$; or (iii) a maximum of 100 outer iterations. Numerical results have shown similar convergence behaviors for all types of stopping criterion, and therefore (i) is used in the whole article (when there is an exact solution). Note that, when using the stopping criterion (ii), numerical

TABLE 1. First test problem. Convergence results and computational costs on the unit square $\Omega = (0,1)^2$. The $\star$ indicates that $\omega = 0.5$ is required to obtain the convergence of the Newton method. The — indicates lack of convergence in the given number of iterations.

Numerical integration (31)			
Algebraic solver 1 (Newton)			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.11480E-06	0.36285E+00	40
0.03535	0.16361E-06	0.18142E+00	97*
0.02357	0.15186E-06	0.12095E+00	114*
0.01767	0.13757E-06	0.90714E-01	125*
0.00883	0.29246E-06	0.45357E-01	115*
Unstructured isotropic mesh			
0.05091	0.28617E-02	0.30375E+00	169
0.02249	0.17258E-02	0.13581E+00	465*
0.01023	0.17046E-02	0.69410E-01	560*
0.00692	0.17472E-02	0.46085E-01	664*
0.00554	0.21852E-02	0.40447E-01	479*
Structured symmetric mesh			
0.05000	0.37465E-01	0.30251E+00	578
0.02500	0.97328E-01	0.80906E+00	4598*
0.01666	—	—	5000*
0.01250	—	—	5000*
0.00833	—	—	5000*

Numerical integration with smoothing (34)			
Algebraic solver 1 (Newton)			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.83577E-02	0.37018E+00	46
0.03535	0.23143E-02	0.18418E+00	132*
0.02357	0.10651E-02	0.12237E+00	148*
0.01767	0.60947E-03	0.91576E-01	169*
0.00883	0.15617E-03	0.45599E-01	256*
Unstructured isotropic mesh			
0.05091	0.10322E-01	0.31579E+00	116
0.02249	0.25425E-02	0.13882E+00	152
0.01023	0.65236E-03	0.69523E-01	391*
0.00692	0.33104E-03	0.45525E-01	495*
0.00554	0.25301E-03	0.38936E-01	321*
Structured symmetric mesh			
0.05000	0.40936E-02	0.21455E+00	100
0.02500	0.12814E-02	0.10627E+00	124
0.01666	0.72963E-03	0.70579E-01	135
0.01250	0.54283E-03	0.52825E-01	178*
0.00833	0.42131E-03	0.35150E-01	187*

Numerical integration (31)			
Algebraic solver 2 ($\mathbf{Q_{min}}$) [33]			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.15704E-05	0.36285E+00	29
0.03535	0.12140E-05	0.18142E+00	35
0.02357	0.10322E-05	0.12095E+00	41
0.01767	0.92135E-06	0.90714E-01	45
0.00883	0.15125E-06	0.45357E-01	66
Unstructured isotropic mesh			
0.05091	0.28615E-02	0.30375E+00	110
0.02249	0.17243E-02	0.13580E+00	164
0.01023	0.17036E-02	0.69406E-01	192
0.00692	0.17463E-02	0.46082E-01	219
0.00554	0.21853E-02	0.40447E-01	243
Structured symmetric mesh			
0.05000	0.37463E-01	0.30251E+00	413
0.02500	0.97332E-01	0.80897E+00	1750
0.01666	0.16735E+00	0.19063E+01	1654
0.01250	0.37988E+05	0.34745E+06	2500
0.00833	—	—	2500

Numerical integration with smoothing (34)			
Algebraic solver 2 ($\mathbf{Q_{min}}$) [33]			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.83561E-02	0.37017E+00	38
0.03535	0.23140E-02	0.18417E+00	55
0.02357	0.10649E-02	0.12236E+00	77
0.01767	0.60939E-03	0.91575E-01	122
0.00883	0.15618E-03	0.45599E-01	278
Unstructured isotropic mesh			
0.05091	0.10320E-01	0.31578E+00	67
0.02249	0.25414E-02	0.13881E+00	98
0.01023	0.65187E-03	0.69518E-01	121
0.00692	0.33072E-03	0.45521E-01	141
0.00554	0.25306E-03	0.38936E-01	161
Structured symmetric mesh			
0.05000	0.40924E-02	0.21455E+00	64
0.02500	0.12810E-02	0.10627E+00	79
0.01666	0.72934E-03	0.70579E-01	87
0.01250	0.54295E-03	0.52825E-01	91
0.00833	0.42142E-03	0.35150E-01	97

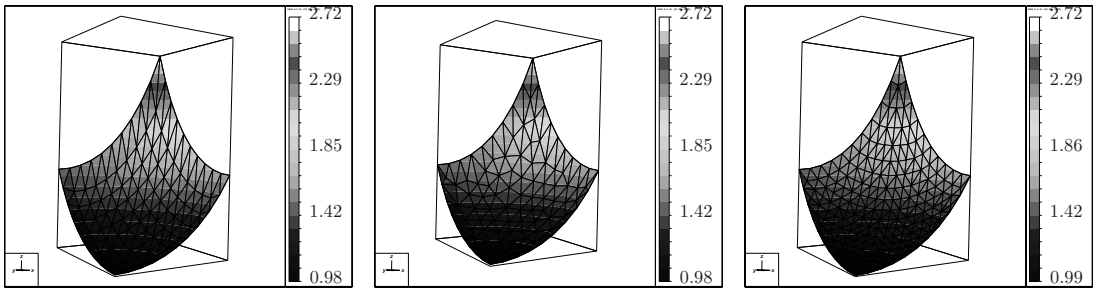

FIGURE 4. Second test problem ($\psi(x_1, x_2) = e^{\frac{1}{2}(x_1^2+x_2^2)}$). Graph of the numerical solution ψ_h. Left: structured asymmetric mesh ($h \simeq 0.0707$) middle: unstructured mesh ($h \simeq 0.0509$); right: structured symmetric mesh ($h = 0.05$).

results show that the residual $\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|_{0h}$ varies like h^2 approximately, which agrees with the results in [11, 13].

Table 2 shows convergence results of the approximation ψ_h (and its first derivatives) towards the exact solution. One observes better performance of the iterative algorithm on structured asymmetric meshes as compared to other types of triangulations. Moreover, the approximations are more accurate since they do not require the use of smoothing techniques. Typically, the CG algorithm converges in $7-10$ iterations, while the $\mathbf{Q_{min}}$ algorithm takes $3-5$ iterations. Figure 5 visualizes the convergence orders of the approximation errors when h decreases to zero. Conclusions are similar to those of the first test problem.

TABLE 2. Second test problem ($\psi(x_1, x_2) = e^{\frac{1}{2}(x_1^2+x_2^2)}$). Convergence results and computational costs.

Numerical integration (31)				Numerical integration with smoothing (34)			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.	h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh				Structured asymmetric mesh			
0.07071	0.26211E-03	0.18269E+00	7	0.07071	0.38793E-01	0.34847E+00	21
0.03535	0.66645E-04	0.91314E-01	8	0.03535	0.11118E-01	0.15425E+00	33
0.02357	0.29859E-04	0.60871E-01	9	0.02357	0.51445E-02	0.93865E-01	51
0.01767	0.16785E-04	0.45652E-01	11	0.01767	0.29478E-02	0.65995E-01	81
0.00883	0.41106E-05	0.22825E-01	20	0.00883	0.75580E-03	0.28775E-01	251
Unstructured isotropic mesh				Unstructured isotropic mesh			
0.05091	0.82765E-03	0.14967E+00	12	0.05091	0.40800E-01	0.35735E+00	24
0.02249	0.19942E-03	0.71111E-01	12	0.02249	0.11210E-01	0.14562E+00	29
0.01023	0.18062E-03	0.35710E-01	16	0.01023	0.27753E-02	0.58341E-01	32
0.00692	0.11006E-03	0.23077E-01	27	0.00692	0.12785E-02	0.34697E-01	34
0.00554	0.15985E-03	0.19824E-01	14	0.00554	0.89411E-03	0.28371E-01	39
Structured symmetric mesh				Structured symmetric mesh			
0.05000	0.92004E-03	0.10655E+00	11	0.05000	0.19129E-01	0.23101E+00	16
0.02500	0.10543E-02	0.53539E-01	13	0.02500	0.52860E-02	0.94312E-01	19
0.01666	0.11588E-02	0.36040E-01	13	0.01666	0.24309E-02	0.56115E-01	20
0.01250	0.11990E-02	0.27397E-01	13	0.01250	0.13930E-02	0.39072E-01	21
0.00833	0.12286E-02	0.18949E-01	13	0.00833	0.63415E-03	0.23746E-01	21

Remark 10.2. The value of the "smoothing parameter" C in (34) has been set to $C = 2$. However this choice is not critical. Figure 6 illustrates, on this exponential example, the influence of the smoothing parameter for

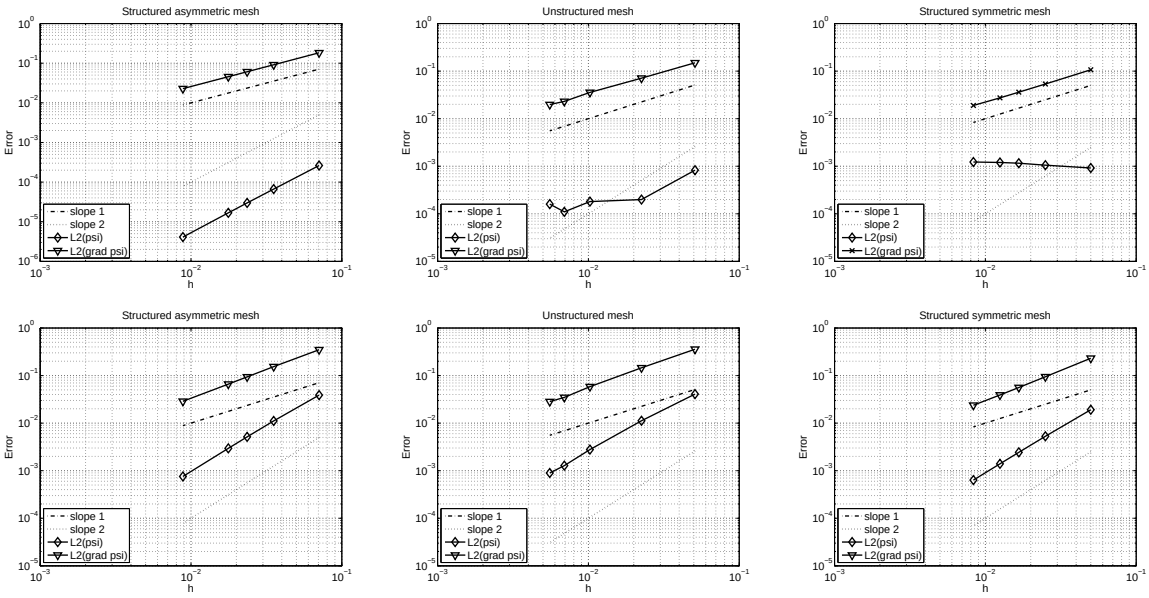

FIGURE 5. Second test problem. Convergence (log-log scale) of the errors $||\psi_h - \psi||_{0h}$, $||\nabla\psi_h - \nabla\psi||_{0h}$; first row: when using non-smoothed approximation of the second derivatives. second row: when using smoothed approximation of the second derivatives. Left: structured asymmetric meshes; middle: unstructured meshes; right: structured symmetric meshes. Stopping criterion $\left|\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|\right| < 10^{-4}$.

unstructured meshes and structured symmetric meshes. The asymmetric case does not require any regularization. These results show that the optimal convergence order for the error $||\psi_h - \psi||_{0h}$ is recovered for any value of $C > 0$. The accuracy of the calculations decreases when C increases. Keeping this remark in mind, we will use $C = 2$ in the sequel.

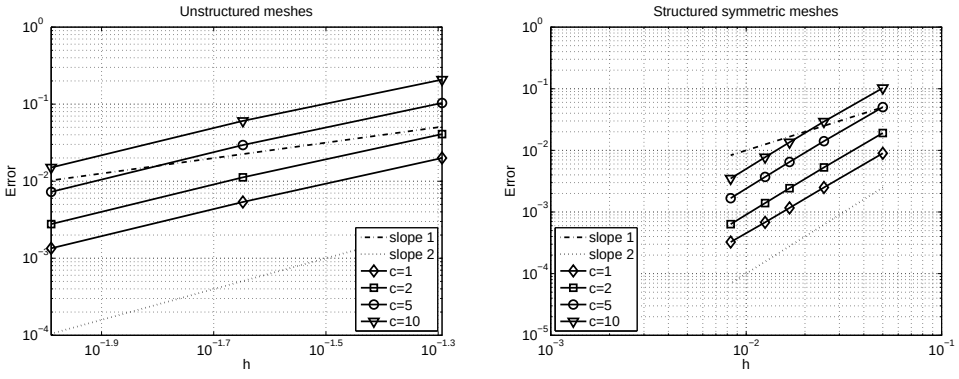

FIGURE 6. Second test problem. Influence of the value of the smoothing parameter C appearing in (34). Convergence (log-log scale) of the errors $||\psi_h - \psi||_{0h}$ for $C = 1, 2, 5$ and 10; Left: unstructured meshes; right: structured symmetric meshes. Stopping criterion $\left|\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|\right| < 10^{-4}$.

10.4. **Third Test Problem**

Let us consider the test problem, defined, for $R \geq \sqrt{2}$, by $f(x_1, x_2) = \frac{R^2}{\left(R^2-(x_1^2+x_2^2)\right)^2}$, and $g(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, whose exact solution is the convex function $\psi(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, $(x_1, x_2) \in \Omega$. When $R > \sqrt{2}$, the exact solution satisfies $\psi \in C^\infty\left(\overline{\Omega}\right)$, while $\psi \in W^{1,p}(\Omega)$, $p \in [1, 4)$, when $R = \sqrt{2}$. Therefore it is interesting to see the performance of the algorithm and the quality of the approximation when R tends to $\sqrt{2}$ from above. In order to highlight this effect, we consider two values of R, namely $R = 2$ (in that case, ψ is smooth), and $R = \sqrt{2} + 0.1$, which is close to the threshold value of $\sqrt{2}$.

Figure 7 shows the graph of ψ_h for $R = 2$. Table 3 and Figure 8 illustrate the computational costs and convergence errors for the three types of triangulations. The numerical experiments show consistent second order accurate approximation of the solution if one smoothes the discrete second order derivatives when employing unstructured isotropic and structured symmetric meshes; they also show that the performances of the method are not altered by the closeness of R to $\sqrt{2}$ (however non-convergence has been observed if $R = \sqrt{2} + 0.01$).

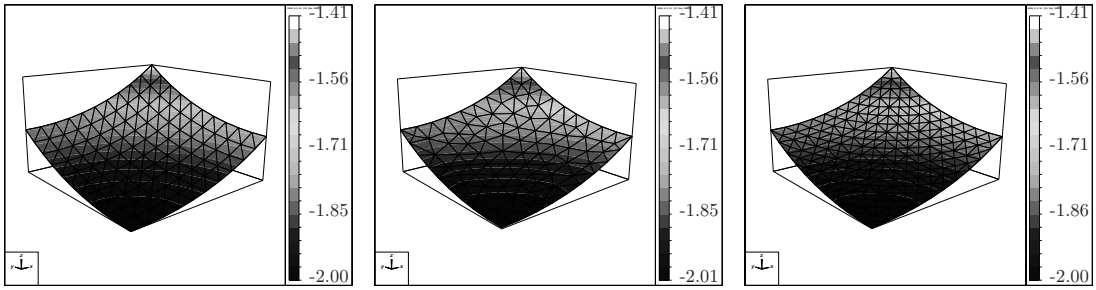

FIGURE 7. Third test problem ($\psi(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, $R = 2$). Graph of the numerical solution ψ_h. Left: structured asymmetric mesh ($h \simeq 0.0707$) middle: unstructured mesh ($h \simeq 0.0509$); right: structured symmetric mesh ($h = 0.05$).

For comparison, Figure 9 shows the graph of ψ_h for $R = \sqrt{2} + 0.1$. Table 4 and Figure 10 illustrate the computational costs and convergence errors for the three types of triangulations. The numerical experiments still show second order accurate approximation of the solution (with the help of smooth approximations of the second derivatives), showing that the performance of the method is not altered by the lack of regularity of the solution. The number of outer iterations is slightly larger when $R \to \sqrt{2}$.

10.5. **Fourth test problem**

The fourth test problem is defined by $f(x_1, x_2) = \frac{1}{\sqrt{x_1^2+x_2^2}}$, and $g(x_1, x_2) = \frac{\left(2\sqrt{x_1^2+x_2^2}\right)^{3/2}}{3}$, whose exact solution is $\psi(x_1, x_2) = \frac{\left(2\sqrt{x_1^2+x_2^2}\right)^{3/2}}{3}$, $(x_1, x_2) \in \Omega$. Figure 11 illustrates the solution ψ_h. Convergence results are given in Table 5 and Figure 12 for the various types of triangulations. Conclusions are similar as in the previous cases, and the importance of the smoothing procedure for the approximation of the second derivatives is again highlighted.

Remark 10.3. The fourth test problem is particularly interesting in the sense that the exact solution $\psi \in H^2(\Omega)$ (in fact $\psi \in W^{2,p}(\Omega)$ for $1 \leq p < 4$) but $\psi \notin C^2\left(\overline{\Omega}\right)$. However, our methodology (which has been constructed to capture solutions with the H^2-regularity) provides optimal order error estimates (without regularization if one uses the uniform asymmetric mesh in Figure 2 (left), and with regularization for the other meshes).

TABLE 3. Third test problem ($\psi(x_1,x_2) = -\sqrt{R^2-(x_1^2+x_2^2)}$, $R=2$). Convergence results and computational costs.

Numerical integration (31)			
h	$\|\|\psi_h-\psi\|\|_{0h}$	$\|\|\nabla(\psi_h-\psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.45221E-04	0.47122E-01	4
0.03535	0.11479E-04	0.23552E-01	5
0.02357	0.50783E-05	0.15700E-01	6
0.01767	0.28570E-05	0.11775E-01	7
0.00883	0.71506E-06	0.58875E-02	12
Unstructured isotropic mesh			
0.05091	0.23140E-03	0.42345E-01	5
0.02249	0.51846E-04	0.19593E-01	5
0.01023	0.23260E-04	0.99121E-02	5
0.00692	0.15681E-04	0.64270E-02	5
0.00554	0.21162E-04	0.55066E-02	5
Structured symmetric mesh			
0.05000	0.29163E-03	0.29676E-01	5
0.02500	0.65022E-04	0.14833E-01	4
0.01666	0.45513E-04	0.98912E-02	5
0.01250	0.47584E-04	0.74211E-02	5
0.00833	0.51990E-04	0.49528E-02	5

Numerical integration with smoothing (34)			
h	$\|\|\psi_h-\psi\|\|_{0h}$	$\|\|\nabla(\psi_h-\psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.12234E-01	0.98603E-01	15
0.03535	0.33987E-02	0.41775E-01	20
0.02357	0.15572E-02	0.25052E-01	34
0.01767	0.88812E-03	0.17481E-01	51
0.00883	0.22620E-03	0.75313E-02	155
Unstructured isotropic mesh			
0.05091	0.13066E-01	0.10374E+00	16
0.02249	0.34240E-02	0.40358E-01	18
0.01023	0.83294E-03	0.15965E-01	18
0.00692	0.38354E-03	0.95234E-02	20
0.00554	0.26919E-03	0.77691E-02	23
Structured symmetric mesh			
0.05000	0.60364E-02	0.66000E-01	11
0.02500	0.16222E-02	0.26370E-01	12
0.01666	0.73754E-03	0.15591E-01	12
0.01250	0.41957E-03	0.10830E-01	12
0.00833	0.18868E-03	0.65727E-02	10

10.6. **Fifth Test Problem**

The last test problem on the unit square is defined, as in the introduction, by $f(x_1,x_2)=1$, and $g(x_1,x_2)=0$. In that case, the Monge-Ampère equation does not have solutions belonging to $H^2(\Omega)$ (it has however viscosity solutions), despite the smoothness of the data. The problem stems from the non-strict convexity of Ω (see [4, 12, 26] for details). Therefore, the solution obtained can only be compared with computational results from the literature, e.g., in [9, 13]. We use the $\mathbf{Q_{min}}$ algorithm in the following discussion, smoothed approximations of the second derivatives (34), and $\omega=1$. The stopping criterion is $\left|\left|\psi_h^n-\psi_h^{n+1}\right|\right|_{0h}<10^{-7}$.

Figure 13 illustrates the solution of the Monge-Ampère equation obtained with all types of triangulations. Figure 14 illustrates the determinant of its computed Hessian. Figure 15 shows a cut of the solution (corresponding respectively to the solutions in Figure 13) for $x_2=1/2$ and $x_1=x_2$ respectively for several mesh sizes. The solution, in particular the solution magnitude, appropriately matches the solution presented in [11–13, 21]. Table 6 shows the values of the residual and the number of iterations for various values of the mesh size h and types of triangulations. A close inspection of the numerical results shows that the curvature of the graph of ψ_h is slightly negative close to the corners of Ω, implying that the Monge-Ampère equation is violated here (indeed the curvature is given by $\det\mathbf{D}^2\psi/(1+|\nabla\psi|^2)^2$). The equation is also violated along the boundary (as emphasized in [21, page 176]) and the Monge-Ampère equation $\det\mathbf{D}^2\psi=1$ is verified with a very high precision sufficiently far away from Γ. For more information on the solutions of $\det\mathbf{D}^2\psi=1$, see [26, Chapter 4] and references therein.

10.7. **A Test Problem on the Unit Disk**

The results in the previous sections have shown the ability of the method based on piecewise linear approximations with arbitrary types of triangulations, including pathological ones. These results agree with the results obtained in the literature, as in, e.g., [12, 17]. In this section, we show that the proposed method based on mixed finite elements applies also to domains with *curved boundaries*. Note that, in the case of curved boundaries, the

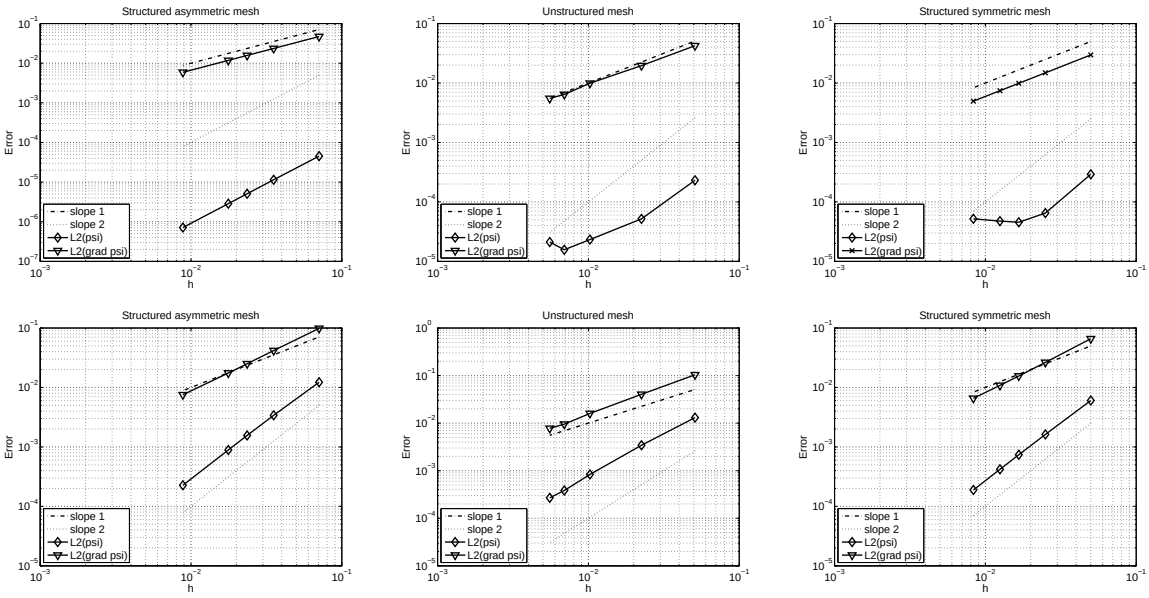

FIGURE 8. Third test problem ($\psi(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, $R = 2$). Convergence (log-log scale) of the errors $||\psi_h - \psi||_{0h}$, $||\nabla\psi_h - \nabla\psi||_{0h}$; first row: when using non-smoothed approximation of the second derivatives. second row: when using smoothed approximation of the second derivatives. Left: structured asymmetric meshes; middle: unstructured meshes; right: structured symmetric meshes. Stopping criterion $\left|\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|\right| < 10^{-4}$.

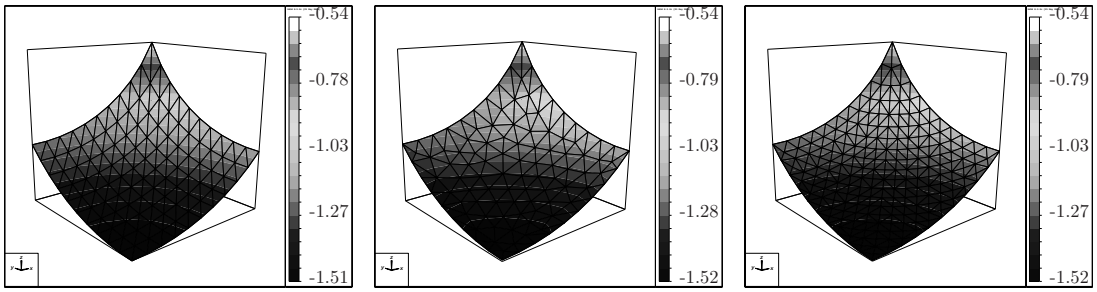

FIGURE 9. Third test problem ($\psi(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, $R = 0.1 + \sqrt{2}$). Graph of the numerical solution ψ_h. Left: structured asymmetric mesh ($h \simeq 0.0707$) middle: unstructured mesh ($h \simeq 0.0509$); right: structured symmetric mesh ($h = 0.05$).

use of mixed piecewise linear finite elements is a substantial simplification compared to using high order finite elements (as in [16] for instance), or finite differences.

We consider the unit disk $\mathcal{S}_1$, with isotropic triangulations built with GMSH [18]. Figure 16 visualizes the solution for $f = 1$ and $g = 0$ on $\mathcal{S}_1$. The exact convex solution is $\psi(x_1, x_2) = 1/2\left[(x_1^2 + x_2^2) - 1\right]$, which is clearly in $C^\infty\left(\overline{\mathcal{S}_1}\right)$. Figure 17 illustrates the convergence results when using the smoothed approximation of the derivatives (34); $||\psi_h - \psi||_{0h}$ exhibits second order convergence.

TABLE 4. Third test problem ($\psi(x_1, x_2) = -\sqrt{R^2 - (x_1^2 + x_2^2)}$, $R = 0.1 + \sqrt{2}$). Convergence results and computational costs.

Numerical integration (31)			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.20303E-03	0.10862E+00	6
0.03535	0.55841E-04	0.54029E-01	7
0.02357	0.24998E-04	0.35983E-01	8
0.01767	0.14130E-04	0.26977E-01	9
0.00883	0.35070E-05	0.13484E-01	20
Unstructured isotropic mesh			
0.05091	0.43561E-03	0.87579E-01	10
0.02249	0.12274E-03	0.41465E-01	15
0.01023	0.77699E-04	0.20497E-01	15
0.00692	0.46390E-04	0.13447E-01	19
0.00554	0.72315E-04	0.11540E-01	22
Structured symmetric mesh			
0.05000	0.45633E-03	0.62474E-01	12
0.02500	0.30236E-03	0.30985E-01	12
0.01666	0.35711E-03	0.20691E-01	11
0.01250	0.38015E-03	0.15588E-01	11
0.00833	0.39758E-03	0.10542E-01	14

Numerical integration with smoothing (34)			
h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.20578E-01	0.18925E+00	19
0.03535	0.59125E-02	0.86538E-01	28
0.02357	0.27314E-02	0.53407E-01	43
0.01767	0.15625E-02	0.37835E-01	70
0.00883	0.39917E-03	0.16700E-01	209
Unstructured isotropic mesh			
0.05091	0.21498E-01	0.19370E+00	21
0.02249	0.58985E-02	0.81153E-01	24
0.01023	0.14525E-02	0.32952E-01	26
0.00692	0.67199E-03	0.19887E-01	29
0.00554	0.46799E-03	0.16297E-01	32
Structured symmetric mesh			
0.05000	0.99552E-02	0.13018E+00	15
0.02500	0.27493E-02	0.54276E-01	18
0.01666	0.12625E-02	0.32501E-01	18
0.01250	0.72238E-03	0.22686E-01	18
0.00833	0.32760E-03	0.13807E-01	16

10.8. **A Test Problem on an Ellipse**

We consider the elliptical domain $\mathcal{E}_{ab} = \left\{(x_1, x_2) \in \mathbb{R}^2\,,\, x_1^2/a^2 + x_2^2/b^2 < 1\right\}$, with corresponding isotropic triangulations built with GMSH [18]. In particular, let us work with the elliptical domain $\mathcal{E}_{1,2}$, and $f = 1/4$, and $g = 0$. In this case, the exact solution to (2) is given by $\psi(x_1, x_2) = \frac{1}{2}(x_1^2 + x_2^2/4 - 1)$. Figure 16 visualizes the solution ψ_h, while Figure 19 illustrates the convergence results when using the smoothed approximation of the derivatives (34); $||\psi_h - \psi||_{0h}$ exhibits again second order convergence.

Remark 10.4. Note that, for both the unit disk and the elliptical domain, convergence properties are lost when using non-smoothed approximations of the second derivatives (31).

10.9. **A Test problem on the Half-Disk**

Finally let us consider now the half-disk domain $\mathcal{S}_{1,-} := \mathcal{S}_1 \cap \{y < 0\}$, and return to the example presented in Section 10.2, namely $f(x_1, x_2) = 1$ and $g(x_1, x_2) = \frac{5}{2}x_1^2 + 2x_1x_2 + \frac{1}{2}x_2^2$. Figure 20 visualizes the contours of the solution on $\mathcal{S}_{1,-}$, while Figure 21 illustrates the convergence results when using the smoothed approximation of the derivatives (34); $||\psi_h - \psi||_{0h}$ exhibits appropriate second order convergence. The non-strict convexity of the domain increases significantly the number of outer iterations, compared to the two previous test cases. Note that, when using the numerical integration method described in (31), convergence is not guaranteed. On the other hand, if one uses the smooth variant (34) (31), the number of iterations decreases.

10.10. **Further Numerical Results**

Finally let us focus on some non-smooth cases with $f = 1$ and $g = 0$, and consider the triangular domain Ω_T defined by $\Omega_T = \left\{(x_1, x_2) \in \mathbb{R}^2 \,:\, x_1, x_2 > 0,\, 4x_1 + 3x_2 < 12\right\}$, and the half-disk $\mathcal{S}_{1,-}$.

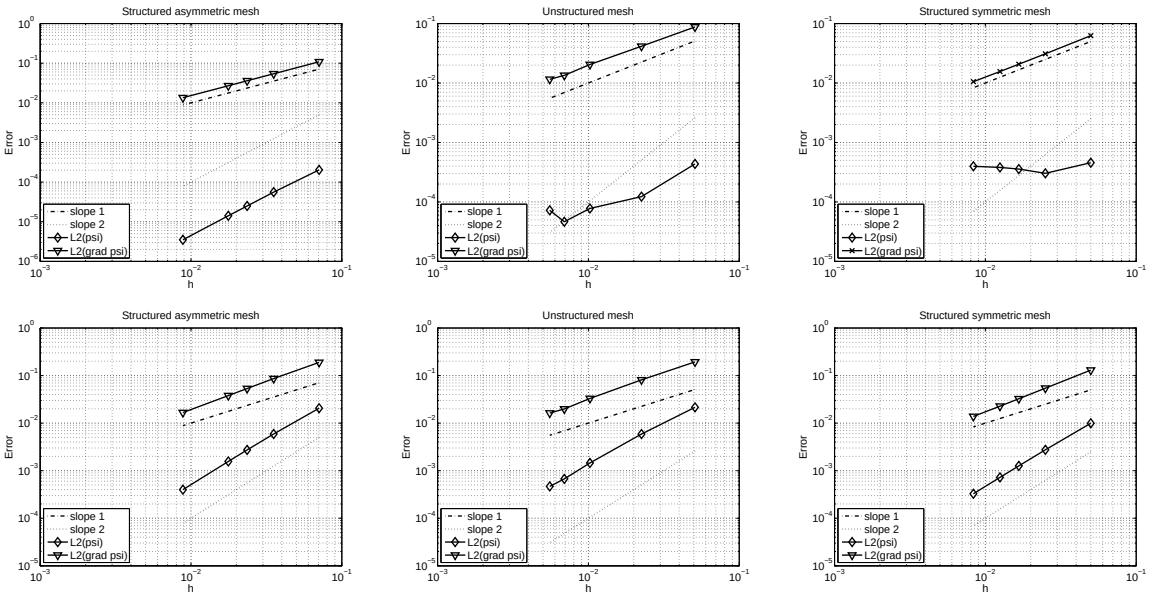

FIGURE 10. Third test problem ($\psi(x_1,x_2) = -\sqrt{R^2-(x_1^2+x_2^2)}$, $R = 0.1+\sqrt{2}$). Convergence (log-log scale) of the errors $||\psi_h - \psi||_{0h}$, $||\nabla\psi_h - \nabla\psi||_{0h}$; first row: when using non-smoothed approximation of the second derivatives. second row: when using smoothed approximation of the second derivatives. Left: structured asymmetric meshes; middle: unstructured meshes; right: structured symmetric meshes. Stopping criterion $\left|\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|\right| < 10^{-4}$.

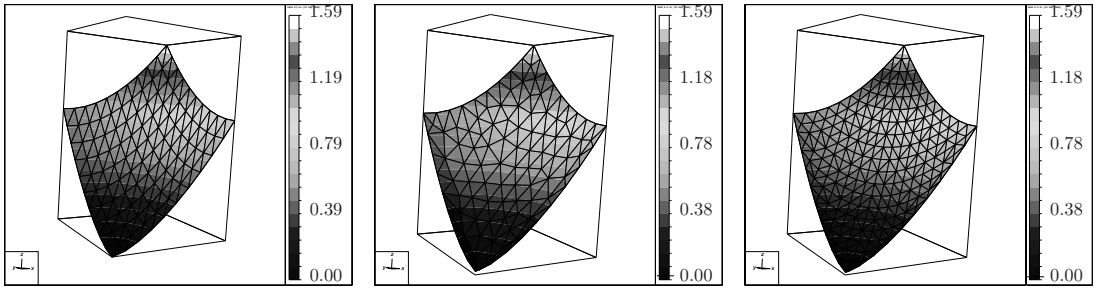

FIGURE 11. Fourth test problem ($\psi(x_1,x_2) = \dfrac{\left(2\sqrt{x_1^2+x_2^2}\right)^{3/2}}{3}$). Graph of the numerical solution ψ_h. Left: structured asymmetric mesh ($h \simeq 0.0707$) middle: unstructured mesh ($h \simeq 0.0509$); right: structured symmetric mesh ($h = 0.05$).

Figure 22 visualizes the approximation of the determinant of the Hessian $\mathbf{D}_h^2(\psi_h^n)$ for these situations, and shows a loss of convexity of the solution in the neighborhood of the corners (and of the parts of the boundary that are not strictly convex) that is similar to the effects observed on the unit square.

11. FURTHER COMMENTS

In this article we have presented a methodology for the numerical solution of the elliptic Monge-Ampère equation in dimension two. The space discretization relies on a stabilized mixed finite element method allowing

TABLE 5. Fourth test problem ($\psi(x_1,x_2)=\frac{\left(2\sqrt{x_1^2+x_2^2}\right)^{3/2}}{3}$). Convergence results and computational costs.

Numerical integration (31)			
h	$\|\|\psi_h-\psi\|\|_{0h}$	$\|\|\nabla(\psi_h-\psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.45388E-03	0.90931E-01	12
0.03535	0.13321E-03	0.45494E-01	12
0.02357	0.63204E-04	0.30334E-01	13
0.01767	0.36721E-04	0.22752E-01	16
0.00883	0.97662E-05	0.11376E-01	33
Unstructured isotropic mesh			
0.05091	0.52042E-03	0.97911E-01	9
0.02249	0.12855E-03	0.45775E-01	10
0.01023	0.95230E-04	0.23044E-01	11
0.00692	0.11528E-03	0.15352E-01	11
0.00554	0.15625E-03	0.13166E-01	11
Structured symmetric mesh			
0.05000	0.11052E-02	0.70733E-01	13
0.02500	0.13901E-02	0.36119E-01	14
0.01666	0.14778E-02	0.24814E-01	15
0.01250	0.15116E-02	0.19345E-01	15
0.00833	0.15370E-02	0.14197E-01	15

Numerical integration with smoothing (34)			
h	$\|\|\psi_h-\psi\|\|_{0h}$	$\|\|\nabla(\psi_h-\psi)\|\|_{0h}$	# iter.
Structured asymmetric mesh			
0.07071	0.25731E-01	0.21046E+00	25
0.03535	0.75209E-02	0.94248E-01	39
0.02357	0.35686E-02	0.57895E-01	67
0.01767	0.20837E-02	0.40811E-01	104
0.00883	0.55681E-03	0.17574E-01	328
Unstructured isotropic mesh			
0.05091	0.28056E-01	0.23726E+00	26
0.02249	0.74447E-02	0.96147E-01	31
0.01023	0.19381E-02	0.40381E-01	39
0.00692	0.90670E-03	0.24606E-01	39
0.00554	0.65328E-03	0.20201E-01	43
Structured symmetric mesh			
0.05000	0.13059E-01	0.15280E+00	19
0.02500	0.37274E-02	0.64631E-01	24
0.01666	0.17583E-02	0.39185E-01	25
0.01250	0.10265E-02	0.27563E-01	25
0.00833	0.47994E-03	0.16901E-01	22

TABLE 6. Fifth test problem ($f=1$, $g=0$). Convergence results and computational costs when the stopping criterion is $\left|\left|\psi_h^n-\psi_h^{n+1}\right|\right|_{0h}<10^{-7}$.

Structured asymmetric mesh			Unstructured isotropic mesh			Structured symmetric mesh		
h	$\left\|\left\|\mathbf{D}^2(\psi_h^n)-\mathbf{p}_h^n\right\|\right\|_{0h}$	# iter.	h	$\left\|\left\|\mathbf{D}^2(\psi_h^n)-\mathbf{p}_h^n\right\|\right\|_{0h}$	# iter.	h	$\left\|\left\|\mathbf{D}^2(\psi_h^n)-\mathbf{p}_h^n\right\|\right\|_{0h}$	# iter.
0.07071	0.10003E-04	43	0.05091	0.52768E-05	48	0.05000	0.10559E-04	48
0.03535	0.78018E-04	140	0.02249	0.42235E-04	105	0.02500	0.64290E-04	146
0.02357	0.19351E-03	337	0.01023	0.22740E-03	203	0.01666	0.15442E-03	261
0.01767	0.29037E-03	763	0.00692	0.47883E-03	259	0.01250	0.28545E-03	377
0.00883	0.80873E-03	2000	0.00554	0.67947E-03	268	0.00833	0.11296E-02	884

the use of piecewise linear approximations for the solution and its second derivatives. This approach is very convenient for domains with curved boundaries. The stabilization procedure provides near optimal orders of convergence for the solution and its gradient. One of the advantages of the least-squares approach discussed here is to provide an alternative to viscosity solutions when the Monge-Ampère problem under consideration has no classical solutions. Actually the solutions obtained by the above least-squares methodology are (kind of) viscosity solutions. To show this property, let us observe first that the least-squares problem (3) is equivalent to the following unconstrained minimization problem

$$\text{Find } (\psi,\mathbf{p})\in V_g\times\mathbf{Q} \text{ such that } J_f(\psi,\mathbf{p})\leq J_f(\varphi,\mathbf{q}),\ \forall(\varphi,\mathbf{p})\in V_g\times\mathbf{Q}, \tag{62}$$

where

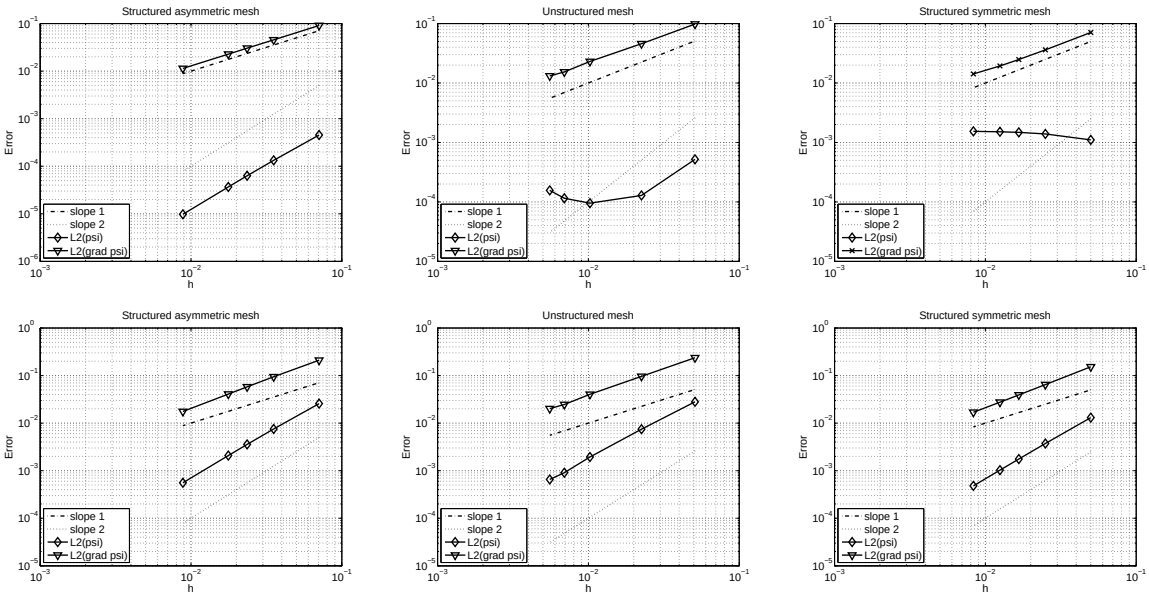

FIGURE 12. Fourth test problem ($\psi(x_1,x_2)=\frac{\left(2\sqrt{x_1^2+x_2^2}\right)^{3/2}}{3}$). Convergence (log-log scale) of the errors $||\psi_h-\psi||_{0h}$, $||\nabla\psi_h-\nabla\psi||_{0h}$; first row: when using non-smoothed approximation of the second derivatives. second row: when using smoothed approximation of the second derivatives. Left: structured asymmetric meshes; middle: unstructured meshes; right: structured symmetric meshes. Stopping criterion $\left|\left|\left|\mathbf{D}_h^2(\psi_h^n)-\mathbf{p}_h^n\right|\right|\right|<10^{-4}$.

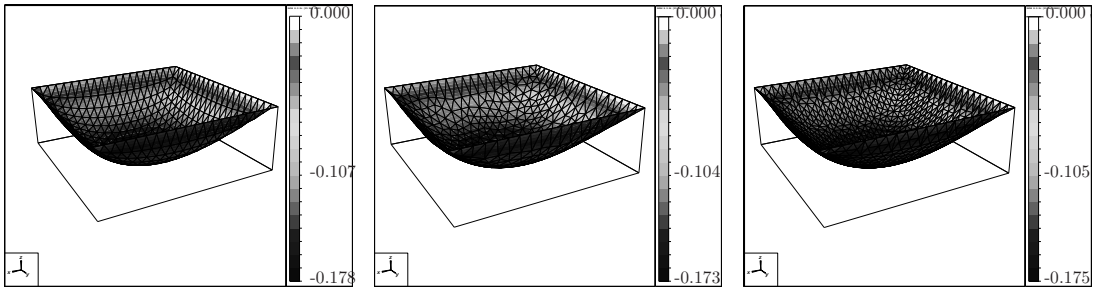

FIGURE 13. Fifth test problem ($f=1$, $g=0$). Graph of the numerical solution ψ_h. Left: structured asymmetric mesh ($h\simeq 0.0353$, after 140 iterations). Middle: unstructured mesh ($h\simeq 0.0225$, after 105 iterations). Right: structured symmetric mesh ($h=0.025$, after 146 iterations).

$$J_f(\varphi,\mathbf{q})=\frac{1}{2}\int_\Omega\left|\mathbf{D}^2\varphi-\mathbf{q}\right|^2 d\mathbf{x}+I_f(\mathbf{q}), \tag{63}$$

with $I_f(\cdot)$ the indicator functional of the set $\mathbf{Q}_f$, that is

$$I_f(\mathbf{q})=\begin{cases}0 & \text{if } \mathbf{q}\in\mathbf{Q}_f,\\ +\infty & \text{if } \mathbf{q}\in\mathbf{Q}\backslash\mathbf{Q}_f,\end{cases}$$

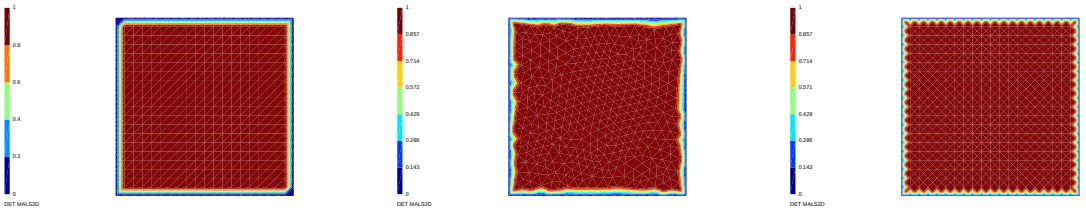

FIGURE 14. Fifth test problem ($f = 1$, $g = 0$). Determinant of the Hessian $\mathbf{D}^2\psi_h$. Left: structured asymmetric mesh ($h \simeq 0.0353$, after 140 iterations). Middle: unstructured mesh ($h \simeq 0.0225$, after 105 iterations). Right: structured symmetric mesh ($h = 0.025$, after 146 iterations).

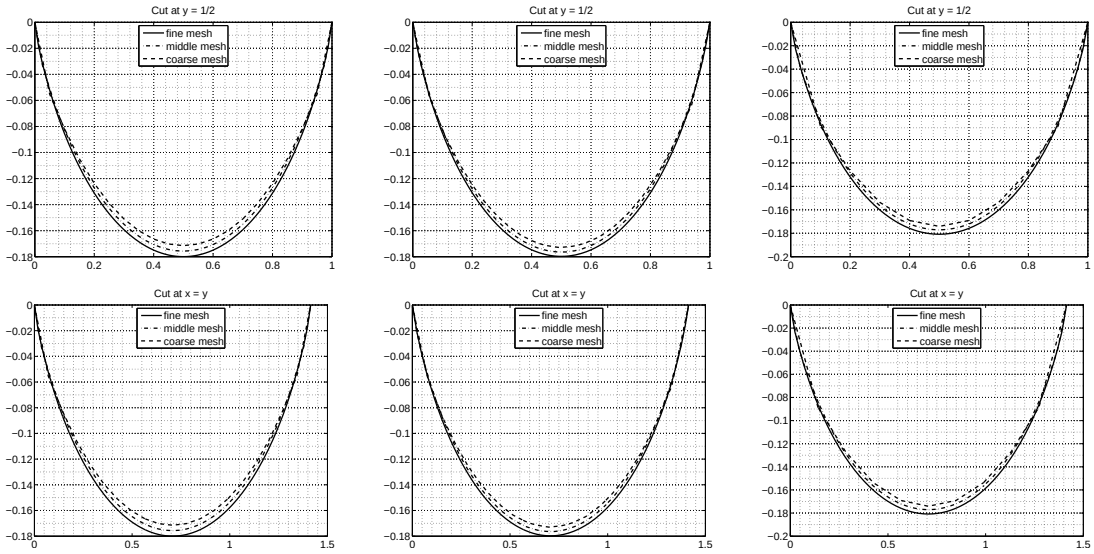

FIGURE 15. Fifth test problem ($f = 1$, $g = 0$). Cut of the graph of the numerical solution ψ_h along the lines $x_2 = 1/2$ (top row) and $x_1 = x_2$ (bottom row). Left: structured asymmetric mesh ($h \simeq 0.0353, 0.0176, 0.0088$). Middle: unstructured mesh ($h \simeq 0.0225, 0.0102, 0.0055$). Right: structured symmetric mesh ($h = 0.05, 0.0166, 0.0083$).

The optimality system associated with (62) reads as

$$\begin{cases} \displaystyle\int_\Omega \mathbf{D}^2\psi : \mathbf{D}^2\varphi d\mathbf{x} = \int_\Omega \mathbf{p} : \mathbf{D}^2\varphi d\mathbf{x}, \quad \forall \varphi \in V_0 \\ \mathbf{p} + \partial I_f(\mathbf{p}) = \mathbf{D}^2\psi. \end{cases} \tag{64}$$

Next, assuming that Ω is simply connected, we introduce

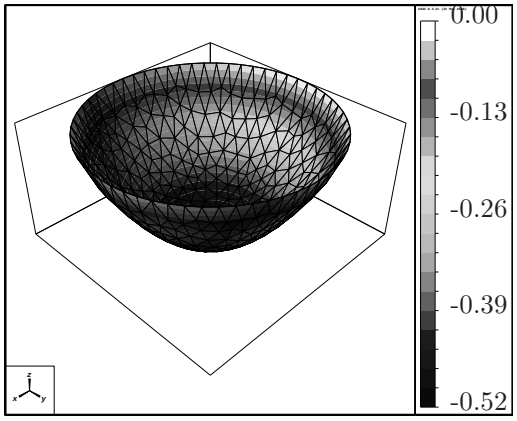

FIGURE 16. Test problem on the unit disk $\mathcal{S}_1$. Graph of the numerical solution ψ_h (for $f = 1$ and $g = 0$) on the unit disk $\mathcal{S}_1$ ($h \simeq 0.04392$, 19 outer iterations).

h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
0.04392	0.35182E-01	0.21176E+00	19
0.02788	0.15247E-01	0.12036E+00	23
0.02083	0.89428E-02	0.84006E-01	19
0.01508	0.58031E-02	0.63335E-01	19
0.01349	0.39951E-02	0.49891E-01	18
0.01028	0.22307E-02	0.35496E-01	21

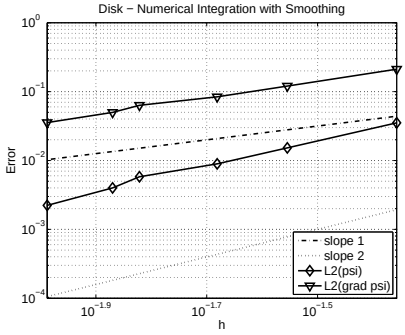

FIGURE 17. Test problem on the unit disk $\mathcal{S}_1$. Convergence of the errors $\|\psi_h - \psi\|_{0h}$ and $\|\nabla\psi_h - \nabla\psi\|_{0h}$ on $\mathcal{S}_1$. The stopping criterion is $\left\|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right\|_{0h} < 10^{-4}$. The algebraic solver is the $\mathbf{Q_{min}}$ algorithm. The second derivatives are approximated with the smoothing technique (33). Left: numerical values; right: Log-log scale plot.

$$
\begin{aligned}
\mathbf{u} &= \{u_1, u_2\} = \left\{\frac{\partial u}{\partial x_2}, -\frac{\partial u}{\partial x_1}\right\}, \quad \mathbf{v} = \{v_1, v_2\} = \left\{\frac{\partial v}{\partial x_2}, -\frac{\partial v}{\partial x_1}\right\}, \\
\mathbf{V}_g &= \left\{\mathbf{v} \in (H^1(\Omega))^2,\, \nabla\cdot\mathbf{v} = 0,\, \mathbf{v}\cdot\mathbf{n} = dg/ds \text{ on } \partial\Omega\right\}, \\
\mathbf{V}_0 &= \left\{\mathbf{v} \in (H^1(\Omega))^2,\, \nabla\cdot\mathbf{v} = 0,\, \mathbf{v}\cdot\mathbf{n} = 0 \text{ on } \partial\Omega\right\}, \\
\mathbf{L} &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},
\end{aligned}
$$

where $\mathbf{n}$ stands for the unit vector of the outward normal at $\partial\Omega$ and s is a counterclockwise curvilinear abscissa on $\partial\Omega$. The formulation (64) is equivalent to the following one: Find $(\mathbf{u}, \mathbf{p}) \in \mathbf{V}_g \times \mathbf{Q}$ such that

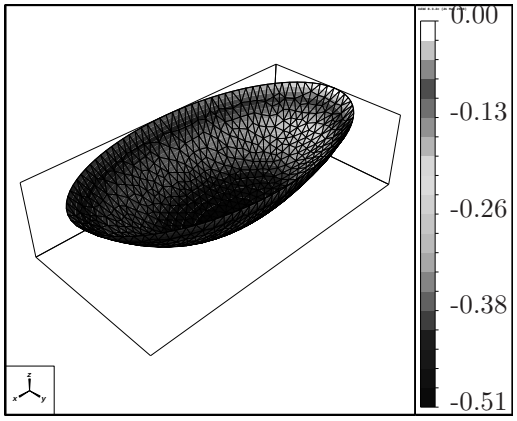

FIGURE 18. Test problem on the elliptical domain $\mathcal{E}_{1,2}$. Graph of the numerical solution ψ_h (for $f = 1/4$ and $g = 0$) on $\mathcal{E}_{1,2}$ ($h \simeq 0.04249$, 72 outer iterations).

h	$\|\|\psi_h - \psi\|\|_{0h}$	$\|\|\nabla(\psi_h - \psi)\|\|_{0h}$	# iter.
0.04249	0.31593E-01	0.18896E+00	72
0.01986	0.80691E-02	0.79596E-01	66
0.01633	0.52340E-02	0.61007E-01	74
0.01377	0.36987E-02	0.48918E-01	68

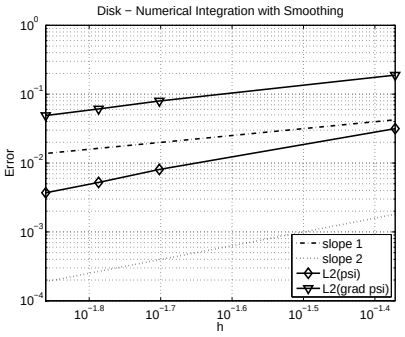

FIGURE 19. Test problem on the elliptical domain $\mathcal{E}_{1,2}$. Convergence of the errors $||\psi_h - \psi||_{0h}$ and $||\nabla\psi_h - \nabla\psi||_{0h}$ on $\mathcal{E}_{1,2}$. The stopping criterion is $\left|\left|\mathbf{D}_h^2(\psi_h^n) - \mathbf{p}_h^n\right|\right|_{0h} < 10^{-4}$. The algebraic solver is the $\mathbf{Q_{min}}$ algorithm. The second derivatives are approximated with the smoothing technique (34). Left: numerical values; right: Log-log scale plot.

$$\begin{cases} \displaystyle\int_\Omega \nabla\mathbf{u} : \nabla\mathbf{v} d\mathbf{x} = \int_\Omega \mathbf{Lp} : \nabla\mathbf{v} d\mathbf{x}, \quad \forall \mathbf{v} \in \mathbf{V}_0 \\ \mathbf{p} + \partial I_f(\mathbf{p}) + \mathbf{L}\nabla\mathbf{u} = \mathbf{0}. \end{cases} \tag{65}$$

The problem (65) has a *visco-elasticity flavor*, $-\mathbf{Lp}$ playing the role of the so-called *elastic stress-tensor*. Similar conclusions can be drawn for the least-squares formulation of other fully nonlinear elliptic equations.

ACKNOWLEDGMENTS

The authors acknowledge the partial support of the National Science Foundation Grants NSF DMS-0412267 and NSF DMS-0913982. The authors thank Prof. E. Dean (Univ. of Houston), Prof. X. Feng (Univ. of Tennessee at Knoxville), Prof. M. Picasso (EPFL) for helpful comments and discussions. The first author gratefully acknowledges the partial support of the Mathematics Institute of Computational Science and Engineering, EPFL, the company Ycoor Systems SA, Switzerland, and the Geneva School of Business Administration (HEG), Switzerland.

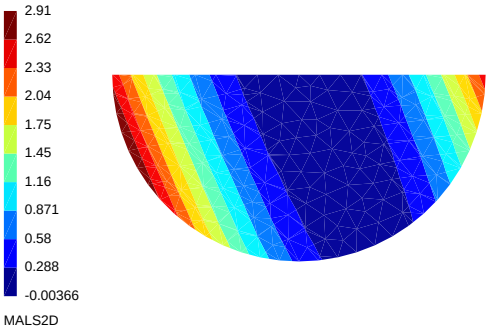

FIGURE 20. Test problem on the half-disk. Contours of the numerical solution ψ_h on $\mathcal{S}_{1,-}$ (for $f(x_1,x_2)=1$ and $g(x_1,x_2)=\frac{5}{2}x_1^2+2x_1x_2+\frac{1}{2}x_2^2$) ($h\simeq 0.04519$, 140 outer iterations).

h	$\|\|\psi_h-\psi\|\|_{0h}$	$\|\|\nabla(\psi_h-\psi)\|\|_{0h}$	# iter.
0.04519	0.95925E-02	0.37198E+00	140
0.02226	0.26230E-02	0.17649E+00	224
0.01009	0.72159E-03	0.84842E-01	281
0.00674	0.34671E-03	0.56310E-01	304

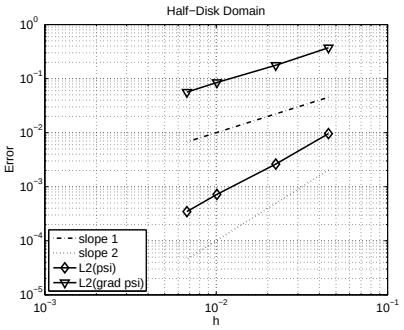

FIGURE 21. Test problem on the half-disk. Convergence of the errors $\|\psi_h-\psi\|_{0h}$ and $\|\nabla\psi_h-\nabla\psi\|_{0h}$ on $\mathcal{E}_{1,2}$. The stopping criterion is $\left\|\mathbf{D}_h^2(\psi_h^n)-\mathbf{p}_h^n\right\|_{0h}<10^{-4}$. The algebraic solver is the $\mathbf{Q_{min}}$ algorithm. The second derivatives are approximated with the smoothing technique (34). Left: numerical values; right: Log-log scale plot.

REFERENCES

[1] J. D. Benamou, B. D. Froese, and A. M. Oberman. Two numerical methods for the elliptic Monge-Ampère equation. *ESAIM: M2AN*, 44(4):737–758, 2010.
[2] M. Bernadou, P.L. George, A. Hassim, P. Joly, P. Laug, A. Perronet, E. Saltel, D. Steer, G. Vanderborck, and M. Vidrascu. Modulef, a modular library of finite elements. Technical report, INRIA, 1988.
[3] K. Boehmer. On finite element methods for fully nonlinear elliptic equations of second order. *SIAM J. Numer. Anal.*, 46(3):1212–1249, 2008.
[4] L. A. Caffarelli. Non linear elliptic theory and the Monge-Ampère equation. In *Proceedings of the International Congress of Mathematicians*, pages 179–187, Beijing, 2002. Higher Education Press.
[5] L. A. Caffarelli and X. Cabré. *Fully Nonlinear Elliptic Equations*. American Mathematical Society, 1995.
[6] L. A. Caffarelli and R. Glowinski. Numerical solution of the Dirichlet problem for a Pucci equation in dimension two. Application to homogenization. *J. Numer. Math.*, 16(3):185–216, 2008.
[7] L. A. Caffarelli, S. A. Kochenkgin, and V. I. Olicker. On the numerical solution of reflector design with given far field scattering data. In *Monge-Ampère Equation: Application to Geometry and Optimization*, pages 13–32, Providence, RI, 1999. American Mathematical Society.
[8] E. J. Dean and R. Glowinski. Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: an augmented Lagrangian approach. *C. R. Acad. Sci. Paris, Sér. I*, 336:779–784, 2003.

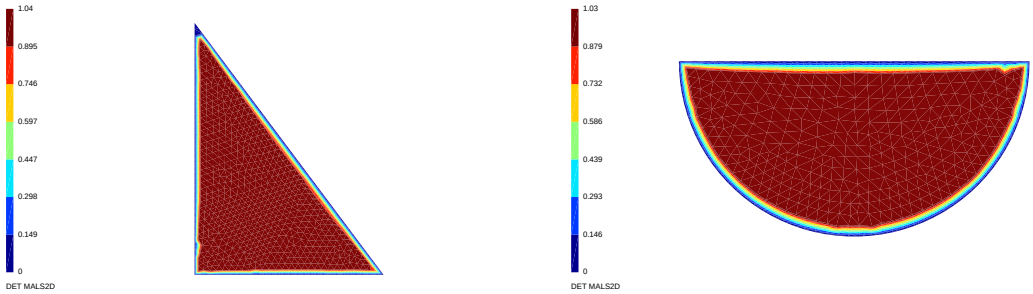

FIGURE 22. Non-smooth example on non strictly convex domains. Determinant of the Hessian of the numerical solution ψ_h ($f = 1$ and $g = 0$). Left: triangular domain Ω_T ($h \simeq 0.0457$, 500 outer iterations). Right: half disk $\mathcal{S}_{1,-}$ ($h \simeq 0.0284$, 500 outer iterations).

[9] E. J. Dean and R. Glowinski. Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: a least-squares approach. *C. R. Acad. Sci. Paris, Sér. I*, 339(12):887–892, 2004.

[10] E. J. Dean and R. Glowinski. Numerical solution of a two-dimensional elliptic Pucci's equation with Dirichlet boundary conditions: a least-squares approach. *C. R. Acad. Sci. Paris, Sér. I*, 341:374–380, 2005.

[11] E. J. Dean and R. Glowinski. An augmented Lagrangian approach to the numerical solution of the Dirichlet problem for the elliptic Monge-Ampère equation in two dimensions. *Electronic Transactions in Numerical Analysis*, 22:71–96, 2006.

[12] E. J. Dean and R. Glowinski. Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type. *Comp. Meth. Appl. Mech. Engrg.*, 195:1344–1386, 2006.

[13] E. J. Dean and R. Glowinski. On the numerical solution of the elliptic Monge-Ampère equation in dimension two: A least-squares approach. In R. Glowinski and P. Neittaanmäki, editors, *Partial Differential Equations: Modeling and Numerical Simulation*, volume 16 of *Computational Methods in Applied Sciences*, pages 43–63. Springer, 2008.

[14] E. J. Dean, R. Glowinski, and T. W. Pan. Operator-splitting methods and applications to the direct numerical simulation of particulate flow and to the solution of the elliptic Monge-Ampère equation. In J.P. Zolésio J. Cagnol, editor, *Control and Boundary Analysis*, pages 1–27. CRC Boca Raton, FLA, 2005.

[15] E. J. Dean, R. Glowinski, and D. Trevas. An approximate factorization/least squares solution method for a mixed finite element approximation of the Cahn-Hilliard equation. *Japan Journal of Industrial and Applied Mathematics*, 13(3):495–517, 1996.

[16] X. Feng and M. Neilan. Mixed finite element methods for the fully nonlinear Monge-Ampère equation based on the vanishing moment method. *SIAM J. Numer. Anal.*, 47(2):1226–1250, 2009.

[17] X. Feng and M. Neilan. Vanishing moment method and moment solutions of second order fully nonlinear partial differential equations. *J. Sci. Comp.*, 38(1):74–98, 2009.

[18] C. Geuzaine and J.-F. Remacle. Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[19] R. Glowinski. *Finite Element Methods For Incompressible Viscous Flow*, volume IX of *Handbook of Numerical Analysis (P.G. Ciarlet, J.L. Lions eds)*, pages 3–1176. Elsevier, Amsterdam, 2003.

[20] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Springer-Verlag, New York, NY, second edition, 2008.

[21] R. Glowinski. Numerical methods for fully nonlinear elliptic equations. In *Invited Lectures, 6th Int. Congress on Industrial and Applied Mathematics, Zürich, Switzerland, 16-20 July 2007*, pages 155–192. EMS, 2009.

[22] R. Glowinski, E. J. Dean, G. Guidoboni, H. L. Juarez, and T. W. Pan. Applications of operator-splitting methods to the direct numerical simulation of particulate and free surface flows and to the numerical solution of the two-dimensional Monge-Ampère equation. *Japan J. Ind. Appl. Math.*, 25(1):1–63, 2008.

[23] R. Glowinski, J.-L. Lions, and J. W. He. *Exact and Approximate Controllability for Distributed Parameter Systems: A Numerical Approach*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2008.

[24] R. Glowinski, D. Marini, and M. Vidrascu. Finite-element approximations and iterative solutions of a fourth-order elliptic variational inequality. *IMA journal of numerical analysis*, 4(2):127–167, 1984.

[25] R. Glowinski and O. Pironneau. Numerical methods for the first bi-harmonic equation and for the two-dimensional Stokes problem. *SIAM Review*, 17(2):167–212, 1979.

[26] C. E. Gutiérrez. *The Monge-Ampère Equation*. Birkhaüser, Boston, 2001.

[27] H. Ishii and P.-L. Lions. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *J. Differential Equations*, 83(1):26–78, 1990.

[28] B. Mohammadi. Optimal transport, shape optimization and global minimization. *C. R. Acad Sci Paris, Sér I*, 351(1):591–596, 2007.

[29] A. Oberman. Wide stencil finite difference schemes for the elliptic Monge-Ampère equations and functions of the eigenvalues of the Hessian. *Discrete and Continuous Dynamical Systems, B*, 10(1):221–238, 2008.

[30] V. I. Oliker and L. D. Prussner. On the numerical solution of the equation $z_{xx}z_{yy} - z_{xy}^2 = f$ and its discretization, I. *Numer. Math.*, 54:271–293, 1988.

[31] M. Picasso, F. Alauzet, H. Borouchaki, and P.-L. George. A numerical study of some Hessian recovery techniques on isotropic and anisotropic meshes. Technical report, INRIA, 2010.

[32] L. Reinhart. On the numerical analysis of the Von Kármán equation: mixed finite element approximation and continuation techniques. *Numer. Math.*, 39:371–404, 1982.

[33] D. C. Sorensen and R. Glowinski. A quadratically constrained minimization problem arising from PDE of Monge-Ampère type. *Numer. Algor.*, 53(1):53–66, 2010.