INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus
Mécanique

Susanne C. Brenner, Li-yeng Sung and Zhiyu Tan

A finite element method for a two-dimensional Pucci equation

Published online: 21 November 2023

https://doi.org/10.5802/crmeca.224

Part of Special Issue: The scientific legacy of Roland Glowinski
Guest editors: Gregoire Allaire (CMAP, Ecole Polytechnique, Institut Polytechnique de Paris, Palaiseau, France), Jean-Michel Coron (Laboratoire Jacques-Louis Lions, Sorbonne Université) and Vivette Girault (Laboratoire Jacques-Louis Lions, Sorbonne Université)

e-ISSN : 1873-7234

Comptes Rendus
Mécanique
Published online: 21 November 2023
https://doi.org/10.5802/crmeca.224

The scientific legacy of Roland Glowinski / *L'héritage scientifique de Roland Glowinski*

A finite element method for a two-dimensional Pucci equation

Méthode des éléments finis pour une équation de Pucci à deux dimensions

Susanne C. Brenner [*, a], **Li-yeng Sung** [a] **and Zhiyu Tan** [b]

[a] Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA

[b] Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA

E-mails: brenner@math.lsu.edu (Susanne C. Brenner), sung@math.lsu.edu (Li-yeng Sung), ztan@cct.lsu.edu (Zhiyu Tan)

Abstract. A nonlinear least-squares finite element method for strong solutions of the Dirichlet boundary value problem of a two-dimensional Pucci equation on convex polygonal domains is investigated in this paper. We obtain a priori and a posteriori error estimates and present corroborating numerical results, where the discrete nonsmooth and nonlinear optimization problems are solved by an active set method and an alternating direction method with multipliers.

Résumé. Cet article étudie une méthode d'éléments finis non linéaire des moindres carrés pour les solutions fortes du problème de la valeur limite de Dirichlet d'une équation de Pucci bidimensionnelle sur des domaines polygonaux convexes. Nous obtenons des estimations d'erreur *a priori* et *a posteriori* et présentons des résultats numériques corroborants, où les problèmes d'optimisation discrets non lisses et non linéaires sont résolus par une méthode d'ensemble active et une méthode de direction alternée avec multiplicateurs.

Keywords. Pucci's equation, finite element method, strong solution, a priori and a posteriori error estimates, nonlinear least-squares, active set method, ADMM.

Mots-clés. équation de Pucci, méthode des éléments finis, solution forte, estimations d'erreur *a priori* et *a posteriori*, moindres carrés non linéaires, méthode des ensembles actifs, ADMM.

Funding. This work was supported in part by the National Science Foundation under Grant No. DMS-22-08404.

Published online: 21 November 2023

---

* Corresponding author.

1. Introduction

Pucci's equation is a fully nonlinear second order partial differential equation that appears in the study of linear uniformly elliptic equations in nondivergence form (cf. [1–3]), with applications to optimal designs (cf. [4]) and population models (cf. [5,6]).

In this paper we consider the following Dirichlet boundary value problem for a two-dimensional Pucci equation:

$$\begin{cases} \alpha\lambda_{\max}\left(D^2u\right)+\lambda_{\min}\left(D^2u\right)=\psi & \text{in } \Omega, \\ u=\phi & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\alpha > 1$, $\lambda_{\max}(D^2u)$ (resp., $\lambda_{\min}(D^2u)$) is the maximum (resp., minimum) eigenvalue of D^2u (the Hessian of u), $\psi \in L^2(\Omega)$, $\phi \in H^2(\Omega)$ and $\Omega \subset \mathbb{R}^2$ is a bounded convex polygon.

Here and below we will follow the standard notation for differential operators, function spaces and norms that can be found for example in [7–10].

The first work on the numerical treatment of Pucci's equation was due to Glowinski and collaborators (cf. [4, 11]). They considered a two-dimensional Pucci equation and investigated a nonlinear least-squares finite element method (which is different from the nonlinear least-squares approach in this paper). Later a nonvariational finite element method for the two-dimensional Pucci equation was studied in [12]. These finite element methods were tested extensively in these cited references without convergence analysis.

In the context of viscosity solutions, a wide stencil monotone finite difference scheme was proposed in [13] for a class of fully nonlinear elliptic partial differential equations in two dimensions that include the Pucci equation. Convergence of the scheme (without convergence rate) was established through the general framework in [14]. Meshless monotone finite difference schemes were developed in [15] and a second order consistent finite difference scheme for the Pucci equation appeared in [16].

We refer interested readers to the survey articles [17,18] for other numerical methods designed for general fully nonlinear elliptic partial differential equations in various solution classes that may also be applied to the Pucci equation.

In this paper we adopt the nonlinear least-squares approach in [19] for the Monge–Ampére equation to construct a finite element method for the problem (1). We establish the existence and uniqueness of a strong solution of (1) in $H^2(\Omega)$, and derive a priori and a posteriori error estimates for the numerical solutions. The nonlinear and nonsmooth discrete least-squares problems are solved by two optimization algorithms. The first one is an active set method and the second one is an alternating direction method with multipliers (ADMM) tailored for the discrete optimization problem resulting from our finite element method. We note that ADMM was pioneered by Glowinski in the 1970's (cf. [20,21]).

The rest of the paper is organized as follows. The existence and uniqueness of the strong solution is established in Section 2. We present the discrete problem in Section 3 together with the error analysis. Numerical results are reported in Section 4 and we end with some concluding remarks in Section 5. Details for the ADMM algorithm is provided in Appendix A.

Throughout the paper we use C (with or without subscripts) to denote generic and strictly positive constants that are independent of the mesh size.

2. Strong Solutions

Let $S_{2\times2}$ be the space of real 2×2 symmetric matrices and $P(M)$ be the Pucci operator defined on $S_{2\times2}$ given by

$$P(M)=\alpha\lambda_{\max}(M)+\lambda_{\min}(M), \tag{2}$$

where the constant α is strictly greater than 1. Then the boundary value problem (1) can be written as

$$\begin{cases} P\left(D^2 u\right) = \psi & \text{in } \Omega, \\ u = \phi & \text{on } \partial\Omega. \end{cases} \tag{3}$$

Remark 1. In the case where $\psi = 0$, we have $\lambda_{\max}(D^2 u) \geq 0 \geq \lambda_{\min}(D^2 u)$ and the operator $P(D^2 u)$ is one of Pucci's extremal operators (cf. [3, Section 2.2]).

Remark 2. We can also write

$$P(M) = \left(\frac{\alpha+1}{2}\right) \operatorname{tr} M + \left(\frac{\alpha-1}{2}\right) \sqrt{(\operatorname{tr} M)^2 - 4 \det M}.$$

Therefore P is nonsmooth at the subset of $S_{2\times 2}$ where $(\operatorname{tr} M)^2 - 4 \det M = 0$.

2.1. *Some Properties of the Pucci Operator*

The properties of $P(M)$ established in this section, which are motivated by results for second order elliptic problems in nondivergence form (cf. [22, 23]), are crucial for the existence and uniqueness of the strong solution for (1) and for the error analysis in Section 3.

We will use the following Hoffman-Wielandt inequality from linear algebra (cf. [24, Theorem 6.3.5, Corollary 6.3.8]):

$$[\lambda_{\max}(M) - \lambda_{\max}(N)]^2 + [\lambda_{\min}(M) - \lambda_{\min}(N)]^2 \leq \|M - N\|_F^2 \qquad \forall\ M, N \in S_{2\times 2}, \tag{4}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

First we have a simple result on the continuity of $P(M)$.

Lemma 3. *We have*

$$|P(M) - P(N)| \leq \sqrt{2}\alpha \|M - N\|_F \qquad \forall\ M, N \in S_{2\times 2}.$$

Proof. It follows from (2), (4) and the Cauchy–Schwarz inequality that

$$|P(M) - P(N)|^2 \leq 2\left(\alpha^2 [\lambda_{\max}(M) - \lambda_{\max}(N)]^2 + [\lambda_{\min}(M) - \lambda_{\min}(N)]^2\right) \leq 2\alpha^2 \|M - N\|_F^2. \qquad \square$$

Next we have a result on the (approximate) monotonicity of $P(M)$.

Lemma 4. *We have, for any $M, N \in S_{2\times 2}$,*

$$\gamma [P(M) - P(N)] \operatorname{tr}(M - N) \geq [\operatorname{tr}(M - N)]^2 - \delta \|M - N\|_F |\operatorname{tr}(M - N)|,$$

where

$$\gamma = \frac{\alpha+1}{\alpha^2+1} \quad \textit{and} \quad \delta = \frac{\alpha - 1}{\sqrt{\alpha^2+1}} < 1. \tag{5}$$

Proof. Let

$$a = \lambda_{\max}(M) - \lambda_{\max}(N) \quad \text{and} \quad b = \lambda_{\min}(M) - \lambda_{\min}(N).$$

It follows that

$$P(M) - P(N) = \alpha a + b \quad \text{and} \quad \operatorname{tr}(M - N) = a + b,$$

and a simple calculation based on the Cauchy–Schwarz inequality yields

$$\begin{aligned} \gamma(\alpha a + b)(a + b) &= (a+b)^2 + \left[\left(\gamma\alpha - 1\right) a + \left(\gamma - 1\right) b\right] (a + b) \\ &\geq (a+b)^2 - \left[\left(\gamma\alpha - 1\right)^2 + \left(\gamma - 1\right)^2\right]^{1/2} \left(a^2 + b^2\right)^{1/2} |a + b| \\ &= (a+b)^2 - \delta \left(a^2 + b^2\right)^{1/2} |a + b|, \end{aligned}$$

where we use the relation $(\gamma\alpha - 1)^2 + (\gamma - 1)^2 = \delta^2$.

Finally we observe that $(a^2 + b^2)^{1/2} \leq \|M - N\|_F$ by the estimate (4). $\square$

Remark 5. The estimate in Lemma 4 becomes a monotonicity result in the case where $\|M - N\|_F \leq |\operatorname{tr}(M - N)|$ (cf. Lemma 6 below).

2.2. *Existence and Uniqueness of Strong Solutions*

We want to show that (1) has a unique strong solution $u \in H^2(\Omega)$, i.e., $P(D^2u) = \psi$ almost everywhere in Ω.

First we observe that this is equivalent to the statement that there exists a unique strong solution $u_0 \in H^2(\Omega)$ such that

$$\begin{cases} P\left(D^2(u_0+\phi)\right) = \psi & \text{in } \Omega, \\ u_0 = 0 & \text{on } \partial\Omega. \end{cases} \tag{6}$$

This will allow us to take advantage of the Miranda-Talenti inequality (cf. [25–27])

$$\left\|D^2 v\right\|_{L^2(\Omega)} \leq \|\Delta v\|_{L^2(\Omega)} \qquad \forall\ v \in H^2(\Omega)\cap H^1_0(\Omega), \tag{7}$$

which is valid for any convex domain Ω. (It is actually an equality when Ω is a polygon (cf. [27, Theorem 4.3.1.4]).)

The following result shows that the operator $P(D^2(v+\phi))$ is monotone with respect to the Laplace operator Δ.

Lemma 6. *The following estimate holds for any $u, v \in H^2(\Omega)\cap H^1_0(\Omega)$:*

$$\int_\Omega \left[P\left(D^2(u+\phi)\right) - P\left(D^2(v+\phi)\right)\right]\Delta(u-v)\,dx \geq \frac{1-\delta}{\gamma}\|\Delta(u-v)\|^2_{L^2(\Omega)}, \tag{8}$$

where γ and δ are the constants defined in (5).

Proof. Let $M = D^2(u+\phi)$ and $N = D^2(v+\phi)$. By Lemma 4, the Cauchy–Schwarz inequality and (7), we have

$$\begin{aligned}\gamma\int_\Omega &\left[P\left(D^2(u+\phi)\right) - P\left(D^2(v+\phi)\right)\right]\Delta(u-v)dx \\ &\geq \|\Delta(u-v)\|^2_{L^2(\Omega)} - \delta\int_\Omega \left\|D^2(u-v)\right\|_F |\Delta(u-v)|dx \\ &\geq \|\Delta(u-v)\|^2_{L^2(\Omega)} - \delta\left\|D^2(u-v)\right\|_{L^2(\Omega)}\|\Delta(u-v)\|_{L^2(\Omega)} \\ &\geq (1-\delta)\|\Delta(u-v)\|^2_{L^2(\Omega)}.\end{aligned}$$

□

We are now ready to establish the main result of this section by applying the theory of Campanato on near operators (cf. [22, 28]).

Theorem 7. *The problem* (1) *has a unique strong solution in $H^2(\Omega)$.*

Proof. It suffices to show that (6) has a unique strong solution for any $\psi \in L^2(\Omega)$.

It follows from Lemma 3 and (7) that

$$\left\|P\left(D^2(u+\phi)\right) - P\left(D^2(v+\phi)\right)\right\|_{L^2(\Omega)} \leq \sqrt{2}\alpha\|\Delta(u-v)\|_{L^2(\Omega)} \quad \forall\ u, v \in H^2(\Omega)\cap H^1_0(\Omega). \tag{9}$$

Let $T_1, T_2 : H^2(\Omega)\cap H^1_0(\Omega) \longrightarrow L^2(\Omega)$ be the operators defined by

$$T_1 u = P\left(D^2(u+\phi)\right) \quad \text{and} \quad T_2 u = \Delta u.$$

The estimates (8) and (9) imply that T_1 and T_2 are near each other in the sense of Companato (cf. [28, Theorem 7]). Since T_2 is a bijection by elliptic regularity on convex domains (cf. [27]), T_1 is also a bijection by [28, Theorem 1]. □

Remark 8. Strong solutions for Pucci type equations were treated in [29, 30] under various assumptions. Theorem 7, which suffices for our purposes, is the simplest result for strong solutions of Pucci's equation.

3. The Discrete Problem

Let $\mathcal{T}_h$ be a quasi-uniform simplicial triangulation of Ω and $V_h \subset H^1(\Omega)$ be the k^{th} degree $(k \geq 2)$ Lagrange finite element space on $\mathcal{T}_h$ (cf. [7,9]). We will denote the piecewise Laplacian operator by Δ_h, the piecewise Hessian operator by D_h^2, the set of the interior edges of $\mathcal{T}_h$ by $\mathcal{E}_h^i$, the length of an edge e by $|e|$, and the jump of the normal derivative of v across an (interior) edge by $[\![\partial v/\partial n]\!]$. The nodal interpolation operator for V_h is denoted by Π_h and $\|\cdot\|_h$ is the mesh-dependent (semi-)norm defined by

$$\|v\|_h^2 = \left\|D_h^2 v\right\|_{L^2(\Omega)}^2 + \sum_{e\in\mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2. \tag{10}$$

The discrete problem is to find

$$u_h = \operatorname*{argmin}_{v_h\in L_h} J_h(v_h), \tag{11}$$

where the constraint set L_h is defined by

$$L_h = \left\{v_h \in V_h : \ v_h = \Pi_h\phi \text{ on } \partial\Omega\right\}, \tag{12}$$

and the cost function J_h is defined by

$$J_h(v_h) = \frac{h^4}{2}\left\|D_h^2 v_h\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2 + \frac{1}{2}\left\|P\left(D_h^2 v_h\right) - \psi\right\|_{L^2(\Omega)}^2. \tag{13}$$

The solvability of the discrete problem is established in the following lemma.

Lemma 9. *The cost function $J_h : L_h \to [0,+\infty)$ has a global minimizer.*

Proof. For $v_h \in L_h$, we have $v_h - \Pi_h\phi \in V_h \cap H_0^1(\Omega)$. According to the Poincaré–Friedrichs inequality for piecewise H^2 functions in [31], we have

$$\left\|v_h - \Pi_h\phi\right\|_{L^2(\Omega)} \leq C\left\|v_h - \Pi_h\phi\right\|_h.$$

Hence, in view of (10) and standard properties of Π_h (cf. [7,9]), we have

$$\begin{aligned}\|v_h\|_{L^2(\Omega)} &\leq \left\|v_h - \Pi_h\phi\right\|_{L^2(\Omega)} + \|\Pi_h\phi\|_{L^2(\Omega)}\\ &\leq C\|v_h - \Pi_h\phi\|_h + \|\Pi_h\phi\|_{L^2(\Omega)} \leq C\big[\|v_h\|_h + \|\phi\|_{H^2(\Omega)}\big],\end{aligned}$$

which together with (13) implies $J_h(v_h) \to +\infty$ as $\|v_h\|_{L^2(\Omega)} \to +\infty$. □

The following discrete Miranda–Talenti estimate (cf. [32, Theorem 1] and [19, A.3]) is useful for the error analysis.

$$\left\|D_h^2(\zeta - v_h)\right\|_{L^2(\Omega)} \leq \|\Delta_h(\zeta - v_h)\|_{L^2(\Omega)} + C_\dagger\left(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac{1}{2}} \tag{14}$$

for any $\zeta \in H^2(\Omega) \cap H_0^1(\Omega)$ and $v_h \in V_h \cap H_0^1(\Omega)$.

3.1. *a priori error estimates*

We begin with a stability estimate.

Lemma 10. *There exists a positive constant C independent of h such that*

$$\|\zeta - v_h\|_h^2 \leq C\left[\left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)}^2 + \sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2 + \|\varphi - \varphi_h\|_h^2\right] \tag{15}$$

for any $v_h, \varphi_h \in L_h$ and any $\zeta, \varphi \in H^2(\Omega)$ that satisfy $\zeta - \varphi \in H_0^1(\Omega)$.

Proof. Let $v_h, \phi_h \in L_h$ be arbitrary. It follows from Lemma 4 and the Cauchy-Schwarz inequality that

$$\|\Delta_h(\zeta - v_h)\|_{L^2(\Omega)} \le \gamma \left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} + \delta \left\|D_h^2(\zeta - v_h)\right\|_{L^2(\Omega)},$$

which then implies through the triangle inequality, the definition of $\|\cdot\|_h$ in (10) and the discrete Miranda–Talenti estimate (14),

$$\begin{aligned}
&\left\|\Delta_h\left[(\zeta-\varphi)-(v_h-\varphi_h)\right]\right\|_{L^2(\Omega)} \\
&\qquad \le \|\Delta_h(\zeta - v_h)\|_{L^2(\Omega)} + \left\|\Delta_h\left(\varphi-\varphi_h\right)\right\|_{L^2(\Omega)} \\
&\qquad \le \gamma \left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} + \delta\left\|D_h^2\left(\zeta - v_h\right)\right\|_{L^2(\Omega)} + \left\|\Delta_h\left(\varphi-\varphi_h\right)\right\|_{L^2(\Omega)} \\
&\qquad \le \gamma \left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} + \delta\left\|D_h^2\left[(\zeta-\varphi)-\left(v_h-\varphi_h\right)\right]\right\|_{L^2(\Omega)} \\
&\qquad\quad + \delta\left\|D_h^2\left(\varphi-\varphi_h\right)\right\|_{L^2(\Omega)} + \|\Delta_h\left(\varphi-\varphi_h\right)\|_{L^2(\Omega)} \\
&\qquad \le \gamma \left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} + \delta\left\|\Delta_h\left[(\zeta-\varphi)-\left(v_h-\varphi_h\right)\right]\right\|_{L^2(\Omega)} \\
&\qquad\quad + \delta C_\dagger\left(\sum_{e\in\mathscr{E}_h^i}|e|^{-1}\,\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac12} + \left(\delta C_\dagger + \delta + \sqrt{2}\right)\left\|\varphi-\varphi_h\right\|_h.
\end{aligned}$$

Consequently we have

$$\begin{aligned}
&\left\|\Delta_h\left[(\zeta-\varphi)-(v_h-\varphi_h)\right]\right\|_{L^2(\Omega)} \\
&\qquad \le \frac{\gamma}{1-\delta}\left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} + C\left[\left\|\varphi-\varphi_h\right\|_h + \left(\sum_{e\in\mathscr{E}_h^i}|e|^{-1}\,\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac12}\right],
\end{aligned}$$

and then, in view of (14),

$$\begin{aligned}
\left\|D_h^2\left[(\zeta-\varphi)-\left(v_h-\varphi_h\right)\right]\right\|_{L^2(\Omega)} &\le \frac{\gamma}{1-\delta}\left\|P\left(D^2\zeta\right) - P\left(D_h^2 v_h\right)\right\|_{L^2(\Omega)} \\
&\quad + C\left[\left\|\varphi-\varphi_h\right\|_h + \left(\sum_{e\in\mathscr{E}_h^i}|e|^{-1}\,\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac12}\right]. \qquad (16)
\end{aligned}$$

The estimate (15) follows from (16) and the estimate

$$\|\zeta - v_h\|_h \le \left\|D_h^2\left[(\zeta-\varphi)-\left(v_h-\varphi_h\right)\right]\right\|_{L^2(\Omega)} + \left\|\varphi-\varphi_h\right\|_h + \left(\sum_{e\in\mathscr{E}_h^i}|e|^{-1}\,\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac12},$$

which is a simple consequence of (10) and the triangle inequality. □

We can now establish a quasi-optimal error estimate for any solution u_h of the discrete problem.

Theorem 11. *Let u be the solution of* (3) *and u_h be a solution of the discrete problem* (11). *We have*

$$\|u - u_h\|_h^2 \le C\left[J_h(v_h) + \|\phi-\phi_h\|_h^2\right] \qquad \forall\ v_h, \phi_h \in L_h. \qquad (17)$$

Proof. In view of the definition of the cost function J_h in (13), we find, by taking $\zeta = u$ and $v_h = u_h$ in (15),

$$\begin{aligned}
\|u - u_h\|_h^2 &\le C\left[\left\|\psi - P\left(D_h^2 u_h\right)\right\|_{L^2(\Omega)}^2 + \sum_{e\in\mathscr{E}_h^i}|e|^{-1}\,\|[\![\partial u_h/\partial n]\!]\|_{L^2(e)}^2 + \|\phi-\phi_h\|_h^2\right] \\
&\le C\left[J_h(u_h) + \|\phi-\phi_h\|_h^2\right]
\end{aligned}$$

for any $\phi_h \in L_h$. The estimate (17) then follows immediately from the optimality of u_h. □

We have concrete error estimates under additional regularity assumptions.

Theorem 12. *Let u be the solution of* (3) *and u_h be a solution of the discrete problem* (11). *Assuming $u, \phi \in H^s(\Omega)$ $(s > 2)$, we have*

$$\|u - u_h\|_h \le C h^{\min\{s-2,k-1,2\}} \tag{18}$$

and

$$\|u - u_h\|_{L^2(\Omega)} + |u - u_h|_{H^1(\Omega)} + \|u - u_h\|_{L^\infty(\Omega)} \le C h^{\min\{s-2,k-1,2\}}. \tag{19}$$

Proof. It follows immediately from (4), (13), Theorem 11 and standard interpolation error estimates that

$$\|u - u_h\|_h \le C\big(h^2 + \|u - \Pi_h u\|_h + \big\|\phi - \Pi_h\phi\big\|_h\big) \le C\Big(h^2 + h^{\min\{s,k+1\}-2}\Big) \le C h^{\min\{s-2,k-1,2\}}.$$

According to the Poincaré–Friedrichs and Sobolev inequalities for piecewise H^2 functions (cf. [31,33]), we have

$$\|\xi\|_{L^2(\Omega)} + |\xi|_{H^1(\Omega)} + \|\xi\|_{L^\infty(\Omega)} \le C\|\xi\|_h$$

for all $\xi \in H_0^1(\Omega)$ that is piecewise H^2 with respect to $\mathcal{T}_h$. Since

$$u - u_h = \big[(u-\phi) - \big(u_h - \Pi_h\phi\big)\big] + \big(\phi - \Pi_h\phi\big)$$

and, in view of the boundary condition in (1) and the definition of L_h in (12),

$$(u-\phi) - \big(u_h - \Pi_h\phi\big) \in H_0^1(\Omega),$$

we can conclude that

$$\begin{aligned}
&\|u - u_h\|_{L^2(\Omega)} + |u - u_h|_{H^1(\Omega)} + \|u - u_h\|_{L^\infty(\Omega)} \\
&\quad \le C\big\|(u-\phi) - \big(u_h - \Pi_h\phi\big)\big\|_h + \big\|\phi - \Pi_h\phi\big\|_{L^2(\Omega)} + |\phi - \Pi_h\phi|_{H^1(\Omega)} + \big\|\phi - \Pi_h\phi\big\|_{L^\infty(\Omega)} \\
&\quad \le C\big(\|u - u_h\|_h + \|\phi - \Pi_h\phi\|_h\big) + \|\phi - \Pi_h\phi\|_{L^2(\Omega)} + |\phi - \Pi_h\phi|_{H^1(\Omega)} + \|\phi - \Pi_h\phi\|_{L^\infty(\Omega)}.
\end{aligned}$$

The estimate (19) follows from (18) and standard interpolation error estimates. □

Remark 13. The numerical results in Section 4 indicate that the orders of convergence in $\|\cdot\|_{L^2(\Omega)}$, $|\cdot|_{H^1(\Omega)}$ and $\|\cdot\|_{L^\infty(\Omega)}$ are better than the one predicted by the estimate (19).

Remark 14. It follows from Theorem 12 that the best error estimate we can expect occurs when $k = 3$, where $h^{\min\{s-2,k-1,2\}} = h^{\min\{s-2,2\}}$. If we use the quadratic Lagrange finite element space, we would have the weaker estimate

$$\|u - u_h\|_h + |u - u_h|_{H^1(\Omega)} + \|u - u_h\|_{L^\infty(\Omega)} + \|u - u_h\|_{L^2(\Omega)} \le C h^{\min\{s-2,1\}}.$$

3.2. *a posteriori error estimates*

Since the discrete problem (11) is a nonlinear optimization problem, in general we do not expect to obtain the global minimizer. However we can monitor the convergence of the computed numerical solutions.

Let $\tilde{u}_h \in L_h$ be an approximate solution of the discrete problem (11). It follows from Lemma 10 and $P(D^2u) = \psi$ that

$$\|u - \tilde{u}_h\|_h^2 \le C\left[\big\|P\big(D_h^2\tilde{u}_h\big) - \psi\big\|_{L^2(\Omega)} + \left(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\,\|[\![\partial\tilde{u}_h/\partial n]\!]\|_{L^2(e)}^2\right)^{\frac{1}{2}} + Osc(\phi)\right] \tag{20}$$

where $Osc(\phi) = \|\phi - \Pi_h\phi\|_h$ is the oscillation term. We note that $Osc(\phi)$ is a higher order term if ϕ is smooth.

On the other hand, according to Lemma 3, we also have

$$\|P(D_h^2\widetilde{u}_h)-\psi\|_{L^2(T)} = \|P(D^2u)-P(D_h^2\widetilde{u}_h)\|_{L^2(T)} \le \sqrt{2}\alpha\,\|D_h^2(u-\widetilde{u}_h)\|_{L^2(T)} \qquad \forall\ T\in\mathcal{T}_h. \quad (21)$$

Let η_h be the residual-based error estimator defined by

$$\eta_h(\widetilde{u}_h) = \|P(D_h^2\widetilde{u}_h)-\psi\|_{L^2(\Omega)} + \Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\,\|[\![\partial\widetilde{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}}. \quad (22)$$

It follows from (20) and (21) that η_h is reliable and locally efficient for the energy error $\|u-\widetilde{u}_h\|_h$. Therefore we can monitor the convergence of $\widetilde{u}_h$ by $\eta_h(\widetilde{u}_h)$. The error estimator η_h can also be used in adaptive mesh refinements. However this will not be pursued in the current paper.

4. Numerical Results

In this section we report some numerical results for the cubic finite element method, where the discrete nonlinear least-squares problem (11) is solved by two optimization algorithms.

The first algorithm is an active set method (cf. [34–36] and [19, Appendix B]) designed for nonlinear smooth optimization problems.

Note that the nonsmooth cost function J_h in (13) can be written explicitly as (cf. Remark 2)

$$\begin{aligned} J_h(v_h) = \frac{h^4}{2}\,\|D_h^2 v_h\|_{L^2(\Omega)}^2 + \frac{1}{2}\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\,\|[\![\partial v_h/\partial n]\!]\|_{L^2(e)}^2 \\ + \frac{1}{2}\Big(\frac{\alpha-1}{2}\Big)^2\Big\|\beta\Delta_h v_h + \sqrt{(\Delta_h v_h)^2 - 4\det D_h^2 v_h} - \widetilde{\psi}\Big\|_{L^2(\Omega)}^2, \end{aligned} \quad (23)$$

where

$$\beta = \frac{\alpha+1}{\alpha-1} \quad \text{and} \quad \widetilde{\psi} = 2\psi/(\alpha-1).$$

Since $(\Delta_h v_h)^2 - 4\det D_h^2 v_h$ is nonnegative, we can apply the active set method to the smooth problem where $\sqrt{(\Delta_h v_h)^2 - 4\det D_h^2 v_h}$ in (23) is replaced by $\sqrt{(\Delta_h v_h)^2 - 4\det D_h^2 v_h + \epsilon}$ for a small ϵ. (We take ϵ to be 10^{-6} in the numerical experiments.

The second algorithm is an ADMM tailored for the discrete minimization problem (11). Details of this ADMM can be found in Appendix A.

We test our finite element method using three examples on $\Omega = (0,1)^2$. For the first two examples, the known exact solutions are smooth and we solve the discrete problems by the active set method with regularization. For the third example, the exact solution is unknown and we solve the discrete problems by both algorithms.

The relative errors of the approximation solution $\widetilde{u}_h$ in various norms are defined by

$$e_{0,h}^r = \frac{\|u-\widetilde{u}_h\|_{L^2(\Omega)}}{\|u\|_{L^2(\Omega)}},\ e_{1,h}^r = \frac{|u-\widetilde{u}_h|_{H^1(\Omega)}}{|u|_{H^1(\Omega)}},\ e_{2,h}^r = \frac{|u-\widetilde{u}_h|_{H^2(\Omega)}}{|u|_{H^2(\Omega)}}, \qquad e_{\infty,h}^r = \frac{\max_{p\in\mathcal{V}_h}|u(p)-\widetilde{u}_h(p)|}{\|u\|_{L^\infty(\Omega)}},$$

where $\mathcal{V}_h$ is the set of all the vertices of the triangulation $\mathcal{T}_h$, and the residual is given by

$$\eta_h(\widetilde{u}_h) = \Big\|\beta\Delta_h\widetilde{u}_h + \sqrt{(\Delta_h\widetilde{u}_h)^2 - 4\det D_h^2\widetilde{u}_h} - \widetilde{\psi}\Big\|_{L^2(\Omega)} + \Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\,\|[\![\partial\widetilde{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}},$$

which is a slight modification of the one in (22).

The numerical experiments were carried out on a MacBook Pro laptop computer with a 2.8GHz Quad-Core Intel Core i7 processor and with 16GB 2133 MHz LPDDR3 memory. We used MATLAB (R2021a v.9.10.0) in our computations.

Example 15. This example is from [11], where $\psi = 0$, $\phi = -\rho^{1-\alpha}$ with

$$\rho = \left[(x+1)^2 + (y+1)^2\right]^{1/2}$$

and $\alpha \in [2,3]$. The exact solution is $u = -\rho^{1-\alpha}$.

We tested the cases $\alpha = 2$ and $\alpha = 3$. In both cases, the exact solutions are smooth. Numerical results for the errors, the error estimator, the value of the cost function and the computational time are reported in Table 1–Table 4 for $\alpha = 2$ and Table 5–Table 8 for $\alpha = 3$.

Our theory predicts that the convergence order of the errors in the energy norm is 2. Overall the convergence results agree with this prediction and the active set algorithm can solve the regularized problem very efficiently (cf. Table 4 and Table 8). By comparing the error estimator $\eta_h(\tilde{u}_h)$ in Table 2 and Table 6 with $e^r_{2,h}$ in Table 1 and Table 5, one can see that the error estimator $\eta_h(\tilde{u}_h)$ is reliable.

As indicated in the tables, the discrete problems are harder to solve when the mesh size is small, which causes some decays of the computed convergence orders.

Table 1. Relative errors versus mesh size h and orders of convergence for Example 15 with $\alpha = 2$.

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
2^0	2.2689e-1	–	4.7301e-2	–	4.9828e-3	–	4.5798e-6	–
2^{-1}	3.1883e-2	2.83	3.0580e-3	3.95	1.9576e-4	4.67	2.4309e-4	-5.73
2^{-2}	3.8655e-3	3.04	1.5444e-4	4.31	1.8540e-5	3.40	2.8734e-5	3.08
2^{-3}	7.7370e-4	2.32	2.0448e-5	2.92	3.3994e-6	2.45	5.1953e-6	2.47
2^{-4}	2.0887e-4	1.89	4.3256e-6	2.24	8.1083e-7	2.07	1.1562e-6	2.17
2^{-5}	7.1042e-5	1.56	1.1062e-6	1.97	2.2113e-7	1.87	3.0758e-7	1.91

Table 2. Value of the error estimator for Example 15 with $\alpha = 2$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$\eta_h(\tilde{u}_h)$	1.5221e-1	2.4366e-2	3.8542e-3	1.0379e-3	3.3148e-4	1.0371e-4
Order	–	2.64	2.66	1.89	1.65	1.68

Table 3. Value of the cost function for Example 15 with $\alpha = 2$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$J_h(\tilde{u}_h)$	2.3935e-2	2.3809e-3	1.5845e-4	1.0128e-5	6.5356e-7	2.9289e-8
Order	–	3.33	3.91	3.97	3.95	4.48

Table 4. CPU time of the active set method for Example 15 with $\alpha = 2$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
CPU time (s)	3.1914e0	1.4785e0	5.2994e-1	6.3343e-1	3.2178e0	1.7387e1

Example 16. In this example, we choose $\alpha = 2$, $\psi = 6$ and $\phi = x^2 + y^2$. The exact solution is $u = x^2 + y^2$.

Table 5. Relative errors versus mesh size h and orders of convergence for Example 15 with $\alpha = 3$.

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
2^0	3.8406e-1	–	1.0392e-1	–	1.4290e-2	–	5.5275e-4	–
2^{-1}	6.3036e-2	2.61	7.0656e-3	3.88	5.8025e-4	4.62	4.8311e-4	0.19
2^{-2}	9.2751e-3	2.76	5.2513e-4	3.75	8.4539e-5	2.78	9.0458e-5	2.42
2^{-3}	2.2242e-3	2.06	1.0001e-4	2.39	2.0319e-5	2.06	2.1078e-5	2.10
2^{-4}	7.0644e-4	1.65	2.2989e-5	2.12	5.2126e-6	1.96	5.1031e-6	2.05
2^{-5}	2.8117e-4	1.33	5.7587e-6	2.00	1.3702e-6	1.93	1.3157e-6	1.96

Table 6. Value of the error estimator for Example 15 with $\alpha = 3$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$\eta_h(\tilde{u}_h)$	2.2278e-1	3.7908e-2	6.4212e-3	1.8748e-3	5.9716e-4	2.0190e-4
Order	–	2.56	2.56	1.78	1.65	1.56

Table 7. Value of the cost function for Example 15 with $\alpha = 3$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$J_h(\tilde{u}_h)$	5.9654e-2	5.6557e-3	3.7854e-4	2.4506e-5	1.6003e-6	1.0859e-7
Order	–	3.40	3.90	3.95	3.94	3.88

Table 8. CPU time of the active set method for Example 15 with $\alpha = 3$.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
CPU time (s)	2.9364e0	1.5116e0	5.2636e-1	6.2461e-1	3.0556e0	1.5558e1

The main purpose here is to test the reliability of the regularization strategy since

$$(\Delta u)^2 - 4 \det D^2 u = 0 \quad \text{in} \quad \Omega.$$

The numerical results are presented in Table 9–Table 12. According to these results, the regularization strategy works well for a small regularization parameter.

Table 9. Relative errors versus mesh size h and orders of convergence for Example 16.

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
2^0	1.5630e-1	–	5.2481e-2	–	1.9371e-2	–	1.1578e-2	–
2^{-1}	7.6790e-2	1.03	2.8312e-2	0.89	1.4077e-2	0.46	1.2540e-2	-0.12
2^{-2}	1.2113e-2	2.66	4.3073e-3	2.72	2.1504e-3	2.71	1.6618e-3	2.92
2^{-3}	2.1715e-3	2.48	7.0744e-4	2.61	3.4824e-4	2.63	2.5854e-4	2.68
2^{-4}	5.4928e-4	1.98	1.7232e-4	2.04	8.4351e-5	2.05	6.0921e-5	2.09
2^{-5}	1.8316e-4	1.58	5.6962e-5	1.60	2.7590e-5	1.61	1.9265e-5	1.66

Table 10. Value of the error estimator for Example 16.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$\eta_h(\widetilde{u}_h)$	9.8952e-1	2.7576e-1	6.0815e-2	1.4328e-2	3.8084e-3	1.0182e-3
Order	–	1.84	2.18	2.09	1.91	1.90

Example 17. This example is also from [11], where $\psi = 0$, $\delta \in (0,1/4)$ and $\phi = \phi_\delta$ is defined as follows: On the edge $\{x = (x_1, x_2) : 0 \le x_1 \le 1,\ x_2 = 0\}$ of Ω:

$$\phi_\delta = \begin{cases} 1 & \text{if } 0 \le x_1 \le 1/4 - \delta, \\ \cos^2\left[1/4(x_1 - 1/4 + \delta)(\pi/\delta)\right] & \text{if } 1/4 - \delta \le x_1 \le 1/4 + \delta, \\ 0 & \text{if } 1/4 + \delta \le x_1 \le 3/4 - \delta, \\ \cos^2\left[1/4(x_1 - 3/4 - \delta)(\pi/\delta)\right] & \text{if } 3/4 - \delta \le x_1 \le 3/4 + \delta, \\ 1 & \text{if } 3/4 + \delta \le x_1 \le 1, \end{cases}$$

and similar definitions on the three other edges. The exact solution of this problem is unknown.

Table 11. Value of the cost function for Example 16.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
$J_h(\widetilde{u}_h)$	3.5978e0	2.5395e-1	1.6816e-2	1.0692e-3	6.8233e-5	4.2705e-6
Order	–	3.82	3.92	3.98	3.97	4.00

Table 12. CPU time of the active set method for Example 16.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}
CPU time (s)	4.8766e0	1.7652e0	1.9735e0	2.8036e0	2.2619e0	1.2670e2

We choose $\delta = 1/16$ and $\alpha = 3$ in our numerical experiments. Since the exact solution of this problem is unknown, we compute the numerical errors by comparing the numerical solutions from consecutive meshes.

The numerical results obtained by the active method are reported in Table 13–Table 16. They indicate first order convergence in the energy norm for small h.

Table 13. Relative errors versus mesh size h and orders of convergence for Example 17 (active set method).

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
2^0	6.3289e-1	–	1.8817e0	–	1.8353e0	–	2.2732e0	–
2^{-1}	6.5455e-1	-0.05	1.0402e0	0.86	7.4838e-1	1.29	9.8616e-1	1.20
2^{-2}	6.5812e-1	-0.01	3.5436e-1	1.55	5.8758e-2	3.67	9.8779e-2	3.32
2^{-3}	5.5569e-1	0.24	1.9872e-1	0.83	1.3194e-1	-1.17	1.6609e-1	-0.75
2^{-4}	2.5840e-1	1.10	5.7753e-2	1.78	4.3745e-2	1.59	5.9671e-2	1.48
2^{-5}	1.4676e-1	0.82	1.5645e-2	1.88	1.1167e-2	1.97	1.5313e-2	1.96

We also solved the discrete problems by an ADMM algorithm (Algorithm A.1 in Appendix A), where we chose $\rho_i = \rho_* \in (0,1)$ in (26). The iteration is stopped when the relative errors of two consecutive iterations are less than or equal to 10^{-6}.

Table 14. Value of the error estimator for Example 17 (active set method).

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
$\eta_h(\tilde{u}_h)$	1.4331e1	5.8551e1	5.1927e1	7.3428e1	3.5648e1	2.1337e1	1.2778e1
Order	–	-2.03	0.17	-0.50	1.04	0.74	0.74

Table 15. Value of the cost function for Example 17 (active set method).

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
$J_h(\tilde{u}_h)$	1.4892e2	1.0550e3	9.7570e2	2.2251e3	5.7678e2	2.1792e2	8.0251e1
Order	–	-2.82	0.11	-1.19	1.95	1.40	1.44

Table 16. CPU time of the active set method for Example 17.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
CPU time (s)	6.9647e-1	2.2712e0	9.9260e-1	4.0610e0	1.5208e1	8.2917e1	5.8819e1

The numerical results are displayed in Table 17–Table 20. They compare favorably with the ones obtained by the active set method. Therefore the ADMM algorithm, which does not require regularization, is also an efficient method for the problem at hand.

Table 17. Relative errors versus mesh size h and orders of convergence for Example 17 (ADMM).

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
2^0	1.0650e0	–	2.5223e0	–	2.2416e0	–	2.3710e0	–
2^{-1}	7.3388e-1	0.54	1.1373e0	1.15	8.0060e-1	1.49	8.5776e-1	1.47
2^{-2}	9.1077e-1	-0.31	5.0596e-1	1.17	1.5866e-1	2.34	2.5905e-1	1.73
2^{-3}	6.8401e-1	0.41	3.9059e-1	0.37	2.4439e-1	-0.62	2.4042e-1	0.11
2^{-4}	2.0324e-1	1.75	7.3586e-2	2.41	5.9711e-2	2.03	8.3211e-2	1.53
2^{-5}	8.6774e-2	1.23	1.3442e-2	2.45	9.5821e-3	2.64	1.3667e-2	2.61

Table 18. Value of the error estimator for Example 17 (ADMM).

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
$\eta_h(\tilde{u}_h)$	1.8941e1	1.0502e2	1.0522e2	8.3472e1	1.7715e1	7.5756e0	3.5026e0
Order	–	-2.47	0.00	0.33	2.24	1.23	1.11

Table 19. Value of the cost function for Example 17 (ADMM).

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
$J^\epsilon_h(\tilde{u}_h)$	1.7906e2	4.5998e3	4.2267e3	2.3878e3	7.9640e1	1.4447e1	3.0829e0
Order	–	-4.68	0.12	0.82	4.91	2.46	2.23

The graphs of the numerical solution computed by the two optimization algorithms with mesh size $h = 2^{-6}$ are depicted in Figure 1. They agree with each other and also with [11, Figure 1(c)].

Table 20. CPU time of ADMM for Example 17.

h	2^0	2^{-1}	2^{-2}	2^{-3}	2^{-4}	2^{-5}	2^{-6}
CPU time (s)	8.6831e-1	1.0133e1	3.3142e0	6.8477e0	5.8732e0	5.8401e1	8.2067e2

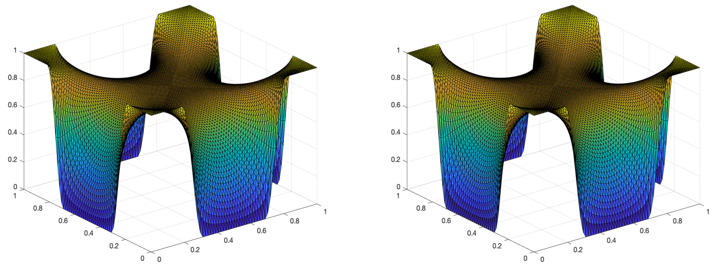

Figure 1. Graphs of solutions at $h = 2^{-6}$: (left) solution by the active set method (right) solution by ADMM.

5. Concluding Remarks

In this paper we designed and analyzed a finite element method for the Dirichlet boundary value problem of a Pucci equation on convex polygonal domains, which is based on a nonlinear least-squares approach. We established the existence and uniqueness of strong solutions and derived a priori and a posteriori error estimates for the finite element approximations. We also constructed an efficient ADMM algorithm for solving the discrete nonsmooth and nonlinear least-squares problems.

Our method can be extended to smooth convex domains and adaptive meshes. It is also interesting and challenging to extend the results to three dimensions. These are some of our ongoing projects.

Appendix A. An ADMM Algorithm

In this appendix we present an ADMM algorithm that can solve the discrete problem (11) efficiently.

The discrete problem (11) is a special case of the following optimization problem:

$$\min_{X\in\mathbb{R}^N} J(X) = \frac{1}{2}X^T AX + \frac{1}{2}\sum_{i=1}^{m} w_i\left(V_i^T X + \sqrt{X^T M_i X} - \Psi_i\right)^2 \tag{24}$$

subject to

$$X_B = b \tag{25}$$

with $X = \binom{X_I}{X_B} \in \mathbb{R}^N$. V_i $(i = 1,\dots,m)$ and Ψ_i are given vectors. We assume that A is symmetric positive definite (SPD) and M_i $(i = 1,\dots,m)$ are symmetric and positive semidefinite.

In order to derive the ADMM algorithm, we first reformulate the general optimization problem (24)–(25) in a separable form.

Since M_i is symmetric and positive semidefinite, there exists a symmetric and positive semidefinite matrix D_i such the that

$$M_i = D_i^T D_i.$$

Let $Z_i = D_i X$. Then we have $\sqrt{X^T M_i X} = \|Z_i\|_2$ and

$$\min_{X\in\mathbb{R}^N} J(X) = \frac{1}{2}X^T AX + \frac{1}{2}\sum_{i=1}^{m} w_i\left(V_i^T X + \|Z_i\|_2 - \Psi_i\right)^2$$

subject to

$$X_B = b \quad \text{and} \quad Z_i = D_i X, \;\; i = 1, 2, \dots, m.$$

Let $\rho = (\rho_1, \dots, \rho_m)$ and

$$\begin{aligned} &L(\rho, \lambda, X, Z) \\ &\quad = \frac{1}{2} X^T A X + \frac{1}{2} \sum_{i=1}^{m} w_i \left(V_i^T X + \|Z_i\|_2 - \Psi_i\right)^2 + \sum_{i=1}^{m} \lambda_i^T (D_i X - Z_i) + \frac{1}{2} \sum_{i=1}^{m} \rho_i \|D_i X - Z_i\|_2^2, \end{aligned} \tag{26}$$

which is the augmented Lagrangian function of the problem related to the constraints $Z_i = D_i X, \;\; i = 1, 2, \dots, m$.

The ADMM algorithm is given as follows:

Algorithm A.1.
1. Given $\rho_i > 0$, $Z_i^{(0)}$ and $\lambda_i^{(0)}$ for $i = 1, 2, \dots, m$.
2. Solve the problem related to X:

$$X^{(k+1)} = \underset{X \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ L\left(\rho, \lambda^{(k)}, X, Z^{(k)}\right) : \;\; X_B = b \right\}. \tag{27}$$

3. Solve the problem related to Z:

$$Z^{(k+1)} = \underset{Z \in \mathbb{R}^{m \times N}}{\operatorname{argmin}} L\left(\rho, \lambda^{(k)}, X^{(k+1)}, Z\right). \tag{28}$$

4. Update λ:

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} + \rho_i \left(D_i X^{(k+1)} - Z_i^{(k+1)}\right) \quad \forall \;\; i = 1, 2, \dots, m.$$

5. Continue until the stopping criterion is satisfied.

Next we consider the solution of the subproblems (27) and (28).

Let

$$A = \begin{pmatrix} A_{II} & A_{IB} \\ A_{BI} & A_{BB} \end{pmatrix} \quad \text{and} \quad X = \begin{pmatrix} X_I \\ X_B \end{pmatrix}.$$

For the first subproblem (27), we need to solve

$$\frac{\partial L\left(\rho, \lambda^{(k)}, X, Z^{(k)}\right)}{\partial X_I} = 0,$$

which is equivalent to

$$H X_I = -F^{(k)},$$

where

$$H = A_{II} + \sum_{i=1}^{m} w_i V_{i,I} V_{i,I}^T + \sum_{i=1}^{m} \rho_i D_{i,I} D_{i,I}^T \tag{29}$$

is SPD, and

$$\begin{aligned} F^{(k)} = A_{IB} X_B + \sum_{i=1}^{m} w_i \left(V_{i,B}^T X_B + \left\| Z_i^{(k)} \right\|_2 - \Psi_i\right) V_{i,I} + \sum_{i=1}^{m} D_{i,I} \lambda_i^{(k)} \\ + \sum_{i=1}^{m} \rho_i \left(D_{i,I} D_{i,B}^T X_B - D_{i,I} Z_i^{(k)}\right). \end{aligned} \tag{30}$$

The second subproblem (28) is equivalent to m independent minimization problems.

$$\operatorname{argmin} \frac{1}{2} w_i \left(V_i^T X^{(k+1)} + \|Z_i\|_2 - \Psi_i\right)^2 + \left[\lambda_i^{(k)}\right]^T \left(D_i X^{(k+1)} - Z_i\right) + \frac{1}{2} \rho_i \left\| D_i X^{(k+1)} - Z_i \right\|_2^2$$

that can be simplified to

$$\operatorname{argmin} \frac{1}{2} (w_i + \rho_i) \|Z_i\|_2^2 + w_i \left(V_i^T X^{(k+1)} - \Psi_i\right) \|Z_i\|_2 - \left(\lambda_i^{(k)} + \rho_i D_i X^{(k+1)}\right)^T Z_i. \tag{31}$$

An explicit solution of (31) is given by

$$Z_i^{(k+1)} = \begin{cases} \dfrac{1}{w_i+\rho_i}\left(1-\dfrac{w_i\left(V_i^T X^{(k+1)}-\Psi_i\right)}{\|\zeta^{(k)}\|_2}\right)\zeta^{(k)} & \text{if } \zeta^{(k)}\neq 0,\\ \bar{Z}_{i,0}^{(k+1)} \in \left\{Z_i:\ \|Z_i\|_2 = \max\left\{0,-\dfrac{w_i\left(V_i^T X^{(k+1)}-\Psi_i\right)}{w_i+\rho_i}\right\}\right\} & \text{if } \zeta^{(k)}=0,\end{cases} \tag{32}$$

where $\zeta^{(k)} = \lambda_i^{(k)} + \rho_i D_i X^{(k+1)}$.

In summary, we have the following results:

- The first subproblem (27) is given by the SPD problem $HX_I^{(k+1)} = -F^{(k)}$ where H (resp., $F^{(k)}$) is given by (29) (resp., (30)).
- The solution of the second subproblem (28) is given by (32).

Note that throughout the iterations, H is unchanged and all the updates in step 3 and step 4 can be implemented in parallel. It is worth pointing out that for our discrete problem the nonzero part of M_i in (24) is at most a 10×10 matrix so that we can obtain D_i easily, and the updates in step 3 and step 4 can be implemented locally. This makes the proposed ADMM algorithm memory efficient.

Remark 18. For convex optimization problems, there is a well-developed convergence theory for ADMM algorithms (cf. [37, 38] and the references therein). On the other hand the convergence results for nonconvex problems are very limited. The ADMM algorithm in this paper, which is designed for a nonconvex and nonsmooth problem, is not covered by the existing theory in the literature. The convergence and efficiency of this algorithm is tested numerically in Example 17 against an active set method that has been thoroughly analyzed.

References

[1] C. Pucci, "Un problema variazionale per i coefficienti di equazioni differenziali di tipo ellittico", *Ann. Sc. Norm. Super. Pisa, Cl. Sci.* **16** (1962), p. 159-172.

[2] C. Pucci, "Operatori ellittici estremanti", *Ann. Mat. Pura Appl.* **72** (1966), p. 141-170.

[3] L. A. Caffarelli, X. Cabré, *Fully Nonlinear Elliptic Equations*, Colloquium Publications, vol. 43, American Mathematical Society, Rhode Island, 1995.

[4] L. A. Caffarelli, R. Glowinski, "Numerical solution of the Dirichlet problem for a Pucci equation in dimension two. Application to homogenization", *J. Numer. Math.* **16** (2008), no. 3, p. 185-216.

[5] V. Quitalo, "A free boundary problem arising from segregation of populations with high competition", *Arch. Ration. Mech. Anal.* **210** (2013), no. 3, p. 857-908.

[6] L. A. Caffarelli, B. Patrizi, V. Quitalo, M. Torres, "Regularity of interfaces for a Pucci type segregation problem", *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **36** (2019), no. 4, p. 939-975.

[7] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Studies in Mathematics and its Applications, vol. 4, North-Holland, 1978.

[8] R. A. Adams, J. J. F. Fournier, *Sobolev Spaces*, second ed., Pure and Applied Mathematics, vol. 140, Academic Press Inc., Amsterdam, 2003.

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

[10] L. C. Evans, *Partial Differential Equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Rhode Island, 2010.

[11] E. J. Dean, R. Glowinski, "On the numerical solution of a two-dimensional Pucci's equation with Dirichlet boundary conditions: a least-squares approach", *C. R. Acad. Sci. Paris, Ser. I* **341** (2005), no. 6, p. 375-380.

[12] O. Lakkis, T. Pryer, "A finite element method for nonlinear elliptic problems", *SIAM J. Sci. Comput.* **35** (2013), no. 4, p. A2025-A2045.

[13] A. M. Oberman, "Wide stencil finite difference schemes for the elliptic Monge–Ampère equation and functions of the eigenvalues of the Hessian", *Discrete Contin. Dyn. Syst.* **10** (2008), no. 1, p. 221-238.

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

[15] B. D. Froese, "Meshfree finite difference approximations for functions of the eigenvalues of the Hessian", *Numer. Math.* **138** (2018), no. 1, p. 75-99.

[16] J. F. Bonnans, G. Bonnet, J.-M. Mirebeau, "Monotone and second order consistent scheme for the two dimensional Pucci equation", in *Numerical mathematics and advanced applications. ENUMATH 2019*, vol. 139, Lecture Notes in Computational Science and Engineering, no. Springer, Cham, 2021, p. 733-742.

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

[18] M. Neilan, A. J. Salgado, W. Zhang, "Numerical analysis of strongly nonlinear PDEs", *Acta Numer.* **26** (2017), p. 137-303.

[19] S. C. Brenner, L.-Y. Sung, Z. Tan, H. Zhang, "A convexity enforcing C^0 interior penalty method for the Monge–Ampère equation on convex polygonal domains", *Numer. Math.* **148** (2021), no. 3, p. 497-524.

[20] R. Glowinski, A. Marrocco, "Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèmes de Dirichlet non linéaires", *Rev. Franc. Automat. Inform. Rech. Operat.* **9** (1975), no. R-2, p. 41-76.

[21] R. Glowinski, "On Alternating Direction Methods of Multipliers: A Historical Perspective, Modeling, simulation and optimization for science and technology", in *Modeling, simulation and optimization for science and technology. Selected contributions based on the presentations at the conferences "Optimization and PDE's with industrial applications"*, Computational Methods in Applied Sciences, vol. 34, Springer, 2014, p. 59-82.

[22] A. Maugeri, D. K. Palagachev, L. G. Softova, *Elliptic and parabolic equations with discontinuous coefficients*, Mathematical Research, vol. 109, John Wiley & Sons, Berlin, 2000.

[23] I. Smears, E. Süli, "Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordès coefficients", *SIAM J. Numer. Anal.* **51** (2013), no. 4, p. 2088-2106.

[24] R. A. Horn, C. R. Johnson, *Matrix Analysis*, second ed., Cambridge University Press, 2013.

[25] C. Miranda, "Su di una particolare equazione ellittica del secondo ordine a coefficienti discontinui", *An. Sti. Univ. Al. I. Cuza Iasi, N. Ser.* **Sect. Ia** (1965), no. 11B, p. 209-215.

[26] G. Talenti, "Sopra una classe di equazioni ellittiche a coefficienti misurabili", *Ann. Mat. Pura Appl.* **69** (1965), p. 285-304.

[27] P. Grisvard, *Elliptic Problems in Non Smooth Domains*, Monographs and Studies in Mathematics, vol. 24, Pitman, Boston, 1985.

[28] S. Campanato, "On the condition of nearness between operators", *Ann. Mat. Pura Appl.* **167** (1994), p. 243-256.

[29] S. Koike, A. Święch, "Existence of strong solutions of Pucci extremal equations with superlinear growth in Du", *Fixed Point Theory Appl.* **5** (2009), p. 291-304.

[30] B. Sirakov, "Solvability of uniformly elliptic fully nonlinear PDE", *Arch. Ration. Mech. Anal.* **195** (2010), no. 2, p. 579-607.

[31] S. C. Brenner, K. Wang, J. Zhao, "Poincaré–Friedrichs inequalities for piecewise H^2 functions", *Numer. Funct. Anal. Optim.* **25** (2004), no. 5-6, p. 463-478.

[32] M. Neilan, M. Wu, "Discrete Miranda–Talenti estimates and applications to linear and nonlinear PDEs", *J. Comput. Appl. Math.* **356** (2019), p. 358-376.

[33] S. C. Brenner, M. Neilan, A. Reiser, L.-Y. Sung, "A C^0 interior penalty method for a von Kármán plate", *Numer. Math.* **135** (2017), no. 3, p. 803-832.

[34] W. W. Hager, H. Zhang, "A new active set algorithm for box constrained optimization", *SIAM J. Optim.* **17** (2006), no. 2, p. 526-557.

[35] W. W. Hager, H. Zhang, "The limited memory conjugate gradient method", *SIAM J. Optim.* **23** (2013), no. 4, p. 2150-2168.

[36] W. W. Hager, H. Zhang, "An active set algorithm for nonlinear optimization with polyhedral constraints", *Sci. China, Math.* **59** (2016), no. 8, p. 1525-1542.

[37] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers", *Found. Trends Mach. Learn.* **3** (2011), no. 1, p. 1-122.

[38] B. He, X. Yuan, "On the $O(1/n)$ convergence rate of the Douglas–Rachford alternating direction method", *SIAM J. Numer. Anal.* **50** (2012), no. 2, p. 700-709.