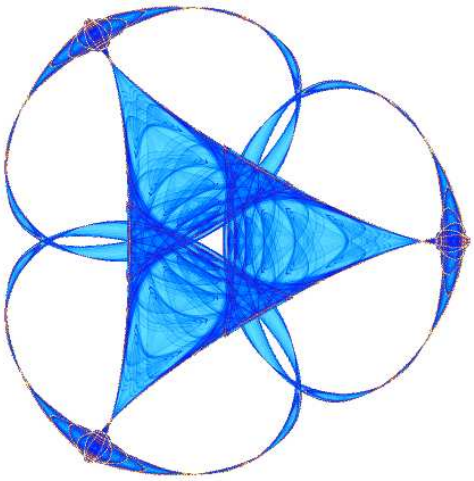
PSEUDO TIME CONTINUATION AND TIME MARCHING METHODS FOR MONGE-AMPÈRE TYPE EQUATIONS

By

Gerard Awanou

IMA Preprint Series # 2350

(November 2010)



INSTITUTE FOR MATHEMATICS AND ITS APPLICATIONS

UNIVERSITY OF MINNESOTA
400 Lind Hall
207 Church Street S.E.
Minneapolis, Minnesota 55455–0436

Phone: 612-624-6066 Fax: 612-626-7370
URL: http://www.ima.umn.edu

PSEUDO TIME CONTINUATION AND TIME MARCHING METHODS FOR MONGE-AMPÈRE TYPE EQUATIONS

GERARD AWANOU

ABSTRACT. We discuss the performance of three numerical methods for the fully nonlinear Monge-Ampère equation. The first two are pseudo time continuation methods while the third is a pure pseudo time marching algorithm. The pseudo time continuation methods are shown to converge for smooth data on a uniformly convex domain. We give numerical evidence that they perform well for the non-degenerate Monge-Ampère equation. The pseudo time marching method applies in principle to any nonlinear equation. Numerical results with this approach for the degenerate Monge-Ampère equation are given as well as for the Pucci and Gauss-curvature equations.

1. INTRODUCTION

We are interested in numerical solutions of equations of type

$$F(x, u(x), Du(x), D^2u(x)) = 0, \tag{1.1}$$

on a bounded domain Ω of $\mathbb{R}^n$ with boundary $\partial\Omega$ and Dirichlet boundary conditions $u = g$ and F real valued. Here u is a real valued function and $Du(x), D^2u(x)$ denote its gradient vector and Hessian matrix respectively.

The fully nonlinear Monge-Ampère equation is given by

$$\det D^2u = f \text{ in } \Omega, \quad u = g \text{ on } \partial\Omega, \tag{1.2}$$

where f, g are given functions with $f > 0$ in the non degenerate case, otherwise we assume that $f \geq 0$. We will also assume that $f \in C(\Omega)$ and $g \in C(\partial\Omega)$.

For a symmetric matrix M with eigenvalues $e_1, \ldots, e_n$, we recall the Pucci extremal operators

$$\mathcal{M}^+[M] = \Lambda \sum_{e_i>0} e_i + \lambda \sum_{e_i<0} e_i, \mathcal{M}^-[M] = \lambda \sum_{e_i>0} e_i + \Lambda \sum_{e_i<0} e_i.$$

One can then consider the Pucci equations

$$\mathcal{M}^+[D^2u(x)] = f(x), \mathcal{M}^-[D^2u(x)] = f(x).$$

The Pucci equations appear in stochastic control where the control variable is the diffusion coefficient. They also play an important role in the theory of fully nonlinear equations. If u is a solution of $F(x, D^2u(x)) = 0$ with F uniformly elliptic [12], then u is a subsolution and a supersolution of equations which do not depend on F but on λ, Λ and f. Any result valid for these classes of equations is also valid for any fully nonlinear uniformly elliptic equation, [12] p. 16. Let us consider the Pucci equation

$$\mathcal{M}^+[D^2u(x)] = f(x),$$

and put $\alpha = \Lambda/\lambda \geq 1$ and denote by $\lambda^+(M), \lambda^-(M)$ the maximum and minimum eigenvalues of M respectively. As there will be no ambiguity, we will write λ^+, λ^- for $\lambda^+(D^2u)$ and $\lambda^-(D^2u)$ respectively. In two dimensions, following [14, 13, 11, 22, 18], we will consider the Pucci equation,

$$\alpha\lambda^+ + \lambda^- = f(x),\ \alpha \geq 1. \tag{1.3}$$

We note that in principle the vanishing moment methodology [16] is applicable to the Pucci equation. However, the operator F is not differentiable in this case as opposed to the Monge-Ampère equation for example, hence Newton's method cannot be used. One can use that methodology following [14, 13, 11], where a Pucci equation is written in terms of the Monge-Ampère operator.

Another equation of interest is the prescribed Gauss curvature equation,

$$\det D^2u - K(x)(1 + |Du|^2)^{\frac{n+2}{2}} = 0. \tag{1.4}$$

Given Ω, K and g, one seeks a convex function solution of the equation.

Starting with [9, 13], interest has grown for finite element methods which are able to capture viscosity solutions of second order fully nonlinear equations. In the context of non-smooth solutions, for proven convergence results, wide stencils finite difference have been used for the Monge-Ampère and the Pucci equations [22] and a class of Bellman-Isaacs equation [15]. This paper is the sequel to [5] where we did a comparative study of three methods suitable for finite element computations of solutions of the Monge-Ampère equation. We refer to the above paper for recent references and an introduction to the spline element method which is also used here to illustrate the numerical results.

In this paper, we consider three numerical methods for the Monge-Ampère equation. The first two are pseudo-time continuation methods. Given a fully nonlinear elliptic equation, $F(u) = 0$ with F differentiable, we consider the sequence of problems

$$(\nu L + F'(u_k))(u_{k+1} - u_k) = -F(u_k),$$

where L is a linear operator which we take in this paper as the identity or the Laplace operator. In the latter case, L may be viewed as a preconditioner and in this case, the computations are speeded up compared to the case where L is the identity operator. In the case of the Monge-Ampère equation, $F(u) = \det D^2u - f$, and the solution of the problem is reduced to solving a sequence of elliptic equations whose solutions are sought here in the spline space of degree d and smoothness $r \geq 1$, Section 2. Note that for the Monge-Ampère equation

$$F'(u_k)(u_{k+1} - u_k) = \operatorname{div}\big((\operatorname{cof} D^2u_k)(Du_{k+1} - Du_k) = (\operatorname{cof} D^2u_k) : (D^2u_{k+1} - D^2u_k).$$

Given an initial guess u_0, we are led to the sequence of approximate solutions

$$\nu L\theta_k + (\operatorname{cof} D^2u_k) : D^2\theta_k = (f - f_k), \quad f_k = \det D^2u_k, \quad \theta_k = u_{k+1} - u_k. \tag{1.5}$$

As initial guess one may take the solution of the Poisson equation $\Delta u = 2\sqrt{f}$ in $\Omega, u = g$ on $\partial\Omega$.

Method 1 corresponds to the case where L is the identity operator and Method 2 to the case where L is the Laplace operator. The third method is a pseudo time marching method. Given $\nu > 0$, we consider the sequence of iterates

$$-\nu\Delta u_{k+1} = -\nu\Delta u_k + F(x, u_k(x), Du_k(x), D^2u_k),\ u_{k+1} = g \text{ on } \partial\Omega. \tag{1.6}$$

It can be interpreted as an Euler discretization of the pseudo-time dependent equation $\frac{\partial\Delta u}{\partial t} + F(x, u(x), Du(x), D^2u) = 0$, or as a Laplacian preconditioner of a simple pseudo time marching algorithm, [18] $u_{k+1} = u_k + \Delta^{-1}(F(x, u_k(x), Du_k(x), D^2u_k(x)))$. See also a remark in [21]. The simple pseudo-time marching algorithm also performs well for numerical solutions in some cases for ν sufficiently large. However the use of the Laplacian preconditioner besides the speed of computation, also helps select a convex solution for the two dimensional Monge-Ampère equation. Indeed, we will see that Method 3 and to some extent Method 2 enforces $\Delta u > 0$, which when combined with $\det D^2u = f \geq 0$ and C^1 continuity implies numerical convexity [20], Lemma 1. The use of the spline element method is also motivated by its higher order of accuracy and its robustness in some limiting situations [4, 5].

It is believed that some methods are able to capture the viscosity solutions and other can't. This paper gives evidence that at least in two dimensions, by relaxation, a method which works only for solutions in $H^2(\Omega)$ can be expected to perform in the non smooth case for a class of problems which include the Monge-Ampère, Pucci, Gauss curvature and a class of Isaacs-Bellman equation.

The paper is organized as follows: we first give a brief description of the spline element method, then we give the convergence results for the iterative methods. Next, we give numerical results for the Monge-Ampère, Pucci and Gauss curvature equation and conclude with some remarks.

2. Spline element discretization

We refer to [2, 6, 7, 8, 19, 4] for a description of the spline element method. In the simplest case, it can be described as follows. Let $u \in V = H_0^m(\Omega), m \geq 1$ solve a variational problem $a(u, v) = f(v)$ with the conditions of the Lax-Milgram lemma satisfied. Take

$$V_h := S_d^r(\mathcal{T}) = \{s \in C^r(\Omega),\ s|_t \in \mathcal{P}_d,\ \forall t \in \mathcal{T}\},$$

for a triangulation $\mathcal{T}$ of the domain and $\mathcal{P}_d$ the space of polynomials of degree less than or equal to d. The space $S_d^r(\mathcal{T})$ is the spline space of smoothness r and degree d. For $r = 0$ and $d = 1$ we have the space of piecewise linear continuous functions. Next choose a representation of polynomials such that $V_h = \{c \in \mathbb{R}^N, Rc = G\}$ for some integer N, matrix R and vector G. The discrete problem is find $c \in V_h, c^TKd = F^Td$ for all $d \in V_h$ for a suitable stiffness matrix K and a load vector F. Introducing a Lagrange multiplier, we are lead to saddle point problems

$$\begin{pmatrix} K & R^T \\ R & 0 \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ \lambda \end{pmatrix} = \begin{bmatrix} F \\ G \end{bmatrix},$$

which are solved by a version of the augmented Lagrangian algorithm

$$(K + \frac{1}{\mu}R^TR)c^{(l+1)} = K^Tc^{(l)} + \frac{1}{\mu}R^TG, \quad l = 1, 2, \ldots \tag{2.1}$$

The convergence properties of the iterative method were given in [3].

3. Convergence of the iterative methods

We first describe the convergence properties of the pseudo-time continuation methods. We consider a damped version of (1.5), namely

$$\begin{aligned} &(\nu L + F'(u_k))(u_{k+1} - u_k) = -\frac{1}{\tau}F(u_k), \\ &\nu L\theta_k + (\text{cof } D^2u_k) : D^2\theta_k = \frac{1}{\tau}(f - f_k), \quad f_k = \det D^2u_k, \quad \theta_k = u_{k+1} - u_k, \end{aligned} \tag{3.1}$$

where $\tau > 0$ is a damping parameter. In the numerical experiments, we used $\tau = 1$. We will need the following global regularity result, [23].

Theorem 3.1. *Let Ω be a uniformly convex domain in R^n, with boundary in C^3. Suppose $g \in C^3(\overline{\Omega})$, inf $f > 0$, and $f \in C^\alpha$ for some $\alpha \in (0,1)$. Then (1.2) has a convex solution u which satisfies the a priori estimate*

$$||u||_{C^{2,\alpha}(\overline{\Omega})} \leq C,$$

where C depends only on n, α, inf $f, \Omega, ||f||_{C^\alpha(\overline{\Omega})}$ and $||g||_{C^3}$.

According to [23], all assumptions in the above theorem are sharp. We have the following analogue of Theorem 2.1 in [21].

Theorem 3.2. *Let Ω be a uniformly convex domain in R^n, with boundary in C^3. Let $0 < m \leq f \leq M, f \in C^\alpha$ for some $m, M > 0$ and $\alpha \in (0,1)$. Assume also that $g \in C^3(\overline{\Omega})$. Then, there exists $\tau \geq 1$ depending on m, f, such that if u_k is the sequence defined by (1.5), it converges in $C^{2,\beta}$ to a solution u of (1.2) for u_0 sufficiently close to u and for every $\beta < \alpha$.*

Proof. We outlined the proof in [5]. See also [21, 17]. One shows by induction that there are constants $C_1, C_2 > 0$ such that

$$\frac{1}{C_1}\det D^2u \leq \det D^2u_n \leq C_1 \det D^2u \text{ and } ||\det D^2u - \det D^2u_n||_{C^\alpha} \leq C_2.$$

Then using Schauder theory, one obtains bounds on the solution of (1.5). Arzela-Ascoli's theorem is then used to prove that the sequence is precompact in $C^{2,\beta}, \beta < \alpha$. Since (1.2) has at most two solutions, by requiring u_0 sufficiently close to u, we assure that the solution is locally unique. □

We also gave an error estimate in [5] for conforming approximations. Recall that

$$\det D^2u = \frac{1}{n}(\text{cof } D^2u) : D^2u = \frac{1}{n}\,\text{div}\big((\text{cof } D^2u)Du\big), \tag{3.2}$$

where $\text{cof}\, A$ is the matrix of cofactors of the matrix A. This gives the following weak formulation, [5], of the Dirichlet problem for the Monge-Ampère equation: find $u \in H^n(\Omega)$, $u = g$ on $\partial\Omega$ such that

$$\int_\Omega (\text{cof } D^2u)Du \cdot Dw \; dx = -n\int_\Omega fw \; dx, \quad \forall w \in H^n(\Omega) \cap H_0^1(\Omega). \tag{3.3}$$

For V^h a conforming finite element subspace of $H^2(\Omega)$, V_0^h be a conforming finite element subspace of $H^2(\Omega) \cap H_0^1(\Omega)$ with approximation properties

$$\inf_{v_h \in V^h} ||v - v_h||_j \leq C_1 h^{p-j} ||v||_4, j = 0, 1, 2 \tag{3.4}$$

for all $v \in H^4(\Omega)$ for some $p \geq 4$, we have

Theorem 3.3. [5] *Assume that problem* (1.2) *has a solution* $u \in H^4(\Omega)$ *hence in* $C^2(\Omega)$ *by Sobolev embedding, then the discrete problem, find* $u_h \in V_h$, $u_h = g$ *on* $\partial\Omega$ *such that*

$$-\frac{1}{2}\int_\Omega (cof\, D^2 u_h) Du_h \cdot Dw_h \ dx = \int_\Omega f w_h \ dx, \quad \forall w_h \in V_0^h, \tag{3.5}$$

has a unique solution in a neighborhood of $I_h u$ *where* I_h *is an interpolation operator associated with* V_h. *Moreover* $||u_h - I_h(u)||_2$ *is at least O(h).*

Next, we give a convergence result for Method 3.

Theorem 3.4. *Let* $u_0 \in C^{2,\alpha}(\Omega)$ *such that* $\Delta u_0 > 0$ *and assume that*

$$F(x, u(x), Du(x), D^2 u(x)) - f(x) \in C^\alpha, \text{ for } u \in C^{2,\alpha}.$$

Assume moreover that $F(x, u_0(x), Du_0(x), D^2 u_0(x)) \leq f(x)$. *Then, there is an increasing sequence* ν_{k+1} *such that the sequence defined by*

$$-\nu_{k+1}\Delta u_{k+1} = -\nu_{k+1}\Delta u_k + F(x, u_k(x), Du_k(x), D^2 u_k) - f(x), \ u_{k+1} = g \text{ on } \partial\Omega. \tag{3.6}$$

is monotonically decreasing.

Proof. We have $-\nu_1 \Delta u_1 = -\nu_1 \Delta u_0 + F(x, u_0(x), Du_0(x), D^2 u_0(x)) - f$, so $-\nu_1 \Delta(u_1 - u_0) \leq 0$. By the maximum principle, since $u_1 - u_0 \in C^0(\overline{\Omega}) \cap C^{2,\alpha}(\Omega)$,

$$u_1 \leq u_0.$$

We next show that for sufficiently large ν_{k+1}, $u_{k+1} \leq u_k$. As $\Delta u_0 > 0$ and $-\nu_1(\Delta u_1 - \Delta u_0) \leq 0$, we have $\Delta u_1 \geq \Delta u_0 > 0$. Assume by induction that ν_k has been chosen such that $\Delta u_k > 0$. Note that (3.6) can be written

$$\begin{aligned} \Delta u_{k+1} &= \Delta u_k - \frac{1}{\nu_{k+1}}\big(F(x, u_k(x), Du_k(x), D^2 u_k) - f(x)\big), \text{ in } \Omega, \\ u_{k+1} &= g \text{ on } \partial\Omega. \end{aligned} \tag{3.7}$$

If $F(x, u_k(x), Du_k(x), D^2 u_k) - f(x) \leq 0$ for all $x \in \Omega$, then $\Delta u_{k+1} \geq \Delta u_k$ and $u_{k+1} \leq u_k$. In general, let $\Gamma = \{x \in \Omega, F(x, u_k(x), Du_k(x), D^2 u_k) - f(x) > 0\}$. Note that the right hand side of the first equation in (3.7) is of the type $A - 1/\nu B$ where A and B are both positive continuous functions hence bounded on $\overline{\Gamma}$, that is there are numbers m_A, M_A, m_B and M_B such that $m_A \leq A(x) \leq M_A$ and $m_B \leq B(x) \leq M_B$ for all x in Ω. Since $\Delta u_k \geq \Delta u_{k-1} \geq \ldots \geq \Delta u_0 > f$, we have $m_A > 0$.

Now, for $\nu \geq \overline{\nu}_{k+1} = M_B/m_A$, $A - 1/\nu B \geq 0$ and we conclude that for $\nu_{k+1} = \max(\nu_k, \overline{\nu}_{k+1})$, $\Delta u_{k+1} \geq \Delta u_k$. This concludes the proof. □

It is not clear whether the sequences ν_k and u_k are bounded from above and below respectively. The assumption $F(x, u_0(x), Du_0(x), D^2 u_0(x)) \leq f(x)$ can be relaxed by choosing ν sufficiently large so that $\Delta u_1 \geq \Delta u_0$.

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d=3$	8	$1.0610\ 10^{-3}$	$1.1101\ 10^{-2}$	$1.6383\ 10^{-1}$
$d=4$	10	$3.5127\ 10^{-5}$	$4.8553\ 10^{-4}$	$9.0596\ 10^{-3}$
$d=5$	13	$4.1572\ 10^{-6}$	$6.5142\ 10^{-5}$	$1.9364\ 10^{-3}$
$d=6$	25	$1.9684\ 10^{-7}$	$3.6401\ 10^{-6}$	$1.4774\ 10^{-4}$
$d=7$	40	$2.2712\ 10^{-8}$	$4.1495\ 10^{-7}$	$2.2424\ 10^{-5}$
$d=8$	157	$1.3867\ 10^{-9}$	$3.1763\ 10^{-8}$	$2.2274\ 10^{-6}$
$d=9$	303	$3.7149\ 10^{-9}$	$1.5998\ 10^{-7}$	$1.1272\ 10^{-5}$

TABLE 1. Test 1, Method 1, $h = 1/2, \nu = 0.015$

4. Numerical results

Unless otherwise indicated, all numerical simulations below are for $r = 1$ and the domain is $[0,1]^n, n = 2, 3$. For $n = 2$, the computational domain is the unit square $[0,1]^2$ which is first divided into squares of side length h. Then each square is divided into two triangles by the diagonal with negative slope. For $n = 3$, the initial tetrahedral partition $\mathcal{T}_1$ consists in six tetrahedra. Each tetrahedron is then uniformly refined into 8 subtetrahedra forming $\mathcal{T}_k, k = 2, 3$. In the tables, n_{int} denote the number of iterations.

4.1. **Monge-Ampère.** We used the following test functions suggested in [13, 10, 16].

Test 1: $u(x,y) = e^{(x^2+y^2)/2}$ so that $f(x,y) = (1 + x^2 + y^2)e^{(x^2+y^2)}$ and $g(x,y) = e^{(x^2+y^2)/2}$ on $\partial\Omega$.

Test 2: $u(x,y) = -\sqrt{2 - x^2 - y^2}$ so that $f(x,y) = 2/(2 - x^2 - y^2)^2$ and $g(x,y) = -\sqrt{2 - x^2 - y^2}$ on $\partial\Omega$.

Test 3: $f(x,y) = 1$ and $g(x,y) = 0$. No exact solution is known.

Test 4: $u(x,y,z) = e^{(x^2+y^2+z^2)/3}$ so that $f(x,y,z) = 8/81(3+2(x^2+y^2+z^2)e^{(x^2+y^2+z^2)}$ and $g(x,y,z) = e^{(x^2+y^2+z^2)/3}$ on $\partial\Omega$.

Test 5: $u(x,y,z) = -\sqrt{3 - x^2 - y^2 - z^2}$ so that $f(x,y,z) = 3(3-x^2-y^2-z^2)^(-5/2)$ and $g(x,y,z) = -\sqrt{3 - x^2 - y^2 - z^2}$ on $\partial\Omega$.

Test 6: $f(x,y,z) = 1$ and $g(x,y,z) = 0$. No exact solution is known.

Test 7: $u(x,y) = |x - 1/2|$ with $g(x,y) = |x - 1/2|$ and $f(x,y) = 0$.

For non-smooth solutions, better results are obtained with L as the Laplace operator, Method 2. Even for smooth solutions, the computations are much faster. For example, for Test 6, one observes a negative curvature of the numerical solution near the corners. This a problem with many numerical simulations of the Monge-Ampère equation, [13, 10, 16]. It is evidenced only when one plots the graph of the solution along the line $y = x$ in $\mathbb{R}^2$. This problem disappears when we use $L = \Delta$ in two dimensions. This can be explained as follows: for plane problems, the scheme enforces element by element $\Delta u > 0$ and for a non degenerate problem $\det D^2 u = f > 0$. This implies that the numerical solution has Hessian positive definite element by element, which when combined with C^1 continuity gives numerical convexity [20], Lemma 1.

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d=3$	10	$1.2809\ 10^{-4}$	$2.6554\ 10^{-3}$	$8.9587\ 10^{-2}$
$d=4$	12	$1.6278\ 10^{-6}$	$4.5619\ 10^{-5}$	$1.7395\ 10^{-3}$
$d=5$	16	$1.1531\ 10^{-7}$	$2.3916\ 10^{-6}$	$1.3444\ 10^{-4}$
$d=6$	24	$1.7609\ 10^{-9}$	$6.8523\ 10^{-8}$	$5.5427\ 10^{-6}$
$d=7$	39	$1.7113\ 10^{-10}$	$4.9437\ 10^{-9}$	$5.3920\ 10^{-7}$
$d=8$	111	$3.9851\ 10^{-10}$	$3.1860\ 10^{-8}$	$4.6123\ 10^{-6}$

TABLE 2. Test 1, Method 1, $h = 1/4, \nu = 0.015$

h	n_{it}	L^2 norm	H^1 norm
$1/2^0$	435	2.195410^{-2}	1.640910^{-1}
$1/2^1$	352	3.609710^{-3}	6.140510^{-2}
$1/2^2$	345	1.068510^{-3}	4.097810^{-2}
$1/2^3$	319	3.766610^{-4}	2.947810^{-2}

TABLE 3. Test 2: Method 1, $d = 3, \nu = 15$, $u(x,y) = -\sqrt{2 - x^2 - y^2} \notin H^2(\Omega)$

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d=3$	1	$1.2338\ 10^{-2}$	$7.6984\ 10^{-2}$	$4.4411\ 10^{-1}$
$d=4$	7	$1.6289\ 10^{-3}$	$1.4719\ 10^{-2}$	$1.3983\ 10^{-1}$
$d=5$	10	$1.5333\ 10^{-3}$	$8.7312\ 10^{-3}$	$6.0412\ 10^{-2}$
$d=6$	18	$1.2324\ 10^{-4}$	$9.7171\ 10^{-4}$	$1.0584\ 10^{-2}$

TABLE 4. Test 4, Method 1, $\mathcal{I}_1, \nu = 0.015$

d	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d=3$	1	$3.1739\ 10^{-3}$	$2.3005\ 10^{-2}$	$2.4496\ 10^{-1}$
$d=4$	14	$3.2786\ 10^{-4}$	$3.5626\ 10^{-3}$	$5.2079\ 10^{-2}$
$d=5$	39	$2.4027\ 10^{-5}$	$3.9210\ 10^{-4}$	$8.8868\ 10^{-3}$

TABLE 5. Test 4, Method 1, $\mathcal{I}_2, \nu = 0.015$

There are three instances where Method 3 outperforms the pseudo-time continuation methods. The concavity property of the concave solution in the case $f(x,y) = 1, g(x,y) = 0$ are better than the one obtained by the vanishing moment methodology, [5, 16]. The second example is its performance for Test 2. And finally, we give results for Test 7, a degenerate Monge-Ampère equation. We obtained similar results for the three dimensional analogue of Test 7.

We note that for smooth solutions, if ν is not high enough, numerical results with Method 3 may be inaccurate. This explains the loss of accuracy for $h = 1/2^6$. However, this is best evidenced by the performance of the method on a smooth solution, Tables 7, 8 and 9. As the number of degrees of freedom increases, it becomes expensive to increase the value of ν.

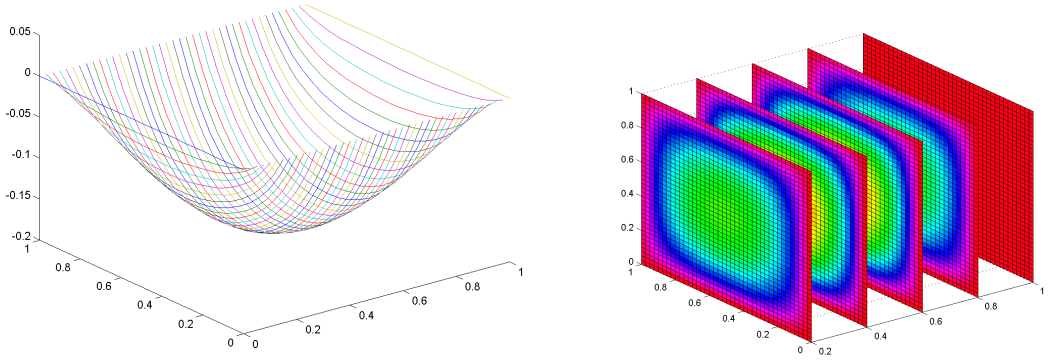


FIGURE 1. Test 6, Method 1, on $\mathcal{I}_3$, $d = 5$, $\nu = 15$ graph $x = 1/2$, slices x-direction

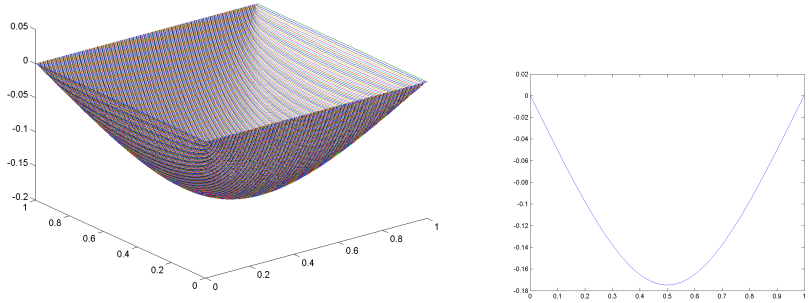


FIGURE 2. Test 3, $L = \Delta$, $h = 1/2^4, d = 5, \nu = 15$.

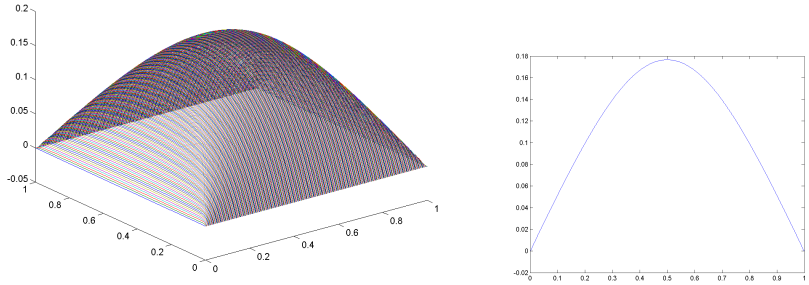


FIGURE 3. Test 3, Method 3, $h = 1/2^4, d = 5, \nu = 50$.

h	n_{it}	L^2 norm	H^1 norm
$1/2^1$	826	$2.0721\ 10^{-2}$	$1.5963\ 10^{-1}$
$1/2^2$	652	$1.8579\ 10^{-3}$	$5.4300\ 10^{-2}$
$1/2^3$	644	$5.0438\ 10^{-4}$	$3.7948\ 10^{-2}$
$1/2^4$	588	$2.1132\ 10^{-4}$	$2.8015\ 10^{-2}$
$1/2^5$	547	$8.4871\ 10^{-5}$	$2.0803\ 10^{-2}$
$1/2^6$	2	$4.4597\ 10^{-3}$	$9.9622\ 10^{-2}$
$1/2^7$	1	NaN	NaN

TABLE 6. Test 2, Method 3, $d = 3, \nu = 50$

h	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d = 3$	24	1.061010^{-3}	1.110110^{-2}	1.638310^{-1}
$d = 4$	24	3.512710^{-5}	4.855310^{-4}	9.059610^{-3}
$d = 5$	40	4.157210^{-6}	6.514210^{-5}	1.936410^{-3}
$d = 6$	2	6.205810^{-4}	7.523610^{-3}	2.128110^{-1}
$d = 7$	2	6.340810^{-4}	8.329910^{-3}	2.669910^{-1}
$d = 8$	2	6.398810^{-4}	8.763610^{-3}	3.060410^{-1}

TABLE 7. Test 1, Method 3, $h = 1/2, \nu = 5$

h	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d = 3$	260	$1.0610\ 10^{-3}$	$1.1101\ 10^{-2}$	$1.6383\ 10^{-1}$
$d = 4$	234	$3.5127\ 10^{-5}$	$4.8553\ 10^{-4}$	$9.0596\ 10^{-3}$
$d = 5$	236	$4.1569\ 10^{-6}$	$6.5142\ 10^{-5}$	$1.9364\ 10^{-3}$
$d = 6$	217	$1.9780\ 10^{-7}$	$3.6411\ 10^{-6}$	$1.4775\ 10^{-4}$
$d = 7$	213	$2.1441\ 10^{-8}$	$4.1381\ 10^{-7}$	$2.2415\ 10^{-5}$
$d = 8$	186	$1.0893\ 10^{-8}$	$6.3270\ 10^{-8}$	$1.6264\ 10^{-6}$

TABLE 8. Test 1, Method 3, $h = 1/2, \nu = 50$

h	n_{it}	L^2 norm	H^1 norm	H^2 norm
$d = 3$	239	$1.2809\ 10^{-4}$	$2.6554\ 10^{-3}$	$8.9587\ 10^{-2}$
$d = 4$	233	$1.6279\ 10^{-6}$	$4.5620\ 10^{-5}$	$1.7395\ 10^{-3}$
$d = 5$	233	$1.1504\ 10^{-7}$	$2.3915\ 10^{-6}$	$1.3444\ 10^{-4}$
$d = 6$	230	$2.1643\ 10^{-9}$	$6.8741\ 10^{-8}$	$5.5412\ 10^{-6}$
$d = 7$	205	$2.9193\ 10^{-9}$	$1.8321\ 10^{-8}$	$4.1951\ 10^{-7}$
$d = 8$	179	$1.5311\ 10^{-8}$	$8.4643\ 10^{-8}$	$6.7135\ 10^{-7}$

TABLE 9. Test 1, Method 3, $h = 1/4, \nu = 50$

Note carefully that the monotone scheme in [22] has a directional error when a fixed stencil is used and has not been tested in the regime where the accuracy is of order 10^{-5} for the singular example of Test 2. Also, the condition number in (2.1) can be

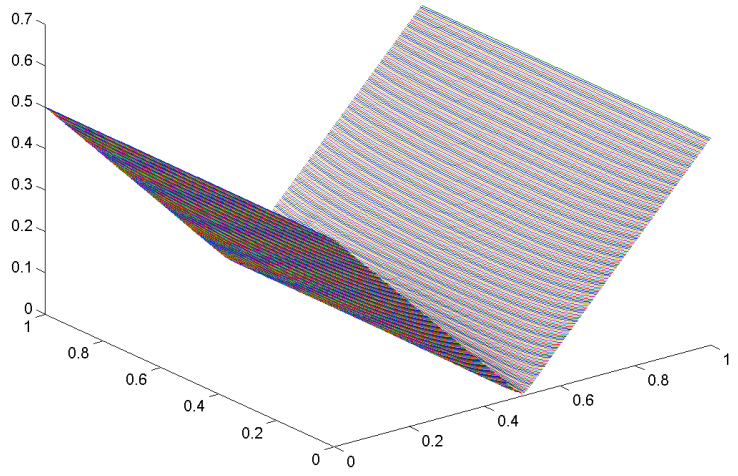


FIGURE 4. Test 7, Method 3, $h = 1/2^4, d = 5, \nu = 50$.

very large and we were not able to compute a numerical solution for $h = 1/2^7$. A factor to take into account is that f is singular and for polynomial interpolation of f, values of f near the singularity point $(1, 1)$ have to be used which inevitably causes overflow. We set $f(1, 1) = 500$ for this test.

For Test 7, the scheme is able again to capture the singularity.

Our schemes did not converge for Test 5. This can be explained as follows: in three dimensions, for a matrix A, $\mathrm{tr}(A) \geq 0$ and $\det A \geq 0$ is not enough to characterize a semi-positive definite Hessian. Using the vanishing moment methodology in the framework of the spline element method, we also observe divergence for Test 5. The paper [16] did not report on that case.

4.2. **Pucci equation.** We use Method 3. Since there is no convexity requirement we could have simply used C^0 splines or Lagrange elements. The zero function was taken as initial guess. Recall, (1.3) that $\alpha \geq 1$. We consider two types of test functions, a smooth one and a problem for which no exact solution is known but cannot be in $H^2(\Omega)$.

Test 8: An exact radial solution, $-((x+1)^2 + (y+1)^2)^{1/2-\alpha/2}$

Test 9: The solution is not in $H^2(\Omega)$ and no exact formula is known. The function $g(x_1, x_2)$ is 1 except on the set $\{(x_1, x_2) \in [0, 1]^2, x_1 = 0, 1, 1/4 < x_2 < 3/4\} \cup \{(x_1, x_2) \in [0, 1]^2, x_2 = 0, 1, 1/4 < x_1 < 3/4\}$ where it is 0.

The structure of Table 10 is similar to Table 4, p. 14 in [22]. The scheme used here is more accurate by several order of magnitude and is conceptually simpler. There does not seem to be a straightforward connection between the number of iterations and the value of ν. For $\alpha = 2, h = 1/2^6$, the number of iterations needed to produce

	h				
α	$1/2^4$	$1/2^5$	$1/2^6$	$1/2^7$	$1/2^8$
2	$1.8717\ 10^{-5}$	$4.6818\ 10^{-6}$	$1.1715\ 10^{-6}$	$2.9291\ 10^{-7}$	$7.3233\ 10^{-8}$
2.5	$3.2042\ 10^{-5}$	$8.0433\ 10^{-6}$	$2.0108\ 10^{-6}$	$5.0288\ 10^{-7}$	$1.2572\ 10^{-7}$
3	$4.7162\ 10^{-5}$	$1.1836\ 10^{-5}$	$2.9584\ 10^{-6}$	$7.3989\ 10^{-7}$	$1.8497\ 10^{-7}$
3.5	$6.2880\ 10^{-5}$	$1.5700\ 10^{-5}$	$3.9248\ 10^{-6}$	$9.8150\ 10^{-7}$	$2.4537\ 10^{-7}$

TABLE 10. Test 8, $d = 3, \nu = 5$

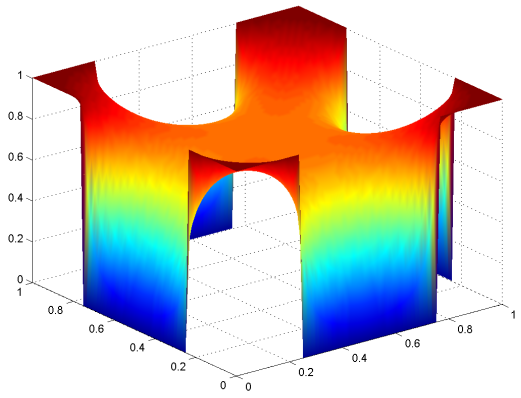


FIGURE 5. Test 9, $d = 3, \nu = 25$ graph of the solution with $\alpha = 2.5$

the same result were for $\nu = 1$, 758, for $\nu = 5$, 71 and for $\nu = 25$, 367. The number of iterations for $\nu = 5$ are less than 80 for the results in Table 1, whereas for $\nu = 25$, some cases took as many as 380 iterations.

However with finite differences, the higher ν, the higher is the number of iterations. Also, as noted in [11], the higher α, the more hyperbolic the problem becomes and more difficult to solve with a Laplacian based solver.

In figure 4.2, we plot the section of the solution on the line $x = 1/2$ with increasing values of $\alpha = 2, 2.5, 3, 3.5$, verifying numerically the discrete comparison principle.

We also found that with random perturbation on the right hand side f, the scheme still reproduces a smooth quadratic solution.

4.3. **Gauss curvature equation.** Following [16], we considered the Gauss curvature equation (1.4) on the domain $[-0.57, 0.57]^2$ with boundary conditions $g(x, y) = x^2 + y^2 - 1$ and ask what is the maximum value of K for which there exists a convex solution. The vanishing moment method breaks down for $K = 2.2$. With Method 3, we are able to capture a convex solution for K as large as 11.2. The initial guess was taken as the solution of the Monge-Ampère equation $\det D^2 u = K$.

Remark 4.1. *The iterations were stopped when the L^∞ norm of the difference between two iterates is less than 10^{-10} or when that value is bigger than the previously computed.*

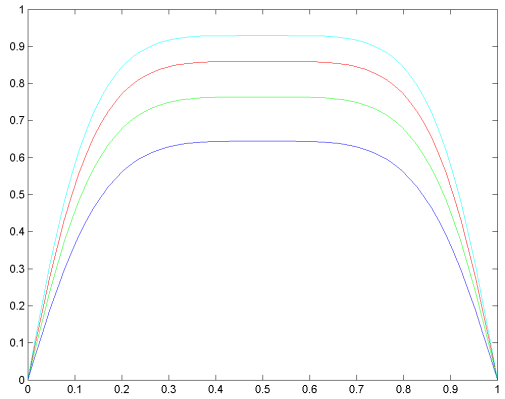


FIGURE 6. Test 9, $d = 3, \nu = 25$ sections of the solutions with increasing values of $\alpha = 2, 2.5, 3, 5$

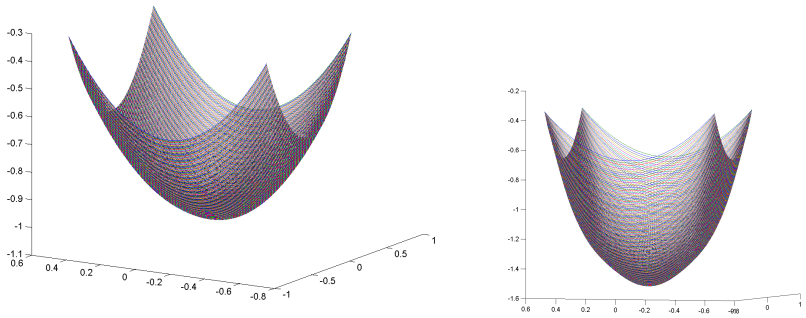


FIGURE 7. Gauss curvature equation by Method 3, $\nu = 50, h = 1/2, d = 3$ and with $K = 2.2$ and $K = 11.2$ respectively

Remark 4.2. *We could have performed the simulations with finite differences even for the two dimensional Monge-Ampère equation due to a remarkable property of the discrete Hessian, that is the symmetric matrix with entries $\mathcal{D}_{k,l}u, k, l = 1, \ldots, 2$, where*

$$\mathcal{D}_{1,1}u_{ij} = \frac{u_{i+1,j} - 2u_{ij} + u_{i-1,j}}{h^2},$$
$$\mathcal{D}_{2,2}u_{ij} = \frac{u_{i,j+1} - 2u_{ij} + u_{i,j-1}}{h^2} \text{ and}$$
$$\mathcal{D}_{1,2}u_{ij} = \frac{u_{i+1,j+1} + u_{i-1,j-1} - u_{i-1,j+1} - u_{i+1,j-1}}{4h^2}.$$

In [1]*, it is shown that the limit of a sequence of grid functions, with semi-positive definite Hessian, which converges in a suitable norm is a convex function. The obvious criticism of finite differences is the difficulty to deal with non rectangular domains. However, for Test 8 with the Monge-Ampère equation, we obtained better results with finite differences.*

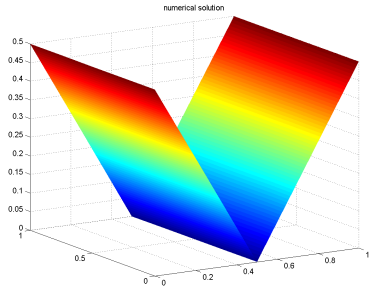


FIGURE 8. By finite differences, Test 8 Monge-Ampère, Method 3, $h = 1/2^2, \nu = 5$

Remark 4.3. *We have also experimented numerically with Method 3 for the Bellman-Isaacs equation using finite differences. The Bellman-Isaacs equations are given by* $F(x, D^2u) := sup_{\beta\in\mathcal{B}} inf_{\alpha\in\mathcal{A}}(L_{\alpha,\beta}u(x) - f_{\alpha,\beta}(x)) = 0$ *for given sets* $\mathcal{A}$ *and* $\mathcal{B}$*, given functions* $f_{\alpha,\beta}$ *and* $L_{\alpha,\beta}$ *a family of second order operators. Take* $L_{\alpha,\beta}u(x) = \mathrm{tr}(\sigma\sigma' D^2u(x))$ *and* $f_{\alpha,\beta}(x)$ *chosen as* $L_{\alpha,\beta}v(x), v(x)$ *exact solution. The scheme performs well with a smooth solution* $u_1(x, y) = \sin(x)\sin(y)$ *even with the symmetric matrix* σ *given by* $\sigma_{11} = x^2, \sigma_{12} = \frac{1}{2}xy, \sigma_{22} = y^2$*, for which no fixed narrow stencil can work in general* [15]*. However, with a non smooth solution such as* $u_2(x, y) = \sin(y/2)\sin(x/2)$ *if* $-\pi \le x < 0$ *and* $u_2(x, y) = \sin(y/2)\sin(x/4)$ *if* $0 \le x \le \pi$ *and the matrix* σ *given by* $\sigma_{11} = \sin(x + y), \sigma_{12} = \sigma_{23} = \beta, \sigma_{13} = \sigma_{22} = 0$*, we could obtain convergent results by choosing the domain as* $[-1/2^m, 1/2^m]$ *for fixed* m *and* ν *high. This suggests the need for a domain decomposition approach.*

Remark 4.4. *The operator* L *in* (1.5) *may be taken as the biharmonic operator. In which case, we add the boundary condition* $\Delta u_{k+1} = 1/\nu^2$*. The resulting algorithm may not be analyzed with the techniques discussed in this paper.*

Remark 4.5. *Have we approximated the viscosity solutions in the non-smooth case? Method 3 approximates the solution of a fully nonlinear equation by* C^2 *functions. This by itself defines a notion of weak solution. It would be interesting to find out if the limit is always unique and the connections with other definitions of weak solutions.*

ACKNOWLEDGMENTS

The author acknowledges discussions with F. Celiker, B. Cockburn, W. Gangbo, R. Glowinski, M.J. Lai, R. Nochetto, A. Oberman and A. Regev. The author was supported in part by NSF grant DMS-0811052 and the Sloan Foundation. This research was supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation.

REFERENCES

1. Néstor E. Aguilera and Pedro Morin, *Approximating optimization problems over convex functions*, Numer. Math. **111** (2008), no. 1, 1–34.
2. G. Awanou, *Energy methods in 3D spline approximations of the Navier-Stokes equations*, Ph.D. Dissertation, University of Georgia, Athens, Ga, 2003.

3. G. M. Awanou and M. J. Lai, *On convergence rate of the augmented Lagrangian algorithm for nonsymmetric saddle point problems*, Appl. Numer. Math. **54** (2005), no. 2, 122–134.
4. Gerard Awanou, *Robustness of a spline element method with constraints*, J. Sci. Comput. **36** (2008), no. 3, 421–432.
5. _____, *Spline element method for the Monge-Ampère equation*, Submitted., 2010.
6. Gerard Awanou and Ming-Jun Lai, *Trivariate spline approximations of 3D Navier-Stokes equations*, Math. Comp. **74** (2005), no. 250, 585–601 (electronic).
7. Gerard Awanou, Ming-Jun Lai, and Paul Wenston, *The multivariate spline method for scattered data fitting and numerical solution of partial differential equations*, Wavelets and splines: Athens 2005, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2006, pp. 24–74.
8. Victoria Baramidze and Ming-Jun Lai, *Spherical spline solution to a PDE on the sphere*, Wavelets and splines: Athens 2005, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2006, pp. 75–92.
9. Jean-David Benamou and Yann Brenier, *A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem*, Numer. Math. **84** (2000), no. 3, 375–393.
10. Jean-David Benamou, Brittany Froese, and Adam Oberman, *Two numerical methods for the elliptic Monge-Ampère equation*, 2010.
11. 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** (2008), no. 3, 185–216.
12. Luis A. Caffarelli and Xavier Cabré, *Fully nonlinear elliptic equations*, American Mathematical Society Colloquium Publications, vol. 43, American Mathematical Society, Providence, RI, 1995.
13. E. J. Dean and R. Glowinski, *Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 13-16, 1344–1386.
14. Edward J. Dean and Roland Glowinski, *On the numerical solution of a two-dimensional Pucci's equation with Dirichlet boundary conditions: a least-squares approach*, C. R. Math. Acad. Sci. Paris **341** (2005), no. 6, 375–380.
15. K. Debradant and E.R. Jakobsen, *Semi-lagrangian schemes for linear and fully nonlinear diffusion equations*, Submitted, 2010.
16. X. Feng and M. Neilan, *Vanishing moment method and moment solutions for second order fully nonlinear partial differential equations*, J. Sci. Comput. **38** (2009), no. 1, 74–98.
17. B.D. Froese and A.M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampere equation in dimensions two and higher*, Submitted, 2010.
18. Roland Glowinski, *Numerical methods for fully nonlinear elliptic equations*, ICIAM 07—6th International Congress on Industrial and Applied Mathematics, Eur. Math. Soc., Zürich, 2009, pp. 155–192.
19. Xian-Liang Hu, Dan-Fu Han, and Ming-Jun Lai, *Bivariate splines of various degrees for numerical solution of partial differential equations*, SIAM J. Sci. Comput. **29** (2007), no. 3, 1338–1354 (electronic).
20. Ming-Jun Lai, *Convex preserving scattered data interpolation using bivariate C^1 cubic splines*, J. Comput. Appl. Math. **119** (2000), no. 1-2, 249–258, Dedicated to Professor Larry L. Schumaker on the occasion of his 60th birthday.
21. Grégoire Loeper and Francesca Rapetti, *Numerical solution of the Monge-Ampère equation by a Newton's algorithm*, C. R. Math. Acad. Sci. Paris **340** (2005), no. 4, 319–324.
22. Adam 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. Ser. B **10** (2008), no. 1, 221–238.
23. Neil S. Trudinger and Xu-Jia Wang, *Boundary regularity for the Monge-Ampère and affine maximal surface equations*, Ann. of Math. (2) **167** (2008), no. 3, 993–1028.

Department of Mathematical Sciences, Northern Illinois University, Dekalb, IL, 60115

E-mail address: awanou@math.niu.edu

URL: http://www.math.niu.edu/~awanou