

AN AVERAGING SCHEME FOR MACROSCOPIC NUMERICAL SIMULATION OF NONCONVEX MINIMIZATION PROBLEMS

ZHIPING LI¹ AND CARSTEN CARSTENSEN²

ABSTRACT. Averaging or gradient recovery techniques, which are a popular tool for improved convergence or superconvergence of finite element methods in elliptic partial differential equations, have not been recommended for nonconvex minimization problems as the energy minimization process enforces finer and finer oscillations and hence at the first glance, a smoothing step appears even counterproductive. For macroscopic quantities such as the stress field, however, this counterargument is no longer true. In fact, this paper advertises an averaging technique for a surprisingly improved convergence behavior for nonconvex minimization problems. Similar to a finite volume scheme, numerical experiments on a double-well benchmark example provide empirical evidence of superconvergence phenomena in macroscopic numerical simulations of oscillating microstructures.

1. INTRODUCTION

Consider a typical problem of minimizing a functional

$$E(u) = \int_{\Omega} W(Du) dx + \int_{\Omega} |u - f|^2 dx, \quad (1.1)$$

in a set admissible functions

$$\mathfrak{A} = \{u \in W^{1,p}(\Omega; R^m) : u|_{\partial\Omega} = u_0\}, \quad (1.2)$$

where $\Omega \subset R^n$ is a connected open set with Lipschitz continuous boundary $\partial\Omega$ and $W(\cdot)$ is non-convex (non-quasiconvex in the vector case, *i.e.* $n \geq 2$ and $m \geq 2$). It is well known that such problems has in general no solutions in the classical sense [1, 2], and the minimizing sequences of $E(\cdot)$ in \mathfrak{A} can develop

Key words and phrases. averaging scheme, nonconvex minimization problem, macroscopic numerical simulation, convexification, adaptive mesh refinement.

The research of Z.-P Li was supported in part by the Special Funds for Major State Basic Research Projects (2005CB321701), NSFC projects (10431050, 10571006 and 10528102) and RFDP of China.

finer and finer oscillations and lead to generalized solutions, or microstructures, represented by gradient Young measures [1, 3, 4]. Even though there are some successful numerical methods based on the direct method for the computation of some special microstructures (see [5, 6, 7] among many others, more references on numerical analysis and computation of microstructures based on the direct method can be found in [8, 9]), it is found that the numerical results for such problems can be strongly mesh dependent [8, 10, 11, 12] and the increasingly fine length scale oscillations are often practically impossible to resolve numerically, so it is generally extremely difficult to solve the problem by direct numerical computation. Thus, naturally, mathematical models and methods based on some kind of relaxation are considered essential in the numerical computations, especially in the macroscopic simulations of such problems [13], and a lot of progress has been made in numerical analysis and computations based on relaxation models and methods (see for example [14, 15, 16, 17, 18, 19], more references can be found in [20]).

On the other hand, it is well known that averaging or gradient recovery techniques are a popular tool for improved convergence or superconvergence of finite element methods in elliptic partial differential equations with applications to a posteriori error control (see for example [21, 22, 23, 24]). For nonconvex minimization problems such techniques have not been recommended as the energy minimization process enforces finer and finer oscillations. Hence, at first glance, a smoothing step appears even counterproductive. For macroscopic quantities such as the stress field, however, this counterargument is no longer true [25]. In fact, this paper advertises averaging techniques for a surprisingly improved convergence behavior. Numerical experiments on a double-well benchmark example [14, 19] provide empirical evidence that there are superconvergence phenomena in macroscopic numerical simulations of oscillating microstructures.

The rest of the paper is organized as follows. In section 2, we establish a numerical algorithm by applying an averaging technique to a relaxation model of the non-convex minimization problem. In section 3, numerical experiments on a double-well benchmark example are shown. Conclusions of the paper are given in section 4.

2. MODEL AND ALGORITHM

For the macroscopic simulation, instead of considering the problem of minimizing the functional $E(\cdot)$ in \mathfrak{A} , we consider the corresponding relaxed problem of minimizing the functional

$$E^{**}(u) = \int_{\Omega} W^{**}(Du) dx + \int_{\Omega} |u - f|^2 dx \quad (2.1)$$

in the set of admissible functions \mathfrak{A} (see (1.2)), where W^{**} is the convex envelope (quasiconvex envelope in the vector case) of the energy density W . It is well known that the relaxed problem is solvable under certain general conditions on W [2].

Let $\mathfrak{T}_h(\Omega)$ be regular triangulations of Ω with mesh size h [26]. Let

$$\mathfrak{A}_h = \{u \in (C(\overline{\Omega}))^m : u|_K \text{ is affine } \forall K \in \mathfrak{T}_h(\Omega)\} \quad (2.2)$$

and

$$\mathfrak{A}_h(u_{0,h}) = \{u \in \mathfrak{A}_h : u|_{\partial\Omega} = u_{0,h}\}, \quad (2.3)$$

where $u_{0,h}$ is the interpolation of u_0 in \mathfrak{A}_h . A standard finite element discretization of the relaxed problem is to minimize the functional $E^{**}(\cdot)$ in the set $\mathfrak{A}_h(u_{0,h})$ of finite element functions. Since Du_h is only piecewise constant for $u_h \in \mathfrak{A}_h$, the approximation is not expected to be good, especially at regions where the solution is non-smooth. In principle, averaging, or gradient recovery techniques are applied in the hope of obtaining a higher order approximation for the gradient field (or stress field), and thus a better approximation of the solution.

Let \mathfrak{N} be the set of nodes in $\mathfrak{T}_h(\Omega)$, and let $\varphi_z, z \in \mathfrak{N}$ be the shape functions defined on $\mathfrak{T}_h(\Omega)$. Define

$$\overline{Du}_h(z) = \frac{1}{|\omega_z|} \int_{\omega_z} Du_h(x) dx \quad (2.4)$$

where $\omega_z = \{x \in \Omega : \varphi_z(x) > 0\}$ and $|\omega_z|$ denotes the measure of ω_z . Then, for a given function $u_h \in \mathfrak{A}_h$, $\{\overline{Du}_h(z) : z \in \mathfrak{N}\}$ defines a piecewise affine gradient field \overline{Du}_h on $\mathfrak{T}_h(\Omega)$. This averaging technique leads to the following discrete problem

$$\begin{cases} \text{Find } u_h \in \mathfrak{A}_h(u_{0,h}) \text{ such that,} \\ \overline{E}_h(u_h) = \min_{v_h \in \mathfrak{A}_h(u_{0,h})} \overline{E}_h(v_h), \end{cases} \quad (2.5)$$

where

$$\bar{E}_h(u_h) = \sum_{z \in \mathfrak{N}} \frac{|\omega_z|}{3} (W^{**}(\overline{Du}_h(z)) + |u_h(z) - f(z)|^2) \quad (2.6)$$

is a discrete functional defined on the finite element function space \mathfrak{A}_h obtained by substituting \overline{Du}_h to Du_h in $E(u_h)$ and by applying a numerical quadrature on the integrals.

Next, we apply the above averaging technique to the following double well benchmark example [14, 19]. Let $\Omega = (0, 1) \times (0, 3/2)$ and let the energy density

$$W(F) = |F - F_1|^2 |F - F_2|^2 \quad \text{for } F \in \mathbb{R}^2, \quad (2.7)$$

where $F_1 = -(3, 2)/\sqrt{13}$ and $F_2 = -F_1$ are the two given distinct wells, and let the boundary condition u_0 be given by

$$u_0(x, y) = \begin{cases} \sqrt{13}(3 - 2y)^3(3(3 - 2y)^2/281216 + 1/507) & \text{for } x = 0; \\ 2y(1 + 4y^2/312)/\sqrt{13} & \text{for } x = 1; \\ \sqrt{13}(1 - x)^3(729(1 - x)^2 + 14976)/281216 & \text{for } y = 0; \\ 3x(1 + 9x^2/13)/\sqrt{13} & \text{for } y = 3/2, \end{cases} \quad (2.8)$$

and finally let $\mathfrak{A} = u_0 + W_0^{1,4}(\Omega)$. It can be shown [14] that the lower convex envelope of W is given by

$$W^{**}(F) = ((|F|^2 - 1)_+)^2 + 4(|F|^2 - (F_1 \cdot F)^2), \quad (2.9)$$

where $(\cdot)_+ = \max\{0, \cdot\}$, and while the infimum of E in \mathfrak{A} is not attainable, the corresponding relaxed problem has a unique solution given by

$$u(x, y) = \begin{cases} f_0(t + \frac{1}{2}) & \text{for } -\frac{1}{2} \leq t \leq 0; \\ f_1(t + \frac{1}{2}) & \text{for } 0 \leq t \leq \frac{1}{2}, \end{cases} \quad (2.10)$$

where $t = (3(x - 1) + 2y)/\sqrt{13}$ and

$$f_0(t + \frac{1}{2}) = -3t^5/128 - t^3/3;$$

$$f_1(t + \frac{1}{2}) = t^3/24 + t.$$

It is not difficult to see [14] that W^{**} is differentiable and

$$DW^{**}(F) = 4((|F|^2 - 1)_+ + 2)F - 8(F_1 \cdot F)F_1. \quad (2.11)$$

The gradients of u is given by

$$Du(x, y) = \begin{cases} t^2(1 + 15t^4/128)F_1 & \text{for } -\frac{1}{2} \leq t \leq 0; \\ (1 + t^2/8)F_2 & \text{for } 0 \leq t \leq \frac{1}{2}, \end{cases} \quad (2.12)$$

which is continuous on the lower left and upper right subdomains

$$\mathfrak{M} = \{\text{the convex hull of } (0, 0), (1, 0), (0, 3/2)\};$$

$$\mathfrak{R} = \{\text{the convex hull of } (1, 0), (1, 3/2), (0, 3/2)\},$$

but is discontinuous on the antidiagonal $\Gamma = \mathfrak{M} \cap \mathfrak{R}$, that is $t = 0$, or $2y + 3(x - 1) = 0$, however the stress field

$$\sigma = DW^{**}(Du) = \begin{cases} 0 & \text{for } -\frac{1}{2} \leq t \leq 0; \\ t^2(1 + 3t^2/16 + t^4/64)F_2 & \text{for } 0 \leq t \leq \frac{1}{2} \end{cases} \quad (2.13)$$

is continuously differentiable on the whole domain Ω .

Since $W^{**}(\cdot)$ is differentiable, $\bar{E}_h(\cdot)$ is also differentiable and the optimization problem (2.5) can be solved by gradient type methods. The conjugate gradient method leads to the following scheme.

Algorithm:

Input an initial $u_h \in \mathfrak{A}_h$ with $u_h = u_0$ on $\partial\Omega$.

(a): Compute $D\bar{E}_h(u_h)$;

(b): If $|D\bar{E}_h(u_h)| < \text{TOL}$ then go to step (f);

(c): Compute the conjugate gradient direction v_h ;

(d): Find t^* to minimize $\bar{E}_h(u_h + tv_h)$;

(e): Set $u_h = u_h + t^*v_h$, go to step (a);

(f): If necessary refine the mesh and go to step (a);

Output u_h .

In step (a) of the algorithm, we apply an incomplete search, in which the step length t is limited to either being doubled or being halved in searching for a "minima". In step (f), either the uniform mesh refinement or adaptive mesh refinement can be applied, and the interpolation of the numerical result obtained on the coarse mesh is taken as the initial guess on the refined mesh. In the adaptive mesh refinements, we use the ZZ-error indicator given in [19] which is defined in the following way: let $\sigma_h = DW^{**}(Du_h)$, and for each element

$T \in \mathfrak{T}_h(\Omega)$, define

$$\eta_{T,z} = \|\sigma_h - A\sigma_h\|_{L^{4/3}(T)}^{4/3},$$

where A is an averaging operator given by

$$A\sigma_h = \sum_{z \in \mathfrak{N}} \left(\frac{1}{|\omega_z|} \int_{\omega} \sigma_h(y) dy \right) \varphi_z,$$

and define the averaging estimator by

$$\eta_z = \left(\sum_{T \in \mathfrak{T}_h(\Omega)} \eta_{T,z} \right)^{3/4}. \quad (2.14)$$

3. NUMERICAL EXPERIMENTS

Noticing that the antidiagonal $\Gamma = \mathfrak{M} \cap \mathfrak{R}$ is a free boundary across which the gradient of the solution is discontinuous, we are easily convinced that a finite element mesh having Γ included in its mesh lines can lead to a much better numerical approximation. Since such free boundaries are not known a priori in practical applications, and the ability to resolve the free boundary is one of the key issues for a successful algorithm, we will use meshes which, even after many adaptive mesh refinements, do not have mesh lines on Γ .

Given integers $M > 1$ and $N > 1$, a uniform mesh is introduced by the following lines

$$\begin{aligned} x &= i/M & 0 \leq i \leq M; \\ y &= 3j/2N & 0 \leq j \leq N; \\ y &= 3(Mx + k)/2N & -M \leq k \leq N. \end{aligned}$$

It is obvious that for an initial mesh with $M \neq N$, neither uniform mesh refinements nor adaptive mesh refinements can produce mesh lines lying on Γ .

First, numerical experiments on uniform meshes are done in which we take $M = 10, 20, 30, 40, 50$ and $N = 3M/2$, or $15, 30, 45, 60, 75$ respectively. For the initial guess, we take $u_h(z) = 0$ for all the interior nodes z . The tolerance TOL in the step (b) of the algorithm is set to 10^{-9} . The conjugate gradient method is restarted in every 25 iterations. The initial step length $t = 0.25$ is taken. Numerical results obtained by the standard finite element method using piecewise affine finite element functions and the numerical results obtained by the method using the averaging technique are compared in Figure 1-4.

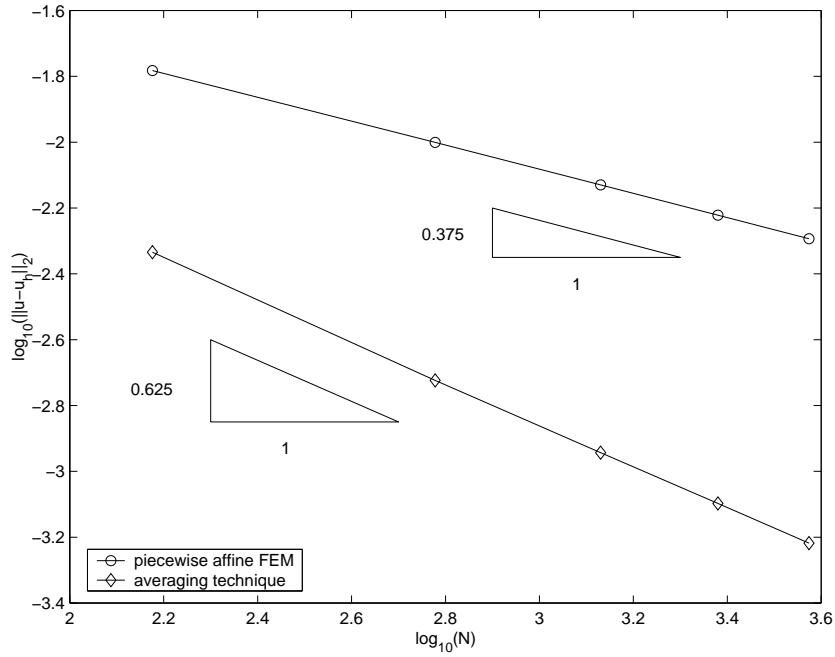


FIGURE 1. Error $\|u - u_h\|_2$ on uniform meshes.

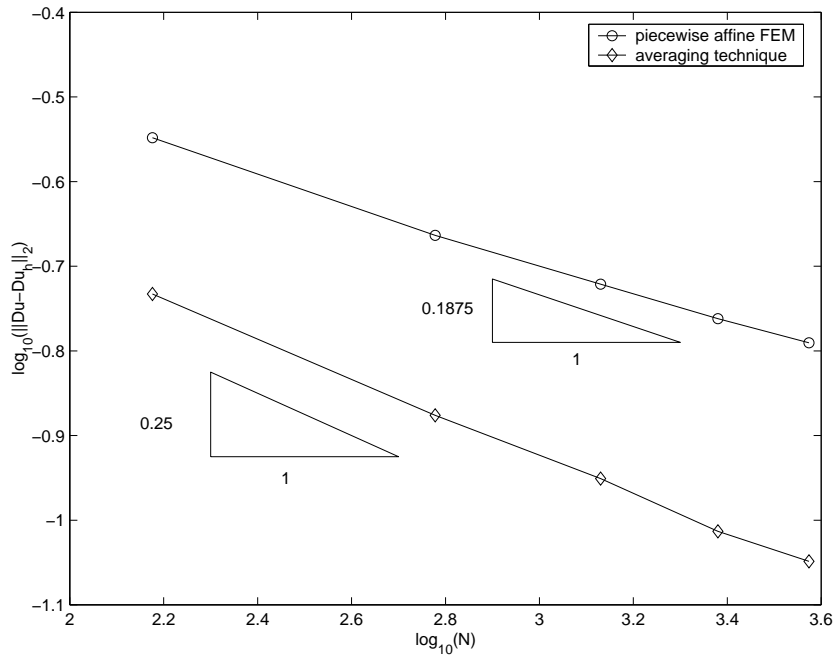


FIGURE 2. Error $\|Du - Du_h\|_2$ on uniform meshes.

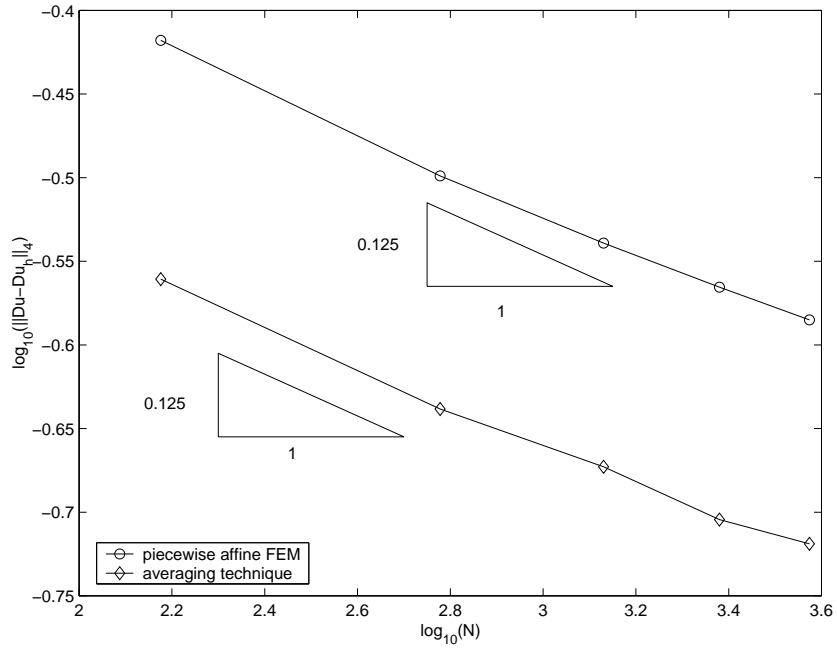


FIGURE 3. Error $\|Du - Du_h\|_4$ on uniform meshes.

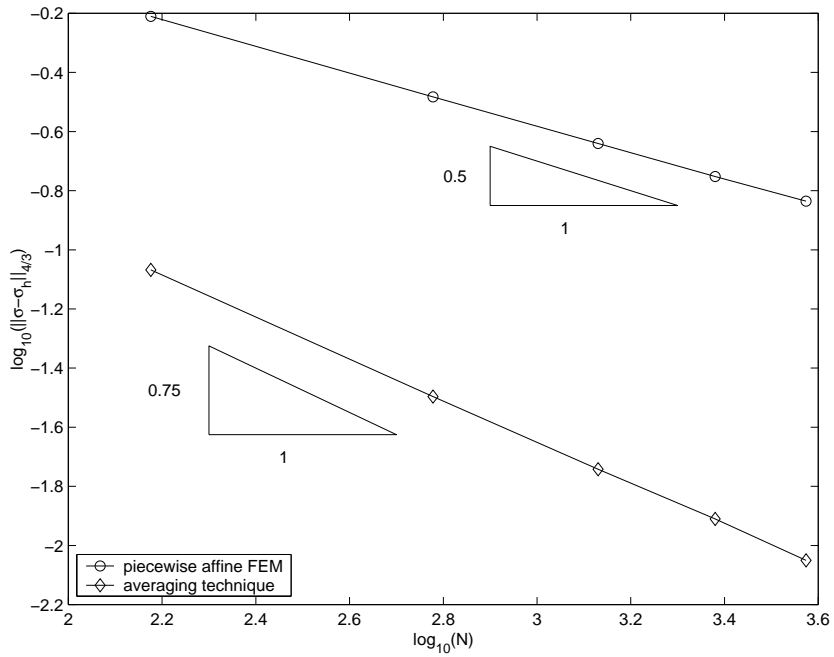


FIGURE 4. Error $\|\sigma - \sigma_h\|_{4/3}$ on uniform meshes.

It is clearly shown in Figure 1-4 that the averaging technique greatly improves the convergence behavior, especially the superconvergence is achieved for u on L^2 and $W^{1,2}$ norms, and for σ on $L^{4/3}$ norm, with the corresponding convergence rates being increased from $3/4$, $3/8$, 1 to $5/4$, $1/2$, $3/2$ respectively.

Next, we show some numerical results obtained by combining the adaptive mesh refinements using η_z (see (2.14)) with the averaging technique. In figure 5, the error in different norms on uniform and adaptively refined meshes are shown, where "*" and "+" represent the adaptive mesh refinements starting from 10×15 and 30×45 uniform meshes respectively. It is clearly seen that the adaptive mesh refinements further improve the convergence. The optimal refinements seem to be when the nodes' number \mathfrak{N} is about 3 to 4 times of the original size, more adaptive mesh refinements further improves the convergence but not as effective (compare the results shown in "*" and "+"). Figure 6 shows a numerical result of u_h on an adaptively refined mesh starting from a 10×15 uniform mesh. In figure 6, we can also see clearly that the antidiagonal $\Gamma = \mathfrak{M} \cap \mathfrak{R}$ is well resolved, however the mesh refinements are too much concentrated around Γ that causes the drop of efficiency of the mesh refinements.

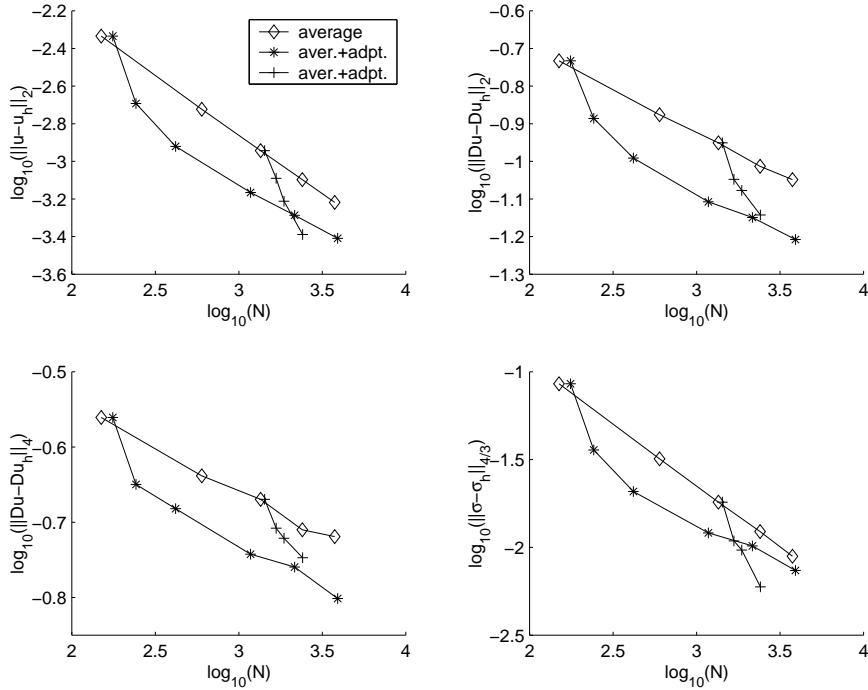


FIGURE 5. Error in different norms on uniform and adaptive meshes.

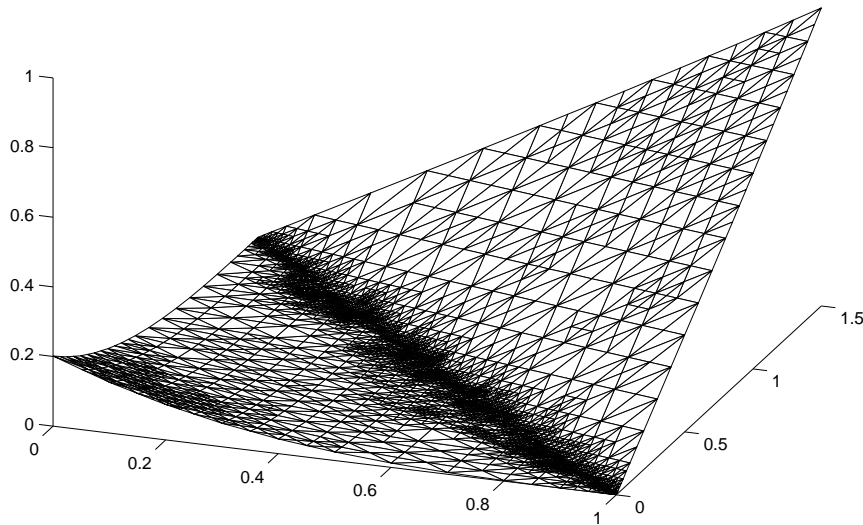


FIGURE 6. u_h on an adaptively refined mesh ($\mathfrak{N} = 5147$).

4. CONCLUSIONS

Even though there is no theoretical justification of any kind, empirical evidences show that averaging techniques can significantly improve the convergence for the macroscopic numerical approximation of nonconvex minimization problems, and further improvement can be achieved by combining adaptive mesh refinements.

For solutions having discontinuous gradient, the alignment of the mesh lines with the set of discontinuity can significantly reduce the numerical error and reduce the adaptive cost. A recent work of Zhou and Li [27] showed that the mesh transformation method can be applied to achieving the alignment. Thus a combination of the averaging techniques, mesh transformation method and adaptive refinements is highly recommended.

REFERENCES

- [1] J. M. Ball and R. D. James, *Fine phase mixtures as minimizers of energy*. **Arch. Rat. Mech. Anal.**, **100**(1987)1, 13-52.

- [2] B. Dacorogna, **Direct Methods in the Calculus of Variations**. Springer-Verlag, Heidelberg, 1989.
- [3] J.M. Ball, *A version of the fundamental theorem for Young measures*. In **Partial Differential Equations and Continuum models of Phase Transitions**, Lecture Notes in Physics, No. 344, M. Rasche, D. Serre and M. Slemrod, eds., Springer-Verlag, New York (1989), pp. 207-215.
- [4] P. Pedregal, **Parameterized Measures and Variational Principles**. Birkhäuser, 1997.
- [5] Z.-P. Li, *Rotational transformation method and some numerical techniques for the computation of microstructures*. **Math. Models Meth. Appl. Sci.**, **8**(1998), 985-1002.
- [6] Z.-P. Li, *A periodic relaxation method for computing microstructures*. **Appl. Numer. Math.**, **32**(2000), 291-303.
- [7] Z.-P. Li, *A mesh transformation method for computing microstructures*. **Numer. Math.**, **89**(2001), 511-533.
- [8] M. Luskin, *On the computation of crystalline microstructure*. **Acta Numerica**, **5**(1996), 191-257.
- [9] M.K. Gobbert and A. Prohl, *A comparison of classical and new finite element methods for the computation of laminated microstructure*, **Appl. Numer. Math.** **36**(2001), pp.155-178.
- [10] C. Collins, *Computation of twinning*, in **Microstructure and Phase Transitions**, IMA Volumes in Mathematics and Its Applications, Vol.54, J. Ericksen, R. James, D. Kinderlehrer and M. Luskin, eds, Springer-Verlag, New York, (1993), pp. 39-50.
- [11] C. Collins, M. Luskin and J. Riordan, *Computational results for a two-dimensional model of crystalline microstructure*, in **Microstructure and Phase Transitions**, IMA Volumes in Mathematics and Its Applications, Vol.54, J. Ericksen, R. James, D. Kinderlehrer and M. Luskin, eds, Springer-Verlag, New York, (1993), pp. 51-56.
- [12] Z.-P. Li, *Laminated microstructure in a variational problem with a non-rank-one connected double well potential*, **J. Math. Anal. Appl.**, **217**(1998), 490-500.
- [13] T. Roubíček, **Relaxation in Optimization Theory and Variational Calculus**. W. de Gruyter, Berlin-New York, 1997.
- [14] C. Carstensen and P. Plecháč, *Numerical solutions of the scalar double-well problem allowing microstructure*. **Math. Comp.**, **66**(1997), 997-1026.
- [15] G. Dolzmann, *Numerical computation of rank-one convex envelopes*. **SIAM J. Numer. Anal.**, **36**(1999), 1621-1635.
- [16] G. Dolzmann and N. J. Walkington, *Estimates for numerical approximations of rank-one convex envelopes*. **Numer. Math.**, **85** (2000), pp. 647-663.
- [17] Z.-P. Li, *Finite order rank-one convex envelopes and computation of microstructures with laminates in laminates*. **BIT**, **40**(4)(2000), pp.745-761.
- [18] E. Adranda and P. Pedregal, *On the computation of the rank-one convex hull of a function*. **SIAM J. Sci. Comput.**, **22**(5)(2001), 1772-1790.

- [19] C. Carstensen and K. Jochimsen, *Adaptive finite element for microstructures? Numerical experiments for a 2-well benchmark*. **Computing**, **71**(2003), 175-204.
- [20] C. Carstensen, *Numerical Analysis of Microstructures, Theory and Numerics of Differential Equations, Durham 2000*, pp. 59-126, (editors J.F. Blowey, J.P. Coleman and A.W. Craig), Springer , Berlin and Heidelberg 2000 (univeritext).
- [21] C. Carstensen and S. Bartles, *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grid, part I: Low order conforming, nonconforming, and mixed FEM*. **Math. Comp.**, **71**(2002), 945-969.
- [22] C. Carstensen and S.A. Funken, *Averaging technique for FE - a posteriori error control in elasticity. Part I: Conforming FEM*. **Comput. Methods Appl. Mech. Engrg.**, **190**(2001), 2483-2498.
- [23] C. Carstensen and S.A. Funken, *Averaging technique for FE - a posteriori error control in elasticity. Part II: λ -independent estimates*. **Comput. Methods Appl. Mech. Engrg.**, **190**(2001), 4663-4675.
- [24] C. Carstensen and S.A. Funken, *Averaging technique for FE - a posteriori error control in elasticity. Part III: Locking-free nonconforming FEM*. **Comput. Methods Appl. Mech. Engrg.**, **191**(2001), 861-877.
- [25] C. Carstensen and S. Müller, *Local stress regularity in scalar non-convex variational problems*. To appear in **SIAM J. Math. Anal.**, **34**(2002), 495-509.
- [26] P.G. Ciarlet, **The Finite Element Method for Elliptic Problems**. North-Holland, Amsterdam, 1978.
- [27] J.-S. Zhou and Z.-P. Li, *Computing non-smooth minimizers with the mesh transformation method*. **IMA J. Numer. Anal.**, **25**(2005), 458-472.

¹ ZHIPING LI, LMAM AND SCHOOL OF MATHEMATICAL SCIENCES, PEKING UNIVERSITY, BEIJING 100871, P.R.CHINA. (lizp@math.pku.edu.cn)

² CARSTEN CARSTENSEN, DEPARTMENT OF MATHEMATICS, HUMBOLDT-UNIVERSITÄT BERLIN, D-10099 BERLIN, GERMANY. (cc@math.hu-berlin.de)