

Fast Solvers for Nonsmooth Optimization Problems in Phase Separation

Pawan Kumar

Department of Mathematics and Computer Science
 Freie Universität Berlin
 Arnimallee 6, 14195 Berlin

Abstract—The phase separation processes are typically modeled by well known Cahn-Hilliard equation with obstacle potential. Solving these equations correspond to a nonsmooth and nonlinear optimization problem. Recently a globally convergent Newton Schur method was proposed for the non-linear Schur complement corresponding to this 2×2 non-linear system. The proposed method is similar to an inexact active set method in the sense that the active sets are first identified by solving a quadratic obstacle problem corresponding to the (1, 1) block of the 2×2 system, and later solving a reduced linear system by annihilating the rows and columns corresponding to identified active sets. For solving the quadratic obstacle problem, various optimal multigrid like methods have been proposed. However solving the reduced system remains a major bottleneck. In this paper, we explore an effective preconditioner for the reduced linear system that allows solving large scale optimization problem corresponding to Cahn-Hilliard and to possibly similar models.

I. INTRODUCTION

THE Cahn-Hilliard equation was first proposed in 1958 by Cahn and Hilliard [1] to study the phase separation process in a binary alloy. Here the term phase stands for the concentration of different components in the alloy. It has been empirically observed that the concentration changes from a possibly mixed state to a distinct spatially separated two phase state when the alloy under preparation is subjected to a rapid cooling at a certain critical temperature. This rapid reduction in the temperature the so-called *deep quench limit* has been found to be modeled efficiently by obstacle potential proposed by Oono and Puri [2] in 1987; also see Blowey and Elliot [3, p. 237, (1.14)]. The phase separation has been noted to be highly non-linear (point nonlinearity to be precise), and the obstacle potential emulates the nonlinearity and non-smoothness that is empirically observed. Dealing with the non-smoothness and designing robust iterative procedure has been the subject of much active research in last decades. Assuming semi-implicit time discretizations [4] to alleviate the time step restrictions, most of the proposed methods essentially differ in the way the nonlinearity and non-smoothness are handled. Two of the main approaches for such problems are: regularization around the non-smooth region [5] or an active set approach [6] i.e., identify the active sets and solve a reduced problem which is linear, [6] unlike [5] also ensures global convergence of the Newton method by using proper damping parameter. The non-linear problem corresponding to Cahn-Hilliard problem with

obstacle potential could be written as a non-linear system in block 2×2 matrix form as follows:

$$\begin{pmatrix} F & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} u^* \\ w^* \end{pmatrix} \ni \begin{pmatrix} f \\ g \end{pmatrix}, \quad u^*, w^* \in \mathbb{R}^n \quad (1)$$

where u^*, w^* are unknowns, $F = A + \partial I_{\mathcal{K}}$, where $I_{\mathcal{K}}$ denotes the indicator functional of the admissible set \mathcal{K} . The matrices A, C are essentially Laplacian with A augmented by a non-local term (a rank one term) reflecting mass conservation. By nonlinear Gaussian elimination of the u^* variables, the system above could be reduced to a nonlinear Schur complement system in w^* variables [6], where the nonlinear Schur complement is given by $-(C + BF^{-1}B^T)$. In [6], a globally convergent Newton method is proposed for this nonlinear Schur complement system which is later interpreted as a preconditioned Uzawa iteration. Note that $F(x)$ is a set valued mapping due to the presence of set-valued operator $\partial I_{\mathcal{K}}$; to solve the inclusion $F(x) \ni y$, or equivalently, $x \in F^{-1}y$ corresponding to the quadratic obstacle problem, many methods have been proposed: projected block Gauss-Seidel [7], monotone multigrid method [8], [9], [10], truncated monotone multigrid [11], and recently introduced truncated Newton multigrid [11]. See the excellent review article [11] that compares all these methods. Solving the quadratic obstacle problem corresponds to identifying the active sets. By annihilating the corresponding rows and columns that belong to the identified active sets, we obtain a reduced linear system as follows:

$$\begin{pmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & -C \end{pmatrix} \begin{pmatrix} \hat{u} \\ \hat{w} \end{pmatrix} = \begin{pmatrix} \hat{f} \\ \hat{g} \end{pmatrix}, \quad \hat{u}, \hat{w} \in \mathbb{R}^n \quad (2)$$

The overall nonlinear iteration is performed in the sense of inexact Uzawa, and the preconditioners are updated with next available active sets.

In this paper our goal is to design effective preconditioner for (2). In particular we consider the Schur complement preconditioner proposed in [5] and adapt it to our linear system. Our linear system differs from [5]; we have to deal with nontrivial kernels at (1,1) and (2,2) block. We study the effectiveness of the preconditioner for various active set configurations that allow solving large scale nonsmooth optimization problem corresponding to similar model problems when similar nonsmooth Newton Schur method is used.

The rest of this paper is organized as follows. In Section III, we describe the Cahn-Hilliard model with obstacle potential,

Funded by EC math and Matheon

we discuss the time and space discretizations and variational formulations. In Section IV, we discuss briefly the solver for Cahn-Hilliard with obstacle problem. In particular, we briefly discuss the recent Nonsmooth Newton Schur method, and the truncated monotone multigrid for the obstacle problem; we describe how the linear system appears. The preconditioners for the reduced linear systems are discussed in Section IV-E. Finally in Section V, we show numerical experiments with the proposed preconditioner.

II. NOTATION

Let SPD and SPSD denote symmetric positive definite and symmetric positive semi-definite respectively. Let $|x|$ denote the absolute value of x , whereas, for a set \mathcal{K} , $|\mathcal{K}|$ denotes the number of elements in \mathcal{K} . Let $Id \in \mathbb{R}^{n \times n}$ denote the identity matrix. For a vector u let $u(i)$ denote the i^{th} entry of vector u . Similarly for a matrix we use the notation $K(i, j)$ to denote the $(i, j)^{\text{th}}$ entry of K . For any matrix K , K^+ shall denote a pseudoinverse of K .

III. CAHN-HILLIARD PROBLEM WITH OBSTACLE POTENTIAL

A. The Model

The Cahn-Hilliard equation in PDE form with inequality constraints reads:

$$\partial_t u = \Delta w, \quad (3)$$

$$w = -\epsilon \Delta u + \psi'_0(u) + \mu, \quad (4)$$

$$\mu \in \partial \beta_{[-1,1]}(u), \quad (5)$$

$$|u| \leq 1, \quad (6)$$

$$\frac{\partial u}{\partial n} = \frac{\partial w}{\partial n} = 0 \text{ on } \partial \Omega, \quad (7)$$

where $\partial \beta_{[-1,1]}(u)$ is the subdifferential of $\beta_{[-1,1]}(u) := \int_{\Omega} I_{[-1,1]}(u)$. The obstacle potential ψ is given as follows:

$$\psi(u) = \psi_0(u) + I_{[-1,1]}(u), \text{ where } \psi_0(u) = \frac{1}{2}(1 - u^2). \quad (8)$$

Here the indicator function $I_{[-1,1]}(u)$ is defined as follows:

$$I_{[-1,1]} = \begin{cases} 0, & \text{if } u(i) \in [-1, 1], \\ \infty, & \text{otherwise.} \end{cases} \quad (9)$$

The subscript $[-1, 1]$ correspond to the fact that u is allowed to take values only between -1 and $+1$, which we sometimes refer to as upper and lower obstacles respectively.

In (3)-(7) the unknowns u and w are called order parameter and chemical potential respectively. For a given $\epsilon > 0$, final time $T > 0$ and initial condition $u_0 \in \mathcal{K}$ where

$$\mathcal{K} = \{v \in H^1(\Omega) : |v| \leq 1\}, \quad (10)$$

the equivalent initial value problem for Cahn-Hilliard equation with obstacle potential interpreted as variational inequality reads:

$$\left\langle \frac{du}{dt}, v \right\rangle + (\nabla w, \nabla v) = 0, \quad \forall v \in H^1(\Omega), \quad (11)$$

$$\epsilon(\nabla u, \nabla(v - u)) - (u, v - u) \geq (w, v - u), \quad \forall v \in \mathcal{K}, \quad (12)$$

where we use the notation $\langle \cdot, \cdot \rangle$ to denote the duality pairing of $H^1(\Omega)$ and $H^1(\Omega)'$. Note that we used the fact that $\psi'_0(u) = -u$ in the second term on the left of inequality (12) above. The existence and uniqueness of the solution of (11), (12) above has been established in Blowey and Elliot [3]. We next consider an appropriate discretization in time and space for the model.

B. Time and Space Discretizations

We consider a fixed non-adaptive grid in time $(0, T)$ and in space $\Omega = (0, 1) \times (0, 1)$. The time step $\tau = T/N$ is kept uniform, N being the number of time steps. We consider the semi-implicit Euler discretization in time and finite element discretization as in Barrett et. al. [7] with triangulation \mathcal{T}_h with the following spaces:

$$\mathcal{S}_h = \{v \in C(\bar{\Omega}) : v|_T \text{ is linear } \forall T \in \mathcal{T}_h\}, \quad (13)$$

$$\mathcal{P}_h = \{v \in L^2(\Omega) : v|_T \text{ is constant } \forall T \in \mathcal{T}_h\}, \quad (14)$$

$$\mathcal{K}_h = \{v \in \mathcal{P}_h : |v|_T \leq 1 \forall T \in \mathcal{T}_h\} = \mathcal{K} \cap \mathcal{S}_h \subset \mathcal{K}, \quad (15)$$

which leads to the following discrete Cahn-Hilliard problem with obstacle potential:

Find $u_h^k \in \mathcal{K}_h, w_h^k \in \mathcal{S}_h$ such that

$$\langle u_h^k, v_h \rangle + \tau(\nabla w_h^k, \nabla v_h) = \langle u_h^{k-1}, v_h \rangle, \quad \forall v_h \in \mathcal{S}_h, \quad (16)$$

$$\epsilon(\nabla u_h^k, \nabla(v_h - u_h^k)) - \langle w_h^k, v_h - u_h^k \rangle \geq \langle u_h^{k-1}, v_h - u_h^k \rangle, \quad (17)$$

$$\forall v_h \in \mathcal{K}_h.$$

holds for each $k = 1, \dots, N$. The initial solution $u_h^0 \in \mathcal{K}_h$ is taken to be the discrete L^2 projection $\langle u_h^0, v_h \rangle = (u_0, v_h), \forall v_h \in \mathcal{S}_h$. Existence and uniqueness of the discrete Cahn-Hilliard equations has been established in [4]. The discrete Cahn-Hilliard equation is equivalent to the set valued saddle point block 2×2 nonlinear system (1) with $F = A + \partial I_{\mathcal{K}_h}$ and

$$A = \epsilon(\langle \lambda_p, 1 \rangle \langle \lambda_p, 1 \rangle + (\nabla \lambda_p, \nabla \lambda_q))_{p,q \in \mathcal{N}_h}, \quad (18)$$

$$B = -(\langle \lambda_p, \lambda_q \rangle)_{p,q \in \mathcal{N}_h}, \quad C = \tau((\nabla \lambda_p, \nabla \lambda_q))_{p,q \in \mathcal{N}_h}. \quad (19)$$

where \mathcal{N}_h stands for the set of vertices in \mathcal{T}_h , and $\lambda_p, p \in \mathcal{N}_h$ denote the standard nodal basis. We write the above in more compact notation as follows

$$A = \epsilon(K + mm^T), \quad B = -M, \quad C = \tau K, \quad (20)$$

where K and M are stiffness and mass matrices respectively.

IV. ITERATIVE SOLVER FOR CAHN-HILLIARD WITH OBSTACLE POTENTIAL

In [6], a nonsmooth Newton Schur method is proposed which is also interpreted as a preconditioned Uzawa iteration. For a given time step k , the Uzawa iteration reads:

$$u^{i,k} = F^{-1}(f^k - B^T w^{i,k}), \quad (21)$$

$$w^{i+1,k} = w^{i,k} + \rho^{i,k} \hat{S}_{i,k}^{-1}(B u^{i,k} - C w^{i,k} - g^k) \quad (22)$$

for the saddle point problem (1). Here i denotes the i^{th} Uzawa step and k denotes the k^{th} time step. Here f^k and g^k are defined in (16) and (17). The time loop starts with an initial value for $w^{0,0}$ which is taken arbitrary, and with the initial value $u^{0,0}$. The Uzawa iteration requires three main computations that we describe below.

A. Computing $u^{i,k}$

The first step (21) corresponds to a quadratic obstacle problem:

$$u^{i,k} = \arg \min_{v \in \mathcal{K}} \left(\frac{1}{2} \langle Av, v \rangle - \langle f^k - B^T w^{i,k}, v \rangle \right). \quad (23)$$

As mentioned in the introduction, this problem has been extensively studied during last decades [7], [8], [9], [11].

B. Computing $\hat{S}_{i,k}^{-1}(Bu^{i,k} - Cw^{i,k} - g^k)$

The descent direction $\hat{d}^{i,k+1} = \hat{S}_{i,k}^{-1}(Bu^{i,k} - Cw^{i,k} - g)$ in (22) is obtained as a solution of the following reduced linear block 2×2 system:

$$\begin{pmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & -C \end{pmatrix} \begin{pmatrix} \hat{u}^{i,k} \\ \hat{d}^{i,k} \end{pmatrix} = \begin{pmatrix} 0 \\ g + Cw^{i,k} - Bu^{i,k} \end{pmatrix}, \quad (24)$$

where

$$\hat{A} = TAT + \hat{T}, \quad \hat{B} = TB. \quad (25)$$

Here T and \hat{T} are defined as follows:

$$T = \text{diag} \begin{pmatrix} 0, \text{if } u^{i,k}(j) \in \{-1, 1\} \\ 1, \text{otherwise} \end{pmatrix}, \quad j = 1, \dots, |\mathcal{N}_h|, \quad (26)$$

$$\hat{T} = Id - T, \quad Id \in \mathbb{R}^{|\mathcal{N}_h| \times |\mathcal{N}_h|}, \quad (27)$$

where $u^{i,k}(j)$ is the j^{th} component of $u^{i,k}$. In other words, \hat{A} is the matrix obtained from A by replacing the i^{th} row and the i^{th} column by the unit vector e_i corresponding to the active sets identified by diagonal entries of T . Similarly, \hat{B} is the matrix obtained from B by annihilating columns, and \hat{B}^T is the matrix obtained from B^T by annihilating rows.

C. Computing Step Length $\rho^{i,k}$

The step length $\rho^{i,k}$ is computed using a bisection method. We refer the reader to [12, p. 88].

D. Algebraic Monotone Multigrid for Obstacle Problem

To solve the quadratic obstacle problem (21), we use the truncated monotone multigrid method proposed in [8]. However, here we use algebraic coarsening [13] that we describe briefly.

1) *Aggregation Based Coarsening*: We first discuss the coarsening for two-grid, the multilevel interpolations are applied recursively. In classical two-grid, a set of coarse grid unknowns is selected and the matrix entries are used to build interpolation rules that define the prolongation matrix P , and the coarse grid matrix A_c is computed from the following Galerkin formula

$$A_c = P^T A P. \quad (28)$$

In contrast to the classical two-grid approach, in aggregation based multigrid, first a set of aggregates G_i is defined. Let $|\mathcal{N}_{h,c}|$ be the total number of such aggregates, then the interpolation matrix P is defined as follows

$$P_{ij} = \begin{cases} 1, & \text{if } i \in G_j, \\ 0, & \text{otherwise,} \end{cases}$$

Here, $1 \leq i \leq |\mathcal{N}_h|$, $1 \leq j \leq |\mathcal{N}_{h,c}|$. Further, we assume that the aggregates G_i are such that

$$G_i \cap G_j = \emptyset, \text{ for } i \neq j \text{ and } \bigcup_i G_i = \{i \in \mathbb{N} : 1 \leq i \leq |\mathcal{N}_h|\}. \quad (29)$$

The matrix P defined above is a $|\mathcal{N}_h| \times |\mathcal{N}_{h,c}|$ matrix, but since it has only one non-zero entry (which are ‘‘one’’) per row, the matrix is compactly represented by a single array of length $\mathcal{N}_{h,c}$ storing the location of the non-zero entry on each row. The coarse grid matrix A_c may be computed as follows

$$(A_c)(i, j) = \sum_{k \in G_i} \sum_{l \in G_j} A(k, l),$$

where $1 \leq i, j \leq |\mathcal{N}_{h,c}|$, and $A(k, l)$ is the $(k, l)^{th}$ entry of A .

Numerous aggregation schemes have been proposed in the literature, but in this paper we consider the standard aggregation based on strength of connection [14, Appendix A, p. 413] where one first defines a set of nodes \mathcal{S}_i to which i is strongly negatively coupled, using the Strong/Weak coupling threshold β :

$$\mathcal{S}_i = \{j \neq i \mid A(i, j) < -\beta \max |A(i, k)|\}.$$

Then an unmarked node i is chosen such that priority is given to a node with minimal M_i , here M_i being the number of unmarked nodes that are strongly negatively coupled to i . For a complete algorithm of aggregation, the reader is referred to Notay [13], [15].

E. Preconditioner for Reduced Linear System

In Bosch et. al. [5], a preconditioner is proposed in the framework of semi-smooth Newton method combined with Moreau-Yosida regularization for the same problem. However, the preconditioner was constructed for a linear system which is different from the one we consider in (24). For convenience of notation, we rewrite the system matrix in (24) as follows

$$\mathcal{A}x = b, \quad (30)$$

where scripted \mathcal{A} above is

$$\mathcal{A} = \begin{pmatrix} \hat{A} & \hat{B}^T \\ \hat{B} & -C \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad (31)$$

where

$$x_1 = \hat{u}^{i,k}, \quad x_2 = \hat{d}^{i,k}, \quad b_1 = 0, \quad b_2 = g + Cw^{i,k} - Bu^{i,k}. \quad (32)$$

The preconditioner proposed in [5] has the following block lower triangular form

$$\mathcal{B} = \begin{pmatrix} \hat{A} & 0 \\ \hat{B} & -S \end{pmatrix}, \quad (33)$$

where $S = C + \hat{B}\hat{A}^{-1}\hat{B}^T$ is the Schur complement. Note that such preconditioners are also called inexact or preconditioned Uzawa preconditioners for the linear saddle point problems. By block 2×2 inversion formula we have

$$\mathcal{B}^{-1} = \begin{pmatrix} \hat{A} & 0 \\ \hat{B} & -S \end{pmatrix}^{-1} = \begin{pmatrix} \hat{A}^{-1} & 0 \\ S^{-1}\hat{B}^T\hat{A}^{-1} & -S^{-1} \end{pmatrix}. \quad (34)$$

Let \hat{S} be an approximation of Schur complement S in \mathcal{B} , then the new preconditioner $\hat{\mathcal{B}}$ and the corresponding preconditioned operator $\hat{\mathcal{B}}^{-1}\mathcal{A}$ are given as follows

$$\hat{\mathcal{B}} = \begin{pmatrix} \hat{A} & 0 \\ \hat{B} & -\hat{S} \end{pmatrix}, \quad \hat{\mathcal{B}}^{-1}\mathcal{A} = \begin{pmatrix} I & \hat{A}^{-1}\hat{B}^T \\ 0 & \hat{S}^{-1}S \end{pmatrix}. \quad (35)$$

Using (35) above, we can note the following trivial result.

Theorem 4.1: Let \mathcal{B} defined in (35) be a preconditioner for \mathcal{A} defined in (31), then there are $|\mathcal{N}_h|$ eigenvalues of $\mathcal{B}^{-1}\mathcal{A}$ equal to one, and the rest are the eigenvalues of the preconditioned Schur complement $\hat{S}^{-1}S$.

Remark 4.1: When using GMRES [16], right preconditioning is preferred. Similar result as for the left preconditioner above Theorem 4.1 holds.

The preconditioned system $\mathcal{B}^{-1}\mathcal{A}x = \mathcal{B}^{-1}b$ is given as follows

$$\begin{pmatrix} I & \hat{A}^{-1}\hat{B}^T \\ 0 & \hat{S}^{-1}S \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \hat{A}^{-1} & 0 \\ S^{-1}\hat{B}^T\hat{A}^{-1} & -S^{-1} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \quad (36)$$

from which we obtain the following set of equations

$$x_1 + \hat{A}^{-1}\hat{B}^T x_2 = \hat{A}^{-1}b_1, \quad (37)$$

$$\hat{S}^{-1}Sx_2 = S^{-1}(\hat{B}^T\hat{A}^{-1}b_1 - b_2). \quad (38)$$

Algorithm 4.1: Objective: Solve $\mathcal{B}^{-1}\mathcal{A}x = \mathcal{B}^{-1}b$

- 1) Solve for $x_2 : \hat{S}^{-1}Sx_2 = \hat{S}^{-1}(\hat{B}^T\hat{A}^{-1}b_1 - b_2)$
- 2) Set $x_1 = \hat{A}^{-1}(b_1 - \hat{B}^T x_2)$

Here if Krylov subspace method is used to solve for x_2 , then matrix vector product with S and a solve with \hat{S} is needed. However, when the problem size i.e. $|\mathcal{N}_h|$ is large, it won't be feasible to do exact solve with \hat{A} , and we need to solve it inexactly, for example, using algebraic multigrid methods. In the later case, the decoupling of x_1 and x_2 as in Algorithm 4.1 is not possible, and we shall need matrix vector product with \mathcal{A} (31) and a solve (forward sweep) with $\hat{\mathcal{B}}$. We discuss at the end of this subsection on how to take advantage of the special structure of \hat{A} in both cases of exact and inexact solves.

As a preconditioner \tilde{S} of S , we choose the preconditioner first proposed in [5]. The preconditioner is given as follows:

$$\tilde{S} = S_1\hat{A}^{-1}S_2 = -(\hat{B} - \tau^{1/2}K)\hat{A}^{-1}(\hat{B}^T - \tau^{1/2}\hat{A}), \quad (39)$$

where K is the stiffness matrix from (19). We observe that the preconditioned Schur complement $\tilde{S}^{-1}S$ is not symmetric, in

particular, not symmetric w.r.t. $\langle \cdot, \cdot \rangle_S$ or w.r.t. $\langle \cdot, \cdot \rangle_{\tilde{S}}$ which is a sufficient condition for the convergence of preconditioned conjugate gradient method [16, p. 262]. Hence we shall use GMRES in Saad [16, p. 269] that allows nonsymmetric preconditioners.

1) *Exact and Inexact Solve with \hat{A} :* In step 1 of Algorithm 4.1, we need to solve with \hat{A} when constructing right hand side, and also in step 2. Let P be a permutation matrix, then solving a system of the form $\hat{A}h = g$ is equivalent to solving $P^T\hat{A}h = P^Tg$ as P^T is nonsingular. With a change of variable $Py := h$, we then solve for y in

$$P^T\hat{A}Py = Pg, \quad (40)$$

and we set $h = Py$ to obtain the desired solution. By choosing P that renumbers the nodes corresponding to the coincidence set, we obtain

$$P^T\hat{A}P = \begin{pmatrix} I & \\ & R^T P^T \hat{A} P R \end{pmatrix}, \quad (41)$$

where R is the restriction operator defined as follows

$$R^T P^T \hat{A} P R = \hat{A}|_{\mathcal{N}_h \setminus \mathcal{N}_h^\bullet}, \quad (42)$$

where

$$\mathcal{N}_h^\bullet = \{i : T(i, i) = 0\} \quad (43)$$

is nothing but set of active nodes. Here R is explicitly given as follows:

$$R = \begin{pmatrix} \begin{pmatrix} 0 & 0 & \dots & 0 \\ \vdots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \ddots & \dots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix} \end{pmatrix}, \quad R \in \mathbb{R}^{|\mathcal{N}_h| \times |\mathcal{N}_h \setminus \mathcal{N}_h^\bullet|} \quad (44)$$

Let $\hat{K} = TKT$, we have

$$\begin{aligned} R^T P^T D P R &= R^T P^T \hat{A} P R \\ &= R^T P^T \epsilon(\hat{K} + \hat{m}\hat{m}^T) P R \\ &= \epsilon(R^T P^T \hat{K} P R + R^T P^T \hat{m}\hat{m}^T P R), \end{aligned}$$

where

$$R^T P^T \hat{K} P R = (P^T K P)|_{\mathcal{N}_h \setminus \mathcal{N}_h^\bullet}, \quad \hat{K} = TKT, \quad \hat{m} = Tm, \quad (45)$$

where m is the rank-one term defined in (20). For convenience of notation, we write

$$R^T P^T \hat{A} P R = \epsilon(\tilde{K} + \tilde{z}\tilde{z}^T), \quad (46)$$

where $\tilde{K} = R^T P^T \hat{K} P R$ and $\tilde{z} = R^T P^T \hat{z}$. In the new notation, we have

$$P^T \hat{A} P = \begin{pmatrix} I & \\ & \epsilon(\tilde{K} + \tilde{z}\tilde{z}^T) \end{pmatrix}. \quad (47)$$

Thus (40) now reads

$$\begin{pmatrix} I \\ \epsilon(\tilde{K} + \tilde{z}\tilde{z}^T) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = Pg =: \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}, \quad (48)$$

which reduces to two-set of equations

$$y_1 = g_1, \quad (49)$$

$$\epsilon(\tilde{K} + \tilde{z}\tilde{z}^T)y_2 = g_2. \quad (50)$$

To solve the latter, we use the Sherman-Woodbury formula

$$(\tilde{K} + \tilde{m}\tilde{m}^T)^+ = \tilde{K}^+ - \frac{\tilde{K}^+ \tilde{m}\tilde{m}^T \tilde{K}^+}{1 + \tilde{m}^T \tilde{K}^+ \tilde{m}}. \quad (51)$$

The AMG discussed before is used to pseudo-invert \tilde{K} , and we thus avoid constructing the dense matrix which would be the case when rank one term is explicitly added.

V. NUMERICAL EXPERIMENTS

All the experiments were performed in double precision arithmetic in MATLAB. The Krylov solver used was GMRES with subspace dimension of 200, and maximum number of iterations allowed was 300. The iteration was stopped as soon as the relative residual was below the tolerance of 10^{-7} .

We consider two samples of active set configurations that occur when a square region evolves as shown in figures 1 and 2. The region between the two squares and the circles is the interface between two bulk phases taking values +1 and -1; initially we chose random values between -0.3 and 0.5 in the interface region. The width of the interface is kept to be 10 times the chosen mesh size. The time step τ is chosen to be equal to ϵ . We compare various mesh sizes leading to number of grid points upto just above 1 million, and compare various values of epsilon for each mesh sizes. We observe that the number of iterations remain independent of the mesh size, however it depends on ϵ . But we observe that for a fixed epsilon, with finer mesh, the number of iterations actually decrease significantly. For example the number of iterations for $h = 2^{-7}, \epsilon = 10^{-6}$ is 84 but the number of iterations for $h = 2^{-10}, \epsilon = 10^{-6}$ is 38, a reduction of 46 iterations! It seems that finer mesh size makes the preconditioner more efficient. We also observe that the time to solve is proportional to number of iterations; the inexact solve for the (1,1) block remains optimal because the (1,1) block is essentially Laplacian for which AMG remains very efficient.

VI. CONCLUSION

For the solution of large scale optimization problem corresponding to Cahn-Hilliard problem with obstacle problem, we proposed an efficient preconditioning strategy that requires two elliptic solves. In our initial experiments upto over million unknowns, the preconditioner remains mesh independent. Although, for coarser mesh, there seems to be strong dependence on epsilon, but as the mesh becomes finer, we observe a significant reduction in iteration count, thus making the preconditioner effective and useful on finer meshes. It is likely that the iteration count continues to decrease on finer meshes.

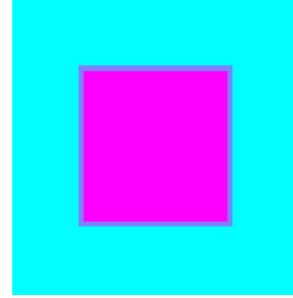


Fig. 1. Square

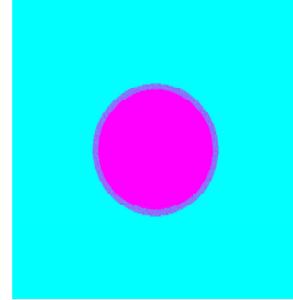


Fig. 2. Circle

TABLE I
COMPARE ITERATIONS COUNT FOR VARIOUS ϵ AND h

h	ϵ	square		circle	
		its	time	its	time
2^{-7}	e-2	6	1.40	6	1.37
	e-3	8	2.00	9	2.16
	e-4	20	4.50	22	4.76
	e-5	41	9.23	45	10.14
	e-6	77	17.83	83	22.02
2^{-8}	e-2	6	6.22	6	4.61
	e-3	5	4.88	5	5.12
	e-4	13	10.56	15	11.91
	e-5	31	24.39	35	27.31
	e-6	60	50.44	67	54.09
2^{-9}	e-2	5	18.13	6	18.50
	e-3	5	16.57	5	18.92
	e-4	8	28.21	9	32.09
	e-5	22	72.59	25	80.03
	e-6	47	166.46	52	180.58
2^{-10}	e-2	5	81.86	6	89.16
	e-3	5	81.69	5	85.23
	e-4	5	98.62	5	97.76
	e-5	14	218.52	15	254.48
	e-6	34	527.49	38	612.41

REFERENCES

- [1] J. W. Cahn and J. E. Hilliard, "Free Energy of a Nonuniform System. I. Interfacial Free Energy," *The Journal of Chemical Physics*, vol. 28, no. 2, 1958. [Online]. Available: <http://dx.doi.org/10.1063/1.1744102>
- [2] Y. Oono and S. Puri, "Study of phase-separation dynamics by use of cell dynamical systems. I. Modeling," *Physical Review A*, vol. 38, no. 1, 1987. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.38.434>
- [3] J. F. Blowey and C. M. Elliott, "The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy Part I: Numerical analysis," *European J. Appl. Math.*, no. 2, pp. 233–280, 1991. [Online]. Available: <http://dx.doi.org/10.1017/S095679250000053X>
- [4] —, "The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy Part II: Numerical analysis," *European J. Appl. Math.*, no. 3, 1992. [Online]. Available: <http://dx.doi.org/10.1017/S0956792500000759>
- [5] J. Bosch, M. Stoll, and P. Benner, "Fast solution of Cahn-Hilliard variational inequalities using implicit time discretization and finite elements," *Journal of Computational Physics*, vol. 262, pp. 38–57, 2014. [Online]. Available: <http://dx.doi.org/10.1016/j.jcp.2013.12.053>
- [6] C. Graeser and R. Kornhuber, "Nonsmooth newton methods for set-valued saddle point problems," *SIAM Journal on Numerical Analysis*, vol. 47, no. 2, pp. 1251–1273, 2009.
- [7] J. Barrett, R. Nurnberg, and V. Styles, "Finite element approximation of a phase field model for void electromigration," *SIAM J. Numer. Anal.*, vol. 42, no. 2, pp. 738–772, 2004. [Online]. Available: <http://dx.doi.org/10.1137/S0036142902413421>
- [8] R. Kornhuber, "Monotone multigrid methods for elliptic variational inequalities I," *Numerische Mathematik*, vol. 69, no. 2, pp. 167–184, 1994.
- [9] —, "Monotone multigrid methods for elliptic variational inequalities II," *Numerische Mathematik*, vol. 72, no. 4, pp. 481–499, 1996.
- [10] J. Mandel, "A Multilevel Iterative Method for Symmetric, Positive Definite Linear Complementarity Problems," *Applied Mathematics and Optimization*, vol. 11, pp. 77–95, 1984.
- [11] C. Graser and R. Kornhuber, "Multigrid Methods for Obstacle Problems," *Journal of Computational Mathematics*, vol. 27, no. 1, pp. 1–44, 2009.
- [12] C. Graser, "Convex Minimization and Phase Field Models," Ph.D. dissertation, FU Berlin, 2011.
- [13] P. Kumar, "Aggregation based on graph matching and inexact coarse grid solve for algebraic two grid," *International Journal of Computer Mathematics*, vol. 91, no. 5, pp. 1061–1081, 2014. [Online]. Available: <http://dx.doi.org/10.1080/00207160.2013.821115>
- [14] U. Trottenberg, C. Oosterlee, and A. Schuller, *Multigrid*. Academic Press, 2001. [Online]. Available: <http://www.academicpress.com>
- [15] Y. Notay, "An aggregation-based algebraic multigrid method," *Electronic Transactions on Numerical Analysis*, vol. 37, pp. 123–146, 2010. [Online]. Available: <http://dx.doi.org/10.1109/ISQED.2007.31>
- [16] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia: SIAM, 2003. [Online]. Available: <http://dx.doi.org/10.1137/1.9780898718003>