# Domain Decomposition Methods for a Complementarity Problem*

Haijian Yang[1] and Xiao-Chuan Cai[2]

[1] College of Mathematics and Econometrics, Hunan University, Changsha, Hunan, 410082, P. R. China, haijian.yang@colorado.edu
[2] Department of Computer Science, University of Colorado at Boulder, Boulder, CO 80309, USA, cai@cs.colorado.edu

Summary. We introduce a family of parallel Newton-Krylov-Schwarz methods for solving complementarity problems. The methods are based on a smoothed grid sequencing method, a semismooth inexact Newton method, and a two-grid restricted overlapping Schwarz preconditioner. We show numerically that such an approach is highly scalable in the sense that the number of Newton iterations and the number of linear iterations are both nearly independent of the grid size and the number of processors. In addition, the method is not sensitive to the sharp discontinuity that is often associated with obstacle problems. We present numerical results for some large scale calculations obtained on machines with hundreds of processors.

## 1 Introduction

Complementarity problems have many important applications (Cottle et al. [1992], Ferris and Pang [1997], Harker and Pang [1990]), and there are growing interests in developing efficient algorithms for solving these semismooth problems on large scale supercomputers. One popular approach is the class of semismooth methods which solves the problem by first reformulating it as a semismooth system of equations and then applying a generalized Newton method to solve this system. There are extensive theoretical and numerical results associated with this approach, see, e.g., (Luca et al. [1996], Fischer [1992], Kanzow [2004]). However, all existing approaches seem to have scalability problems in the sense that when the degree of freedoms in the problem increases the number of nonlinear or linear iterations increases drastically.

In this paper, we introduce a class of general purpose two-grid Newton-Krylov-Schwarz (NKS) algorithms for complementarity problems associated

* The research was supported in part by DOE under DE-FC-02-06ER25784, and in part by NSF under grants CCF-0634894 and CNS-0722023.

with partial differential equations. The methods are based on an inexact semismooth Newton method, a smoothed grid sequencing method and a two-level cascade restricted overlapping Schwarz preconditioning technique. It turns out, with an appropriate grid sequencing, the convergence rate of the semismooth Newton method can be made nearly independent of the number of unknowns of the system using either the Fischer-Burmeister function or the minimum function. Using the two-level restricted Schwarz preconditioner with sufficient overlap, the number of linear iterations also becomes nearly independent of the number of unknowns of the system. More importantly, both the linear and nonlinear iterations are nearly independent of the number of processors in our experiments on machines with hundreds of processors.

## 2 Semismooth function approaches for complementarity problems

Let $\Omega \in R^2$ be a bounded open domain on which a linear or nonlinear differential operator $L(u)$ is defined. Many problems can be described as finding a function $u(x)$ defined in certain space such that

$$\begin{cases} L\mathrm{u}(x) \geq 0, & x \in \Omega \\ \mathrm{u}(x) \geq \Phi, & x \in \Omega \\ (\mathrm{u}(x) - \Phi)L\mathrm{u}(x) = 0, & x \in \Omega \end{cases} \tag{1}$$

with some boundary conditions assumed for $u(x), x \in \partial\Omega$. Here $\Phi$ is given and often called an obstacle. Consider the following complementarity problem:

$$\begin{array}{ll} \text{find} & u_h \in R^n, \\ \text{such that} & u_h \geq \phi, \quad F(u_h) \geq 0, \quad (u_h - \phi)^T F(u_h) = 0, \end{array} \tag{2}$$

where $F(u_h) = (F_1(u_h), \cdots, F_n(u_h))^T : R^n \to R^n$ denotes a continuously differentiable function from the discretized version of $L(u)$, and $\phi \in R^n$ denotes the obstacle from the discretization of $\Phi$.

### 2.1 Semismooth Newton methods

Let $a_i = (u_h - \phi)_i$ and $b_i = F_i(u_h)$, the reformulations of the complementarity problem based on the Fischer-Burmeister function Fischer [1992] and the minimum function Kanzow [2004] are as follows:

$$\mathcal{F}_{FB}(a, b) := a + b - \sqrt{a^2 + b^2} = 0, \tag{3}$$

$$\mathcal{F}_{MIN}(a, b) := \min\{a, b\} = 0. \tag{4}$$

In fact, the Fischer-Burmeister function is differentiable everywhere except at the point $(a, b) = (0, 0)$, while the minimum function is piecewise smooth with its non-differentiable points forming the line $\{(a, b)^T \in R^2 : a = b\}$.

If we apply a Newton-type method to (3) and (4), respectively, then it leads to the class of inexact semismooth Newton methods, in which we need to solve a right-preconditioned Jacobian system

$$\|\mathcal{F}(u_h^k) + J_k M_k^{-1}(M_k s_k)\| \leq \max\{\eta_r \|\mathcal{F}(u_h^k)\|, \eta_a\},$$

where $J_k$ is a generalized Jacobian of $\mathcal{F}(u_h^k)$ to be introduced below, $\eta_r \in [0,1)$ is a relative tolerance, $\eta_a \in [0,1)$ is an absolute tolerance, and $M_k^{-1}$ is an overlapping Schwarz preconditioner (Smith et al. [1996], Toselli and Widlund [2005]).

For both the Fischer-Burmeister function and the minimum function, the generalized Jacobian matrix $J_k$ is of the form

$$J_k = D_a^k + D_b^k F'(u_h^k) \tag{5}$$

with diagonal matrices (depending on the iteration index $k$)

$$D_a^k = \mathrm{diag}(d_{a_1}, \ldots, d_{a_n}), \qquad D_b^k = \mathrm{diag}(d_{b_1}, \ldots, d_{b_n}). \tag{6}$$

The values of $D_a^k$ and $D_b^k$ in (6) corresponding to the Fischer-Burmeister function take the form

$$d_{a_i} := \begin{cases} 1 - a_i/\sqrt{a_i^2 + b_i^2}, & \text{if } a_i^2 + b_i^2 \neq 0, \\ 1, & \text{if } a_i^2 + b_i^2 = 0, \end{cases}$$

and

$$d_{b_i} := \begin{cases} 1 - b_i/\sqrt{a_i^2 + b_i^2}, & \text{if } a_i^2 + b_i^2 \neq 0, \\ 1, & \text{if } a_i^2 + b_i^2 = 0. \end{cases}$$

Similarly, when using the minimum function, the values of $D_a^k$ and $D_b^k$ in (6) assume the form

$$d_{a_i} := \begin{cases} 1, & a_i < b_i, \\ 0, & a_i \geq b_i, \end{cases}$$

and

$$d_{b_i} := \begin{cases} 0, & a_i < b_i, \\ 1, & a_i \geq b_i. \end{cases}$$

When using a Newton type method to solve complementarity problems, one of the major problems is the deterioration of the convergence rate when the mesh is refined. We here propose a smoothed grid sequencing technique: First, compute the solution $u_H^*$ of the nonlinear system $\mathcal{F}_H(u_H) = 0$ on a coarse grid. Second, interpolate the coarse solution to obtain $\tilde{u}_h^0 = I_H^h u_H^*$, which is then smoothed by replacing its value at each grid point with the following weighed average of its neighboring values:

$$\begin{array}{ccccc} \frac{1}{16} & & \frac{1}{8} & & \frac{1}{16} \\ & \searrow & \downarrow & \swarrow & \\ \frac{1}{8} & \rightarrow & \frac{1}{4} & \leftarrow & \frac{1}{8} \\ & \nearrow & \uparrow & \nwarrow & \\ \frac{1}{16} & & \frac{1}{8} & & \frac{1}{16}. \end{array}$$

The smoothed vector is then used as the initial guess for the fine grid Newton iteration. In Fig. 1, we show the surface plots of the nonlinear system $\mathcal{F}_h(I_H^h u_H^*)$ on a fine grid without smoothing (left figure), and $\mathcal{F}_h(u_H^0)$ with one sweep of smoothing (right figure) for an obstacle problem. More details of this problem will be discussed in the numerical experiments section.

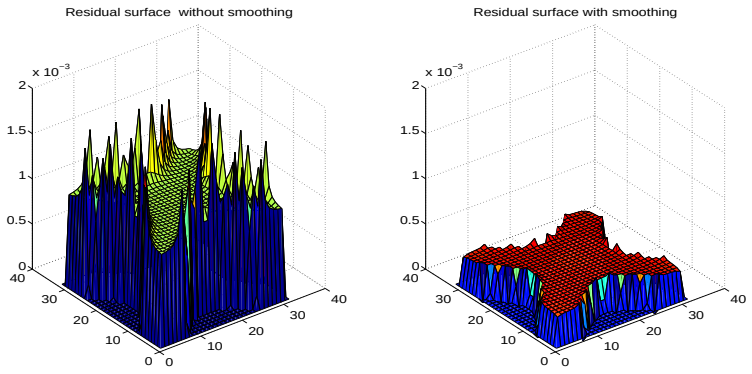


**Fig. 1.** The effect of smoothing of the interpolated coarse grid solution on the fine grid.

### 2.2 Schwarz preconditioner

Let $J$ be the Jacobian matrix on the fine grid and $R_i^\delta$ and $R_i^0$, the restriction operator from $\Omega$ to its overlapping and non-overlapping subdomains, respectively. Then the one-level restricted additive Schwarz (RAS) preconditioner Cai and Sarkis [1999] is

$$M_{RAS}^{-1} = \sum_{i=1}^{N_s} (R_i^0)^T J_i^{-1} R_i^\delta. \tag{7}$$

with $J_i = R_i^\delta \; J \; (R_i^\delta)^T$ and $N_s$ is the number of subdomains, which is the same as the number of processors.

Let $J_c$ be the Jacobian matrix on the coarse grid and $I_h^H$ a restriction operator from the fine grid to the coarse grid. Then the two-level restricted Schwarz preconditioner is

$$M^{-1} = M_c^{-1} + M_{RAS}^{-1} - M_{RAS}^{-1} J M_c^{-1}$$

with $M_c^{-1} = (I_h^H)^T J_c^{-1} I_h^H$. We refer to Smith et al. [1996], Toselli and Widlund [2005] for further analysis and examples of Schwarz preconditioning techniques.

## 3 Numerical experiments

We report some results of our numerical experiments. Our solver is implemented using PETSc (Balay et al. [2009]). We consider an obstacle problem: find $u(x)$ such that

$$\begin{cases} -\Delta u(x) + C \geq 0, & x \in \Omega, \\ u(x) \geq -\mathrm{d}(x, \partial\Omega), & x \in \Omega, \\ (u(x) + \mathrm{d}(x, \partial\Omega))(-\Delta u(x) + C) = 0, & x \in \Omega, \\ u(x) = 0, & x \in \partial\Omega, \end{cases} \tag{8}$$

where the $\mathrm{d}(x, \partial\Omega)$-operator measures the distance from a point $x$ to the domain boundary $\partial\Omega$, and the parameter $C = 5$.

For the discretization we use the standard second-order five-point finite difference method on a uniform grid. The initial guess $u_h^0$ for the global Newton iteration is the obstacle from the discretization of $-\mathrm{d}(x, \partial\Omega)$ in (8). We stop the fine grid Newton iteration if

$$\|\mathcal{F}(u_h^k)\| \leq \max\{10^{-6}\|\mathcal{F}(u_h^0)\|, 10^{-10}\}.$$

The fine grid Jacobian system is solved with GMRES(30), and the iteration is stopped if the tolerance

$$\|\mathcal{F}(u_h^k) + J_k s_k\| \leq \max\{10^{-4}\|\mathcal{F}(u_h^k)\|, 10^{-10}\}$$

is satisfied. The subdomain problems are solved with LU factorization. Throughout this section, "*np*" stands for the number of processors which is the same as the number of subdomains, "INB' the number of inexact Newton iterations, "RAS" the number of RAS preconditioned GMRES iterations, and "Time" the total compute time in seconds.

### 3.1 One-level results

We first study the one-level method with overlap $\delta = 3$. As shown in Table 1, on a fixed grid, the number of Newton iterations is independent of the number of processors, but the number of GMRES iterations increases as the number of processors increases for both the Fischer-Burmeister function and the minimum function. The major problem with the one-level method shows up, if we look at the scalability for a fixed number of processors. For each row in the table, every time we refine the grid by a factor of 2, the number of Newton iterations doubles. This problem prohibits the use of the method for high resolution applications.

**Table 1.** Results for the one-level method with overlap $\delta = 3$.

| Mesh | $256 \times 256$ | | | $512 \times 512$ | | | $1024 \times 1024$ | | | $2048 \times 2048$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $np$ | INB | RAS | Time | INB | RAS | Time | INB | RAS | Time | INB | RAS | Time |
| The Fischer-Burmeister function | | | | | | | | | | | | |
| 64 | 82 | 11.4 | 3.3 | 162 | 14.7 | 32.5 | 320 | 19.1 | 384.1 | 639 | 24.4 | 4781.1 |
| 128 | 82 | 13.6 | 2.1 | 162 | 17.5 | 17.4 | 320 | 22.2 | 180.4 | 639 | 30.8 | 2236.3 |
| 256 | 82 | 14.4 | 1.5 | 162 | 18.9 | 10.2 | 320 | 24.3 | 95.9 | 639 | 34.1 | 1110.0 |
| 512 | 82 | 17.2 | 1.1 | 162 | 22.6 | 7.5 | 320 | 32.3 | 62.3 | 639 | 38.5 | 568.9 |
| The minimum function | | | | | | | | | | | | |
| 64 | 80 | 11.7 | 2.9 | 159 | 15.3 | 29.5 | 319 | 19.9 | 361.9 | 637 | 26.4 | 4673.0 |
| 128 | 80 | 14.0 | 1.9 | 159 | 18.3 | 16.1 | 319 | 23.7 | 173.5 | 637 | 33.7 | 2201.8 |
| 256 | 80 | 14.9 | 1.4 | 159 | 19.7 | 9.7 | 319 | 26.1 | 94.4 | 637 | 36.5 | 1104.8 |
| 512 | 80 | 17.7 | 1.3 | 159 | 23.8 | 7.3 | 319 | 34.5 | 62.0 | 637 | 41.1 | 567.5 |

### 3.2 Two-level results

In this subsection, we present the numerical results using the two-level approach in which a coarse grid is used in the nonlinear solver for generating a better initial guess and also in the linear solver for generating part of the Schwarz preconditioner. In the test, the initial guess for the global Newton iteration on the coarse grid is the obstacle $\phi$ in (2), and the tolerance conditions on the fine grid are the same as in the one-level method. We stop the coarse grid Newton iteration if

$$\|\mathcal{F}_H(u_H^k)\| \leq \max\{10^{-4}\|\mathcal{F}_H(u_H^0)\|, 10^{-10}\}.$$

In the test, the Jacobian system on the coarse grid is solved with a one-level RAS preconditioned GMRES(30) with the following stopping condition

$$\|\mathcal{F}_H(u_H^k) + J_H^k M_{H,RAS}^{-1}(M_{H,RAS} s_k)\| \leq \max\{10^{-4}\|\mathcal{F}_H(u_H^k)\|, 10^{-10}\},$$

where $M_{H,RAS}^{-1}$ is defined similar to (7) on the coarse grid. The subdomain problems are solved with LU factorization.

Using $\delta = 6$ and $\delta_c = 3$, we solve the test problem on several different fine grids with the two-level method and the results are summarized in Table 2, for both the Fischer-Burmeister function and the minimum function. The main concern is the size of the coarse grid $H$, which is taken as $h/2$, $h/4$ and $h/8$, where $h$ is the size of the fine grid. In terms of the total number of Newton iterations, $H = h/2$ is certainly the best, but $H = h/8$ offers the best results in terms of the total compute time. Note that some cases, marked as "$*$", for the fine grid $256 \times 256$ are not available because the corresponding coarse grids are too coarse and the coarse Newton may not converge. The compute time includes the coarse grid calculation of the initial guess, the smoothing of the coarse solution, and the solving of the fine grid problem. Note that

**Table 2.** Results with different fine and coarse grids. The overlapping sizes of the coarse grid and the fine grid are $\delta_c = 3$ and $\delta = 6$, respectively. The preconditioner is the two-level RAS. $h$ and $H$ are the fine and coarse grid sizes, respectively.

| Mesh | $256 \times 256$ | | | $512 \times 512$ | | | $1024 \times 1024$ | | | $2048 \times 2048$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $np$ | INB | RAS | Time | INB | RAS | Time | INB | RAS | Time | INB | RAS | Time |
| The Fischer-Burmeister function | | | | | | | | | | | | |
| $H = h/2$ | | | | | | | | | | | | |
| 64 | 6 | 10.8 | 2.4 | 5 | 15.8 | 11.5 | 4 | 21.8 | 76.0 | 4 | 26.3 | 848.5 |
| 128 | 6 | 13.0 | 2.2 | 5 | 18.8 | 8.5 | 4 | 26.3 | 49.0 | 4 | 35.8 | 536.1 |
| 256 | 6 | 13.8 | 1.8 | 5 | 20.8 | 6.0 | 4 | 30.0 | 33.6 | 4 | 37.8 | 291.9 |
| 512 | 6 | 19.3 | 2.8 | 5 | 24.4 | 5.9 | 4 | 34.6 | 32.9 | 4 | 43.0 | 206.5 |
| $H = h/4$ | | | | | | | | | | | | |
| 64 | $*$ | | | 7 | 15.3 | 6.2 | 7 | 19.0 | 33.9 | 6 | 25.8 | 201.7 |
| 128 | $*$ | | | 7 | 18.4 | 4.8 | 7 | 22.6 | 21.9 | 6 | 32.7 | 120.8 |
| 256 | $*$ | | | 7 | 20.3 | 3.6 | 7 | 25.4 | 25.4 | 6 | 38.7 | 71.6 |
| 512 | $*$ | | | 7 | 23.9 | 4.3 | 7 | 33.9 | 12.1 | 6 | 43.7 | 56.6 |
| $H = h/8$ | | | | | | | | | | | | |
| 64 | $*$ | | | 9 | 15.7 | 5.9 | 9 | 19.7 | 31.0 | 8 | 26.4 | 169.9 |
| 128 | $*$ | | | 9 | 19.1 | 4.6 | 9 | 23.7 | 18.8 | 8 | 33.9 | 99.2 |
| 256 | $*$ | | | 9 | 21.0 | 3.5 | 9 | 26.6 | 10.7 | 9 | 36.6 | 54.3 |
| 512 | $*$ | | | $*$ | | | 9 | 34.3 | 10.8 | 8 | 45.3 | 34.4 |
| The minimum function | | | | | | | | | | | | |
| $H = h/2$ | | | | | | | | | | | | |
| 64 | 2 | 14.0 | 1.3 | 3 | 14.3 | 8.3 | 3 | 17.7 | 60.4 | 2 | 31.5 | 777.7 |
| 128 | 2 | 16.5 | 1.2 | 3 | 16.7 | 6.0 | 3 | 21.7 | 40.6 | 2 | 39.0 | 446.3 |
| 256 | 2 | 17.5 | 1.0 | 3 | 18.3 | 4.3 | 3 | 26.3 | 26.3 | 2 | 49.5 | 260.6 |
| 512 | 2 | 25.0 | 1.5 | 3 | 20.0 | 3.9 | 3 | 29.7 | 22.2 | 2 | 57.0 | 178.5 |
| $H = h/4$ | | | | | | | | | | | | |
| 64 | $*$ | | | 4 | 15.5 | 3.9 | 4 | 18.8 | 21.1 | 5 | 19.6 | 160.3 |
| 128 | $*$ | | | 4 | 18.8 | 2.9 | 4 | 24.8 | 14.6 | 5 | 23.2 | 93.1 |
| 256 | $*$ | | | 4 | 20.3 | 2.3 | 4 | 28.0 | 9.2 | 5 | 24.4 | 50.8 |
| 512 | $*$ | | | 4 | 24.5 | 2.8 | 4 | 33.0 | 7.7 | 5 | 30.6 | 39.7 |
| $H = h/8$ | | | | | | | | | | | | |
| 64 | $*$ | | | 7 | 15.6 | 4.6 | 6 | 21.3 | 21.7 | 6 | 25.7 | 126.9 |
| 128 | $*$ | | | 7 | 18.6 | 3.5 | 6 | 24.8 | 13.1 | 6 | 33.3 | 74.2 |
| 256 | $*$ | | | 7 | 20.7 | 2.7 | 6 | 26.7 | 7.3 | 6 | 37.0 | 37.0 |
| 512 | $*$ | | | $*$ | | | 6 | 34.8 | 7.4 | 6 | 42.5 | 25.4 |

the minimum function approach is always faster than the Fischer-Burmeister function approach in terms of all measures.

We should mention that the use of smoothed grid sequencing plays an important role in the two-level methods. In Fig. 1, the surface plots of the residual function before and after the smoothing are shown and they are quite different. The cost of the smoothing step is very small and fewer number of Newton iterations is needed as a result of the smoothing.

## 4 Some final remarks

We have developed a family of highly scalable, two-grid algorithms for solving general complementarity problems. In addition to the fine grid, on which the PDE is discretized and the complementarity problem is solved, a coarse grid is introduced to accelerate the nonlinear convergence, and to precondition the linear Jacobian solver in a semismooth Newton iteration. With the help of a smoothed grid sequencing, a semismooth Newton method and a two-level restricted Schwarz preconditioner, we have showed numerically that the family of two-grid Newton-Krylov-Schwarz algorithms has a fast and robust convergence and that the rate of convergence is nearly independent of the number of unknowns of the problem and the number of processors. Surprisingly good results were obtained for solving some rather difficult obstacle problems with millions of unknowns and on machines with up to 512 processors.

## References

S. Balay, K. Buschelman, W. D. Gropp, D. Kaushik, M. Knepley, L. C. McInnes, B. F. Smith, and H. Zhang. *PETSc Users Manual*. Argonne National Laboratory, 2009.

X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM J. Sci. Comput.*, 21:792–797, 1999.

R. Cottle, J.-S. Pang, and R. Stone. *The Linear Complementarity Problem*. Academic Press, Boston, MA, 1992.

M. C. Ferris and J.-S. Pang. Engineering and economic applications of complementarity problems. *SIAM Rev.*, 39:669–713, 1997.

A. Fischer. A special newton-type optimization method. *Optimization*, 24: 269–284, 1992.

P. T. Harker and J.-S. Pang. Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications. *Math. Prog.*, 48:161–220, 1990.

C. Kanzow. Inexact semismooth newton methods for large-scale complementarity problems. *Opt. Meth. Software*, 19:309–325, 2004.

T. De Luca, F. Facchinei, and C. Kanzow. A semismooth equation approach to the solution of nonlinear complementarity problems. *Math. Prog.*, 75: 407–439, 1996.

B. Smith, P. Bjørstad, and W. Gropp. *Domain Decomposition: Parallel Multilevel Methods for Eliptic Partial Differential Equations*. Cambridge University Press, New York, NY, 1996.

A. Toselli and O. Widlund. *Domain Decomposition Methods – Algorithms and Theory*. Springer-Verlag, Berlin, 2005.