

# AN EFFICIENT MULTIGRID STRATEGY FOR LARGE-SCALE MOLECULAR MECHANICS OPTIMIZATION

JINGRUN CHEN AND CARLOS J. GARCÍA-CERVERA

ABSTRACT. Static mechanical properties of materials require large-scale nonlinear optimization of the molecular mechanics model under various controls. This paper presents an efficient multigrid strategy to solve such problems. This strategy approximates solutions on grids in a *quasi-atomistic* and *inexact* manner, transfers solutions on grids following a coarse-to-fine (oneway) schedule, and finds physically relevant minimizers with linear scaling complexity. Compared to the full multigrid method which has the same complexity, the prefactor of this strategy is orders of magnitude smaller. Consequently, the required CPU time of this strategy is orders of magnitude smaller than that of the full multigrid method, and is smaller than that of the brute-force optimization for systems with more than 200,000 atoms. Considerable savings are found if the number of atoms becomes even larger due to the super-linear scaling complexity of the brute-force optimization. For systems with 1,000,000 atoms (over three million degrees of freedom), on average a more than 70% reduction of CPU time is observed regardless of the type of defects, including vacancies, dislocations, and cracks. In addition, linear scalability of the proposed strategy is tested in the presence of a dislocation pair for systems with more than 100 million atoms (over 400 million degrees of freedom).

## 1. INTRODUCTION

In the classical picture, atoms constitute matter and their interaction is described by a potential energy function. For a given system of atoms at zero (low) temperature, its equilibrium configuration is obtained by minimizing the potential energy with respect to atomic positions, known as the molecular mechanics (atomistic) model (see, e.g., [8]). The potential energy is highly nonlinear and the number of minimizers grows exponentially [46]

---

*Date:* May 21, 2017.

*2000 Mathematics Subject Classification.* 65K05, 65N55, 74G65, 90C06.

*Key words and phrases.* nonconvex optimization, molecular mechanics, multigrid, linear scaling, efficiency.

We thank Professor Weinan E for suggesting the topic studied here and Professor Pingbing Ming for helpful discussions. This work was supported by National Science Foundation grant DMS-1217315. JC also acknowledges support from National Natural Science Foundation of China via grant 21602149. CJGC also acknowledges support from the Bizkaia Talent program through the Basque Center for Applied Mathematics (BCAM).

with respect to the number of atoms. This provides an intrinsically discrete and nonconvex optimization problem. It is known that the stable configuration of the system without external loading is given by the global minimizer of the potential energy. However, under external loading, the equilibrium configuration of the system is actually a local minimizer (see, e.g., [44, 15]). How to efficiently find the equilibrium configuration in the latter case is the primary objective of the current work.

Due to the intrinsic discreteness of the model, a natural idea to solve it is to apply a minimization technique of the conjugate-gradient type or the (quasi-)Newton type in search of local minimizers. This is known as brute-force optimization. For many atomistic systems of interest, atoms only interact with nearby atoms in a short range. Typically only local adjustments are performed by each minimization step. Even though the computational complexity of one step minimization is proportional to the number of atoms, the number of steps to achieve the convergence may depend on the number of atoms, resulting a super-linear scaling complexity overall. Moreover, the presence of dislocations and cracks also induces long-range effects on the displacement field, which further increases the cost of brute-force optimization approaches.

Over the past two decades, there has been a significant amount of research devoted to this problem; see for example [14]. One of the most popular approaches is the Quasicontinuum method (QC) introduced by Tadmor, Ortiz, and Phillips [41]. In this approach, the atomistic model is kept only over the region where atomic details are essential while a coarse-grained model is adopted elsewhere in the framework of the domain decomposition method [42]. Typical coarse-grained models are elasticity models [30], which have much fewer degrees of freedom (dofs) and result in sub-linear scaling methods. The elasticity model part can capture the long-range elastic displacement field effectively while the atomistic model part can resolve atomic details of interest. However, due to the lack of consistency at the atomistic/continuum interface, solutions in this case may not converge to the solution of the fully atomistic model; see for example [32, 11, 12] for explicit error estimates of the discrepancy from 1D to 3D. A significant amount of work has been done to solve this issue; see for example [24, 25] for the early analysis results of the QC method, [35, 34, 38, 39, 40, 36, 37] for energy based approaches, [45, 23, 27, 28] for force based approaches, [1, 2] for complex lattices, and review papers [31, 29] and references therein. However, an energy-based and consistent method for the general case is still lacking. One more issue along with this work is the implicit assumption that the system can be well described by the elasticity models except near a localized region described by an atomistic model. Therefore, prior knowledge of the system is required and then regions with different models can be preselected. In this paper, instead, we search along a different direction and focus on fully atomistic approaches. We adopt a multigrid methodology, and illustrate how it can be used for the efficient numerical simulation of fully atomistic

systems. Coarse-grained descriptions at different coarse levels are obtained via averaging, thereby eliminating the need for an analytically constructed continuum model.

The multigrid method was introduced as a way to solve the linear systems of equations arising from the discretization of linear elliptic partial differential equations (PDEs) with optimal (linear) scaling complexity [5, 18]. To solve the discretized problem on a fine grid, a sequence of grids ranging from the coarsest level to the fine level are constructed with associated discretized problems at each level. Roughly speaking, each level can efficiently capture the grid-dependent information of the fine solution using the smoothing step. This information can then be transferred to a finer grid using interpolation and to a coarser grid using restriction. There are several types of multigrid methods with optimal scaling, depending on how the grid-dependent information is transferred. Popular choices include V-cycle multigrid, W-cycle multigrid, oneway (cascadic) multigrid, and full multigrid (FMG); see [7, 43] for example. For nonlinear problems, depending on how the linearization and grids are combined, there are global linearization based approaches [7, 43] where a multigrid method is used for each linearization step, and local linearization based approaches, such as the full approximation scheme (FAS) [5, 7, 43] where linearization is imbedded in a multigrid method. These two approaches may find different solutions if the given problem has multiple solutions [19].

The application of multigrid methods for the molecular mechanics model was first proposed by Brandt [6]. However, to the best of our knowledge, no specific illustration of the idea or its efficiency has been presented. An interesting idea, connected to the approach presented here, was developed by Goedecker *et al.* [17]. In their work, atomic forces are transferred to a regular mesh, where additional displacements are computed using a linear elasticity model, and these displacements are transferred back to the atomic locations. The efficiency of the approach presented in [17] was illustrated by studying silicon crystals with point defects. A more general approach was proposed by Chen and Ming [9]. The coarse-grid operator was constructed by the Cauchy-Born rule [3], and then transferred to the atomistic level in the framework of the oneway multigrid method. Since the local linearization was first done at the coarse level, many unphysical local minimizers were automatically bypassed. Unfortunately the efficiency of this approach relies heavily on the validity of the Cauchy-Born rule [15]. Even though a constrained Cauchy-Born elasticity can be used to improve its efficiency for nanoindentations [10], a generic and efficient multigrid method for the molecular mechanics model is still lacking.

Two main issues need to be solved for an efficient and robust approach. First, there should not be any assumption on defects in a system of interest. This is important from the application perspective for systems where many defects coexist in high concentration; such is the case of semiconductor doping and high-density dislocations. Second, there should not be any

assumption on the validity of coarse-grained models. The usage of elasticity models will limit the applicability of a method. In this paper, we propose an efficient multigrid strategy to solve the molecular mechanics model by incorporating the oneway multigrid [9] and the discretized coarse-grid problems [6]. This strategy inherits the advantage of the oneway multigrid method, including automatic bypassing of many unphysical minimizers, insensitivity to parameters of nonlinear iterative solvers, and the optimal scaling complexity [9], and does not rely on any kind of continuum theory because of the usage of discretized coarse-grid problems [6]. Compared to the oneway multigrid method in [9], it retains optimal scaling complexity even for cases where vacancies, dislocations, and cracks are present. Compared to the full approximation scheme - full multigrid method [6], it has a prefactor in terms of CPU time which is several orders of magnitude smaller and outperforms brute-force optimization algorithms of the conjugate-gradient type or the quasi-Newton type if the number of atoms is greater than 200,000. For systems with one million atoms, on average a more than 70% reduction of CPU time is observed regardless of the type of defects, including vacancies, dislocations, and cracks.

The remainder of the paper is organized as follows: In Section 2 we describe the molecular mechanics model, the full approximation scheme - full multigrid method, the oneway multigrid strategy, and present detailed comparisons of these two multigrid methods using two examples: a perfect crystal under a shear deformation and a crystal with a sole vacancy under a shear deformation. Examination of the oneway multigrid strategy for more general cases, including vacancies under a shear deformation, dislocations under a shear deformation, and cracks under a tensile deformation is presented in Section 3. Possible additional applications are discussed in Section 4.

## 2. MULTIGRID METHODS

**2.1. Molecular mechanics model.** Consider a system with  $N$  atoms and the potential energy function  $V$ , and let  $\mathbf{f}_i$  be the external force on the  $i$ th atom. The total energy of the system can be written as

$$E^{\text{tot}}(\mathbf{y}) = V(\mathbf{y}_1, \dots, \mathbf{y}_N) - \sum_{i=1}^N \mathbf{f}_i \cdot \mathbf{y}_i$$

with  $\mathbf{y}_i$  the position of the  $i$ th atom in the deformed state. The atomic configuration is then given by the following minimization problem:

$$(1) \quad \{\mathbf{y}_1, \dots, \mathbf{y}_N\} = \operatorname{argmin} E^{\text{tot}}(\mathbf{y})$$

with  $\mathbf{y}$  subject to certain boundary condition.

We restrict ourselves to the case where the underlying systems are crystals and leave possible applications of the multigrid strategy we present to

general molecular systems as a future project. Atoms in a crystal are arranged regularly over a lattice. For a simple lattice, any lattice site  $\mathbf{x}$  takes the following form

$$\mathbf{x} = \sum_{i=1}^3 \nu^i \mathbf{e}_i + \mathbf{o},$$

where  $\{\nu^i\}_{i=1}^3 \in \mathbb{Z}$ ,  $\{\mathbf{e}_i\}_{i=1}^3$  are the basis vectors, and  $\mathbf{o}$  is a particular lattice site which can be taken as the origin. In general, any lattice can be regarded as a union of congruent simple lattices with shift vectors among them.

The displacement of the  $i$ th atom is defined as

$$\mathbf{u}_i = \mathbf{y}_i - \mathbf{x}_i,$$

where  $\mathbf{x}_i$  is the position of the  $i$ th atom in the undeformed configuration. (1) can be reformulated as

$$(2) \quad \{\mathbf{u}_1, \dots, \mathbf{u}_N\} = \operatorname{argmin} E^{\text{tot}}(\mathbf{u})$$

with

$$E^{\text{tot}}(\mathbf{u}) = V(\mathbf{x}_1 + \mathbf{u}_1, \dots, \mathbf{x}_N + \mathbf{u}_N) - \sum_{i=1}^N \mathbf{f}_i \cdot (\mathbf{x}_i + \mathbf{u}_i).$$

Since the number of minimizers scales exponentially with respect to  $N$  [46], multigrid methods with global linearization may find distinct minimizers [9]. These minimizers may have multiple fractures inside the system, which is considered to be physically irrelevant. In what follows, we shall consider multigrid methods with local linearization: full approximation scheme - full multigrid (FAS-FMG) and oneway multigrid.

Following [6], we first give a general description of the atomistic model (2) at any coarse level (denoted by  $k$  with grid size  $h$ ). We consider a domain  $\Omega$  that contains the atomistic system, and a triangulation, defined by  $n$  nodes with reference coordinates  $\{\boldsymbol{\xi}_1^k, \dots, \boldsymbol{\xi}_n^k\}$ . We associate to each node a displacement,  $\{\mathbf{u}_1^k, \dots, \mathbf{u}_n^k\}$ . Note that the points  $\{\boldsymbol{\xi}_1^k, \dots, \boldsymbol{\xi}_n^k\}$  may not necessarily be atomic positions  $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ . We define basis functions centered on each node by

$$w_j^k(\mathbf{x}) = w\left(\frac{\mathbf{x} - \boldsymbol{\xi}_j^k}{h}\right),$$

where for simplicity we have fixed a function  $w$  with compact support. The atomic displacement of an atom at  $\mathbf{x}$  can be constructed by

$$(3) \quad \mathbf{u}(\mathbf{x}) = \sum_{j=1}^n w_j^k(\mathbf{x}) \mathbf{u}_j^k = \sum_{j=1}^n w\left(\frac{\mathbf{x} - \boldsymbol{\xi}_j^k}{h}\right) \mathbf{u}_j^k \quad \forall \mathbf{x} \in \Omega.$$

The atomic position of atom  $i$ , for  $i = 1, \dots, N$  in the atomistic model (1) then can be written as

$$\mathbf{y}_i = \mathbf{x}_i + \sum_{j=1}^n w \left( \frac{\mathbf{x}_i - \boldsymbol{\xi}_j^k}{h} \right) \mathbf{u}_j^k.$$

The discrete coarse-grid model reads as

$$(4) \quad \{\mathbf{u}_1^k, \dots, \mathbf{u}_n^k\} = \operatorname{argmin} E^{\text{tot}}(\mathbf{u}^k)$$

with

$$(5) \quad \begin{aligned} E^{\text{tot}}(\mathbf{u}^k) &= E^{\text{tot}}(\mathbf{y}_1, \dots, \mathbf{y}_N) \\ &= V \left( \mathbf{x}_1 + \sum_{j=1}^n w \left( \frac{\mathbf{x}_1 - \boldsymbol{\xi}_j^k}{h} \right) \mathbf{u}_j^k, \dots, \mathbf{x}_N + \sum_{j=1}^n w \left( \frac{\mathbf{x}_N - \boldsymbol{\xi}_j^k}{h} \right) \mathbf{u}_j^k \right) \\ &\quad - \sum_{i=1}^N \mathbf{f}_i \cdot \left( \mathbf{x}_i + \sum_{j=1}^n w \left( \frac{\mathbf{x}_i - \boldsymbol{\xi}_j^k}{h} \right) \mathbf{u}_j^k \right). \end{aligned}$$

In this coarse-grid energy, the contribution of each atom is explicitly accounted while only the displacements at the grid-points are treated as unknowns in the coarse-grid minimization problem (4). This makes the computational complexity of the coarse-grid model  $\mathcal{O}(N) + \mathcal{O}(n)$ . To distinguish between this and the QC method [41], we refer it to as the *quasi-atomistic* description.

The basis function is normally chosen such that

$$\sum_j w \left( \frac{\mathbf{x} - \boldsymbol{\xi}_j^k}{h} \right) = 1 \quad \forall \mathbf{x} \in \Omega,$$

which means a constant displacement field is preserved after interpolation. Since hexahedron elements are used for all simulations, a typical choice is the tri-linear interpolation, i.e., the linear interpolation in each coordinate direction. One can easily check it preserves a constant displacement field.

While equilibrium states of the atomistic system under external loading are only local minimizers, we shall use conjugate-gradient type or Newton type nonlinear solvers. Since Hessians of considered models are difficult to compute, we only consider conjugate-gradient methods and quasi-Newton methods where the explicit evaluation of the Hessian matrix is not required. The gradient of the energy at atomic location  $i$ , for  $i = 1, \dots, N$  in the atomistic model is

$$(6) \quad \mathbf{g}_i(\mathbf{y}) = \frac{\partial E^{\text{tot}}(\mathbf{y})}{\partial \mathbf{y}_i} = \frac{\partial V}{\partial \mathbf{y}_i}(\mathbf{y}_1, \dots, \mathbf{y}_N) - \mathbf{f}_i.$$

The gradient of the coarse-grained energy at grid point  $i$ , for  $i = 1, \dots, n$  can be calculated by the chain rule:

$$\begin{aligned}
\mathbf{g}_i^k &= \mathbf{g}_i(\mathbf{u}^k) = \frac{\partial E^{\text{tot}}(\mathbf{u}^k)}{\partial \mathbf{u}_i^k} \\
&= \sum_{m=1}^N \frac{\partial V}{\partial \mathbf{y}_m}(\mathbf{x}_1 + \mathbf{u}_1, \dots, \mathbf{x}_N + \mathbf{u}_N) w\left(\frac{\mathbf{x}_m - \boldsymbol{\xi}_i^k}{h}\right) - \sum_{m=1}^N \mathbf{f}_m w\left(\frac{\mathbf{x}_m - \boldsymbol{\xi}_i^k}{h}\right) \\
(7) \quad &= \sum_{m=1}^N w\left(\frac{\mathbf{x}_m - \boldsymbol{\xi}_i^k}{h}\right) \mathbf{g}_m(\mathbf{u}).
\end{aligned}$$

In (7), the gradient at the  $i$ th grid point can be computed as the weighted average of atomic gradients over the support of the basis function  $w$  centered at  $\boldsymbol{\xi}_i^k$ , and the approach provides a way to systematically upscale the atomistic system to a coarse-grained model. By choice, the tri-linear interpolation has a finite support, and the total cost for evaluating (7) is only proportional to  $n$  provided the atomic gradient is given.

**2.2. Multigrid methods.** Consider a nested sequence of triangulations  $\mathcal{T}_0 \subset \mathcal{T}_1 \subset \dots \subset \mathcal{T}_\ell$  of  $\Omega$ , which can be constructed by the bisection procedure. Without loss of generality, we assume  $\mathcal{T}_\ell$  to be the atomic positions and the discrete model on  $\mathcal{T}_\ell$  to be the atomistic model. Let  $T$  be an element in  $\mathcal{T}_i$ . The mesh size  $h_i \equiv \max_{T \in \mathcal{T}_i} \text{diam} T$  satisfies

$$h_i = h_{i-1}/2 \quad \text{for } i = 1, \dots, \ell - 1.$$

The associated finite element spaces  $X_i$  are also nested

$$X_0 \subset X_1 \subset \dots \subset X_{\ell-1}.$$

Denote  $I_i^j$  the interpolation operator from level  $i$  to level  $j$  if  $i < j$  and the restriction operator if  $i > j$ .

To make ease of the description of multigrid methods, we define the following surrogate model on grid  $\mathcal{T}_k$  ( $k < \ell$ ) for a given force  $\mathbf{f}^k$  on level  $k$  as

$$(8) \quad \mathbf{u}^k = \text{argmin } E^{\text{tot}}(\mathbf{u}^k, \mathbf{f}^k)$$

with

$$E^{\text{tot}}(\mathbf{u}^k, \mathbf{f}^k) = E^{\text{tot}}(\mathbf{u}^k) - \sum_{j=1}^n \mathbf{f}_j^k \cdot \mathbf{u}_j^k.$$

Here the first term is exactly the same as (5), and the second term accounts for the residual term that comes from other levels.  $\mathbf{f}_j^k$  and  $\mathbf{u}_j^k$  denote the force and displacement at the  $j$ -th grid point on level  $k$ , respectively.

In order to describe the FMG, we first state the recursive form of FAS V-cycle multigrid method in Algorithm 1. Note that the discrete problem

---

**Algorithm 1** V-cycle multigrid method ( $\mathbf{u}^k \leftarrow V^k(\mathbf{u}^k, \mathbf{f}^k)$ )

---

**Step 1** Relax the discrete problem (8) on  $\mathcal{T}_k$   $\mu_1$  times to obtain  $\mathbf{u}^k$  with a given initial guess.

**Step 2** If  $k = 0$ , then go to **Step 4**.

Else

$$\mathbf{f}^{k-1} \leftarrow -I_k^{k-1} \mathbf{g}^k(\mathbf{u}^k),$$

$$\mathbf{u}^{k-1} \leftarrow I_k^{k-1} \mathbf{u}^k,$$

$$\mathbf{u}^{k-1} \leftarrow V^{k-1}(\mathbf{u}^{k-1}, \mathbf{f}^{k-1}).$$

**Step 3**  $\mathbf{u}^k \leftarrow \mathbf{u}^k + I_{k-1}^k(\mathbf{u}^{k-1} - I_k^{k-1} \mathbf{u}^k)$ .

**Step 4** Relax (8) on  $\mathcal{T}_k$   $\mu_2$  times with the initial guess  $\mathbf{u}^k$ .

---

(8) will be replaced by the atomistic model (2) if  $k = \ell$  in Algorithm 1. Then FAS-FMG in the recursive form is given by Algorithm 2. It is clear that Algorithm 1 is part of the method recursively; see also Figure 1 for illustration. It is recommended that higher order interpolations should be used in FMG if the approximate solution is interpolated to a new level, denoted by dashed lines in FMG in Figure 1. The oneway multigrid method

---

**Algorithm 2** FAS - FMG ( $\mathbf{u}^k \leftarrow \text{FMG}^k(\mathbf{f}^k)$ )

---

**Step 1** Initialize  $\mathbf{f}^{k-1} \leftarrow I_k^{k-1} \mathbf{f}^k$  for  $k = \ell, \dots, 1$ .

**Step 2** If  $k = 0$ , set  $\mathbf{u}^k \leftarrow \mathbf{0}$  and go to **Step 3**.

Else

$$\mathbf{u}^{k-1} \leftarrow \text{FMG}^{k-1}(\mathbf{f}^{k-1}).$$

**Step 3** Correct  $\mathbf{u}^k \leftarrow I_{k-1}^k \mathbf{u}^{k-1}$ .

**Step 4**  $\mathbf{u}^k \leftarrow V^k(\mathbf{u}^k, \mathbf{f}^k)$   $\mu_0$  times.

---

is described in Algorithm 3.

---

**Algorithm 3** Oneway multigrid method

---

**Step 1** Relax the discrete problem (4) on  $\mathcal{T}_0$   $\mu^0$  times to obtain  $\mathbf{u}_0$  with a trivial initial guess.

**Step 2** For  $k = 1, \dots, \ell - 1$ , relax the discrete problem (4) on  $\mathcal{T}_k$   $\mu_k$  times to obtain  $\mathbf{u}^k$  with the initial guess  $I_{k-1}^k \mathbf{u}^{k-1}$ .

**Step 3** Solve the atomistic problem (1) until convergence with the initial guess  $I_{\ell-1}^\ell \mathbf{u}^{\ell-1}$ .

---

Figure 1 shows the schedule of grids in the order where a multigrid method visits. Solid points denotes levels and lines denotes interpolations and restrictions. From left to right, we have V-cycle multigrid, see also Algorithm 1; FAS-FMG, see also Algorithm 2 with  $\mu_0 = 1$ ; oneway multigrid, see also Algorithm 3.



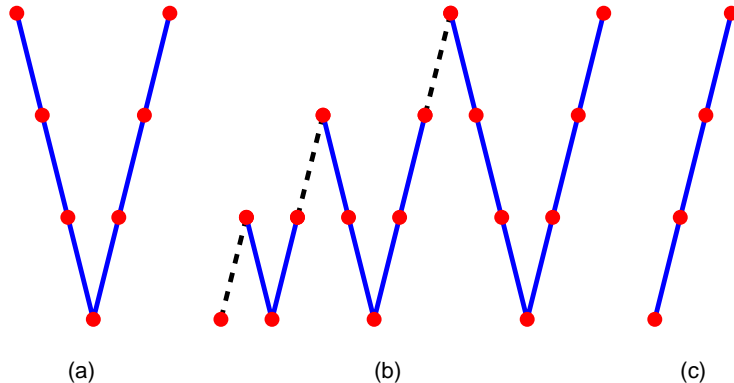


FIGURE 1. Schedule of multigrid methods for  $\ell = 3$ . Solid points denotes levels and lines denotes interpolations and restrictions. From left to right: (a) V-cycle, see also Algorithm 1; (b) FMG, see also Algorithm 2 with  $\mu_0 = 1$ ; (c) oneway, see also Algorithm 3.

We have used a conjugate-gradient method: CG\_DESCENT [20, 21] and a limited BFGS (L-BFGS) method [33, 26] for the relaxation (smoothing) steps in these algorithms. A line search multigrid method [47] is also tested. Similar performance is found for many examples, but CG\_DESCENT seems to give the best compromise between accuracy and efficiency. Moreover, the linesearch implementation in CG\_DESCENT appears to be the most robust on coarse levels. Therefore, all results reported here are computed by CG\_DESCENT.

**2.3. Model illustration.** For brevity, we first consider a one-dimensional chain (Figure 2) to illustrate the difference between the Cauchy-Born elasticity model [3], the quasi-atomistic description (5), and the atomistic model. Let  $V$  be a pairwise potential and consider up to the second nearest neighbor interaction. Denote  $u_i$  the displacement of the  $i$ -th atom and  $\varepsilon$  the lattice constant. The total energy of the atomistic model is

$$(9) \quad E_{\text{at}}(u_1, \dots, u_9) = \sum_{j=1}^2 \sum_{i=1}^{9-j} V\left(\frac{u_{i+j} - u_i}{\varepsilon}\right).$$

There are two simplifications in the Cauchy-Born elasticity model compared to the atomistic model. One is the kinematic reduction: Only the displacements of rep-atoms are used and the displacements of the other atoms are obtained by interpolation. This is illustrated in Figure 2: only  $u_1$ ,  $u_5$ , and  $u_9$  are used in the Cauchy-Born elasticity model. The other reduction in this model is energetic: The energy of each element is approximated by the energy of a unit cell multiplied by the length (area or volume)

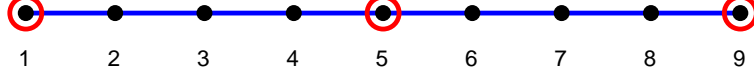


FIGURE 2. A one-dimensional chain consisting of nine atoms. Solid points are atoms, encircled points are grid points (rep-atoms).

of the element. In Figure 2, the total energy in the Cauchy-Born elasticity model equals

$$(10) \quad E_{\text{cb}}(u_1, u_5, u_9) = 4\varepsilon \left( \frac{1}{\varepsilon} \left( V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + V\left(2\frac{u_5 - u_1}{4\varepsilon}\right) \right) \right) + 4\varepsilon \left( \frac{1}{\varepsilon} \left( V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + V\left(2\frac{u_9 - u_5}{4\varepsilon}\right) \right) \right)$$

In the quasiautomistic description, we use the kinematic reduction, but not the energetic reduction. The total energy is still done by summing up all sites energies associated to each atom. Since site energies of atoms 1 to 9 are  $\frac{1}{2}V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_5 - u_1}{4\varepsilon}\right)$ ,  $V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + V\left(2\frac{u_5 - u_1}{4\varepsilon}\right)$ ,  $V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + V\left(2\frac{u_5 - u_1}{4\varepsilon}\right)$ ,  $V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_5 - u_1}{4\varepsilon}\right) + \frac{1}{2}V\left(\frac{u_9 + 3u_5}{4\varepsilon} - \frac{3u_5 + u_1}{4\varepsilon}\right)$ ,  $\frac{1}{2}V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_5 - u_1}{4\varepsilon}\right) + \frac{1}{2}V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ ,  $\frac{1}{2}V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ ,  $\frac{1}{2}V\left(\frac{u_9 + 3u_5}{4\varepsilon} - \frac{3u_5 + u_1}{4\varepsilon}\right) + V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ ,  $V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ ,  $V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ ,  $\frac{1}{2}V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + \frac{1}{2}V\left(2\frac{u_9 - u_5}{4\varepsilon}\right)$ . the total energy (5) therefore equals

$$(11) \quad E_{\text{qa}}(u_1, u_5, u_9) = 4V\left(\frac{u_5 - u_1}{4\varepsilon}\right) + 3V\left(2\frac{u_5 - u_1}{4\varepsilon}\right) + V\left(\frac{u_9 - u_1}{4\varepsilon}\right) + 4V\left(\frac{u_9 - u_5}{4\varepsilon}\right) + 3V\left(2\frac{u_9 - u_5}{4\varepsilon}\right).$$

Consider the case where the deformation gradient is a constant. Since atom 2 does not have a second-neighbor atom on the left-hand side and atom 8 does not have a second-neighbor atom on the right-hand side, ghost atoms must be added. Otherwise, instead of a uniform deformation, boundary layers will occur near the left and right end points. This contribution is not explicitly considered in the atomistic model as well as the quasiautomistic description. Despite this difference, one can easily check that a uniform deformation is a minimizer of the Cauchy-Born elasticity model (10), the quasiautomistic description (11), and the atomistic model (9).

If we now remove atom 6 from the chain, then its site energy will be removed from the atomistic model (9) and the quasiautomistic model (11). However, the absence of an atom will not change the energy in the Cauchy-Born elasticity model. Therefore the quasiautomistic model will certainly be better than the Cauchy-Born elasticity model to describe systems with defects, not just vacancies, but also dislocations and cracks. On the other hand, the inhomogeneous deformation induced by defects may not be well approximated by coarse elements if the inhomogeneity is localized and the

nonlocal part of the deformation can be well approximated by coarse elements if the inhomogeneity is delocalized. An adaptive mesh refinement may improve the approximation accuracy, but a robust refinement may require solving surrogate problems with significant cost. Overall, such a refinement may not be computationally more efficient.

To test the robustness of the full multigrid method and the oneway multigrid method we use the same example as in [11]: a one-dimensional chain with Lennard-Jones potential and with random initial guesses. Both methods produce the same results as the oneway multigrid method with the Cauchy-Born elasticity model; see Figure 3 in [11]. This is not surprising since both methods start from the coarsest mesh.

**2.4. Comparison.** We use a face-centered cubic (FCC) crystal aluminum for demonstration in this paper. Atoms interact through the embedded-atom method (EAM) potential of the form [13]

$$E = \frac{1}{2} \sum_{i,j} \phi_{ij}(r_{ij}) + \sum_i U_i \left[ \sum_j \rho_i(r_{ij}) \right],$$

where  $\phi_{ij}$  is the pairwise potential,  $r_{ij}$  is the distance between the  $i$ th and  $j$ th atoms,  $U_i$  is the glue function and  $\rho_i$  is the atomic density function of the  $i$ th atom. Parameters in the potential function are taken from [16]. The cut-off radius  $r_{\text{cut}} = 5.56\text{\AA}$  [16], which includes up to the 3rd neighbor interaction. The  $[110]$ ,  $[1\bar{1}0]$  and  $[001]$  directions are chosen as  $x$ ,  $y$  and  $z$  axes, respectively. We simulate the mechanical response in a quasi-static manner, i.e., the stable configuration from the previous step is used as the initial guess for the current step with a small increment (1% typically) added to the boundary loading. In the presence of defects, we first relax the system to get the corresponding equilibrium state, which is later set to be the initial configuration.

We first consider a shear deformation in the  $[110](001)$  direction for the perfect aluminum crystal. Dirichlet boundary conditions are used in three directions. To avoid the boundary effect, we introduce ghost atoms (outside of the domain of interest) with thicknesses greater than  $2r_{\text{cut}}$  according to the boundary loading. This is in the same spirit of the fixed region used in [30]. The smallest system contains  $8 \times 8 \times 8$  unit cells, 2048 atoms in total. The system is then doubled in each direction successively, with total atoms increased to 16384, 131072, and 1048576. Since each atom has three coordinates, the largest system we test involves more than 3,000,000 dofs. For FAS-FMG,  $\mu_0 = 1$  on all coarse levels in Algorithm 2 and  $\mu_1 = \mu_2 = 5$  in Algorithm 1. The stopping criterion is  $|\mathbf{g}|_\infty < \text{tol}^\ell$  in (6) where  $\text{tol}^\ell$  is the given tolerance at the atomistic level.  $\text{tol}^\ell = 10^{-6}$  in this example. For the oneway multigrid method, the number of relaxations on all coarse levels is set to be 5. For comparison, we also use CG\_DESCENT to solve the atomistic model (1).

Figure 3(a) plots the CPU time of the brute-force optimization, FAS-FMG, and the oneway multigrid method in terms of number of atoms in the log-log scale. Least squares approximation suggests a linear scaling complexity for both FAS-FMG and the oneway multigrid method, while a super-linear (1.28) scaling complexity for the brute-force optimization. Asymptotically, both FAS-FMG and the oneway multigrid method are superior to the brute-force optimization. Moreover, prefactors of both FAS-FMG and the oneway multigrid method are much smaller than that of the brute-force optimization, which means both approaches are computationally cheaper than the brute-force optimization, as supported by Figure 3(b). For the largest system tested here, the savings of FAS-FMG are 91% and the savings of the oneway multigrid method are 99%. It is worth mentioning that the cost of the oneway multigrid method is almost the same as that of the oneway multigrid method with the Cauchy-Born elasticity model [11] in the case of uniform deformations.

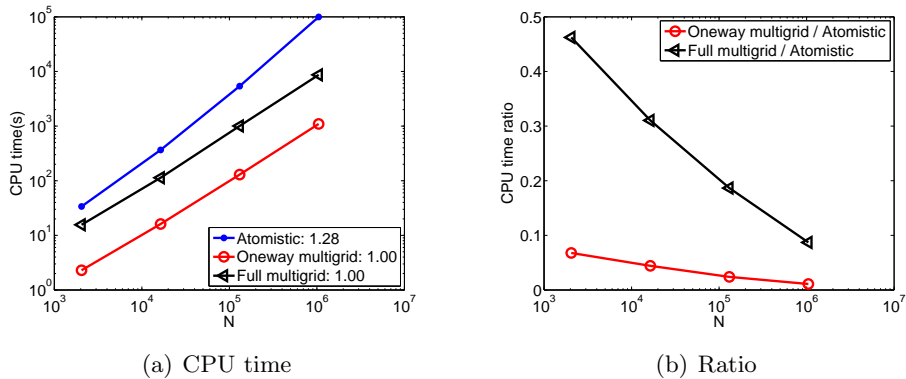


FIGURE 3. Costs of the brute-force optimization, FAS-FMG, and the oneway multigrid method for a perfect aluminum crystal under a shear deformation. (a) CPU time in the log-log scale; (b) Ratio of CPU time between FAS-FMG, the oneway multigrid method, and the brute-force optimization in the semilog scale.

While the parameters in FAS-FMG and the oneway multigrid method can be tuned to produce even better results, we stress this will not change the above observations qualitatively. Besides, the oneway multigrid method seems to be superior to FAS-FMG. This will be further confirmed by the vacancy example later. It is known that for a uniform (smooth) deformation, (4) is a good approximation of (1), thus both FAS-FMG and the oneway multigrid method work well. For a smooth deformation, boundary conditions and ghost atoms need to be used properly to avoid the surface effect.

The second example is a shear deformation in the  $[110](001)$  direction for a perfect aluminum crystal with a centered vacancy, i.e., one aluminum atom is removed from the system. For FAS-FMG,  $\mu_0 = 1$  on all coarse levels in Algorithm 2 and  $\mu_1 = \mu_2 = 10$  in Algorithm 1. For the oneway multigrid

method, the number of relaxations on all coarse levels is set to be 3. Figure 4(a) plots the CPU time of the brute-force optimization, FAS-FMG, and the oneway multigrid method in terms of number of atoms in the log-log scale. The oneway multigrid method still preserves linear scaling complexity, while FAS-FMG has a slightly larger scaling (1.09). The brute-force optimization has a super-linear (1.26) scaling complexity again. Asymptotically, both FAS-FMG and the oneway multigrid method are superior to the brute-force optimization. However, the prefactors of both FAS-FMG and the oneway multigrid method are not very small anymore. Increasing the order of the interpolation used, or adjusting the iteration parameters for FAS-FMG did not seem to lower significantly the cost of this approach. In fact, FAS-FMG was more costly than brute-force optimization for all the cases we tested except the uniform deformation case. The oneway multigrid method costs more compared to the brute-force optimization only when  $N = 2048$ , and the crossover happens when the number of atoms is around 10,000; see Figure 4(b). Therefore, the oneway multigrid method is computationally cheaper than the brute-force optimization if the system size is greater than 10,000. Considerable savings are achieved if the system size is further enlarged. For systems with more than one million atoms, the savings are over 50%.

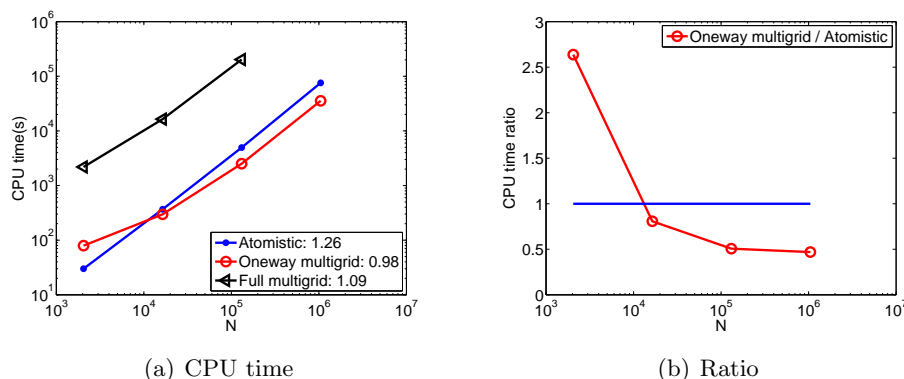


FIGURE 4. Costs of the brute-force optimization, FAS-FMG, and the oneway multigrid method for the vacancy case under a shear deformation. (a) CPU time in the log-log scale; (b) Ratio of CPU time between the oneway multigrid method and the brute-force optimization in the semilog scale.

Again, parameters in FAS-FMG and the oneway multigrid method can be tuned to produce even better results. However, for all systems tested here, FAS-FMG is computationally more expensive than the brute-force optimization. To check the inefficiency of FAS-FMG, we plot the atomic displacement  $(u, v, w)$  over a centered slice in the  $y$  direction of a system with 131072 atoms for the vacancy case under 1% shear deformation in Figure 5. Since a vacancy has a localized effect, the coarse-grid model (4) cannot give a good approximation to the atomistic model (1), though the

missing atom does change the potential energy function on the coarse level (5). The approximation becomes better only when the grid is sufficiently close to the atomic resolution. Figure 6 plots displacement differences of  $u$  between the solution of the coarse-grid model (4) on a  $2 \times 2 \times 2$  mesh and a  $32 \times 16 \times 32$  mesh, and the solution of the atomistic model (1). The other two displacement components  $v$  and  $w$  behavior similarly. No significant improvement of (4) if a fine mesh is used. Note that there are  $32 \times 32 \times 32$  unit cells in this case. Therefore, the V-cycle part in FMG is inefficient for the vacancy case, resulting in the inefficiency of FAS-FMG. If the V-cycle part is removed from FMG, then it basically recovers the oneway multigrid and is computationally more efficient.

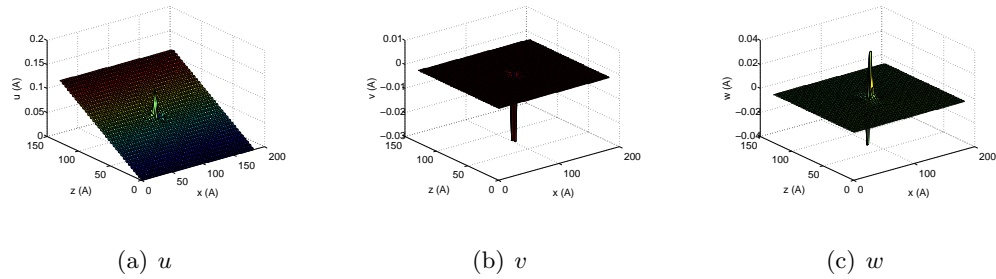


FIGURE 5. Atomic displacement  $(u, v, w)$  over the centered slice in the  $y$  direction of a system with 131072 atoms for the vacancy case under 1% shear deformation. (a)  $u$ ; (b)  $v$ ; (c)  $w$ .

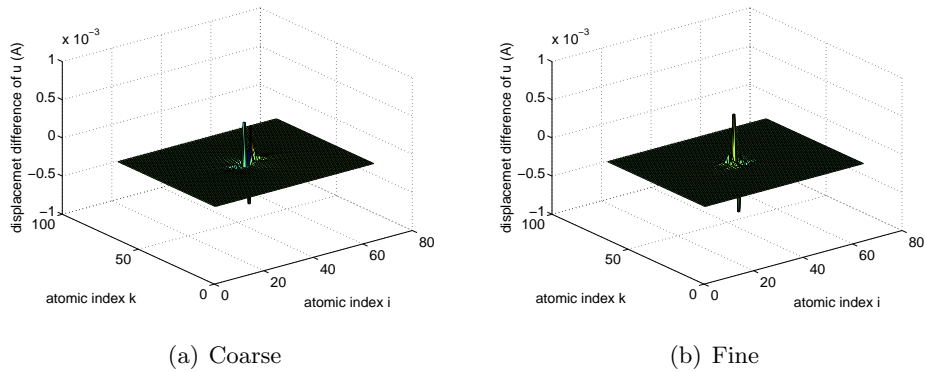


FIGURE 6. Displacement differences of  $u$  between the solution of the coarse-grid model (4) on a  $2 \times 2 \times 2$  mesh and a  $32 \times 16 \times 32$  mesh, and the solution of the atomistic model (1) for the vacancy case under a shear deformation. (a) coarse; (b) fine.

To study the parameters that may affect the performance of the oneway multigrid method, we list CPU time (in seconds) of the oneway multigrid versus the number of relaxations  $\mu$  on each level (except the atomistic level).

For comparison, CPU time of the brute-force optimization, the oneway multigrid with accuracy control on each level, and the oneway multigrid with number of relaxations control on each level is also listed. Denote  $\text{tol}^k$  the tolerance at the  $k$ -th level and set the tolerance at the atomistic level  $\text{tol}^\ell = 10^{-4}$ . The oneway multigrid with accuracy control on each level is done by setting  $\text{tol}^k = \text{tol}^\ell / 2^{\ell-k}$ . The oneway multigrid with number of relaxations control on each level is done by setting  $\mu^k = 2^{\ell-k}$ . Compared to the brute-force optimization, the oneway multigrid with a fixed number of relaxations  $\leq 5$  is quite efficient, and is better than the oneway multigrid with accuracy control. We use this throughout the work without any tuning. This is somewhat contrary to the classical convergence theory of the oneway multigrid method for elliptic PDEs [4, 48]. For both linear [4] and nonlinear [48] problems, sufficient relaxations must be done on coarse levels in order to achieve the optimal computational complexity with desired accuracy on the fine level. Moreover, approximation accuracy of the coarse-grid model on coarse levels is required in order to have the convergence. The presence of a vacancy introduces a localized and inhomogeneous displacement which only can be well described by the atomistic model (1), and the resolution of the solution of the coarse-grid model (4) is poor. Therefore, any attempt to solve the coarse-grid model to high accuracy will slow down the convergence of multigrid overall; see Table 1. The same behavior is observed in the presence of dislocations and cracks. This also explains why we do not impose any accuracy requirement on coarse levels in Algorithm 3. Since coarse-grid approximations cannot provide very good initial guesses in the vicinity of defects, quite a few iterations are required to achieve the convergence. The number of iterations on the atomistic level stabilizes around 50 for the vacancy case with  $\text{tol}^\ell = 10^{-6}$ . The main savings come from the fact that the long-range displacement is captured in an *inexact* manner on coarse levels; see Figure 5 and Figure 6.

Due to the localized effect of a vacancy, as discussed earlier, (5) cannot provide a very good approximation. In this case, the oneway multigrid works with optimal scaling but is only superior to the brute-force optimization when the system size is greater than 10,000. On the other hand, other defects, such as dislocations and cracks, have nonlocal effects on the displacement field. Let  $r$  be the distance to the dislocation core or the crack tip; then their effects typically follow a  $1/r$  or  $1/\sqrt{r}$  decay law away from the dislocation core or the crack tip. In this scenario, (5) is able to capture the long range deformation caused by not just the boundary loading, but also the defects. Even though (5) cannot provide a very good approximation near the defect, the oneway multigrid again works with optimal scaling and is superior to the brute-force optimization when the system size is greater than 25,000; see Section 3.

$\mu$	CPU time ( $N = 16384$ )	CPU time ( $N = 131072$ )
1	61	537
2	75	713
3	90	829
4	95	860
5	94	861
Number control	86	833
Accuracy control	168	1403
Brute-force	232	2957

TABLE 1. CPU time (in seconds) of the oneway multigrid method versus the number of relaxations  $\mu$  on each level (except the atomistic level). For comparison, CPU time of the brute-force optimization, the oneway multigrid with accuracy control on each level, and the oneway multigrid with number of relaxations control on each level is also listed. Here  $\text{tol} = 10^{-4}$ .

### 3. RESULTS

In this section, we apply the oneway multigrid method for three examples: multiple vacancies to mimic the doping effect in semiconductors, a Lomer dislocation pair, and a crack. The same stopping criteria  $|\mathbf{g}|_\infty < \text{tol}^\ell = 10^{-4}$  is used for these examples if it is not specified. In order to make a meaningful comparison, we check all atomic configurations in the oneway multigrid method with those of the atomistic model. The maximum displacement error is smaller than  $0.005\text{\AA}$  in all examples.

**3.1. Doping.** Doping is a standard way of modulating the electronic property of a pure semiconductor by intentionally introducing impurities. In semiconductor production, typical doping concentration ranges from  $10^{19} \sim 10^{24}/\text{m}^3$ . To mimic this, we create vacancies (almost) homogeneously in space for the aluminum system with 1048576 atoms under a shear deformation. Displacement field  $(u, v, w)$  is plotted in Figure 7 over a slice in the  $y$  direction where multi vacancies are created for a system with 1048576 atoms in the absence of deformation. The number of relaxations on coarse levels is set to be 3. Figure 8(a) plots the CPU time of the brute-force optimization and the oneway multigrid method in terms of doping concentration with the ratio between these two in Figure 8(b). Performances of both the brute-force optimization and the oneway multigrid method seem to be independent of the doping concentration. Compared to the brute-force optimization, the oneway multigrid saves around 79% in terms of CPU time. Random distributions of vacancies are also tested and our method still saves over 70%. A weak growth of the CPU time is found when vacancies are generated in adjacent neighbors. In the extreme case, if all the vacancies are adjacent, a crack will be created and the performance of the oneway multigrid is given in Section 3.3.



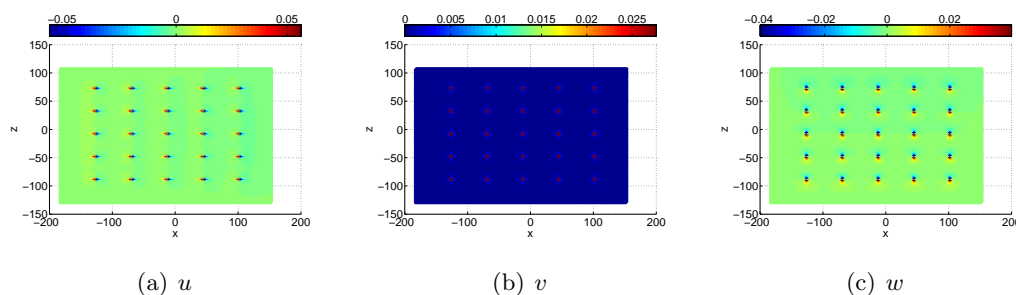


FIGURE 7. Atomic displacement  $(u, v, w)$  over a slice in the  $y$  direction where multi vacancies are created for a system with 1048576 atoms with any deformation. (a)  $u$ ; (b)  $v$ ; (c)  $w$ .

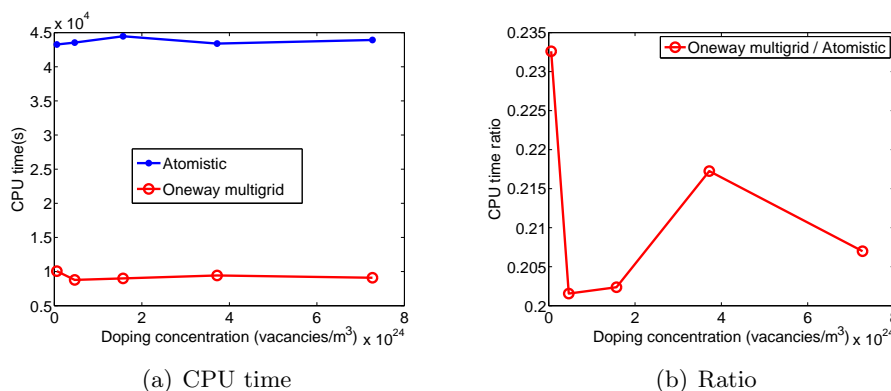


FIGURE 8. Costs of the brute-force optimization and oneway multigrid for the doping case under a shear deformation. (a) CPU time versus doping concentration; (b) Ratio of CPU time between oneway multigrid and the brute-force optimization.

**3.2. Dislocation.** Systems used for the dislocation case have  $32 \times 8 \times 32$ ,  $64 \times 8 \times 64$ ,  $128 \times 8 \times 128$  and  $128 \times 8 \times 256$  unit cells. Dirichlet boundary conditions are used in the  $x$  and  $z$  directions with ghost atoms, and periodic boundary condition is used in the  $y$  direction. A pair of dislocations is created in the same way as that in [30]. Figure 9 plots the  $u$  and  $w$  components of the displacement field for the  $64 \times 8 \times 64$  system without external loading. Around dislocation cores, one can see the nonlocal effect on the displacement field in the presence of dislocations. This nonlocal effect typically decays like  $1/r$  where  $r$  is the distance to the core, and thus is more difficult to solve than the vacancy case.

Three relaxations on coarse levels are used in the oneway multigrid method. Figure 10(a) plots the CPU time of the brute-force optimization and the oneway multigrid method in the log-log scale with the ratio between these two in Figure 10(b). Due to the non-locality of the displacement field in the presence of dislocations, the brute-force optimization scales  $\mathcal{O}(N^{1.66})$

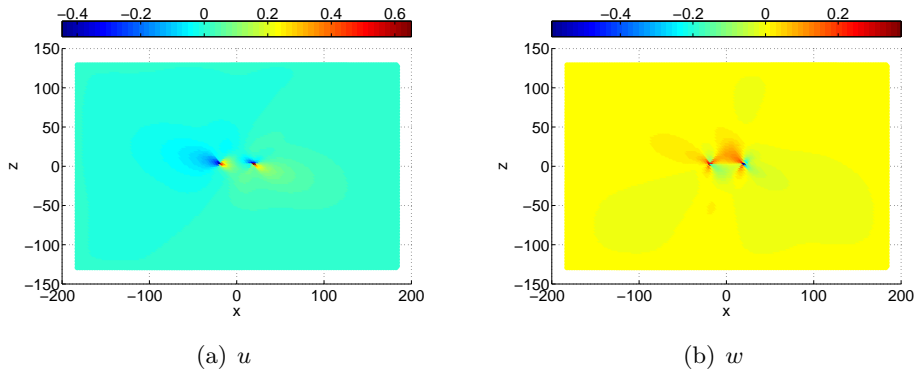


FIGURE 9.  $u$  and  $w$  components of the displacement field over a centered slice in the  $y$  direction for the dislocation case without external loading. (a)  $u$ ; (b)  $w$ .

asymptotically, while the oneway multigrid method scales linearly (0.97) asymptotically. Moreover, for the results in Figure 10(a) and Figure 10(b), the crossover between the oneway multigrid method and the brute-force optimization happens when the system contains around 150,000 atoms, which is only twice larger than that for the vacancy case. Compared to the brute-force optimization, the savings of the oneway multigrid method are around 77% for systems with one million atoms. Due to the pair interaction between dislocations, scaling of the brute-force optimization is a bit larger than that in the crack case; see Section 3.3. On the other hand, the displacement field of a dislocation decays faster than that of a crack, so the crossover for the dislocation case is smaller than that for the crack case. The linear scalability of the oneway multigrid method is further tested for systems with atoms up to 134,217,728 atoms (over 400 million dofs); see Figure 10(c). We use a relaxed stopping criteria  $|\mathbf{g}|_\infty < 10^{-2}$ , which is found to reduce the displacement error (Equation (57) in [30]) to  $\leq 0.1\%$  for small systems with exact configurations provided by the brute-force optimization.

We further visualize the displacement field with atomic resolution by Atomeye [22] in Figure 11(a). The number of atoms in the nearest-neighbor of an atom is defined as its coordinate number. For a perfect FCC crystal, atoms with coordinate number 12 are regular atoms, while the others are irregular atoms. In Figure 11(a), atoms in lightgoldenrod are regular atoms, atoms in mediumvioletred have coordinate number 11, and atoms in green have coordinate number 13. These represent a pair of dislocations. Since the oneway multigrid method is a fully atomistic approach, it produces exactly the same result as that of the brute-force optimization. In details, the core spacing is initially set to be  $40\text{\AA}$  and is kept until the applied shear strain is 0.05. The system then stabilizes with core spacing around  $50\text{\AA}$  until the applied shear strain is raised to 0.057. Afterwards, dislocations move toward the outer edges of the crystal. Computationally, the oneway multigrid

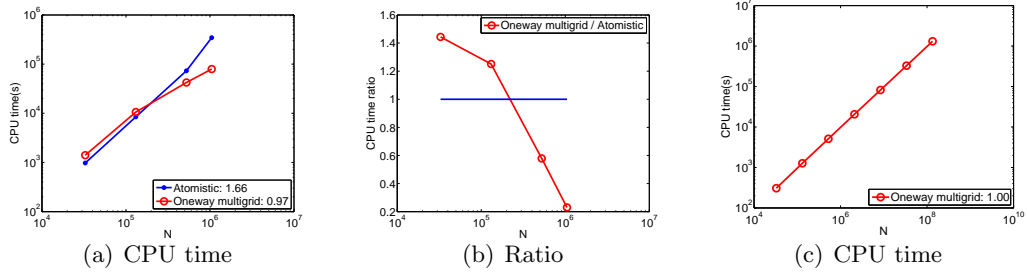


FIGURE 10. Costs of the brute-force optimization and the oneway multigrid method for the dislocation case under a shear deformation. (a) CPU time in the log-log scale; (b) Ratio of CPU time between the oneway multigrid method and the brute-force optimization in the semilog scale; (c) CPU time in the log-log scale with a relaxed stopping criteria  $|\mathbf{g}|_\infty < 10^{-2}$ .

method saves over 70% on average for the sequence of applied shear deformations for the largest system. Even when dislocations move apart from each other dramatically, 70% saving is observed.

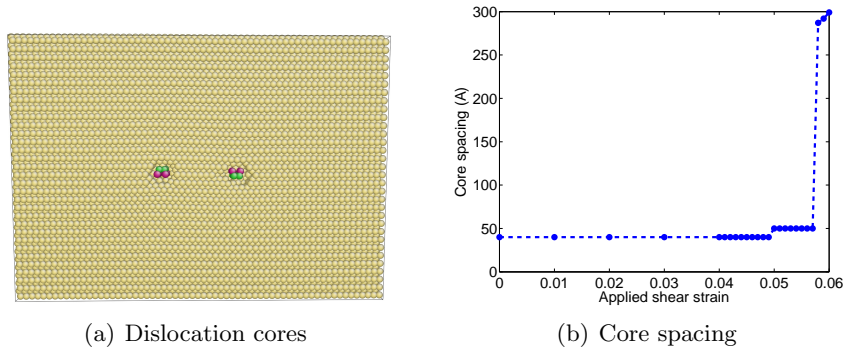


FIGURE 11. Visualization of dislocation cores by AtomEye [22] and the core spacing under a sequence of successfully increasing shear deformation. (a) Dislocation cores; (b) The core spacing versus the applied shear strain.

**3.3. Crack.** Systems used for the crack case have  $32 \times 8 \times 32$ ,  $64 \times 8 \times 64$ ,  $128 \times 8 \times 128$  and  $128 \times 8 \times 256$  unit cells. Dirichlet boundary conditions are used in the  $x$  and  $z$  directions with ghost atoms, and periodic boundary condition is used in the  $y$  direction. To create a crack for the aluminum system, we remove tens of atoms in the center of the  $x-z$  plane along the  $y$  direction. Figure 12 plots the  $u$  and  $w$  components of the displacement field for the  $32 \times 8 \times 32$  system without external loading. Around crack tips, one can see the nonlocal effect on the displacement field in the presence of cracks. Compared to Figure 9, we see that the nonlocal effect for a crack is stronger

than that of a dislocation, which follows  $1/\sqrt{r}$  law typically. Therefore, the crack case is even more difficult than the dislocation case. We test this example only under small strains. If the strain further increases, dislocations will appear near the crack tip and multiple stable configurations may occur with small energy differences. In this case, the physical relevant minimizer depends on the thermal noise, and this is out of the scope of the current work.

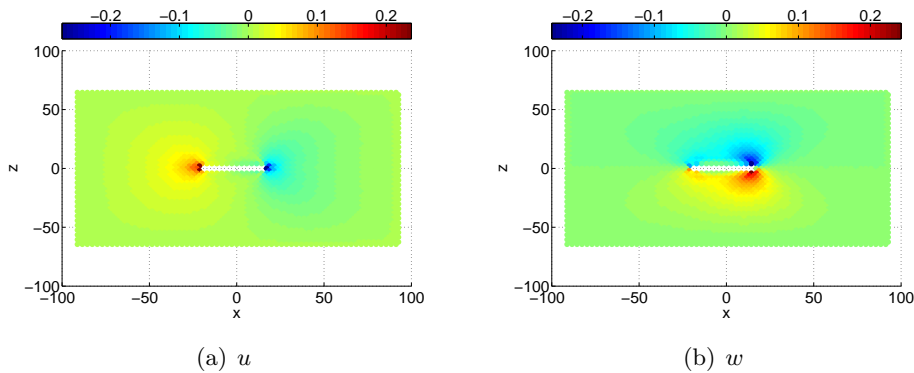


FIGURE 12.  $u$  and  $w$  components of the displacement field over a centered slice in the  $y$  direction for the crack case without external loading. (a)  $u$ ; (b)  $w$ .

Three relaxations on coarse levels are used in the oneway multigrid method. Figure 13(a) plots the CPU time of the brute-force optimization and the oneway multigrid method in the log-log scale with the ratio between these two in Figure 13(b). Due to the non-locality of the displacement field in the presence of cracks, the brute-force optimization scales  $\mathcal{O}(N^{3/2})$  asymptotically, while the oneway multigrid method scales a bit worse than linear scaling (1.07) asymptotically. Linear scaling complexity for the oneway multigrid method can be recovered if the number of relaxations on coarse levels is tuned. However, even for the results plotted in Figure 13, the crossover between the oneway multigrid method and the brute-force optimization happens when the system contains around 200,000 atoms, which is slightly larger than that in the dislocation case. Moreover, compared to the brute-force optimization, the savings of the oneway multigrid method are around 70% for systems with one million atoms.

#### 4. CONCLUSION

This paper presents an efficient multigrid strategy for large-scale molecular mechanics optimization. The coarse-grid problem is constructed from the atomistic model systematically on-the-fly using a quasi-atomistic approximation, which does not require analytic knowledge of a continuum model, such as the Cauchy-Born rule. The oneway multigrid method is used with

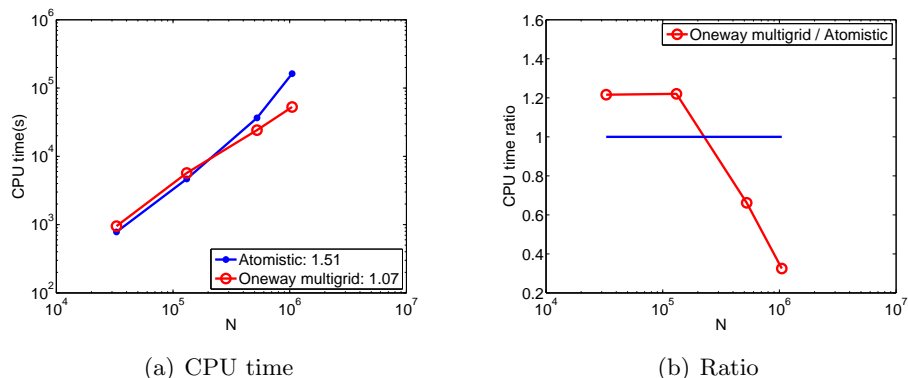


FIGURE 13. Costs of the brute-force optimization and the oneway multigrid method for the crack case under a tensile deformation in the  $z$  direction. (a) CPU time in the log-log scale; (b) Ratio of CPU time between the oneway multigrid method and the brute-force optimization in the semilog scale.

inexact approximations at coarse levels. Optimal scaling complexity is observed for perfect crystals and crystals with defects, like vacancies, dislocations, and cracks (with a minor adjustment of relaxation numbers). Moreover, this strategy outperforms the brute-force optimization for all examples if more than 200,000 atoms are involved. For systems with more than one million atoms, this strategy saves more than 70% CPU time on average.

Mesh refinements based on *a posteriori* estimates could be incorporated into this methodology in order to obtain better approximations on the coarse-grained levels. But a robust refinement strategy may require significant cost. Therefore a compromise between the quality of a mesh refinement strategy and its solution cost will be important. The approach could also be combined with a coarse-grained model in a domain decomposition way, similarly to what is done in the QC method. Moreover, the strategy presented is not limited to crystals or solids. How it would work for mechanical responses of polymers and large biological molecules without knowing internal variables for optimization a priori will be of great interest. Results in these directions will be presented elsewhere.

## REFERENCES

1. A. Abdulle, P. Lin, and A. V. Shapeev, *Numerical methods for multilattices*, Multiscale Model. Simul. **10** (2012), 696–726.
2. A. Abdulle, P. Lin, and A. V. Shapeev, *A priori and a posteriori  $W^{1,\infty}$  error analysis of a QC method for complex lattices*, SIAM J. Numer. Anal. **51** (2013), 2357–2379.
3. M. Born and K. Huang, *Dynamical theory of crystal lattices*, Oxford University Press, 1954.
4. F. Bornemann and P. Deuffhard, *The cascadic multigrid method for elliptic problems*, Numer. Math. **75** (1996), 135 – 152.
5. A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), 333–390.

6. ———, *Multigrid methods in lattice field computations*, Nuclear Physics B (Proc. Suppl.) **26** (1992), 137–180.
7. W. L. Briggs, V. E. Henson, and S. F. McCormick, *A multigrid tutorial: Second edition*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2000.
8. M. J. Buehler, *Atomistic modeling of material failure*, Springer, New York, 2008.
9. J. Chen and P. B. Ming, *An efficient multigrid method for molecular mechanics modeling in atomic solids*, Commun. Comput. Phys. **10** (2011), 70–89.
10. J. Chen, P. B. Ming, and J. Z. Yang, *A constrained Cauchy-Born elasticity accelerated multigrid method for nanoindentation*, Commun. Comput. Phys. **15** (2014), 470–486.
11. J. Chen and P. B. Ming, *Ghost force influence of a quasicontinuum method in two dimension*, J. Comput. Math. **30** (2012), 657–683.
12. L. Cui and P. B. Ming, *The effect of ghost forces for a quasicontinuum method in three dimension*, Sci. in China Ser. A: Math. **56** (2013), 2571–2589.
13. M. S. Daw and M. I. Baskes, *Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals*, Phys. Rev. B **29** (1984), 6443.
14. W. E, *Principles of multiscale modeling*, Cambridge University Press, 2011.
15. W. E and P. B. Ming, *Cauchy-born rule and the stability of the crystalline solids: Static problems*, Arch. Rational Mech. Anal. **183** (2007), 241–297.
16. F. Ercolessi and J. B. Adams, *Interatomic potentials from first-principles calculations: The force-matching method*, Europhys. Lett. **26** (1994), 583.
17. S. Goedecker, F. Lançon, and T. Deutsch, *Linear scaling relaxation of the atomic positions in nanostructures*, Phys. Rev. B **64** (2001), 161102.
18. W. Hackbusch, *On the multi-grid method applied to difference equations*, Computing **20** (1978), 291–306.
19. W. Hackbusch, *Multi-grid methods and applications*, Springer, 1985.
20. W. W. Hager and H. Zhang, *A new conjugate gradient method with guaranteed descent and an efficient line search*, SIAM J. Optim. **16** (2005), 170–192.
21. W. W. Hager and H. Zhang, *Algorithm 851: CG\_DESCENT, a conjugate gradient method with guaranteed descent*, ACM Trans. Math. Soft. **32** (2006), 113–137.
22. J. Li, *Atomeye: an efficient atomistic configuration viewer*, Model. Simul. Mater. Sci. Eng. **11** (2003), 173–177.
23. X. Li, M. Luskin, C. Ortner, and A. V. Shapeev, *Theory-based benchmarking of the blended force-based quasicontinuum method*, Comput. methods Appl. Mech. Engrg. **268** (2014), 763–781.
24. P. Lin, *Theoretical and numerical analysis for the quasi-continuum approximation of a material particle model*, Math. Comp. **72** (2003), no. 242, 657–675.
25. ———, *Convergence analysis of a quasi-continuum approximation for a two-dimensional material*, SIAM J. Numer. Anal. **45** (2007), 313–332.
26. D. C. Liu and J. Nocedal, *On the limited memory BFGS method for large scale optimization*, Math. Program. **45** (1989), 503–528.
27. J. Lu and P. B. Ming, *Convergence of a force-based hybrid method in three dimensions*, Comm. Pure Appl. Math. **66** (2013), 83–108.
28. ———, *Convergence of a force-based hybrid method with planar sharp interface*, SIAM J. Numer. Anal. **52** (2014), 2005–2026.
29. M. Luskin and C. Ortner, *Atomistic-to-continuum coupling*, Acta Numerica **22** (2013), 397–508.
30. R. E. Miller and E. B. Tadmor, *A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods*, Modelling Simul. Mater. Sci. Eng. **17** (2009), no. 5, 053001–053051.
31. R. E. Miller and E. B. Tadmor, *A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods*, Model. Simul. Mater. Sci. Eng. **17** (2009), 053001.

32. P.B. Ming and J.Z. Yang, *Analysis of a one-dimensional nonlocal quasi-continuum method*, Multiscale Model. Simul. **7** (2009), 1838–1875.
33. J. Nocedal, *Updating quasi-Newton matrices with limited storage*, Math. Comput. **35** (1980), no. 151, 773–782.
34. J. T. Oden, S. Prudhomme, A. Romkes, and P. Bauman, *Multiscale modeling of physical phenomena: Adaptive control of models*, SIAM J. Sci. Comput. **28** (2006), 2359–2389.
35. J. T. Oden and K. Vemaganti, *Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials: Part I. error estimates and adaptive algorithms*, J. Comput. Phys. **164** (2000), no. 1, 22–47.
36. C. Ortner and L. Zhang, *Construction and sharp consistency estimates for atomistic/continuum coupling methods with general interfaces: a 2D model problem*, SIAM J. Numer. Anal. **50** (2012), 2940–2965.
37. C. Ortner and L. Zhang, *Energy-based atomistic-to-continuum coupling without ghost forces*, Comput. methods Appl. Mech. Engrg. **279** (2014), 29–45.
38. S. Prudhomme, P. T. Bauman, and J. T. Oden, *Error control for molecular statics problems*, Int. J. Multiscale Comput. Eng. **4** (2006), no. 5–6, 647–662.
39. A.V. Shapeev, *Consistent energy-based atomistic/continuum coupling for two-body potentials in one and two dimensions*, Multiscale Model. Simul. **9** (2011), 905–932.
40. A.V. Shapeev, *Consistent energy-based atomistic/continuum coupling for two-body potentials in three dimensions*, SIAM J. Sci. Comput. **34** (2012), no. 3, B335–B360.
41. E. B. Tadmor, M. Ortiz, and R. Phillips, *Quasicontinuum analysis of defects in solids*, Phil. Mag. A **73** (1996), 1529–1563.
42. A. Toselli and O. Widlund, *Domain decomposition methods – algorithms and theory*, Springer Series in Computational Mathematics, 2004.
43. U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press Inc., San Diego, CA, 2001, with contributions by A. Brandt and P. Oswald and K. Stüben.
44. L. Truskinovsky, *Fracture as a phase transition*, Contemporary Research in the Mechanics and Mathematics of Materials (R. C. Batra and M. F. Beatty, eds.), CIMNE, Barcelona, 1996, pp. 322–332.
45. B. Van Koten and M. Luskin, *Analysis of energy-based blended quasi-continuum approximations*, SIAM J. Numer. Anal. **49** (2011), no. 5, 2182–2209.
46. D. J. Wales, *Energy landscapes*, Cambridge University Press, 2003.
47. Z. Wen and D. Goldfarb, *A line search multigrid method for large-scale nonlinear optimization*, SIAM J. Optim. **20** (2009), 1478–1503.
48. J. Xu, *Two-grid discretization techniques for linear and nonlinear PDEs*, SIAM J. Numer. Anal. **33** (1996), 1759–1777.

MATHEMATICAL CENTER FOR INTERDISCIPLINARY RESEARCH AND SCHOOL OF MATHEMATICAL SCIENCES, SOOCHOW UNIVERSITY, SUZHOU, 215006, CHINA AND MATHEMATICS DEPARTMENT, SOUTH HALL 6705, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA93106, USA

*E-mail address:* jingrunchen@suda.edu.cn

MATHEMATICS DEPARTMENT, SOUTH HALL 6707, UNIVERSITY OF CALIFORNIA, SANTA BARBARA, CA93106, USA AND BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS, MAZARREDO 14, E48009 BILBAO, BASQUE COUNTRY, SPAIN

*E-mail address:* cgarcia@math.ucsb.edu