Adaptive Multigrid Strategy for Large-scale Molecular Mechanics Optimization *

Mingjie Liao †, Yangshuai Wang ‡, Kejie Fu §, Jianjun Chen ¶, and Lei Zhang /uni2016

Abstract

In this paper, we present an efficient adaptive multigrid strategy for large-scale molecular mechanics optimization. The one-way multigrid method is used with inexact approximations, such as the quasi-atomistic (QA) approximation [8] or the blended ghost force correction (BGFC) approximation [36] on each coarse level, combined with adaptive mesh refinements based on the gradient-based a posteriori error estimator. For crystalline defects, like vacancies, micro-crack and dislocation, sublinear complexity is observed numerically when the adaptive BGFC method is employed. For systems with more than ten millions atoms, this strategy has a fivefold acceleration in terms of CPU time.

1 Introduction

In materials modeling [39, 40, 41], for a given atomistic system at zero (low) temperature, the equilibrium configuration can be obtained by minimizing the potential energy with respect to the positions of the atoms, which is called the molecular (atomistic) mechanics model. Efficient algorithms to compute equilibrium configurations have attracted considerable attentions from both the engineering community and the mathematical community in recent years [4, 5, 8, 9, 29, 42].

Due to the inherent discreteness of the model, the natural idea to solve this problem is to apply a (conjugate) gradient type or (quasi) Newton type minimization technique to find a local minimizer [5, 19], which is the so-called brute-force optimization. The computational complexity of one single step minimization is at least proportional to the number of atoms, and the number of steps to achieve the convergence may depend on the number of atoms, which leads to a super-linear scaling complexity overall. In addition, the existence of dislocations and cracks imposes a long-range effect on the displacement field, which further increases the cost of brute-force optimization.

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 [6, 17]. Popular choices of multigrid methods include V-cycle

---

*This work was partially supported by NSFC grant 11471214, 11871339, 11571314.
† mingjieliao@sjtu.edu.cn. Institute of Natural Sciences, School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University.
‡ yswang2016@sjtu.edu.cn. Corresponding author. Institute of Natural Sciences, School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University.
§ kjfu@zju.edu.cn. Center for Engineering and Scientific Computation, and School of Aeronautics and Astronautics, Zhejiang University.
¶ chenjj@zju.edu.cn. Center for Engineering and Scientific Computation, and School of Aeronautics and Astronautics, Zhejiang University.
/uni2016 lzhang2012@sjtu.edu.cn. Institute of Natural Sciences, School of Mathematical Sciences, and MOE-LSC, Shanghai Jiao Tong University.
multigrid, W-cycle multigrid, oneway (cascadic) multigrid, full multigrid (FMG) and full approximation scheme (FAS) for nonlinear problems. The application of multigrid methods for the non-linear optimization was studied recently by [16, 45, 30]. As for the usage in the molecular mechanics, the idea was first proposed by [6]. A more general approach was proposed by Chen and Ming [9]. The coarse-grid operator was constructed by the Cauchy–Born rule [10], and then transferred to the atomistic level in the framework of the oneway multigrid method. Unfortunately the efficiency of this approach relies heavily on the validity of the Cauchy–Born rule. Recently, the so-called quasi-atomistic approximation was proposed by [8], which was used on the nested uniform triangulation. However, it is unnecessary to use the uniform mesh in the region that far away from the defect core. To the best of our knowledge, a generic and efficient multigrid method (sub-linear scaling of complexity) for the molecular mechanics model is still lacking.

In our proposed adaptive multigrid strategies (Algorithm 3 and Algorithm 5), the oneway multigrid method (Algorithm 1) is used due to its efficiency [8]. At the last step of the oneway multigrid method, the finest atomistic problem still need to be solved, which means the coarse-grid problems only provide “good” initial guesses for the finest atomistic problem. This makes the resulting total computational complexity at least a linear scaling of complexity. However, the sub-linear scaling of complexity is surprisingly observed in our numerical experiments (see Figure 5 and Figure 10) if the BGFC method is solved adaptively at the coarse levels. We explain that although the linear scaling of finest atomistic problem still exists, the BGFC method can scale sub-linearly, and the number of steps in solving the finest atomistic problem is small since a very “good” initial guess is given by the BGFC method (see Figure 6(a) and its discussion in §5 for more details).

Atomistic/continuum (a/c) coupling method is a class of computational multiscale methods [20, 23, 24, 27, 36, 41] attempting to achieve sub-linear scaling of complexity, it aims to combine the accuracy of fine scale models and the efficiency of coarse scale models for crystalline defects. Namely, fine scale models can be applied in a small neighborhood of the localized defects such as point defects and dislocations, while coarse scale models can be employed away from the defect cores where elastic deformation occurs. Energy based methods and force based methods are two major classes of a/c coupling methods, we refer to [24, 27, 41] for reviews of many existing a/c coupling methods. One of the major challenges for energy based a/c coupling methods is to eliminate the so-called “ghost forces” near the atomistic/continuum interface. The blended ghost force correction (BGFC) scheme was first constructed in [36] by integrating two popular ideas: blending [23] and ghost force correction [37]. It is quasi-optimal in the sense that it yields the same convergence rate as the force based a/c coupling schemes [20].

The purpose of this paper is to improve the results of [8] by combining different coarse-grained models including the quasi-atomistic approximation and the BGFC approximations with corresponding adaptive algorithms. Even though the adaptive mesh refinement is used, the computational complexity of the quasi-atomistic approximation is of the order $O(N)$, where $N$ is the total number of atoms in the system. The a/c coupling methods, especially BGFC method, can obtain optimal accuracy with sublinear complexity, due to the usage of continuum model in far field. Given a posteriori error estimators, the triangulation in each step can be generated adaptively. We design the adaptive multigrid algorithms and verify the algorithms by some prototypical benchmark examples such as the single-vacancy, micro-crack and edge dislocation in 3D.

Outline. The paper is organized as follows: In §2 we introduce the reference molecular mechanics optimization problem and the multigrid strategy. In §3 we propose several coarse grained models including the quasi-atomistic approximation given by [8], the truncated-atomistic approximation and the BGFC method [36]. We then design the corresponding adap-
tive multigrid algorithms for large-scale molecular mechanics optimization with different coarse-grid problems in §4 and we implement the algorithms for several defect cases. Possible additional applications are discussed in §6.

2  Model Setup

In this section, we setup the atomistic problem for crystal defects and introduce the multigrid strategy for atomistic simulations. For simplicity of notations, we mainly discuss the point defects in this section, but we note that all algorithms discussed in this work can be developed to straight dislocations and multi-lattice crystals (see [12, 31]).

2.1 Atomistic problem

Let \( \Lambda^{\text{hom}} \subset \mathbb{R}^d \) be a perfect single lattice crystal possessing no defects and \( \Lambda \subset \mathbb{R}^d \) is the single lattice with some local defects. The mismatch between \( \Lambda \) and \( \Lambda^{\text{hom}} \) represents possible defects, which are contained in some localized defect cores. For simplicity, we assume the defects are near the origin and

\[
\Lambda \setminus B_{R_{\text{DEF}}} = \Lambda^{\text{hom}} \setminus B_{R_{\text{DEF}}},
\]

with \( B_{R_{\text{DEF}}} \) being a ball centered at the origin with radius \( R_{\text{DEF}} \geq 0 \).

Given \( d \in \{1, 2, 3\} \), we denote the set of vector-valued lattice functions by

\[
\mathcal{U} := \{ v : \Lambda \to \mathbb{R}^d \}.
\]

Recall that the deformed configuration of \( \Lambda \) is a map \( y \in \mathcal{U} \) which can be decomposed as

\[
y(\ell) = x_0(\ell) + u_0(\ell) + u(\ell) =: x(\ell) + u(\ell), \quad \forall \ell \in \Lambda
\]

with identity map \( x_0 \in \mathcal{U}, x_0(\ell) = \ell, \forall \ell \in \Lambda \), \( u_0 \in \mathcal{U} \) a predictor prescribing the far-field boundary condition, and \( u \in \mathcal{U} \) the corrector. We denote the positions of the atomistic reference configuration as \( x(\ell) := x_0(\ell) + u_0(\ell), \ell \in \Lambda \), which is usually a far-field crystalline environment or an explicit linearised elasticity solution. For point defects, we simply take \( u_0 = 0 \). For dislocations, \( u_0 \) can be constructed following [12, 7].

For each atom \( \ell \in \Lambda \), we define the finite difference stencil for \( v \in \mathcal{U} \)

\[
Dv(\ell) := \{ D_\rho v(\ell) \}_{\rho \in R_\ell} := \{ v(\ell + \rho) - v(\ell) \}_{\rho \in R_\ell},
\]

where \( R_\ell := \{ \ell' - \ell \mid \ell' \in N_\ell \} \) is the interaction range with interaction neighborhood \( N_\ell := \{ \ell' \in \Lambda \mid 0 < |\ell' - \ell| \leq r_{\text{cut}} \} \) with some cut-off radius \( r_{\text{cut}} \).

Let \( \tilde{\zeta} \in W^{1,\infty}(\Lambda^{\text{hom}}; \mathbb{R}) \) be the \( \mathcal{P}_1 \) nodal basis function associated with the origin satisfying \( \tilde{\zeta}(0) = 1 \) and \( \tilde{\zeta}(\ell) = 0 \) for \( \ell \neq 0 \) and \( \ell \in \Lambda^{\text{hom}} \). The nodal interpolant of \( v \in \mathcal{U} \) can be written as

\[
\tilde{v}(x) := \sum_{\ell \in \Lambda} v(\ell)\tilde{\zeta}(x - \ell).
\]

We then introduce the discrete homogeneous Sobolev spaces via the nodal interpolant [36, 21, 12]

\[
\mathcal{U}^{1,2} := \{ u \in \mathcal{U} \mid \nabla \tilde{u} \in L^2 \},
\]

with semi-norm \( \| \nabla \tilde{u} \|_{L^2} \).
In this paper, we consider the general multibody interaction potential of the generic pair functional form, which includes the widely used potentials such as EAM (Embedded Atom Method) potential \cite{11} and Finnis-Sinclair model \cite{15}. Namely, the potential is a function of the distances between atoms within interaction range and has no angular dependence. For example, for each \( \ell \in \Lambda \), let \( E_\ell(y) \) denote the site energy associated with the lattice site \( \ell \in \Lambda \), the potential reads,
\[
E_\ell(y) := \sum_{\ell' \in N_\ell} \Phi(|y(\ell) - y(\ell')|) + F\left(\sum_{\ell' \in N_\ell} \varphi(|y(\ell) - y(\ell')|)\right),
\]
\[
= \sum_{\rho \in \mathcal{R}_\ell} \Phi(|D_\rho y(\ell)|) + F\left(\sum_{\rho \in \mathcal{R}_\ell} \varphi(|D_\rho y(\ell)|)\right),
\]
with the pair potential \( \Phi \), the electron density function \( \varphi \) and the embedding function \( F \).

We assume that the potential \( E_\ell(y) \in C^k((\mathbb{R}^d)\mathcal{R}_\ell), k \geq 2 \) and \( E_\ell(y) \) is homogeneous outside the defect region \( \mathcal{D}^{\text{def}} \), namely, \( E_\ell = E \) and \( \mathcal{R}_\ell = \mathcal{R} \) for \( \ell \in \Lambda \setminus \mathcal{D}^{\text{def}} \). Furthermore, \( E \) and \( \mathcal{R} \) have the following point symmetry: \( \mathcal{R} = -\mathcal{R} \), and \( E(\{-g_{-\rho}\}_{\rho \in \mathcal{R}}) = E(g) \). We define \( V_\ell : (\mathbb{R}^d)\mathcal{R}_\ell \rightarrow \mathbb{R} \) by
\[
V_\ell(Du_0 + Du) := E_\ell(x_0 + u_0 + u).
\]
We can formally define the energy-difference functional
\[
\mathcal{E}(u) = \sum_{\ell \in \Lambda} \left( E_\ell(x_0 + u_0 + u) - E_\ell(x_0 + u_0) \right)
\]
\[
= \sum_{\ell \in \Lambda} \left( V_\ell(Du_0 + Du(\ell)) - V_\ell(Du_0(\ell)) \right).
\]
It was shown in \cite{12} that \( \mathcal{E}(u) \) is well-defined on the space \( \mathcal{U}^{1,2} \).

Throughout this paper, we project the atomistic problem to a finite-dimensional subspace. Due to the decay estimates \cite[Theorem 1]{13}, it is natural to simply truncate to a finite domain. We choose a computational domain \( \Omega_R \subset \Lambda \) satisfying \( (B_{R_l} \cap \Lambda) \subset \Omega_R \subset (B_{R_0} \cap \Lambda) \) for given constants \( 0 < R_0 < R_l \). For simplicity we will use \( \Omega \) instead of \( \Omega_R \) when there is no confusion in the context. We also denote \( \Lambda^\Omega \) as \( \Lambda \cap \Omega \) and \( N := \#\Lambda^\Omega \). We then modify the displacement space \( \mathcal{U}^{1,2} \) as
\[
\mathcal{U}^{1,2}_N := \{ u \in \mathcal{U}^{1,2} \mid u = 0 \text{ in } \Lambda \setminus \Omega \},
\]
which indicates the Dirichlet boundary condition is applied.

We can now rigorously formulate the atomistic problem for the equilibrium state as
\[
P_N^\Lambda : \ u_N^\lambda \in \arg \min \{ \mathcal{E}(u) \mid u \in \mathcal{U}^{1,2}_N \},
\]
where “arg min” is understood as the set of local minima. Our paper focuses on the adaptive multigrid algorithm for large-scale (the system size \( N > 10^5 \)) molecular optimization.

### 3 Coarse-grid Problems

In this section, we introduce two coarse-grid problems, whose computational costs are much less than the finest atomistic problem, including the quasi-atomistic approximation \cite{8} and the recently developed blended ghost force correction (BGFC) method \cite{36} in the field of atomistic-to-continuum coupling methods.
3.1 Quasi-atomic approximation

Given \( k \in \mathbb{N} \), let \( T_k \) be a shape-regular simplicial partition of \( \Omega \) for \( k \)-th level and \( N_k := \{ \xi_1, ..., \xi_{n_k} \} \) be the set of all nodes of \( T_k \), where \( n_k := \#N_k \). We note that the coarse-grid points \( \xi_1, ..., \xi_{n_k} \) may not necessarily be the atomistic reference positions \( x(\ell) \), for \( \ell \in \Lambda^\Omega \). We first define the solution space for \( k \)-th level

\[
U_{QA}^k := \left\{ u_k : N_k \rightarrow \mathbb{R}^d \mid \nabla \bar{u}_k \in L^2, \ u_k = 0 \text{ on } \partial \Omega \right\}.
\]

We define the basis function centered on the \( j \)-th node of given partition \( T_k \) by \( \psi(\cdot - \xi_j) \), where for simplicity we assume \( \psi \) has compact support. The basis function \( \psi \) is normally chosen such that

\[
\sum_{j=1}^{n_k} \psi(\omega - \xi_j) = 1, \quad \forall \omega \in \Omega,
\]

which means a constant displacement field is preserved after interpolation and it can be chosen as the nodal interpolation defined in (2.3).

For \( k \)-th level, given \( u_k \in U_{QA}^k \), the atomistic displacement at \( \omega \) can be constructed by

\[
u(\omega) = \sum_{j=1}^{n_k} \psi(\omega - \xi_j)u_k(\xi_j), \quad \forall \omega \in \Omega.
\]

Then, the deformation of atom \( \ell \), for any \( \ell \in \Lambda^\Omega \), can be written as

\[
y(\ell) = x(\ell) + \sum_{j=1}^{n_k} \psi(x(\ell) - \xi_j)u_k(\xi_j).
\]

By inserting (3.11) into (2.6), we can obtain the energy functional of quasi-atomic approximation

\[
E_{QA}(u_k) := \mathcal{E} \left( \sum_{j=1}^{n_k} \psi(x(\ell) - \xi_j)u_k(\xi_j) \right),
\]

where \( x := x_0 + u_0 \) is the reference atomistic configuration defined in § 2.1.

The coarse-grid problem for quasi-atomic approximation on \( k \)-th level reads

\[
P_{QA}^k : u_{qa}^k \in \text{argmin} \{ E_{QA}(u_k) \mid u_k \in U_{QA}^k \}.
\]

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. This makes the computational complexity of (3.13) as \( O(N) + O(n_k) \).

As for the choice of optimization solver, according to [8], CG_DESCENT [18] gives the best compromise between accuracy and efficiency compared with L-BFGS [22]. Moreover, the line search implementation in CG_DESCENT appears to be the most robust on coarse levels. Therefore, (3.13) is solved by CG_DESCENT, which is implemented by an open source Julia package, Optim.jl [28]. We denote \( u_{qa}^{(\mu_k)} \) as its approximated solution after \( \mu_k \) steps of CG_DESCENT.

Remark 3.1. We note that the computational complexity for the quasi-atomic approximation in one step minimization is still unaffordable since it depends on \( N \) linearly, where \( N \) is the total number of atoms in \( \Omega \). The continuous approximation such as Cauchy-Born approximation [44] at the region that far away from the defect core should be considered if our goal is the sub-linear scaling of complexity. We will introduce the atomistic-to-continuum (a/c) coupling methods as the coarse-grid problem to reduce the degrees of freedom significantly in the next section. The a/c coupling methods need a priori know the probable location of defects [8], however throughout this paper we assume that the location has already been known.
3.2 BGFC Method

In this section, we first introduce the continuum approximation and then present the blended ghost force correction (BGFC) method. Though they have already been discussed rigorously in many multiscale modeling literatures [14, 21, 36], we adapt them to the setting of our paper for the sake of completeness.

3.2.1 Continuum approximation. As a starting point of BGFC method, from the atomistic model (2.6), a continuum model can be derived by coarse graining, and computationally it allows for the reduction of degrees of freedom where the deformation is smooth, e.g. the region that far away from the defect core. A typical choice in the multi-scale context is the Cauchy-Born continuum model [33, 44]. Let $W: \mathbb{R}^{d \times d} \rightarrow \mathbb{R}$ be a strain energy density function, the Cauchy-Born energy density $W$ is defined by

$$ W(F) := V(F \cdot R), $$

and the Cauchy-Born energy difference is defined by

$$ W'(F) := W(F + l) - W(l), \quad \forall F \in \mathbb{R}^{d \times d}, $$

where $l \in \mathbb{R}^{d \times d}$ is the identity matrix.

We note that Cauchy-Born approximation can not be used as coarse-grid problem directly since it can not capture the local defects [10, 9], which then leads to low efficiency of the corresponding algorithm. We will use the blended ghost force correction (BGFC) method as our coarse-grid problem instead to achieve the quasi-balance between accuracy and computational cost [36].

3.2.2 BGFC model. We first decompose the computational domain $\Omega = \Omega^a \cup \Omega^b \cup \Omega^c \subset \mathbb{R}^d$ into three regions, consists of the atomistic region $\Omega^a$ with width $R^a$, the blending region $\Omega^b$ with width $R^b$ and the continuum region $\Omega^c$. Given the reference lattice $\Lambda$ with some local defects, we define the core atomistic set $\Lambda^a := \Lambda \cap \Omega^a$ and the blended set $\Lambda^b := \Lambda \cap \Omega^b$. Let $\mathcal{T}_k^a$ be the canonical triangulation induced by $\Lambda^a \cup \Lambda^b$, and $\mathcal{T}_k^c$ be a shape-regular simplicial partition of the continuum region. For $k$-th level, we denote $\mathcal{T}_k = \mathcal{T}_k^a \cup \mathcal{T}_k^c$ as the triangulation of the a/c coupling configuration. See Figure 1 for an illustration of $\mathcal{T}_k$.

![Image](image.png)

Figure 1: Illustration of $\mathcal{T}_k$, red mesh denotes $\mathcal{T}_k^a$ while green mesh for $\mathcal{T}_k^c$. 


We introduce the blending function $\beta \in C^{2,1}(\mathbb{R}^d)$ with $\beta = 0$ in $\Omega^a$ and $\beta = 1$ in $\Omega^c$. The blending region $\Omega^b$ has width $R^b - R^a \sim R^a$. The choice of the blending functions $\beta$ will heavily influence the convergence and it has already been discussed in detail in [25], so we omit it here and choose the optimal (in the sense of optimal convergence rate) $\beta$ in practice.

The space of coarse-grained displacements is,

$$U_{BGFC}^k := \{ u_k : \Omega \to \mathbb{R}^d \mid u_k \text{ is continuous and p.w. affine w.r.t. } T_k, \quad u_k = 0 \text{ in } \Lambda \setminus \Omega \}.$$ 

The energy functional of the BGFC approximation is based on a second renormalization of the potential [36], for $\ell \in \Lambda$, $u \in U^{1,2}$,

$$V''_\ell(Du) := V_\ell(Du) - V_\ell(Du_0) - \langle \delta V_\ell(Du_0), Du \rangle,$$

where $u_0$ is the predictor introduced in §2.1. The corresponding second renormalized Cauchy-Born energy density is

$$W''(F) := W(F + I) - W(I) - \langle \delta W(I), F \rangle, \quad \text{for } F \in \mathbb{R}^{d \times d}. \quad (3.16)$$

We then obtain the BGFC energy functional

$$E_{BGFC}(u_k) = \sum_{\ell \in \Lambda^a \cup \Lambda^b} (1 - \beta(\ell)) V''_\ell(Du_k) + \int_\Omega Q_h \left[ \beta(x) W''(\nabla u_k) \right] dx,$$

where $Q_h$ is the $P_0$ midpoint interpolation operator, so that $\int_\Omega Q_h f dx$ is the mid-point rule of $\int_\Omega f dx$. The coarse-grid problem for BGFC approximation on $k$-th level reads

$$P_{BGFC}^k : u_{bgfc}^k \in \arg\min \{ E_{BGFC}(u_k) \mid u_k \in U_{BGFC}^k \}. \quad (3.18)$$

The error analysis of BGFC method in 3D is given by [14, Theorem 2.1]. It shows that if the blending function $\beta$ and the triangulation $T_k$ satisfy certain assumptions [14, Assumption 1], the following convergence can be achieved, $\| \nabla \bar{u} - \nabla \bar{u}_{bgfc}^k \|_{L^2} \lesssim n_k^{-5/6}$, where $\bar{u}$ and $\bar{u}_{bgfc}^k$ are the exact solution of (2.7) and (3.18) respectively after the interpolation (2.3), and $n_k$ is the degree of freedom of space $U_{BGFC}^k$.

Similar with (3.13), the minimization problem (3.18) on $k$-th level is also solved by CG_DESCENT [18] and we denote $u_{bgfc}^{k, \mu_k}$ as its approximated solution after $\mu_k$ steps of CG_DESCENT.

Remark 3.2. The resulting BGFC method combines the efficiency of the blending methods as well as the accuracy of the quasi-nonlocal (QNL) type consistent methods [32]. Its efficiency will be shown in §4. Moreover, compared to other energy-based consistent a/c coupling methods (e.g. geometry reconstruction atomistic-to-continuum (GRAC) coupling method [34]), BGFC method is more friendly for implementation and easier to extend to three dimensional problems, hence we choose it as the coarse-grid problem in this paper.

4 Numerical Algorithm

In this section, we first introduce the mesh generation and multigrid strategy that needed in our numerical experiments in §4.1 and §4.2. We then design the adaptive multigrid algorithms with the coarse-grid problems introduced above in §4.3.
4.1 Mesh generation

In practice, for the triangulation $T_k$ on $k$-th level, we use the cubic triangulation for quasi-atomistic approximation and the tetrahedral triangulation in continuum region for BGFC method. The reason that we choose different triangulation is as following. In [8] the quasi-atomistic approximation with uniform cubic triangulation was used, to make a comparison, we apply the cubic triangulation for our adaptive quasi-atomistic approximation. However, in the literature of $a/c$ coupling methods [27, 33], the tetrahedral triangulation is more natural than the cubic one for its convenience of evaluating the strain $\nabla u_k$ in Cauchy-Born energy functional (3.17). Hence, the tetrahedral triangulation is used to obtain $T_k^c$, combing with the canonical triangulation $T_k^a$, the triangulation $T_k$ is then constructed. We note that though the different kinds of triangulations are used in different algorithms (e.g. Algorithm 3 and Algorithm 5) but they will not influence the numerical results in principle.

If the coarse-grid problem is chosen to be $P_{QA}^k$ defined by (3.13) on $k$-th coarse level, we use the package Deal II [1, 3] to generate the cubic triangulation (see Figure 2 for illustration). We note that in practice the adaptive mesh refinements are also done by Deal II only if the local error estimators are given, and we will discuss the error estimators in the next section. The advantage of Deal II is open-sourced and well-developed. The cubic elements will lead to hanging nodes, which are the nodes on the interfaces of cells which belong to one side, but are unbalanced on the other. The Deal II package solves the problem of hanging nodes automatically [1, 3].

Figure 2: Illustration of the cubic triangulation for quasi-atomistic approximation generated by Deal II [1, 3].

If the coarse-grid problem is chosen to be $P_{BGFC}^k$ defined by (3.18) on the given $k$-th coarse level, the tetrahedral triangulation is used for BGFC method. The implementation of BGFC method is available as an open source Julia package, JuAC.jl [26]. Its mesh generator originally comes from Tetgen [38], which is a C++ program for generating good quality tetrahedral meshes. Compared with the triangulation used in [14] which was only available for FCC crystal, our triangulation can handle more general case. See Figure 1 for an illustration of the triangulation.

As shown in Figure 1 the atomistic regions contains those atoms in the center of the computational domain. The tetrahedrons in the continuum region are constructed by graded tesselation from the atomistic scale to the boundary length scale.

4.2 Multigrid strategy

In this section, we briefly introduce the idea of multigrid strategy for the molecular mechanics model. Given $L \in \mathbb{N}$, we first consider a nested sequence of triangulations $T_0 \subset T_1 \subset \ldots \subset T_L$ of $\Omega$. Its construction has already been discussed in the last section. Let $T_N$ be the canonical
triangulation induced by ΛΩ and hN be the lattice spacing. For T ∈ Tk, k = 1, · · · , L, we define the diameter of T, diam(T) := sup{|x − y|, x, y ∈ T}. The mesh size hk = minT∈Tk diam(T) which satisfies hk = hk−1/2 for k = 1, · · · , L. From current level to next level, we define a operator to pass the data. We denote Ij as the interpolation operator (e.g. tri-linear interpolation operator, i.e., the linear interpolation in each coordinate direction) from level i to level j if i < j and as the restriction operator if i > j. We note that the oneway multigrid strategy is utilized throughout this paper, hence the restriction operator will not be used in practice.

In this paper, we make full use of the oneway multigrid strategy for the reasons of its efficiency. In [8], they have shown that FMG strategy is even computationally more expensive than the reference atomistic optimization. Since defects always have localized effects, the coarse-grid model cannot give a very good approximation to the atomistic model. The approximation becomes better only when the grid is sufficiently close to the atomistic resolution. Therefore, the V-cycle part in FMG is inefficient, resulting in the inefficiency of FMG. If the V-cycle part is removed from FMG, then it basically recovers the oneway multigrid and is computationally more efficient. Hence we use the oneway multigrid throughout this paper, which is given by the following algorithm.

**Algorithm 1 Oneway multigrid strategy**

1. Relax the coarse-grid problem PQA0 (PBGFC0) on the initial mesh T0 to obtain u0 with a trivial initial guess.

2. For k = 1, · · · , L, relax the coarse-grid problem PQAk (PBGFCk) on Tk to obtain uk with the initial guess Ik−1uk−1.

3. Solve the atomistic problem PAn (2.7) until convergence with the initial guess INL uL.

### 4.3 Algorithms

In this section, we propose our main algorithms including the oneway adaptive multigrid strategy with quasi-atomistic approximation (Algorithm 3), and the oneway adaptive multigrid strategy with BGFC method (Algorithm 5).

Before we introduce the algorithms of adaptive multigrid, we first note that the partition Tk in the quasi-atomistic approximation introduced in §3.1 and the domain decomposition in the BGFC method introduced in §3.2 are all adaptively constructed by using the following mesh refinement strategy with the well-known Dörfler strategy from the initial triangulation T0.
Algorithm 2 Mesh refinement strategy.
Prescribe $0 < \tau_1 < \tau_2 < 1$ and $k \in \mathbb{N}$.

1. Given a partition $\mathcal{T}_k$, according to the approximate solution $u_k$, compute the local error estimator $\eta_T = ||\nabla u_k||_{L^2(T)}$ for each element $T \in \mathcal{T}_k$ and the global error estimator $\eta_k = \sum_{T \in \mathcal{T}_k} \eta_T$.

2. Choose the maximum sets $\mathcal{T}_c \subset \mathcal{T}_k$ and the minimum sets $\mathcal{T}_r \subset \mathcal{T}_k$ such that the following Dörfler properties are satisfied

\[
\sum_{T \in \mathcal{T}_c} \eta_T \leq \tau_1 \eta_k, \quad \sum_{T \in \mathcal{T}_r} \eta_T \geq \tau_2 \eta_k.
\] (4.19)

3. Mark all the elements in $\mathcal{T}_r$ for refinement (bisection) and the elements in $\mathcal{T}_c$ for coarsening to obtain $\mathcal{T}_{k+1}$.

Remark 4.1. We note that the local a posteriori error estimator for each coarse-grid problem, is simply chosen as $\eta_T = ||\nabla u_k||_{L^2(T)}$, where $T \in \mathcal{T}_k$ and $u_k$ is the approximated solution of the coarse-grid problems $P_{QA}^k$ or $P_{BGFC}^k$ on $k$-th level. We will numerically show the reliability of the estimator and its good performance (see §5), but the theoretical analysis is not involved since it is not the key point in this work.

We then give the following adaptive multigrid strategy by using the coarse-grid problem $P_{QA}^k$ defined by (3.13) on $k$-th coarse level. Here we recall that the total number of atoms in atomistic problem and the degree of freedom of the coarse-grid problem on $k$-th level are denoted as $N$ and $n_k$, respectively.

Algorithm 3 Oneway adaptive multigrid strategy with quasi-atomistic approximation.

1. Prescribe the parameters $\mu_k$ and $c$.

2. Relax the quasi-atomistic problem $P_{QA}^0$ defined by (3.13) on the initial mesh $\mathcal{T}_0$ for $\mu_0$ times to obtain $u_0$ with a trivial initial guess, compute $n_0$, set $k = 1$.

3. Carry out the mesh refinement based on Algorithm 2 to obtain $\mathcal{T}_k$, relax the quasi-atomistic problem $P_{QA}^k$ defined by (3.13) on $\mathcal{T}_k$ for $\mu_k$ times to obtain $u_k$ with the initial guess $I_{k-1}^k u_{k-1}$, compute $h_k$, if $h_k \leq c \ast h_N$, $k = k + 1$, repeat Step 3; Otherwise, goto Step 4.

4. Solve the atomistic problem $P_A^N$ defined by (2.7) until it is convergence with the initial guess interpolated from the solution on $\mathcal{T}_k$.

We first give a guidance of how to generate the atomistic region and to refine the continuum region based on the local error estimators, which basically follows the idea of [43].
Algorithm 4 Mesh refinement for BGFC method.
Prescribe $0 < \tau_3, \tau_4 < 1$.

1. Given a partition $\mathcal{T}$ and the approximation solution $u$, compute the local error estimator
   $\rho_T = ||\nabla u||_{L^2(T)}$ for each element $T \in \mathcal{T}$.
2. Choose a minimal subset $\mathcal{M} \subset \mathcal{T}$ such that
   $$\sum_{T \in \mathcal{M}} \rho_T \geq \tau_1 \sum_{T \in \mathcal{T}} \rho_T.$$
3. Find the interface elements $\mathcal{M}_i := \{T \in \mathcal{M} : T \cap \Lambda_i \neq \emptyset\}$. If
   $$\sum_{T \in \mathcal{M}_i} \rho_T \geq \tau_2 \sum_{T \in \mathcal{M}} \rho_T,$$
   let $\mathcal{M} = \mathcal{M} \setminus \mathcal{M}_i$.
4. Expand the atomistic region $\Lambda_a$ outward by one layer and bisect all elements $T \in \mathcal{M}$ to
   obtain $\mathcal{T}_{k+1}$.

Finally, according to Algorithm 4, we give the following adaptive multigrid strategy by using
the coarse-grid problem $P_k^{BGFC}$ defined by (3.18) on $k$-th coarse level.

Algorithm 5 Oneway adaptive multigrid strategy with BGFC method.

1. Prescribe the relaxation step $\mu_k$, the tolerance $tol_k$ on $k$-th level and $c$.
2. Relax the BGFC problem $P_0^{BGFC}$ defined by (3.18) on the initial mesh $\mathcal{T}_0$ until reaching
   the tolerance $tol_0$ or reaching $\mu_0$ times to obtain $u_0$ with a trivial initial guess, compute
   $n_0$, set $k = 1$.
3. Carry out the mesh refinement based on Algorithm 4 to obtain $\mathcal{T}_k$, solve the $P_k^{BGFC}$
   defined by (3.18) with the initial guess interpolated from the solution on $\mathcal{T}_k$ until reaching
   the tolerance $tol_k$ or reaching $\mu_k$ times, compute $n_k$, if $n_k \leq c \cdot N$, $k = k + 1$, then goto
   Step 3 again; Otherwise, goto Step 4.
4. Solve the atomistic problem $P_N^A$ defined by (2.7) until convergence with the initial guess
   interpolated from the solution on $\mathcal{T}_k$.

5 Numerical Experiments

test our algorithms for single vacancy, micro-crack and edge dislocation in §5. The numerical
results first show that all strategies can outperform the brute-force optimization for all examples
if more than 200000 atoms are involved. Moreover, all three cases show that the combination
of BGFC method and the adaptive multigrid strategy can achieve the sub-linear computational
complexity.

In this paper, we use pure tungsten for all numerical experiments, which is a common body-
centered cubic (BCC) structure. The cut-off radius is chosen as $r_{cut} = 5.6$, which includes up to
the third neighbour interaction. All numerical experiments are tested with Dirichlet boundary
condition used in three directions. To avoid the boundary effects, we introduce several layers of ghost atoms whose thickness are greater than $2r_{\text{cut}}$ outside of the domain of interest.

Recall the EAM potential defined in (2.4). Let

$$\phi(r) = \exp(-2a(r - 1)) - 2 \exp(-a(r - 1)), \quad \varphi(r) = \exp(-br),$$

$$F(\rho) = C \left[ (\rho - \rho_0)^2 + (\rho - \rho_0)^4 \right],$$

with parameters $a = 4.4, b = 3, c = 5$ and $\rho_0 = 6 \exp(-b)$, which is the same setting as that in [35].

We apply our oneway adaptive multigrid strategies (Algorithm 3 and Algorithm 5) for three cases: single vacancy, micro-crack and [100](001) edge dislocation. We use the same stopping criteria $|g|_{\infty} < \text{tol} = 10^{-4}$ for single vacancy and micro-crack while $|g|_{\infty} < \text{tol} = 10^{-3}$ for edge dislocation. In order to show the efficiency of our strategies, we compare several methods including the brute-force optimization, oneway multigrid strategy with quasi-atomistic approximation introduced in [8], oneway multigrid strategy with quasi-atomistic approximation by using adaptive mesh refinements (Algorithm 3), and oneway multigrid strategy with BGFC method (Algorithm 5).

5.1 Parameters study

We note that the stopping criteria $\mu_k$ or $\text{tol}_k$ of the coarse-grid problems on $k$-th coarse level will heavily influence the performance of algorithms. Hence, in this section, we mainly study the parameters in Algorithm 3 and Algorithm 5.

As [8] shows, the presence of defect introduces a localized and inhomogeneous displacement which can only be well described by the atomistic model, and the resolution of the solution of the coarse-grid model is poor. Therefore, any attempt to solve the coarse-grid model to high accuracy will slow down the convergence of multigrid overall. Since coarse-grid approximations cannot provide very good initial guesses in the vicinity of defects, only quite a few iterations are required to achieve the convergence. Throughout this paper, we choose the iteration step $\mu_k$ in Algorithm 3 to be less than 5, which is the same as the setting in [8].

As for the stopping criterions in Algorithm 5, we observe that the BGFC method can already capture the effects of defects, which is a natural advantage compared to the quasi-atomistic approximation. To study the parameters that may affect the performance in Algorithm 5, we use single vacancy (its detailed setting will be introduced in the following section) to test the Algorithm. We first set $R^a, R^b$, and $R_{\Omega}$ to be $4r_0, 2r_0$ and $20r_0$ respectively, where $r_0$ is the lattice constant of tungsten. We compute the total error $\|u - u_k^{\text{BGFC}}\|_{L^2}$ and the algebraic error $\|u_k^{\text{BGFC}} - u_k^{\text{BGFC}}(\mu_k)\|_{L^2}$ with $u$ the exact solution of (3.13), $u_k^{\text{BGFC}}$ the exact solution of (3.18) and $u_k^{\text{BGFC}}(\mu_k)$ the approximated solution of (3.18) after $\mu_k$ steps of relaxation. Figure 3 shows that the algebraic error decays continuously as the iteration $\mu_k$ increases. However, the total error is bounded since the modeling error between the atomistic and BGFC method dominant in this case.

We then still use single vacancy case for $N = 4.32 \times 10^5$ and $N = 3.456 \times 10^6$ to test different stopping criteria as [8] did. We list the CPU time (in seconds) of the oneway adaptive multigrid strategy with BGFC method versus several stopping criterions on each level (except the atomistic level), including fixed accuracy, fixed the number of relaxations, number of relaxations control on each level and accuracy control on each level. For comparison, CPU time of the brute-force optimization for corresponding system is also listed. Denote $\text{tol}_k$ as the tolerance at $k$-th level and set the tolerance at the atomistic level as $\text{tol}_L = 10^{-4}$. The oneway multigrid with accuracy control on each level is done by setting $\text{tol}_k = \text{tol}_L \times 2^{L-k}$, which is required to achieve

[12]
the convergence. Table 5.1 shows that compared to the brute-force optimization, the one-way multigrid with accuracy control is quite efficient, while for fixed accuracy ($\text{tol}_k = 10^{-2}, 10^{-3}$) or for fixed number of relaxations ($\mu_k = 3, 10$), there are no heavy differences compared to the brute-force optimization. Moreover, the good sub-linear scaling of complexity can be achieved when the accuracy control on each level is used, while others can still get linear scaling of complexity. The results are different from that in [8], where fixed number of relaxations $\leq 5$ is more efficient than the accuracy control. The reason is natural since the quasi-atomistic approximation on coarse-grid level can not capture the defect while BGFC method expertize it. We then use this stopping criterion in Algorithm 5 throughout the work without any tuning.

| Strategy vs Time (s) | CPU time ($N = 4.32 \times 10^7$) | CPU time ($N = 3.456\times 10^6$) |
|----------------------|----------------------------------|----------------------------------|
| fixed accuracy 1e-2  | 2847                             | 20242                            |
| fixed accuracy 1e-3  | 2321                             | 14888                            |
| fixed number 3       | 2593                             | 15847                            |
| fixed number 10      | 2065                             | 13418                            |
| accuracy control     | 1832                             | 9275                             |
| brute-force          | 2215                             | 21842                            |

Table 1: Stopping criteria

5.2 Single vacancy

The atomistic systems used for the single vacancy case have $30^3, 60^3, 120^3, 180^3, 300^3$ and $370^3$ unit cells such that the degree of freedom ($N$) of the reference atomistic problem ranges from $5.4 \times 1^{-4}$ to $1.014 \times 10^8$. Dirichlet boundary conditions are used in all three directions with ghost atoms. To create a single vacancy for the tungsten system and preserve symmetry, we simply remove the center atoms in each system. Figure 4 plots the displacement field in $L_2$ norm over a centered slice in the $z$ direction of $30^3$ system, showing that a vacancy has a localized effect. According to [13, Theorem 1], the effect of point defects follows $1/r^2$ decay law typically, where $r$ is the distance to the defect core.

Three relaxations on coarse levels are used in the adaptive one-way adaptive multigrid method with quasi-atomistic approximations. The parameters in the adaptive one-way multigrid with BGFC method are determined in § 5.1. For $30^3$ system, we compute the total error...
∥u − u_{bgfc,(i)}^{k}\parallel_{L^2}$, where $u$ is the reference solution and $u_{bgfc,(i)}^{k}$ is algebraic errors.

Figure 5(a) plots the CPU time of the brute-force optimization (the red line), the oneway multigrid method in [8] (the blue line), the adaptive oneway multigrid method with quasi-atomistic (the green line), and the global refined adaptive oneway multigrid with BGFC method (the purple line) in the log–log scale with the ratio between these methods and the brute-force optimization in Figure 5(b). Except the global refined adaptive oneway multigrid with BGFC method, the brute-force optimization and other methods scale $O(N^{1.05})$ asymptotically. The global refined adaptive oneway multigrid with BGFC method scales sub-linearly $O(N^{0.78})$.

Moreover, compared to the brute-force optimization, the savings of this method are around 80% for systems with ten millions atoms. Figure 7 and Figure 8 show the mesh evolution in the adaptive process for quasi-atomistic and BGFC approximations correspondingly for $30^3$ system.
5.3 Micro-crack

Systems used for the crack case have $30^3, 60^3, 120^3, 180^3, 300^3$ and $370^3$ unit cells. Dirichlet boundary conditions are used in all three directions with ghost atoms. To create a micro-crack for the tungsten system, we remove seven adjacent atoms in the center of the xy plane along the y direction. Figure 9 plots the displacement field in $L_2$ norm over a centered slice in the z direction of $30^3$ system. Around crack tips, one can see the nonlocal effect on the displacement field in the presence of cracks. Compared to Figure 4, we see that the nonlocal effect for a crack is stronger than that of single vacancy. Therefore, the crack case is more difficult than the single vacancy case to find the local minimizer.

Figure 8: Adaptive mesh refinement (DOF from left to right: 942, 1414, 2420, 4296) of BGFC approximation for single vacancy.

Figure 7: Adaptive mesh refinement (DOF from left to right: 125, 272, 618, 1536) of quasi-atomistic approximation for single vacancy.

Figure 6: Error in Adaptive BGFC approximation.

Figure 9: Plots the displacement field in $L_2$ norm over a centered slice in the z direction of $30^3$ system.

Figure 10: Around crack tips, one can see the nonlocal effect on the displacement field in the presence of cracks.

Figure 11: Compared to Figure 4, we see that the nonlocal effect for a crack is stronger than that of single vacancy.
Figure 9: Displacement for micro-crack case.

Figure 10: Costs of the brute-force optimization, oneway multigrid methods with uniform mesh and adaptive multigrid method for the micro-crack case.

Three relaxations on coarse levels are used in the adaptive one way adaptive multigrid method with quasi-atomistic and truncated atomistic approximations. The parameters on coarse levels in the adaptive one way multigrid with BGFC method are determined in §5.1. Figure 10(a) plots the CPU time of the brute-force optimization, the one way multigrid method in [8], the adaptive one way multigrid method with quasi-atomistic (Algorithm 3), the global refined adaptive one way multigrid with BGFC method and the local refined adaptive one way multigrid with BGFC method (Algorithm 5) in the log–log scale with the ratio between these methods and the brute-force optimization in Figure 10(b). Due to the non-locality of the displacement field in the presence of cracks, the brute-force optimization scales $O(N^{1.27})$ asymptotically, while other methods except the global refined adaptive one way multigrid with BGFC method one way multigrid method scales a bit worse than linear scaling (1.1) asymptotically. Sub-linear scaling (0.84) can still be recovered if the accuracy control on each coarse level for BGFC method is tuned. The crossover between the adaptive multigrid methods and the brute-force optimization happens when the system contains around 5000,000 atoms, which is slightly larger than that in the single vacancy case. Moreover, compared to the brute-force optimization, the savings of the adaptive one way multigrid with BGFC method are around 80% for systems with ten millions atoms.
5.4 Edge dislocation

Dirichlet boundary conditions are used in all three directions with ghost atoms. We construct an (001)[100] edge dislocation for tungsten system (see Figure 11(a)). Figure 11(b) plots the displacement field in $L_2$ norm over a centered slice in the $z$ direction of $30^3$ system. Around dislocation core, one can see the nonlocal effect on the displacement field in the presence of dislocations. This nonlocal effect typically decays like $1/r^2$ where $r$ is the distance to the core, and thus is more difficult to solve than the point defects case.

![Figure 11: Configuration and displacement for (001)[100] edge dislocation in tungsten.](image)

6 Conclusions

This paper presents several efficient multigrid strategies for large-scale molecular mechanics optimization. The coarse-grid problems can be constructed from the atomistic model systematically on-the-fly using a quasi-atomistic approximation or the blended ghost force correction.
Figure 13: Adaptive mesh refinement (DOF from left to right: 125, 272, 618, 1536) of quasi-atomistic approximation for single vacancy.

Figure 14: Adaptive mesh refinement (DOF from left to right: 942, 1414, 2420, 4296) of BGFC approximation for single vacancy.

The adaptive oneway multigrid method is used with inexact approximations at coarse levels. Optimal scaling complexity is observed for crystals with defects, like vacancies, micro-cracks and dislocations. Moreover, the strategies introduced in this paper outperform the brute-force optimization for all examples if more than 200,000 atoms are involved. For systems with more than ten million atoms, this strategy saves more than 80% CPU time on average.

In our future work, we will consider the complex multigrid strategies like full multigrid (FMG) and full approximation scheme (FAS) in molecular mechanics optimization, and test with more realistic defected atomistic systems. We will also consider the fast solver for molecular dynamics in finite temperature, and the efficient multigrid strategy in quantum mechanics optimization.

References

[1] G. Alzetta, D. Arndt, W. Bangerth, V. Boddu, B. Brands, D. Davydov, R. Gassmoeller, T. Heister, L. Heltai, K. Kormann, M. Kronbichler, M. Maier, J.-P. Pelteret, B. Turcksin, and D. Wells. The deal.II library, version 9.0. *Journal of Numerical Mathematics*, 26(4):173–183, 2018.

[2] Peter M Anderson, John P Hirth, and Jens Lothe. *Theory of dislocations*. Cambridge University Press, 2017.

[3] W. Bangerth, R. Hartmann, and G. Kanschat. deal.II – a general purpose object oriented finite element library. *ACM Trans. Math. Softw.*, 33(4):24/1–24/27, 2007.

[4] Malthe K Bisbo and Bjørk Hammer. Efficient global structure optimization with a machine-learned surrogate model. *Physical review letters*, 124(8):086102, 2020.

[5] Erik Bitzek, Pekka Koskinen, Franz Gähler, Michael Moseler, and Peter Gumbsch. Structural relaxation made simple. *Physical review letters*, 97(17):170201, 2006.
[6] Achi Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of computation*, 31(138):333–390, 1977.

[7] Julian Braun, Maciej Buze, and Christoph Ortner. The effect of crystal symmetries on the locality of screw dislocation cores. *SIAM Journal on Mathematical Analysis*, 51(2):1108–1136, 2019.

[8] Jingrun Chen and Carlos J Garcia-Cervera. An efficient multigrid strategy for large-scale molecular mechanics optimization. *Journal of Computational Physics*, 342:29–42, 2017.

[9] Jingrun Chen and Pingbing Ming. An efficient multigrid method for molecular mechanics modeling in atomic solids. *Communications in Computational Physics*, 10(1):70–89, 2011.

[10] Jingrun Chen, Pingbing Ming, and Jerry Zhijian Yang. A constrained cauchy-born elasticity accelerated multigrid method for nanoindentation. *Communications in Computational Physics*, 15(2):470–486, 2014.

[11] 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(12):6443 – 6453, 1984.

[12] V. Ehrlacher, C. Ortner, and A. V. Shapeev. Analysis of boundary conditions for crystal defect atomistic simulations. *Arch. Rat. Mech. Anal.*, 222:1217–1268, 2016.

[13] V Ehrlacher, C Ortner, and AV Shapeev. Analysis of boundary conditions for crystal defect atomistic simulations. *Archive for Rational Mechanics and Analysis*, 222(3):1217–1268, 2016.

[14] L. Fang and L. Zhang. Blended ghost force correction method for 3d crystalline defects. arXiv:2006.06501.

[15] M. W. Finnis and J. E. Sinclair. A simple empirical n-body potential for transition-metals. *Philos.Mag.A*, 50(1):45 – 55, 1984.

[16] Serge Gratton, Annick Sartenaer, and Philippe L Toint. Recursive trust-region methods for multiscale nonlinear optimization. *SIAM Journal on Optimization*, 19(1):414–444, 2008.

[17] Wolfgang Hackbusch. *Multi-grid methods and applications*, volume 4. Springer Science & Business Media, 2013.

[18] William W Hager and Hongchao Zhang. Algorithm 851: Cg descent, a conjugate gradient method with guaranteed descent. *ACM Transactions on Mathematical Software (TOMS)*, 32(1):113–137, 2006.

[19] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dulak, Jesper Friis, Michael N Groves, Björk Hammer, Cory Hargus, et al. The atomic simulation environment—a python library for working with atoms. *Journal of Physics: Condensed Matter*, 29(27):273002, 2017.

[20] X. H. Li, M. Luskin, C. Ortner, and A. V. Shapeev. Theory-based benchmarking of the blended force-based quasicontinuum method. *Comput. Methods Appl. Mech. Engrg*, 268:763–781, 2014.

[21] X. H. Li, C. Ortner, A. Shapeev, and B. Van Koten. Analysis of blended atomistic/continuum hybrid methods. *Numer. Math.*, 134, 2016.

[22] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1-3):503–528, 1989.

[23] M. Luskin, C. Ortner, and B. Van Koten. Formulation and optimization of the energy-based blended quasicontinuum method. *Comput. Methods Appl. Mech. Engrg.*, 253, 2013.

[24] Mitchell Luskin and Christoph Ortner. Atomistic-to-continuum coupling. *Acta Numerica*, 22:397, 2013.
[25] Mitchell Luskin, Christoph Ortner, and Brian Van Koten. Formulation and optimization of the energy-based blended quasicontinuum method. *Computer Methods in Applied Mechanics and Engineering*, 253:160–168, 2013.

[26] Y. Wang *et al*. M. Liao. JuAC.jl.git. https://github.com/mingjie-liao/JuAC.

[27] Ronald E Miller and Ellad B Tadmor. A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods. *Modelling and simulation in materials science and engineering*, 17(5):053001, 2009.

[28] Patrick Kofod Mogensen and Asbjørn Nilsen Riseth. Optim: A mathematical optimization package for Julia. *Journal of Open Source Software*, 3(24):615, 2018.

[29] Letif Mones, Christoph Ortner, and Gábor Csányi. Preconditioners for the geometry optimisation and saddle point search of molecular systems. *Scientific reports*, 8(1):1–11, 2018.

[30] Stephen G Nash. A multigrid approach to discretized optimization problems. *Optimization Methods and Software*, 14(1-2):99–116, 2000.

[31] Derek Olson and Christoph Ortner. Regularity and locality of point defects in multilattices. *Applied Mathematics Research eXpress*, 2017(2):297–337, 2017.

[32] C. Ortner. A priori and a posteriori analysis of the quasi-nonlocal quasicontinuum method in 1D. *Math. Comp.*, 80(275):1265–1285, 2011.

[33] C. Ortner and F. Theil. Justification of the Cauchy-Born approximation of elastodynamics. *Arch. Ration. Mech. Anal.*, 207(3):1025–1073, 2013.

[34] 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.

[35] C. Ortner and L. Zhang. Energy-based atomistic-to-continuum coupling without ghost forces. *Comput. Methods Appl. Mech. Engrg.*, 279:29–45, 2014.

[36] C. Ortner and L. Zhang. Atomistic/continuum blending with ghost force correction. *SIAM Journal on Scientific Computing*, 38(1):A346–A375, 2016.

[37] V. B. Shenoy, R. E. Miller, E. B. Tadmor, D. Rodney, R. Phillips, and M. Ortiz. An adaptive finite element approach to atomic-scale mechanics?the quasicontinuum method. *J.Mech.Phys.Solids*, 47(3):611 – 642, 1999.

[38] Hang Si. Tetgen, a delaunay-based quality tetrahedral mesh generator. *ACM Transactions on Mathematical Software (TOMS)*, 41(2):1–36, 2015.

[39] EB Tadmor, R Phillips, and M Ortiz. Hierarchical modeling in the mechanics of materials. *International Journal of solids and Structures*, 37(1-2):379–389, 2000.

[40] Ellad B Tadmor and Ronald E Miller. *Modeling materials: continuum, atomistic and multiscale techniques*. Cambridge University Press, 2011.

[41] Erik Van Der Giessen, Peter A Schultz, Nicolas Bertin, Vasily V Bulatov, Wei Cai, Gábor Csányi, Stephen M Foiles, Marc GD Geers, Carlos González, Markus Hütter, et al. Roadmap on multiscale materials modeling. *Modelling and Simulation in Materials Science and Engineering*, 28(4):043001, 2020.

[42] David Wales et al. *Energy landscapes: Applications to clusters, biomolecules and glasses*. Cambridge University Press, 2003.

[43] H. Wang, M. Liao, P. Lin, and L. Zhang. A posteriori error estimation and adaptive algorithm for the atomistic/continuum coupling in two dimensions. *SIAM J. Sci. Comput.*, 40(4):A2087–A2119, 2018.
[44] E Weinan and Pingbing Ming. Cauchy-born rule and the stability of crystalline solids: Static problems. *Arch. Rat. Mech. Anal.*, 183(2):241–297, 2007.

[45] Zaiwen Wen and Donald Goldfarb. A line search multigrid method for large-scale nonlinear optimization. *SIAM Journal on Optimization*, 20(3):1478–1503, 2010.