Quantum mechanical simulations of nanoindentation of Al thin film

Qing Peng *, Xu Zhang, Gang Lu

Department of Physics and Astronomy, California State University Northridge, Northridge, CA, USA

ARTICLE INFO

Article history:
Received 20 September 2009
Received in revised form 30 October 2009
Accepted 1 November 2009
Available online 3 December 2009

PACS:
71.15.Mb
62.20.Mk
71.15.Dx

Keywords:
Quantum mechanics/molecular mechanics
Nanoindentation
First-principles electron structure theory
Embedding theory
Multiscale

ABSTRACT

QCDFT is a multiscale modeling approach that can simulate multi-million atoms effectively via density functional theory (DFT). The method is based on the framework of quasicontinuum (QC) approach with DFT as its sole energetics formulation. The local QC energy is calculated by DFT with Cauchy–Born hypothesis and the nonlocal QC energy is determined by a self-consistent embedding approach, which couples nonlocal QC atoms to the vertices of the finite elements at the local QC region. The QCDFT method is applied to a nanoindentation study of an Al thin film in the presence and absence of Mg impurities. The results show that the randomly distributed Mg impurities can significantly increase the ideal and yield strength of the Al thin film.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The ability to perform quantum mechanical simulations of materials properties over length scales that are relevant to experiments represents a grand challenge in computational materials science. If one could treat multi-millions or billions of electrons effectively at micron scales, such first-principle quantum simulations could revolutionize materials research and pave the way to the computational design of advanced materials.

In this paper, we propose a multiscale approach that is based entirely on density functional theory (DFT) and allows quantum simulations at micron scale and beyond. The method, termed QCDFT [12], combines the coarse graining idea of the quasicontinuum (QC) approach and the coupling strategy of the quantum mechanics/molecular mechanics (QM/MM) method, and represents a major advance in the quantum simulation of materials properties. It should be stated at the outset that QCDFT is not a brute-force electronic structure method, but rather a multiscale approach that can treat large systems – effectively up to billions of electrons. Therefore, some of the electronic degrees of freedom are reduced to continuum degrees of freedom in QCDFT. On the other hand, although QCDFT utilizes the idea of QM/MM coupling, it does not involve any classical/empirical potentials (or force fields) in the formulation – the energy calculation of QCDFT is entirely based on DFT. This is an important feature and advantage of QCDFT, which qualifies it as a bona fide quantum mechanical simulation method.

Since QCDFT is formulated within the framework of the QC method, we shall give a brief introduction to QC in Section 2.1 to set up the stage of QCDFT. In Section 2.2, we briefly explain the local QC calculations. In Section 2.3, we introduce a DFT-based QM/MM approach that can treat the nonlocal QC region accurately and efficiently. In Section 3, we apply QCDFT to the study of nanoindentation of an Al thin film in the presence and absence of Mg impurities. We present the nanoindentation results in Section 4 and finally our conclusions in Section 5.

2. QCDFT methodology

2.1. Quasicontinuum method

The goal of the QC method is to model an atomistic system without explicitly treating every atom in the problem [18,13]. This is achieved by replacing the full set of N atoms with a small subset of N_t, “representative atoms” or repatoms (N_t \ll N) that approximate the total energy through appropriate weighting. The energies of individual repatoms are computed in two different ways depending on the deformation in their immediate vicinity. Atoms
experiencing large variations in the deformation gradient on an atomic scale are computed in the same way as in a standard atomistic method. In QC these atoms are called nonlocal atoms. In contrast, the energy of atoms experiencing a smooth deformation field on the atomic scale is computed based on the deformation gradient \( G \) in their vicinity as befitting a continuum model. These atoms are called local atoms because their energy is based only on the deformation gradient at the point where it is computed. In a classical system where the energy is calculated based on classical/empirical interatomic potentials, the total energy \( E_{\text{tot}} \) can be written

\[
E_{\text{tot}} = \sum_{i=1}^{n_{\text{l}}} E_i(R) + \sum_{j=1}^{n_{\text{w}}} h_j E_{\text{loc}}^j(G).
\]  

The total energy has been divided into two parts: an atomistic region of \( n_{\text{al}} \) nonlocal atoms and a continuum region of \( n_{\text{loc}} \) local atoms (\( n_{\text{al}} + n_{\text{loc}} = N \)). The calculation in the nonlocal region is identical to that in atomistic methods with the energy of the atom depending on the coordinates \( R \) of the surrounding atoms. Rather than depending on the positions of neighboring atoms, the energy of a local repatom depends on the deformation gradients \( G \) characterizing the finite strain around its position.

2.2. Local QC calculation with DFT

In the local QC region, a finite element mesh is constructed with each repatom on the vertices of surrounding finite elements. The energy and force of each local repatom can be obtained from the strain energy density and the stress tensor of the finite elements that share the same repatom. More specifically, according to the Cauchy–Born rule, the deformation gradient \( G \) is the same for a given finite element, therefore the local energy density \( \varepsilon \) and the stress tensor for each finite element can be calculated as a perfect infinite crystal undergoing a uniform deformation specified by \( G \). In other words, one could perform a DFT-based energy/stress calculation for an infinite crystal by using periodic boundary conditions with the primitive lattice vectors of the deformed crystal, \( h_i \), given by

\[
h_i = G h_i, \quad i = 1, 2, 3.
\]

Here \( H \) are the primitive lattice vectors of the perfect undeformed crystal and \( V_0 \) is the volume of the primitive unit cell. The deformed crystal can be derived from the perfect crystal via the deformation gradient \( G \) as shown in Fig. 1.

For the deformation gradient \( G_k \) associated with the \( k \)th element, a periodic DFT calculation can be performed to determine the strain energy per unit cell \( E_{\text{DFT}}(G_k) \). The Cauchy stress tensor can be defined as follows:

\[
\sigma_{ab} = \frac{1}{\Omega} \sum_{v} \left. \frac{\partial E_{\text{DFT}}(G_k)}{\partial h_{av}} \right|_{h_{bv}} h_{bv}.
\]

with \( \Omega \) being the volume of the deformed unit cell and \( h_{ij} \) denoting the component of the deformed lattice vector \( h_i \) in Cartesian coordinate \( i \).

Once the strain energy \( E_{\text{DFT}}(G_k) \) is determined, the energy contribution of the \( j \)th local repatom is given as

\[
E_{\text{loc}}^j(G) = \sum_{k=1}^{M_j} W_k E_{\text{DFT}}(G_k),
\]

where \( M_j \) is the total number of finite elements shared by the \( j \)th repatom, and \( W_k \) is the weight associated with the \( k \)th finite element for the \( j \)th local repatom. The force on the \( j \)th local repatom is defined as the gradient of the total energy with respect to its coordinate \( R^j \). In practice, the nodal force on each finite element is calculated from the stress tensor of the finite element by using the Principle of Virtual Work [25]. The force on the repatom is then obtained by summing the nodal force contributions from each surrounding finite elements.

2.3. Nonlocal QC calculation with DFT

The nonlocal QC is modeled at the atomistic level with a QM/MM approach. In a typical QM/MM calculation, the system is partitioned into two domains: a QM region and an MM region. In QCDFT, the QM atoms refer to the nonlocal repatoms and the MM atoms refer to the buffer atoms which are the combination of both dummy atoms and local repatoms in QC terminology. The so-called dummy atoms are in the local region and are not independent degrees of freedom, but rather slaves to the local repatoms. In other words, the position of a dummy atom is determined by the finite element interpolation from the relevant local repatom positions [18,13]. The dummy atoms provide the appropriate boundary conditions for nonlocal DFT calculation while the energy of the dummy atoms is still treated with the Cauchy–Born rule, consistent with their status. The self-consistent embedding theory [1,22,23] is employed for the QM/MM calculations. More specifically, both the energy of the nonlocal atoms and the interaction energy between the nonlocal atoms and the buffer atoms are calculated by DFT. To simply the notation, we denote the nonlocal region as region I, and the buffer region as region II, as shown in Fig. 2. Typically, the buffer region consists of several

Fig. 1. The deformed crystal derived from the perfect crystal via the deformation gradient \( G \).

Fig. 2. The schematic diagram of domain partition in QCDFT with a dislocation in Al lattice as an example. The black and white spheres represent the nonlocal and buffer atoms, respectively. The dotted box represents \( \Omega^I \) and the solid box represents the periodic box \( \Omega^B \). The volume \( \Omega^I \) and \( \Omega^B \) is 2.8 Å and 8 Å beyond the region I in \( x \) and \( y \) directions, respectively.
atomic layers surrounding the nonlocal region. We associate each buffer atom in region II with a valence electron density $\rho^i$ and a pseudopotential: both of them are constructed a priori and remain fixed during a QM/MM simulation [22]. The nonlocal energy $E^{\text{nl}}$ as defined in Eq. (1) can be expressed as

$$E^{\text{nl}} = \min_{\rho^i} \{ E_{\text{QDFT}}[\rho^i, \mathbf{R}^{\text{ii}}] + E_{\text{OF}}[\rho^i, \rho^i, \mathbf{R}^{\text{ii}}] \}. \tag{5}$$

Here $\mathbf{R}^{\text{I}}$ and $\mathbf{R}^{\text{ii}}$ denote atomic coordinates in regions I and II, respectively. The charge density of region I, $\rho^i$, is the degree of freedom and is determined self-consistently by minimizing the nonlocal energy functional. The charge density of region II, $\rho^\text{II}$, is defined as the superposition of atomic-centered charge densities $q_j$ [1,22]. This significant computational saving is due to the cancellation in evaluating the first and second term of Eq. (9), and it is rendered as orbital-free DFT (OFDFT) [20,19].

A basic ansatz of the nonlocal energy functional (Eq. (5)) is that $\rho^i$ must be confined within a finite volume ($\mathcal{Q}^i$) that is necessarily larger than region I but much smaller than the entire QM/MM region. In addition, since some terms in the formulation of Eq. (5) could be more efficiently computed in reciprocal space [22], we also introduce a volume $\mathcal{Q}^i$ over which the periodic conditions are applied. The periodic box $\mathcal{Q}^i$ should be large enough to avoid the coupling errors induced by the implementation of periodic boundary condition [22].

The interaction energy, $E^{\text{int}}_{\text{OF}}$, formulated by OFDFT is defined as following:

$$E^{\text{int}}_{\text{OF}}[\rho^i, \rho^\text{II}, \mathbf{R}^{\text{II}}] = E_{\text{OF}}[\rho^i, \mathbf{R}^{\text{II}}] - E_{\text{OF}}[\rho^i, \mathbf{R}^{\text{II}}] - E_{\text{QDFT}}[\rho^i, \mathbf{R}^{\text{II}}] + E_{\text{QDFT}}[\rho^i, \mathbf{R}^{\text{II}}]. \tag{6}$$

where $\mathbf{R}^{\text{II}} \equiv \mathbf{R}^{\text{I}} \cup \mathbf{R}^{\text{II}}$ and $\rho^{\text{tot}} = \rho^i + \rho^\text{II}$. In addition to its computational efficiency, OFDFT allows Eq. (9) to be evaluated over $\mathcal{Q}^i$ rather than over the entire QM/MM system as Eq. (9) appears to suggest [1,22]. This significant computational saving is due to the cancellation in evaluating the first and second term of Eq. (9), and it is rendered as orbital-free nature of OFDFT and the localization of $\rho^i$.

A single-particle embedding potential $\mu_{\text{emb}}(\mathbf{r})$ can be defined as a functional derivative of the interaction energy with respect to $\rho^i$

$$\mu_{\text{emb}}(\mathbf{r}) = \frac{\delta E^{\text{int}}_{\text{OF}}[\rho^i, \rho^\text{II}, \mathbf{R}^{\text{II}}]}{\delta \rho^i}, \tag{7}$$

which represents the effective potential that region I electrons feel due to the presence of region II [1,22]: it is through $\mu_{\text{emb}}(\mathbf{r})$ that the QM/MM coupling is achieved quantum mechanically at the level of OFDFT. The embedding potential provides rigorous boundary conditions for $\rho^i$ and is updated self-consistently during the minimization of the nonlocal energy functional.

Since $\rho^i$ is a key quantity for the accurate calculation of the interaction energy and the embedding potential, it is crucial to construct an appropriate representation of $\rho^i$. In fact, the construction of an appropriate charge density distribution in region II represents a common challenge to many QM/MM methods [11]. In this paper, we represent $\rho^i$ as a superposition of spherical atomic-like charge densities centered on each ions in region II, which is a good approximation for metallic systems. Ideally, the constructed $\rho^i(\mathbf{r})$ should reproduce the bulk (or solid) charge density obtained by a DFT calculation of the perfect lattice. That is to say, one needs to determine $\rho^\text{II}(\mathbf{r})$ by minimizing the function

$$\int_{V_u} [\rho^i(\mathbf{r}) - \rho^\text{solid}(\mathbf{r})]^2 \, d\mathbf{r}. \tag{8}$$

with $\rho^\text{solid}(\mathbf{r}) = \sum_{\text{ions}} \rho^\text{II}(\mathbf{r} - \mathbf{R}_\text{ion})$. Here $V_u$ represents the volume of the unit cell, and $\rho^\text{solid}$ is the solid charge density obtained by a periodic DFT calculation for the perfect reference system. The summation of $\mu$ includes all the ions which have contribution to the charge density in the unit cell.

In this paper, we employ the parameterized multiple Slater-type orbitals (MSTO) [2] for the expansion of $\rho^\text{II}(\mathbf{r})$. With MSTO, the atomic wave function $\Phi$ of many-electrons is the superposition of all relevant atomic orbitals: $\Phi(\mathbf{r}, \theta, \varphi) = \sum c_i \phi_i(\mathbf{r}, \theta, \varphi)$, where $c_i$ is the weight of orbital $i$ in the expansion and the $i$th atomic orbital can be written as

$$\phi_i(\mathbf{r}, \theta, \varphi) = Ar^{-\eta} e^{-\zeta r} Y_{lm}(\theta, \varphi), \tag{9}$$

where $n$, $l$, $m$ are the principal, angular momentum and magnetic quantum number of the orbital. $Y_{lm}(\theta, \varphi)$ is spherical harmonic function and $\zeta$ is related to the effective charge of the ion. $A$ is a normalization constant and is expressed as $A = (2\zeta)^{3/2}/\sqrt{(2n)!}$. With this expansion, the atomic-centered charge density can be calculated as

$$\rho^\text{II}(\mathbf{r}) = \int_{\bar{\mathcal{X}}} \int_{\bar{\mathcal{Y}}} \Phi(\mathbf{r}, \theta, \varphi) \Phi(\mathbf{r}, \theta, \varphi) \, d\mathbf{r} \, d\mathbf{r}. \tag{10}$$

The parameters $c_i$ and $\zeta$ are determined by minimizing Eq. (8) with the constraint of preserving the correct number of valence electrons. In Fig. 3, we present the solid charge density $\rho^\text{solid}$ determined from $\rho^\text{II}$ and from the periodic DFT calculation for a perfect Al lattice. It can be seen that the constructed $\rho^\text{solid}$ reproduces very well the solid charge density calculated by DFT calculations for the same perfect lattice.

3. Computational details

3.1. Model setup

The present QCDFT approach is applied to nanoindentation of an Al thin film resting on a rigid substrate with a rigid knife-like indenter. The QCDFT method is appropriate for the problem because it allows the modeling of system dimensions on the order of microns and thus minimizes the possibility of contaminating the results by the boundary conditions arising from small model sizes typically used in MD simulations. The reason we chose this particular system is because there exists a good kinetic energy functional and an excellent EAM potential [4] for Al. In this paper, we have rescaled the "force-matching" EAM potential of Al [4] so that it matches precisely the DFT value of the lattice constant and bulk modulus of Al [12].

The crystallographic orientation of the system is displayed in Fig. 4. The size of the entire system is $2 \mu m \times 1 \mu m \times 4.9385 \AA$ along the [111] (x direction), the [110] (y direction), and the
4. Results and analysis

The load–displacement curve is the typical observable for nanoindentation, and is widely used in both experiment and theory, often serving as a link between the two. In particular, it is conventional to identify the onset of incipient plasticity with the full relaxed configuration of a previous loading step, until the onset of the plasticity occurs at \( d = 8.1 \, \text{Å} \).

The parameters of the OFDFT density-dependent kernel are chosen from reference [21], and Al ions are represented by the Good–Win–Needs–Heine local pseudopotential [5]. The high kinetic energy cutoff for the plane wave basis of 1600 eV is used to ensure the convergence of the charge density. For the nonlocal calculation, the grid density for the volume \( \Omega^L \) is 5 gridpoints per Å. The \( \Omega^L \) box goes beyond the nonlocal region by 8 Å in \( \pm x \) and \( \pm y \) directions so that \( \rho^L \) decays to zero at the boundary of \( \Omega^L \), as shown in Fig. 2. The relaxation of all repatoms is performed by a conjugate gradient method until the maximum force on any repatom is less than 0.03 eV/Å.

![Figure 4](image4.png)

**Fig. 4.** Schematic representation of the nanoindentation of Al thin film showing the relevant dimensions and orientations.

![Figure 5](image5.png)

**Fig. 5.** Schematic diagram of the randomly distributed Mg impurities in the Al thin film. The red spheres and blue pentagons represent nonlocal Al and Mg atoms, respectively. The green triangle represents Al buffer atoms. The dimensions are given in Å. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

![Figure 6](image6.png)

**Fig. 6.** Load–displacement plot for the nanoindentation of the Al thin film with a rigid rectangular indenter: pure Al (red squares) and randomly distributed Mg impurity system (green circles). The corresponding lines are the best fit to the data points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
the first drop in the load–displacement curve during indentation
[3,15,7,17,14,24,10,8,6,12]. In the present work, the load is given
in N/m, normalized by the length of the indenter in the out-of-
plane direction.

For pure Al, the load–displacement \( P - d \) curve shows a linear
relation followed by a drop at \( d = 8.2 \) Å, shown by the dashed line
in Fig. 6. The drop corresponds to the homogeneous nucleation of
dislocations beneath the indenter – the onset of plasticity. A pair
of straight edge dislocations are nucleated at \( x = \pm 13 \) Å, and
\( y = -50 \) Å. In Fig. 7, we present the out-of-plane (or screw) dis-
placement \( u_z \) of the nonlocal repatoms. The non-zero screw dis-
placement of the edge dislocations suggests that each dislocation
is dissociated into two \( \frac{1}{6} \) Shockley partials bound by a
stacking fault with a width of about 19 Å. The activated slip planes
are those \{111\} planes that are adjacent to the edges of the inden-
ter. The slope for the linear part of the curve is 27.1 GPa, which is
less than the shear modulus \( l = 33.0 \) GPa and \( C_4 = 29.8 \) GPa.
The critical load, \( P_{cr} \), for the homogeneous dislocation nucleation
is 18.4 N/m, corresponding to a hardness of 7.3 GPa (the critical load
normalized by the area of the indenter), which is 0.22 \( \mu \). The drop
in applied load due to the nucleation of dislocations is \( \Delta P = 6.8 \) N/m, agreeing with the load drop estimated by the elastic
model [17] which is \( \Delta P = 7.7 \) N/m.

For randomly distributed impurities in the Al thin film, the
load–displacement curve shows a linear relation up to a depth of
8.0 Å, followed by a drop at \( d = 8.1 \) Å, as shown by the solid line
in Fig. 6. The slope of initial linear part of the load–displacement
curve is 26.7 GPa, rather close to the corresponding pure Al value.
The maximum load in linear region is \( P_{cr} = 19.2 \) N/m, correspond-
ing to a hardness of 7.6 GPa, which is 0.3 GPa greater than the pure
Al system. A pair of Shockley partial dislocations is nucleated at
\( x = -13 \) Å, \( y = -25 \) Å and \( x = 13 \) Å, \( y = -22 \) Å respectively as shown
in the right panel of Fig. 7. The drop in applied load due to the
dislocation nucleation is 5.9 N/m. The estimated load drop by the
elastic model is \( \Delta P = 7.6 \) N/m. The smaller drop of the load for
the random case than the elastic model is probably due to the pres-
ence of the Mg impurities, which is not accounted for in the elastic
model [17]. The fact that the critical load and the hardness of
the Al–Mg alloy are greater than that of the pure Al system demon-
strates that the Mg impurities are responsible for the solid solution
strengthening of the Al thin film. The presence of Mg impurities
also hinders the formation of full edge dislocations and as a result,
only partial dislocations are nucleated and they are pinned near
the surface as shown in Fig. 7.

Finally we point out the possibility that the emitted dislocations
may be somewhat constrained by the local/nonlocal interface from
going further into the bulk. Because the critical stress to move an
dislocation in Al is vanishingly small (\( 10^{-3} \) \( \mu \)) comparing to that
to nucleate a dislocation (\( 10^{-1} \) \( \mu \)), a small numerical error in
stress could easily lead to a large difference in the equilibrium dis-
location position. The four-order-of-magnitude disparity poses a
significant challenge to all atomistic simulations in predicting dis-
location nucleation site, QCDFT method included. One can only
hope to obtain a reliable critical load for the incipient plasticity,
rather than for the equilibrium position of dislocations. The same
problem has been observed and discussed by others [16]. However,
deeply the problem, the dramatic difference observed in the two
panels of Fig. 7 unambiguously demonstrates the strengthening ef-
fect of Mg impurities. Therefore the conclusion is still valid.

5. Conclusion

In summary, we propose a concurrent multiscale method that
makes it possible to simulate multi-million atoms based on the
density functional theory. The method – QCDFT – is formulated
within the framework of the QC method, with DFT as its sole en-
ergy input. The full-blown DFT and DFT-based elasticity theory
would be the two limiting cases corresponding to a fully nonlocal
or a fully local version of QCDFT. The QCDFT method is applied to
nanoindentation of an Al thin film in the presence and absence of
randomly distributed Mg impurities. The Mg impurities are found
to strengthen the hardness of Al and hinder the dislocation nuc-
leation. The results suggest that QCDFT is a promising method for
quantum simulation of materials properties at length scales rele-
vant to experiments.

Acknowledgements

The work at California State University Northridge was
supported by NSF PREM Grant DMR-0611562 and DoE SciDAC
Grant DE-FC02-06ER25791.
References