

PROGRESS SUMMARY
NSF DMR-9980015
“MULTI-SCALE SIMULATION INCLUDING CHEMICAL REACTIVITY
IN MATERIALS BEHAVIOR THROUGH
INTEGRATED COMPUTATIONAL HIERARCHIES”

MULTI-SCALE, MULTI-PASS SIMULATIONS WITH CHEMICAL REALISM:
FROM MOLECULES TO MATERIALS
(Revised April 19, 2002)

Computer modeling of materials potentially can have an enormous impact in identifying new materials, and understanding or controlling their behavior. Issues include how the material fractures or decomposes, and its optical, electrical, magnetic, and thermal properties. Controlling and modifying materials behavior implies several challenges to materials simulations: predicting bond breaking and formation, handling chemical reactions, and characterizing optical behavior (spectroscopy) are among the most critical.

A way to state the goal of this project is this: to achieve technologically useful multi-scale simulations on complex materials in which chemo-mechanical processes are important. Three key themes must be developed for such simulations. They must be (a) *predictive*, (b) *chemically realistic*, and (c) *state-specific*. Of necessity they must be multi-scale: treating all the atoms with quantum mechanics is impossible.

THEMES

PREDICTIVE

By now it is well-established that ab initio quantum chemistry has achieved the quality of being ‘predictive’ to within established small error bars for most properties of isolated, relatively small molecules. This property makes it far easier to obtain requisite information about molecules from applications of theory than from complicated and expensive experimental observation. In fact, applied quantum chemistry, as implemented in many widely used computer programs, has now attained the status of a tool that is complementary to X-ray structure determination and NMR and IR spectra in the routine determination of the geometry and spectra of molecules. To a considerable extent the same is true for first-principles treatment of crystals and their surfaces.

The vastly greater opportunity – and challenge - is for simulations of complex materials to be equally predictive. Unlike molecules, which can usually be characterized in detail by spectra, real materials are far more complex and cannot usually be studied experimentally under similarly simplified conditions. Often their most interesting features must be studied at elevated temperatures and under non-equilibrium conditions. Frequently, the intended application of the material involves an extreme situation that is de facto inaccessible in a laboratory. Hence, if we intend to use more economical computer models to learn how to suitably modify a material to achieve an objective, our materials simulations

must be ‘predictive,’ in order to trust both the qualitative and quantitative consequences of the simulations.

CHEMICALLY REALISTIC

Besides the predictive aspect, another theme that permeates our work is ‘chemical realism’. Chemical realism means that, unlike the idealized systems that have been the focus of most of the simulation work in materials sciences, we aim to consider the essential interactions among many different molecular species. Our particular focus is on systems under stress. As an example, a long unsolved problem in materials is why water will cause forms of silica to weaken by several orders of magnitude compared to their dry forms. However, ammonia with silica shows a different behavior. A proper, quantum mechanically based simulation should reflect these differences, qualitatively and quantitatively. This is chemical realism.

STATE-SPECIFIC

The third theme of our work is that by virtue of using a quantum mechanical core in multi-scale simulations, unlike simulations based upon classical potentials, we have quantum state specificity. In a problem like etching silica with CF_3 , a classical potential cannot distinguish between CF_3^+ , CF_3^- , and CF_3^* , yet obviously the chemistry will be very different. Furthermore, we also have developed the capability of using excited electronic states in our simulations, to include species like CF_3^* , eg, or to distinguish between different modes of fractures of the silica target, such as radical dissociation opposed to ionic dissociation.

MULTI-SCALE

Simulations can be no better than the forces provided by the models of interaction among the atoms. Today, classical two or three body potentials are most common, because only such simplified forms permit the very rapid generation of forces required by large scale molecular dynamics. Furthermore, while such potentials have been laboriously developed over many years, adding new species frequently demands another long-term effort to generate potentials for the new interactions. Many simulations also depend upon crystalline symmetry, so are not directly applicable to the often more technologically important amorphous materials. *Moreover*, bond-breaking and chemical reactivity require a quantum mechanical basis for the simulations. This requires a multi-scale philosophy, in which a quantum mechanical core is tied to a classical atomistic region, which in turn is embedded in a continuum (e.g., a reaction field or a finite-element region). Most multi-scale work uses a tight-binding core; a much smaller amount uses a density functional theory (DFT) core that with far fewer atoms.

SPECIFIC RESULTS AND STUDIES

MULTI-PASS

Including real chemistry and real surface physics has led us to a reanalysis of the whole multi-scale simulation enterprise. The outcomes in brief, are these. First, no tight-binding quantum mechanical core is adequate to the task. This is because tight-binding enforces non-reactive, atomic character on the material. It is not obvious that DFT is adequate,

since bond breaking is not a strength of many commonly used DFT approximations. Second, the location of a failure region or a crucial reaction region is not necessarily known a priori. Example: if the location of a crack tip is already known, then the QM core must be roughly centered on the tip. But what if the failure region location is not known a priori? This is by far the more realistic case. For it, *multi-pass* simulation capability is *essential*: first an all-classical simulation to locate the more likely reactive regions, followed by a multi-scale re-simulation to treat such regions realistically. Third, huge numbers of particles in the QM region are not necessarily needed. Most chemical and surface phenomena are *local*, so the main criterion is the *quality* of the QM description, not the gross number of atoms simulated. Another benefit is that the multi-pass strategy provides upper bounds to the proper number of atoms for the QM region. Remarkably, the whole multi-pass aspect has gone almost completely untreated in the multi-scale literature.

TRANSFER HAMILTONIAN FOR QM

Prior to our work, the only quantum mechanically based multi-scale simulations that would permit as many as 500 to 1000 atoms in the QM region were based upon the tight-binding method. (DFT has been used for much smaller QM regions.) TB is an oversimplified, non-self-consistent description so it cannot properly describe bond breaking, where charge transfer is absolutely essential. Instead, we have appealed to the highest level of ab initio quantum chemistry, namely coupled-cluster (CC) theory, to use as a basis for a ‘transfer Hamiltonian’. This is a ‘zero-differential overlap’ type Hamiltonian that replicates accurate, predictive quality CC forces (energy gradients) taken from suitable clusters, yet is an operator of very low rank. The combination makes it possible to do fully self-consistent QM calculations on ~500-1000 atoms. Testing on SiO₂ tensile fracture confirms that the transfer Hamiltonian retains the accuracy of CC theory. Hence, for any phenomenon accessible to MD, we should be able to retain the predictive quality of the CC method in materials simulations. Similarly, we have no problem with changing atoms or adding new molecules to our simulations, as our transfer Hamiltonian is applicable to any system once trained to ensure its proper description.

Even the transfer Hamiltonian is too demanding to be used at every MD step however. Therefore we have introduced a *force-shift* procedure. In it we run the MD for some moderate number of steps N using the classical potential. Then we compute the forces for that configuration from the transfer Hamiltonian, hence the shifts relative to the classical forces. For the ensuing N steps, the classical potential forces are augmented by those shifts, whence a new set of shifts is computed and so on. Initial testing shows that the scheme works reliably for N of order 10 when published classical potentials are used. Larger N values should be readily attainable with more carefully adapted potentials.

CONSISTENT EMBEDDING

Compatible embedding of the transfer Hamiltonian region (and the force shifts) in a CM region requires rethinking the calibration of classical potentials. Customarily the embedding potentials have been fit either to experimental data alone or to a mixture of ab initio calculated results and experimental data. This approach poses two major problems that have gone un-examined in the multi-scale literature: (i) a priori the simulation will not be

predictive and (ii) the classical potential most consistent with the QM techniques may well not be close to empirically calibrated potentials.

These two problems are aspects of the larger issue of information transfer at the QM/MD interface. To explore it, we studied the effects of cluster or crystal choice, basis set size, methodology ("level" of approximation), and fitting schemes on the quality of classical embedding with focus on the popular TTAM and BKS potentials for silica. We studied 14 $H_mSi_nO_q$ clusters plus α -quartz. Methods ranged from restricted Hartree-Fock to highly correlated CCSD(T) for the clusters; α -quartz was done only with DFT. ACES-II (for molecules) and GTOFF (periodic systems) were used, with basis sets ranging from 6-31G to very rich sets (aug-cc-pVDZ). We reparameterized on a grid of geometries distorted from calculated equilibrium. Our test for the recalibrated potentials was an MD calculation of tensile strain to failure of a 36 formula unit silica nanorod. We find: (a) Contrary to TTAM and BKS, there are large methodological, basis set, and fitting dependences; (b) There is no parameterization of the CCSD(T) results for the BKS cluster H_4SiO_4 that will yield a stable equilibrium nanorod; (c) Three versions of BKS-type and TTAM-type reparameterizations from the α -quartz DFT data yield a stable nanorod; two of these fracture at about 25% lower strain than the empirical potential, one fractures at about 25% higher strain. Following the logic of empirical potential calibration therefore is not a promising route for proper QM embedding. We are exploring alternative approaches, including parameterization to calculated gradients and inclusion of pre-calculated reactivity indices.

INFORMATION AT THE INTERFACE

A consistent embedding is no better than the method used to link the regions. For the QM/MD interface, we have explored two variants, "sparkle atoms" and "link atoms"

In silica, for example, "sparkle atoms" are O^* atoms that have one valence to use to terminate bonds in a QM region. They are described by a simple pseudopotential, which so far simply means that the effective nuclear charge Z_{eff} is adjusted. (More sophisticated pseudopotentials could be constructed of course.) Thus they are built to have the correct charge on an O in the middle of a QM cluster; and/or to get the correct QM forces involving that atom. Since their mass is correct, they also participate in the MD region being described by a BKS-style classical potential, hence those forces also have to be right, at least in the sense that they are consistent.

The link atom scheme takes a different approach. It terminates the QM region with special atoms that are constrained to lie on the bond lines that connect near neighbor pairs that have one member each in the QM and MD regions. The distance between the QM atom at the boundary and the link atom is held at a constant value that is pre-calculated from a full quantum treatment of a model ("training") cluster. In the simulation, the Born-Oppenheimer electronic energy of the QM region is calculated using the nuclear positions and electronic coordinates of the link atoms. The electronic energy of interaction of all the electrons (QM and link) with the classical ions $U_{\text{elec}}(\mathbf{R}_C, \rho)$, with ρ the electron density including link atoms, is combined with the interaction potential for the QM ions with the classical ions $U'(\mathbf{R}_C, \mathbf{R}_Q)$ in a single classical function that, for consistency,

is chosen to be the *same* as for the classical-classical ion interaction. As a result, the link atoms terminate the QM region without making a direct, potentially spurious, contribution to the dynamics.

Although the MD/continuum interface is not a focus of our effort, we have developed and used a 2D Finite Element code to provide a simple, swift continuum environment as part of our overall software suite. A second MD/continuum coupling technique that we have under investigation is the Finite Difference Time Domain/ MD bridging between an elastic continuum and an atomistic MD region. To date the method has performed well in the study of an acoustic pulse along a long chain with an embedded MD region. This is a form of temporal embedding.

Another temporal embedding is Kinetic Monte Carlo (KMC). MD typically becomes quite slow in the thermodynamic vicinity of a phase change. We have begun exploration of KMC techniques to see if they can provide time-step acceleration. The problem explored is dislocation etch pit formation, an important example of a long-time-scale chemo-mechanical process. Irreducible removal of atoms is governed by a coordination-dependent probability that is a function of distance from the dislocation. The dislocation itself is modeled as a stress source. Parameterization is via a coulombic bond strength (coordination dependent), and a bond and etch-strength exponential. Reasonable semi-quantitative agreement for Si (100) is found. This opens the possibility of computing the requisite probabilities from the multi-scale, multi-pass methods above, then “jumping” the system forward via KMC, then recomputing the probabilistic inputs.

Finally a small subset of our effort has worked on formal theoretical justification of the multi-scale decomposition. The objective, admittedly difficult, is to find rigorous means for identifying the various regions. A related objective is to understand and rationalize the content of the sparkle and link atom methods. The approach is via partial traces of the total system density operator; some interesting but inconclusive formal results have been obtained so far.

SOFTWARE

At the outset of this project, we projected a need for Universal Modeling Language and techniques implicit in it for software integration. However, the rethinking of the simulation enterprise described above led to the realization that the required physical science could not be done in finite time with finite resources if a “stop the world, I want to get off” approach of drastic software re-engineering were adopted. In fact, it becomes clear that the real need is for a comprehensive set of versatile tools that are linked well-enough to be used flexibly and readily on new problems. Additional reflection on the problem led to the realization the linking software itself *must* be open source, even if some of the tools linked are not. This is because only the open-source approach can be general and inclusive with regard to all the tools.

KDI-Mol is the multi-scale, multi-pass modeling software we have undertaken to fulfill this need. KDI-Mol is based on the existing, open-source project JMol, a Java molecular viewer for 3D chemical structures. See <http://sourceforge.net/projects/jmol/>. Features include reading a variety of file types and output from widely-used quantum chemistry

codes, animation of multi-frame files and molecular normal modes computed from those quantum chemistry codes. In keeping with our open-source approach, the distribution of KDI-Mol is governed by LGPL (Lesser General Public License), just as is JMol.

KDI-Mol extensions and enhancements to JMol include

- ability to run external programs
- ability to create input files to and parse output files from several molecular dynamics packages
- Z-plane clipping
- ability to specify information necessary for an MD run (initial velocities, potential choice, methodological parameters, etc.)
- enhanced number of parsers for molecular data
- optimization of the rendering code for faster display of large systems

Planned work includes:

- attaching a high-quality renderer for publication purposes (Atomic Eye and/or PovRay)
- further optimization of the code
- adding data display tools for quantities like radial distribution function, temperature and energy averages adding a crystal builder
- adding a movie creation capability and linking to VMD

The present version of KDI-Mol is based on JMol ver. 1.0, while the current JMol version is 2.1. We plan to incorporate the most important and relevant changes from Ver. 2.1 into the KDI-Mol source soon.

Also in the spirit of open source, the Florida participants obtained a group license for DL_POLY, (see http://www.cl.ac.uk/TCSC/Software/DL_POLY/main.html), the Daresbury open source classical MD package. The Arizona collaborators already were using this package. For development and prototyping, a simpler, faster MD code was written at Florida as well.

EXPERIMENT

One member of our team is an experimenter. His role has been to inform the rest of the team on various kinds of chemo-mechanical experiments (AFM, with and without etchant, cracking) both his own and others and to interpret those in terms of theory, computation, and modeling.

CONCLUSION

This synopsis demonstrates that the multi-disciplinary nature of our KDI effort has paid solid dividends, some unexpected at the outset. We see the materials simulation problem from a distinct, complementary perspective to that typical in materials science. Our emphasis on predictability, chemical realism, and state specificity is novel in the field. The tools we have developed, all tied together with highly flexible software, set the stage for reliable materials design based on *predictive simulations*.