

Multiscale Modeling of Organic-Inorganic Semiconductor Materials: Opportunities and Challenges

Paulette Clancy*

*School of Chemical and Biomolecular Engineering,
Cornell University, Ithaca NY 14853, U.S.A., Paulette.Clancy@cornell.edu

ABSTRACT

We discuss our recent advances in the multiscale modeling of organic-inorganic semiconductor materials, highlighting the impact that has been achieved using a combination of methods at different length- and time-scales. We also discuss the challenges and limitations that remain for the community to address. Test cases related to all-organic solar cells composed of a COF framework of pores filled with fullerene molecules, heteroepitaxial crystal growth of C60 on pentacene, and an accelerated Bayesian search to find energetically stable polymorphs of a small molecule organic semiconductor. We propose the advantage of embracing a new field that we call “physical analytics” that combines information from simulations, modeling and experiments with machine learning-informed Bayesian search techniques to help us tackle high dimensional, often combinatorially complex problems in materials science.

Keywords: molecular simulation, organic semiconductors, *ab initio* simulation, multiscale modeling.

1 INTRODUCTION

There has been growing appreciation of the accuracy and value of using molecular simulation techniques, including electronic structure calculations using *ab initio* approaches, to represent the equilibrium and non-equilibrium behavior of increasingly exotic materials. The purpose and point of engagement of molecular-level simulations has also undergone a slow transition from its roots to explain the underlying mechanisms for an experimentally observed or predicted phenomenon. While the proven ability of molecular simulation to *explain* observed phenomena will surely always remain, there is increasing recognition of the growing ability of molecular-level simulation to *predict* ahead of experiments, especially in situations where the number of variables in the space of the investigation becomes overwhelmingly large for experiments to conduct (in time and resources) in a sequential manner. Similarly, the computational expense of the molecular simulations is also sufficient to preclude large combinatorial investigations of materials in which the ligand, the solvent, the processing conditions, *etc.*, are all

varying. We will suggest that a new approach is needed, which we call “physical analytics” that brings together researchers who understand the physical science of the system with those who have the mathematical tools to accelerate the search for (non-sequentially investigated) combinations of variables using methods such as Bayesian optimization and machine learning. Such an approach is in its infancy and is hampered by the lack of forums to bring these investigators together and cross-train their students to have skills in both disciplines. This presentation will explore the successes and challenges of the field of multiscale modeling of materials, here focusing on semiconductor materials, especially in systems of increasing importance to the semiconductor industries in which both organic and inorganic moieties are present.

2 MULTISCALE MODELING HISTORY AND APPROACHES

2.1 Brief History of Molecular Simulation

The field of molecular simulation, largely defined to include Monte Carlo and Molecular Dynamics methods, is about 60 years old: The first Monte Carlo simulation is generally ascribed to a 1953 paper by Metropolis *et al.* [1], while the Molecular Dynamics method, a numerical solution to Newtonian (or other) equations of motion, dates back to the pioneering work of Alder and Wainwright in 1957 [2]. Quantum chemistry has a similarly long history, dating back to a 1935 book by Pauling and Wilson, who were early proponents of the practical application of quantum mechanics to problems in chemistry [3]. Since those beginnings, the field of molecular simulation has burgeoned in terms of the number of engaged researchers, the scope of the algorithms that are available, and the breadth of intermolecular force models. These models are the key user-supplied input and provide a vital link between the inherent physics and the manifestation of statistical mechanics.

2.2 Multiscale Modeling

Multiscale modeling is a more recent phenomenon and arose from a desire to solve problems whose component aspects require understanding from the smallest to the largest length and time scales. This led to two Nobel Prizes in Chemistry, the latest in 2013 to Martin Karplus, Michael

Levitt and Arieh Warshel for their role in "the development of multiscale models for complex chemical systems." [4]

A well known example is the modeling of crack formation which requires modeling from the atomic to the continuum scale in length and from picoseconds to hours. In the field of semiconductor simulation and modeling, the length and time scales range from electronic scale representations to at least the micron scale. Typical categories for the scales that comprise multiscale modeling are: (1) the electron scale (requiring quantum mechanical models); (2) the molecular scale (using Molecular Dynamics and Monte Carlo simulation approaches), which may be sub-divided into an explicit all-atom representation or some degree of coarse graining (*e.g.*, using a "bead" or "united atom" to represent a functional unit, *i.e.*, one bead representing a $-CH_3$ group); (3) mesoscale modeling that can reach into experimentally accessible length- and time scales. This is usually the province of Kinetic Monte Carlo [5]; but the recent addition of "ultra-coarse graining" [6] could arguably fall into this category, and finally continuum models (*e.g.*, finite element modeling). Each level hence covers phenomena that occur over a particular length and time scale. Multiscale modeling has been particularly successful in situations where the properties of a material or a system of materials investigates the triad of process-structure-property relationships.

2.3 Physical Analytics- call for a new field

Traditional theoretical and computational approaches to modeling materials systems typically involve hand-tailored simplified models of system behavior, augmented by experimental studies that fit a few model parameters. This works well when the relevant scales separate (in time, length, or energetically), or when the renormalization group yields a scale-invariant macroscopic behavior. Often, however, no systematic simplification is known; yet simulating the full complex model is computationally slow. What is needed to circumvent this bottleneck is to incorporate machine learning, statistics, information geometry, and optimization techniques.

Our group is combining the strengths of physical models with advances in data analytics using learning techniques for uncertainty quantification and calibration and using approximate simulation results as "features" that guide predictions. Our approach, which we call "**physical analytics**," (Fig. 1) will be distinguished from most traditional machine learning and optimization by its focus on problems where data are both expensive to generate and highly heterogeneous in structure. This necessitates new mathematical approaches. Physical analytics will be distinguished from physics-based computational modeling by recognizing that modeling is a *process*. Models are created using an adaptive closed-loop approach: information generated by computations and experiments informs how models can be improved. Physical analytics will control and optimize this model-building process. Our

research creates a collection of mathematical methods for prediction, optimization, model reduction, parameter estimation/calibration, and uncertainty quantification.

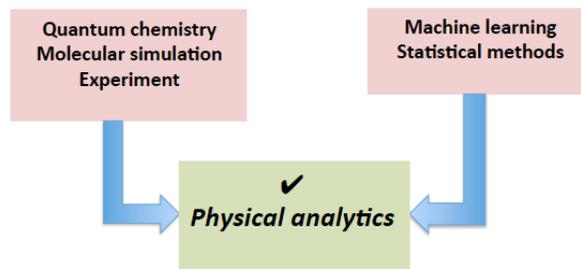


Figure 1: Cartoon depicting the combination of skill sets that comprise this new field of Physical Analytics, proposed here.

3 MULTISCALE MODELING OF SEMICONDUCTOR MATERIALS

In this section, we provide three recent examples from our own group's research to illustrate the challenges and successes of a multiscale modeling approach.

3.1 Multiscale Modeling of All-Organic Solar Cells: Covalent Organic Framework pores filled with C_{60} molecules

While no longer considered a viable route to a highly efficient solar cell in the light of the advent of metal halide perovskites, the idea of filling the open, predictably sized, pores of a Covalent Organic Framework (COF) material [6] like Pc-PBBA (a *p*-type material) with (*n*-type) C_{60} molecules (Fig. 2) has other interesting ramifications for the use of COFs for purposes such as sensing or separation. Our approach involved using a combination of Monte Carlo, Kinetic Monte Carlo and Molecular Dynamics simulations. Fast, straightforward to set up, Monte Carlo (MC) simulations were able to quickly determine the maximum packing fraction and show that experimental packing of a similar COF with fullerenes was close to the theoretical maximum packing that we predicted (49% vs. our predicted 51% maximum packing) [7]. However, a more realistic KMC simulation showed that the optimal packing would occur with large enough inter-fullerene separations that electron transport would be poor; borne out by the extremely low efficiency of an experimental solar cell that was fabricated from a similar COF. [8]

While the COF is well suited to a lattice representation in the KMC, our follow-up Molecular Dynamics simulations starting from the end-point of the KMC, shows that while KMC overestimates the extent of defects inside the COF pores, it confirmed the KMC prediction that fullerene filling in the pores is insufficiently good to create efficient solar cells in that particular COF. Computational

design could explore the viability of other COF designs for solar cell design; however, other applications of COF pore-filling seem more attractive.

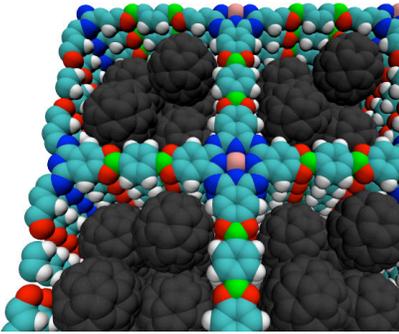


Figure 2: PC-PBBA COF (colorful outer framework) enclosing nanoscale diameter pores that are filled with (black) C_{60} fullerene molecules.

3.2 Multiscale Modeling of Crystal Growth: C_{60} molecules growing on a pentacene surface

The growth of one material on top of the surface of a different material in which neither of them share a common equilibrium crystal structure is known as heteroepitaxial growth. It was difficult to understand *a priori* which of the multiscale approaches will be the best to capture the nuances of the three-dimensional growth that ensues. Multilayer growth with its need to study 100s of molecules effectively rules out both *ab initio* and explicit all-atom MD, and hence some form of “coarse-graining” is necessary. We have recently showed [9] that Kinetic Monte Carlo does an adequate job of capturing that the growth is three-dimensional (Fig. 3), but it does not reproduce the correct growth mode (Volmer-Weber) seen by the experiments. In addition, it requires that an extensive library of events (moves) and their associated energy barriers before the KMC can be attempted. These energy barriers have to be calculated by MD or DFT which requires considerable time and computational resources to determine. For this system at least, with its quasi-spherical C_{60} molecules, a coarse-grained MD approach has the advantage over KMC. While the time scale that CG-MD can cover (order of ns) is much shorter than that of KMC (order microseconds), the CG-MD showed that the ability to access off-lattice events was enough to tip the balance towards the prediction of the correct growth habit in comparison to experiments. Comparison to a novel continuum approach showed that the otherwise hard to validate KMC produced the correct island size distribution relative to the continuum model for sub-monolayer growth.

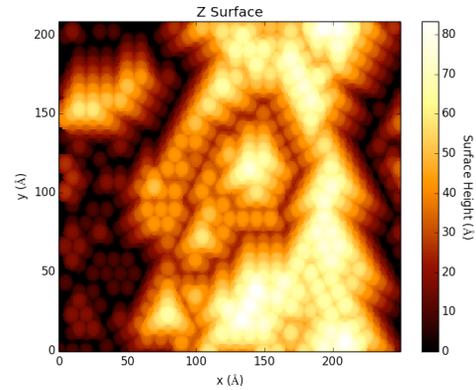


Figure 3: Top down view of the surface roughness of multiple layers of C_{60} molecules deposited on a smooth (but anisotropic) pentacene surface.

3.3 Organic Semiconductor Polymorph Identification using Physical Analytics

Our experimental collaborators at Stanford showed that there were more polymorphs (4-5) of TIPS-pentacene, a commonly studied organic semiconductor, than originally known. Our MD simulations using the MM3 intermolecular potential energy function due to Allinger *et al.* [10] showed that, in fact, the energy landscape is full of a fractal-like set of energy minima, or metabasins, separated by only tens of meV that can easily be overcome by the thermal energy of the system during processing (see Fig. 4) [11]. It was a triumph for the efficacy of the models that the simulation was able to match the unit cell dimensions of the four lowest energy polymorphs to experiments with a high degree of accuracy. This is especially true since the polymorphs were only ~ 30 meV apart; a result matched quantitatively by simulation. We also demonstrated the need for a suitably accurate intermolecular force field to see such nuances.

While these results were extremely encouraging, they are slow to achieve using a traditional systematic search across all possible values of the unit cell dimensions. As an alternative, we (Peter Frazier, ORIE Cornell) used a Bayesian search approach to make better guesses at “fitter” (better) solutions through an accelerated search. Such a search reduced the time to find the global minimum solution by a factor of ~ 20 and with less user oversight, as illustrated in Fig. 5. This is an excellent proof of concept for a “physical analytics” approach.

4 SUCCESSES AND CHALLENGES

There have been a number of clear successes in our use of multiscale modeling and simulation for semiconductor materials. Featured here, we have shown that the contributions of *ab initio* simulations, Molecular Dynamics, Kinetic Monte Carlo, and continuum modeling combine to

paint a comprehensive picture of some complex physical phenomena and materials design. The widespread use of well-managed MD codes like Sandia's LAMMPS has made

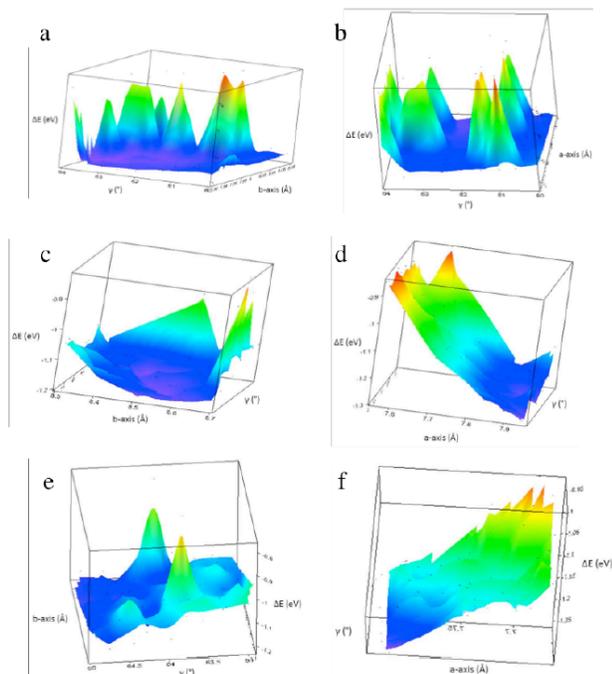


Figure 4: Energy landscapes for the response of a TIPS-pentacene crystal unit cell to changes in the unit cell dimensions (a, b, and γ) showing (in different views a-f) that some changes produce flat energy plateaux where changes in these dimensions have little effect (“sloppy” directions), like (c) and other changes that produce strong changes in the energy (“stiff” directions), like (a) and (b) and parts of the other views.

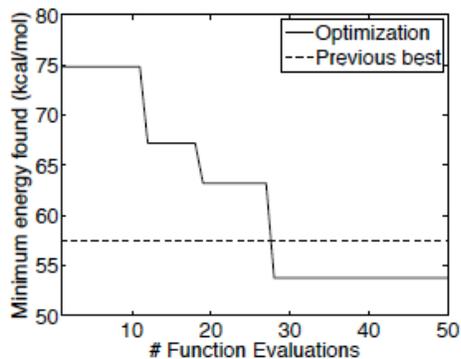


Figure 5: Bayesian search for the global energy minimum structure showing the rapid attainment of the final solution (solid lines) and a comparison to the previous best result obtained by a systematic (manual) search of the parameter space.

molecular simulation more accessible, although the usual disclaimer applies about the potential for untrustworthy

results if the code and models are used without due diligence by amateur users.

There are, however, other challenges. Multiscale simulation will need improvements in at least these areas: First, intermolecular potential models and force fields will need to expand to cover non-traditional materials systems like organic-inorganic hybrid systems that may include interactions of organic moieties with species such as Cs, Pb, Cd, Te, Se, *etc.* Most of these interactions do not currently exist in common databases. Second, many of the processes that are applicable to electrochemistry, energy materials, combustion, crystal nucleation and growth, *etc.*, involve a suitable treatment of chemical reactions. While there are a number of excellent reactive force fields [12] (of which REAX-FF and COMB are exemplars), they are complex to parameterize and simpler reactive models that are easy to parameterize are needed [13].

5 ACKNOWLEDGMENTS

The authors acknowledge the financial support of the NSF IGERT on interfacial surface science, the NSF IGERT on earth-energy systems, the NSF GK12 “Grass Roots” program at Cornell, and the Intel Foundation for an MS Scholarship.

REFERENCES

- [1] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, *J. Chem. Phys.*, 21, 1087, 1953
- [2] B.J. Alder and T. Wainwright, *J. Chem. Phys.*, 31, 459, 1959
- [3] L. Pauling and E. Bright Wilson, *Introduction to Quantum Mechanics with Applications to Chemistry*, McGraw Hill, 1935.
- [4] Brooks et al., *J. Comp. Chem.*, 30, 1545, 2009.
- [5] A. F. Voter, in *Radiation Effects in Solids*, ed. by K. E. Sickafus and E. A. Kotomin (Springer, NATO Publishing Unit, Dordrecht, The Netherlands, 2005).
- [6] J. F. Dama, A.V. Sinitskiy, M. McCallagh, J. Weare, B. Roux, A.R. Dinner, G. Voth, *J. Chem. Theory Comp.*, 9, 2466, 2013
- [6] Côté, A. P. *et al.*, *Science*, 310, 1166, 2005.
- [7] B.T. Koo, P. G. Berard and P. Clancy, *J. Chem. Theor. and Comp.*, 11, 1172, 2015
- [8] M. Dogru, *et al.*, *Angew. Chem.* 125, 2992, 2013.
- [9] Y. M. Acevedo, R. A. Stehl, P. Berard, D. L. Koch and P. Clancy, *Langmuir*, 2016, 10.1021/acs.langmuir.5b04500
- [10] N. L. Allinger, Y. H. Yuh, J.H. Lil, *J. Am. Chem. Soc.* 111, 8551, 1989
- [11] Y. Diao, *et al.*, *JACS*, 136, 17046-17057 (2014)
- [12] T. Liang, *et al.*, *Ann. Rev. Mat. Sci.*, 43, 109, 2013
- [13] J. Andrejevic, J.M. Stevenson and P. Clancy, *J. Chem. Theor. Comp.*, 2016, 10.1021/acs.jctc.5b00918