NASA/TM-2008-215118

State of the Art Assessment of Simulation in Advanced Materials Development

Kristopher E. Wise

National Institute of Aerospace, Hampton, Virginia

May 2008

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counterpart of peer-reviewed formal professional papers, but having less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

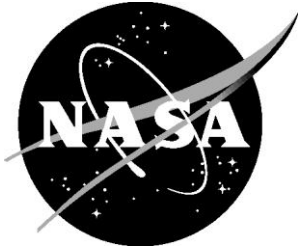- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results ... even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at *http://www.sti.nasa.gov*

- E-mail your question via the Internet to help@sti.nasa.gov

- Fax your question to the NASA STI Help Desk at (301) 621-0134

- Phone the NASA STI Help Desk at (301) 621-0390

- Write to:
  NASA STI Help Desk
  NASA Center for AeroSpace Information
  7115 Standard Drive
  Hanover, MD 21076-1320

NASA/TM-2008-215118

State of the Art Assessment of Simulation in Advanced Materials Development

Kristopher E. Wise

National Institute of Aerospace, Hampton, Virginia

May 2008

Available from:

I. Introduction

Research and development efforts with the objective of designing and optimizing new materials for use in advanced aircraft are a critical endeavor in enabling the production of next generation aircraft. One need look no further than Boeing's new 787 airliner for evidence of this. Without the substantial effort and investment made in the development of carbon fiber reinforced composite materials during the last several decades of the 20th century, there is little doubt that this aircraft would have been built with aluminum, just as its predecessors had been. The non-negligible cost of this research and development work will be more than justified by the expected improvements it will enable in aircraft efficiency, safety, and durability. Materials research and development work has, of course, continued to advance at a rapid pace. While carbon fiber composite materials provide a number of key benefits relative to aluminum, the gains which may be achievable with nanocomposite materials are even more impressive. Whereas filling polymers with carbon fibers yields primarily mechanical property improvement, filling them with nanoscale fillers offers the hope of achieving true multifunctional materials. Next generation nanocomposite materials will incorporate, in addition to mechanical reinforcement, properties such as electrical conductivity and sensing and actuating capabilities.

Making significant advances in the development of materials demands concomitant advances in the experimental and theoretical tools used in the development process. The objective of this document is to summarize a number of theoretical and computational approaches currently being developed and applied in the field of materials modeling. In view of the enormous range of materials systems and theoretical frameworks which could be covered, it would be impossible to provide a comprehensive summary of all work being conducted in this area. Rather, the focus will be placed the techniques most relevant to modeling work supporting materials development in the Advanced Materials and Processing branch with support from the Subsonic Fixed Wing aircraft program. In particular, attention will be given primarily to atomistic scale classical (non-quantum mechanical) modeling of advanced polymer and polymer nanocomposite materials.

II. Challenges in Nanocomposite Materials Modeling

In assessing the state of the art in this field, one quickly realizes that it is still in a fairly primitive state relative to the accomplishments which have been achieved in other areas of modeling. For example, quantum mechanical

calculations have been routinely performed on small molecules, defined as containing on the order of ten heavy atoms, for approximately 50 years. The accuracy with which energies, structures, and spectroscopic properties can be calculated for these systems is often within the experimental error of the corresponding experimental measurements.[1,2] Molecular simulations of the thermodynamic properties of simple liquids and mixtures thereof, as well as crystalline solids have also advanced to a similar level of accuracy. There are examples in the literature in which both quantum mechanical calculations and molecular simulations have called into question, and led to the correction of, erroneous experimental results.[3]

Unfortunately, the application of these same methods to polymers and nanocomposites is not nearly as advanced. The reason for this lies primarily in the spatial and temporal scales which are relevant in these systems. In the field of quantum mechanical calculations, well-developed methods exist for calculating the properties of both small molecules and, by taking advantage of their periodic symmetry, effectively infinitely large crystalline materials.[4] In between these extremes, however, no reliable approaches have been developed. Likewise, in the case of classical molecular simulations, relatively small systems (~100,000 atoms) may be simulated for relatively short times (~10 – 100 ns) with acceptable accuracy on current supercomputers. While simulations of this size are an impressive accomplishment that would have seemed unobtainable as recently as a decade ago, they still fall tens of orders of magnitude short of the experimentally relevant size and time scales. The unfortunate reality is that polymer molecules and nanoscale fillers are very large and move very slowly on the atomic size and time scales, which are the basis of both quantum mechanical calculations and classical simulations.

Realizing that the brute force approach of throwing ever more powerful computers at these challenges will achieve no more than a few orders of magnitude improvement in either the spatial or time scales which may be addressed, the physics and chemistry communities have instead sought more innovative approaches to circumventing them. A number of these ideas will be described in what follows. Most, if not all, of these methods are still very much the subject of active research and should be regarded as such. The present discussion will focus on classical simulation methods and a description of how they are being revised and extended to address nanocomposite materials. The particular topics selected for discussion are not exhaustive, but should provide a sense of where the limits currently lie and where substantial gaps and challenges remain.

2

III. Overview of Classical Simulations

Dating from the dawn of the digital computer era, methods such as Monte Carlo (MC) and molecular dynamics (MD) simulations have been essential methods in computational chemistry and physics, as well as materials and biomolecular modeling. In their common forms, both MC and MD simulations describe interactions between atoms using an empirical force field which is derived by fitting to either experimental or accurate quantum mechanically derived data. This includes interactions between chemically bonded atoms, such as bond stretching, angle bending, and torsional twisting interactions, as well as between non-bonded atoms, such as electrostatic and van der Waals interactions.

In both cases, one specifies the initial positions of all atoms and a potential energy function that specifies the mathematical form of the various interactions described above. The force field parameters are the constants, specific to each atom type, which allow the energy of a particular geometrical arrangement of atoms to be calculated using the potential energy function. These potential energy functions are designed to maximize both the physical accuracy of the computed quantities and the computational efficiency of their evaluation. This process involves a compromise between accuracy and speed which must be balanced by the demands of the study being conducted. A number of these potentials have been devised and tested extensively over the years and their behavior and reliability in a variety of situations is well understood.

Another common feature of MC and MD simulations is that they are capable of simulating physical systems in specified thermodynamic ensembles. In particular, one may specify some combination of thermodynamic state variables, such as the total number of particles (N), temperature (T), pressure (P), or chemical potential (μ). By holding combinations of these variables constant during a simulation, different thermodynamic ensembles may be considered. Commonly encountered examples include the canonical (constant NVT), grand canonical (constant μVT), and isothermal–isobaric (NPT) ensembles. This capability is critical when attempting to model their corresponding experimental configurations.

The essential difference between MC and MD methods lies in the inclusion or absence of time dependence in the simulation. In MD simulations, the simulation is conducted in a deterministic, time dependent manner by numerically integrating Newton's equations of motion. The mass weighted acceleration experienced by each particle at a particular time step is calculated by summing the forces acting on it, which originate at each of the other atoms and from any externally applied perturbations. The accelerations are used to calculate velocities, and in turn displacements which are applied to the particle positions. This process is iterated until a predetermined simulation duration is reached. This explicit time dependence permits the study of dynamic properties of the system by observing the time evolution of system properties.

In contrast, MC simulations are not time dependent. As in the case of MD simulations, one specifies the initial positions and connectivity of the atoms, but rather than following a deterministic trajectory through time, one explores the conformational space of the system by accepting or rejecting proposed moves according to the Metropolis algorithm given below:

Metropolis Algorithm :

1. Create new configuration from existing configuration by random move
2. Calculate energy of new configuration
3. Calculate $\Delta E = E_{new} - E_{old}$
4. If $\Delta E < 0$, accept move and continue
5. If $\Delta E > 0$, draw a random number R between 0 and 1

  If $e^{-\frac{\Delta E}{kT}} > R$, accept move and continue

  If $e^{-\frac{\Delta E}{kT}} < R$, reject move and continue with old configuration

---

The new configurations described in Step 1 are generated by choosing a type of move from a predefined catalog of moves for either single atoms (translations) or groups of atoms (eg. changes in bond length, bond angle, torsional angles etc). The energy of the proposed configuration is calculated and compared with the energy of the previous configuration and the new configuration is accepted or rejected (Steps 4 and 5). Monte Carlo simulations are often used in materials simulations, particularly for polymeric systems, because they permit more rapid exploration of the available conformational space due to their ability to make large changes in the structure of a complex molecule in a

4

single step, provided the energy of the new structure is not too different from the current structure. This ability is particularly important for systems with rugged potential energy surfaces in which a number of states with similar energies are separated by high energy states that would require a large activation energy to surmount. In effect, a Monte Carlo simulation can 'tunnel' through these barriers in a single step while a molecular dynamics simulation might require thousands or millions of small steps to cross them. While this is an important advantage of the Monte Carlo method over molecular dynamics, it is certainly still possible for Monte Carlo simulations to become trapped in a local potential energy basin when this region of configuration space is surrounded by large and broad potential energy barriers. Methods for addressing this problem are discussed below.

In Monte Carlo simulations, one calculates ensemble averages of some set of thermodynamic observables by averaging over a large number of uncorrelated measurements of those observables. These averages become statistically meaningful when the algorithm used to generate the system configurations uniformly samples the systems phase space and when the number of corresponding measurements becomes large. In molecular dynamics, on the other hand, time averages of a set of thermodynamic observables are calculated by averaging over the individual measurements taken over the course of a sufficiently long simulation. Again, the statistical reliability of the time averages depends on the ability of the simulation algorithms to fully sample phase space within the duration of the simulation. The formal equivalence between Monte Carlo and molecular dynamics simulations arises from the ergodic hypothesis. This states that the ensemble average value of some thermodynamic observable becomes equal to the time average of that variable, provided the ensemble average is taken over a large enough sample of configurations and that the time average is taken over a sufficiently long simulation. Satisfying these requirements is difficult for all but the simplest systems, and almost impossible, in practical terms, for polymer melts and nanocomposites. This reality represents one of the major gaps in the current state of the art of molecular simulations of these materials.

The following two sections of this overview will describe the two key issues that must be addressed to advance the state of the art in classical simulations of advanced materials being considered for use in aerospace applications. The first of these, discussed in Section IV, is the scaling of the computational cost, typically measured in CPU hours, for performing simulations as a function of the size of the simulated system and of the duration of the

simulation. This section will explain why we are currently limited to $\sim 10^{-5}$ seconds (for a small number of atoms) and 10^{-15} grams (for very short simulations), and some approaches that are currently being developed to address these limits. Secondly, the issue of the accuracy and reliability of the predicted materials properties is discussed in Section V. The primary focus of this section is on the assumptions and simplifications that are inherent in current generation potential energy expressions and the force field parameters that are used in them. Of particular interest are the various tactics that are being developed to improve the flexibility and transferability of the potentials and force fields. It is impossible, of course, to completely disentangle these two topics. Improving the quality of the representation of the interatomic and intermolecular potentials used in these simulations invariably increases their computational cost. Conversely, using progressively more approximate, or coarse grained potentials can dramatically reduce the cost of the calculation, but at the expense of reduced fidelity. Part of the *art* of molecular and materials simulations is balancing the two countervailing objectives of computational economy and accuracy.

IV. Scaling of Computational Cost with Size and Duration of Simulations

Prior to addressing the new methods being developed in this field, it is useful to understand what the current practical limits are for classical simulations. Carrying out long duration simulations is particularly important in the biomolecular sciences for studying processes like protein folding and the diffusion and binding of drug molecules to protein active sites. This is particularly true because the experimental timescales for many of these processes are tantalizingly close to the outer limits of current simulation methods. The limiting factor in carrying out long duration simulations is the size of the time step that may be used. To maintain numerical stability in the integration algorithm, the time step used must be shorter than the period of the fastest motion occurring in the system. For molecular systems, this is the length of the covalent bond stretching vibration. Accurately resolving these motions limits the magnitude of the time step to about 1 fs (10^{-15} sec). To the best of the author's knowledge, the longest simulation of a realistic system, which included the surrounding solvent molecules, was a one microsecond simulation of the headpiece of the Villin protein solvated in water.[5,6] This is the equivalent of 10^9 one fs time steps. By using an effective solvent model, which replicates the dielectric environment produced by the real solvent without actually including solvent molecules in the simulation, another group was able to study a similarly sized polypeptide containing 36 residues (amino acid repeat units) for 10 microseconds (10^{-5} sec), which is 10^{10} one fs time steps.[7] It should be noted that in both of these studies techniques were used to increase the time step size,

6

resulting in fewer steps actually being covered in the simulations. This was done by freezing the vibrations of bonds involving hydrogen atoms. This permits a sizable decrease in computational time while incurring a minimal loss of accuracy.

While lengthy simulations like these are becoming more common in the biomolecular simulation subfield, they are still quite rare in studies of polymer melts or nanocomposites. There is a large literature demonstrating that researchers in these areas make extensive use of classical simulations to study the dynamics of short timescale processes and the equilibrium properties of these systems.[8-12] It is widely recognized, however, that it is futile to attempt to simulate the longer relaxation time processes of dense, entangled polymer melts which can range from seconds to days. Physical aging processes present an even greater challenge in which the timescales can stretch into years. It is obvious that brute force application of current simulation methodologies will not solve these problems.

If long duration simulations are most commonly performed in the biomolecular community, large system size simulations, in terms of the number of atoms, are more the domain of the materials modelers. At the time of this writing, the largest classical simulation reported in the literature was of a cube of crystalline material containing approximately 20 billion (2×10^{10}) atoms.[13] This was a very short simulation aimed at modeling the initial stages of crack propagation. While this enormous calculation require the dedicated usage of a state of the art DOE supercomputer, it is somewhat less impressive when one realizes that 20 billion silicon atoms is only ~10^{-12} g, about 14 orders of magnitude smaller than an experimentally relevant mass.

The two key issues controlling the computational cost of a simulation are (1) the time spent per time step to evaluate the energy (and forces in a molecular dynamics simulation) of the ensemble of atoms at that step, and (2) the number of steps that must be taken in that simulation to insure that all of the statistically relevant configurations of the ensemble of atoms have been visited enough times to produce a meaningful set of ensemble averaged thermodynamic properties for the system being simulated. This second issue is of critical importance, as any results produced by a simulation that is too short to adequately sample the phase space of the system will produce incorrect and meaningless results. There are, unfortunately, many published simulation studies in which this key point is either not understood or ignored.

IV. 1. Cost per Simulation Step

There are two groups of interactions which contribute to the energy and forces that must be calculated at each time step. The bonded interactions, which include covalent bonds, bond angles, and torsional or dihedral terms, are relatively inexpensive to calculate. These terms are localized, involving 2, 3, and 4 atoms, respectively, and generally have a simple quadratic or exponential form. As a result, the computational cost scales linearly with the number of atoms in the simulation. The second set of interactions are those acting between atoms which are not chemically bonded. This generally means van der Waals interactions and charge – charge electrostatic interactions, although some potential energy functions include the interaction between higher order electrostatic moments (eg. dipoles, quadrapoles, etc). These terms are substantially more expensive to calculate because they are nonlocal, long range interactions that act between all pairs of atoms in the simulation. This leads to a computational cost that scales as the square of the number of atoms, in the most straight-forward implementation. This problem was recognized very early in the development of classical simulation methods and a number of algorithms have been developed to address it. The best of these currently achieve a high degree of accuracy while reducing the scaling of the computational cost to (N log N), where N is the number of atoms. This is clearly a significant improvement over the N^2 scaling of the naive approach. This area is still a topic of ongoing research, however, as concerns about accuracy persist and because the size of the prefactor in the scaling relation can be quite large, depending on the quality of the algorithm.

The first and most obvious method for expanding the duration and size of classical simulations is to reduce the level of detail in the simulated system, generally referred to as coarse graining, which reduces the cost associated with evaluating the energy and forces at each time step. The simplest and most widely used type of coarse graining is referred to as the united atom model. In this approach hydrogen atoms are combined with the atom they are chemically bound to, typically a carbon atom, to produce a new united atom with mass and diameter increased appropriately. For systems composed primarily of aliphatic carbons, a common situation in studies of synthetic polymers, the number of particles to be considered in the simulation is reduced by about a factor of three. This also permits the use of a slightly longer time step due to the elimination of all C-H bond stretching vibrations, which are typically the highest frequency, or shortest period, vibrations in a simulation. Of course this approach can be

extended further by, for example, collapsing all atoms in each phenyl ring of an aromatic polymer into a single bead. In this case ten atoms (C_6H_4) are reduced to one effective interaction site.

In essence, coarse graining amounts to removing degrees of freedom from the fully atomistic molecule or polymer. So long as the group of atoms being replaced by a coarse grain bead is reasonably rigid, both geometrically and electronically, the substitution will have little impact on the fidelity of the simulation while dramatically increasing its efficiency. For the cases described in the preceding paragraph, aliphatic and aromatic hydrocarbons, this is likely to be the case as both functional groups are structurally rigid, having no internally rotatable bonds, and have straightforward electronic properties, with no (or minimal) net charges or permanent dipole moments. The gaps in coarse graining are exposed, however, when one attempts to construct representations for groups of atoms with structural flexibility or more complex electronic structure. In these cases, the reduction of the fully atomistic structure to a corresponding coarse grained representation relies on the 'chemical intuition' of the individual constructing the reduced model, due to the lack of any generally accepted method for algorithmically performing this operation. As a result, two researchers working on the same system may arrive at somewhat different coarse grained representations, neither of which may be optimal or unbiased.

Resolving this problem of subjectivity in reducing atomistic representations to their coarse grained equivalents is clearly critical to achieving optimal efficiency and correctness. Doing so will require that the community reach some consensus, at least within families of related molecules or polymers, on the most physically correct approach. It is likely that a consensus will emerge organically after enough studies are reported in the literature to allow for an empirical evaluation of their relative merits. This process would be greatly accelerated by the development and broad distribution of an open source implementation of the various state of the art algorithmic procedures for constructing coarse grained models. This would be most useful if it were integrated with a similar package for developing force field parameters as discussed in Section V.

While the discussion thus far has focused on methods for creating coarse grained representations of atomistically detailed chemical structures, the converse process of reverse mapping from a coarse grained to atomistic model is of equal importance. Once a sufficiently long coarse grained simulation has been performed, for example to reach an

equilibrium configuration or to reach some other region of the systems phase space, it is often desirable to revert to a more detailed model of the system to accurately sample the physical properties of the system at this new configuration. Doing so requires an accurate and reliable method for mapping the atomistic details of the system onto the coarse grain representation. This turns out to be more difficult than one might initially suspect due to the reversed mapped system having highly strained or overlapping structures which are difficult to relax to correct local equilibrium structures. Another gap in this area lies in reconstructing systems with the correct chirality or stereochemistry on the molecular scale and the correct tacticity on the scale of polymer backbones.

This section can best be summarized by noting that, beyond the united atom model, coarse graining (and reverse mapping) techniques are still relatively new in the molecular simulation field. It is to be expected that it will take a period of time for researchers to explore and publish a variety of different approaches before the community will be able to make a good judgment on what are the best practices, in terms of accuracy and efficiency. It is the author's view that this process would be greatly accelerated by the development of an open source software package with user modifiable implementations of these methods. This would enable a broader community of researchers to utilize and improve on the state of the art techniques.

IV. 2. Number of Simulation Steps Required to Reach Equilibrium

The second determining factor for the computational cost of a simulation, as mentioned above, is the number of simulation steps that must be taken to ensure that a statistically meaningful sampling of the configurational space of the system has been achieved. Calculating ensemble or time averaged thermodynamic properties of the system, such as the entropy, free energy, or heat capacity, requires that all regions of the configurational phase space of the system be sampled. Proper sampling of the system's phase space is much more difficult for systems containing long, stiff polymer molecules and massive filler particles than it is for a simple liquid or gas. In simulations of simple, low molecular weight systems, the molecules move about relatively freely which allows them to reach equilibrium with the imposed thermodynamic conditions quickly and increases the probability that they will sample all of the statistically relevant regions of phase space. For long polymers, on the other hand, mobility is drastically reduced by their mass and by the fact that their connectivity does not allow for them to pass through one another. In fact, the dominant mode of molecular motion in dense polymer systems is referred to as reptation, which describes the

polymer's snake-like motion. When the polymer system is sufficiently concentrated that the individual polymers are entangled with one another, reptation becomes the only possible means of moving, as any translation or rotation will result in an unfavorable collision with another polymer. When filler particles, even those of nanoscopic dimensions, are added to the simulation, matters become even more difficult. Moving the filler particle through the dense polymer matrix requires displacing all of the entangled polymer molecules in its path and filling in the void left by its movement. Increasing the aspect ratio of the particle in either one (e.g. carbon nanotubes) or two (e.g. graphite oxide nanosheets) further increases the number of polymer molecules that must be displaced to permit movement of the filler particle.

One approach to accelerating these simulations, simplifying or coarse graining the atomistic representation, has already been discussed in the previous subsection. This strategy is limited by the fact that the extent of coarse graining required to reach the relevant timescales is so great that the physical meaning or relevance of the resulting simulation would be questionable. The underlying problem is that while one has substantially reduced the number of degrees of freedom being simulated, the simulation is still preferentially sampling the low energy portions of phase space. What is needed is a way of preferentially sampling the high energy, and therefore improbable, conformational transitions that form barriers between the physically interesting regions of phase space. This Section will describe two approaches for accomplishing this goal. Both of these methods are relatively recent developments which are just beginning to be employed in polymer and composite simulations after having been created for simulations of quite different physical systems.

The first approach is variously referred to as parallel tempering or replica exchange simulation.[17] This approach attempts to increase the sampling of high-energy structures, thereby improving the quality of the overall simulation. This is done by simultaneously performing multiple simulations on the same system over a span of temperatures. The replicas being simulated at higher temperatures have greater internal kinetic energy and are thus able to escape potential energy barriers that would trap lower temperature simulations. At the same time, the replicas being simulated at lower temperatures are sampling the potential energy surface that the simulation is intended to explore. The key to this method lies in periodically interchanging the particle configurations between simulations being conducted at adjacent temperatures. This allows the structure in each individual simulation to explore a range of

temperatures, and by extension, a range of regions of phase space. There are still unresolved problems with this method, for example the quality of the sampling is very dependent on how the temperature ranges are discretized. Too broad or narrow temperature ranges can affect the efficiency with which the overall temperature range is explored, leading to less efficient simulations. Ways of automatically and adaptively controlling these parameters are currently being explored.[18,19]

A second method for increasing the efficiency with which phase space is explored is referred to as generalized ensemble simulations.[20-22] In this approach a set of statistical weights are used to normalize the sampling of the phase space explored in a simulation. The underlying idea is that by flattening the potential energy surface of the system, it will be sampled more evenly and more efficiently than by sampling it in an unbiased manner. This is accomplished by deemphasizing very probable configurations, and enhancing low probability configurations by assigning to them low and high weights, respectively. At the end of the simulation, the statistical values are restored to their proper values by reweighting, or removing the effect of the weights applied during the simulation. As in the case of parallel tempering, this is a relatively new technique that is still an area of active research, particularly its application to simulations of polymers and composites.[23,24]

V. Accuracy and Reliability of Materials Properties Prediction

While the focus of Section IV was primarily on the practical and algorithmic factors that limit one's ability to carry out a simulation over a duration and size scale that is most physically appropriate for calculating properties of interest, this Section will focus on the ability of the potential energy function and force field parameters to yield accurate results. As the interest here is in the accuracy of the results, the focus will be on the most detailed (atomistic) representation of the system.

V. 1. Physical Correctness of Potential Energy Functions

As alluded to above, the accuracy and reliability of any computed results will depend both on the physical correctness of the potential energy function used to describe the form of the interactions between the atoms and the system as well as on the quality of the force field parameters used to evaluate the potential. Of course the most accurate method of evaluating the forces acting on each atom at each time step is to carry out a full quantum

12

mechanical calculation on the system. This approach is actually being used, and with increasing frequency, with formulations such as *ab initio* molecular dynamics (AIMD) and Carr-Parinello molecular dynamics (CPMD).[25] As might be expected, these calculations are quite expensive and are practically limited to systems containing on the order of 100 heavy atoms for durations on the order of 100 ps. While these limits will expand as larger and more capable computers are developed, they are unlikely to ever be useful for the size of systems and duration of simulations being discussed in this document. Instead, empirical approximations have been developed that attempt to model the various physical interatomic interactions that contribute most significantly to the total interaction energy.

Classical potential energy functions, originally developed for molecular dynamics and Monte Carlo simulations of molecular gases and liquids and biomolecules such as proteins and DNA, are typically composed of terms which describe covalently linked atoms (bonded terms) and terms which describe through space interactions between atoms which are not covalently bonded (non-bonded terms). The first two terms of the bonded component of the potential energy function describe the stretching of chemical bonds between two atoms and the bending of the angle between three adjacent atoms. Both of these terms are generally quadratic in form, yielding a parabolic potential energy surface, although a Morse potential is sometimes used for the bond stretching term to reflect anharmonicity. The third bonded term describes the torsional potential along four consecutively bonded atoms and is usually expanded in a series of cosine terms to reflect the rotational periodicity of this energy.

The nonbonded terms in classical force fields account for van der Waals and Coulombic interactions. The van der Waals term accounts for both the attractive and repulsive interactions between non-bonded atoms which arise from the London dispersion and Pauli exclusion interactions, respectively. In this term the repulsive component goes as the inverse 12th power of the interatomic distance, which yields a very steep but short ranged interaction. The attractive term, on the other hand, goes as the inverse 6th power of the interatomic distance, allowing for the repulsive term to dominate at very short separations while providing an attractive interaction at moderate ranges. The Coulombic term describes the interaction between permanent charges on the atoms and goes as the product of the charges divided by the square of the distance between them. This term falls off much more gradually than either

of the terms in the van der Waals interaction, which results in a much longer ranged attraction or repulsion, depending on the signs of the two charges.

Several research groups have developed classical potential energy functions, along with the necessary force field parameters, using the general approach described above.[26,27] They have been used extensively by the modeling community and shown to be quite effective in predicting accurate physical properties. Most of these groups work in the field of biomolecular modeling, which is reflected in the types of physical systems used in the development and validation of the potential energy functions. Adapting them to simulations of synthetic polymers is generally straightforward, usually requiring only adjustment of the force field parameters to reproduce the properties of the polymer being studied.

There are, however, at least two situations in which classical potential energy functions are inadequate for modeling the polymeric and composite systems that are of interest in this paper. The first of these involves the formation (and breaking) of chemical bonds during the course of a simulation. In simulations using classical potential energy functions all covalent bonds are declared in the input at the beginning of the calculation and do not change. While this is a reasonable restriction when simulating, for example, biomolecules, molecular liquids, and even polymers in many cases, there are other situations where it is unacceptable. Some of these include simulating the crosslinking of thermoset polymers, chemisorption to a surface, or chain backbone scission under conditions of mechanical deformation. The capability to model bond formation and breakage has been explored and implemented in advanced potential energy functions, most notably for the simulation of carbon and silicon nanostructures. Unfortunately, these are very specialized techniques that are narrowly focused on the systems of interest to their developers. Generalizing these approaches to novel bonding motifs and to a broader range of chemical elements has not been extensively explored in the literature. In their absence, modelers have had to resort to more expensive quantum mechanical or tight binding approaches which significantly limit the practical size and duration of simulations.

A second gap in the quality of current generation potential energy functions is their limitation, in almost all cases, to the static point charge representation of interatomic electrostatic interactions. Again, this is a relatively good

approximation for typical simulations of most synthetic and biopolymers. The reason for this is that these systems feature primarily non-aromatic bonding (single or isolated double bonding). Electrons involved in this sort of bonding tend to be localized and non-polarizable. For polymers with very polarizable functional groups, conjugated polymers, and carbon nanotubes, however, the use of static point charges misses the essential physics that make these systems so interesting. As in the case of incorporating bond formation and breakage, a number of research groups are actively working on this deficiency.[28,29] A number of diverse approaches to this problem are being studied, including utilizing higher order multipole moments, adding dynamically polarizable dipole moments, and including the possibility of dynamic charge redistribution, which responds to both internal structural changes as well as external electrostatic perturbations. These developments will be particularly important for simulations of the highly aromatic polymers often encountered in aerospace materials, and their composites with carbon nanotubes and other highly polarizable nanoscale fillers.

V. 2. Accuracy and Transferability of Force Field Parameterizations

For systems in which the classical form of the potential energy function is sufficient (those not requiring the treatment of changing bond topology or polarizability), a key determinant of the accuracy of the simulation is the quality of the force field parameterization. Developing a force field for a particular molecular or polymeric structure entails establishing numerical parameters for the various terms in the potential function. For the harmonic bonds and angles these are the Hooke's law constants, while for the torsional angles they are a set of constants describing the periodicity of the potential and the relative energy barriers between the various local minima. For the electrostatic term the parameters are a set of point charges, generally positioned at the location of each atom. Finally, the van der Waals term requires constants for the radii of each atom type and maximal attractive energy between each possible pair of atom types.

Developing an accurate set of parameters for a particular system is a labor intensive and time consuming task. Parameters for the bonded interactions and the atomic charges are generally derived from accurate quantum mechanical calculations. The van der Waals parameters, on the other hand, are generally adjusted to match experimental data, such as density, by carrying out a series of short simulations. Due to subtle correlations between the various terms in the force field, however, the whole process is usually iterative in nature.

One approach to reduce the effort of developing an entire new force field for each new structure to be studied is to assume some degree of transferability of parameters between similar molecules and atoms in similar local environments. While new atomic charges are almost always required when changing a structure, in many cases the bonded terms and van der Waals parameters can be reused with only minor modification. Knowing which parameters are likely to be transferable is largely a matter of experience and chemical understanding of the system being studied, and must always be carefully validated. The difficulty in developing new parameters and reliably transferring parameterizations between chemically analogous molecules are two key gaps hindering greater use of molecular simulations. There are a few groups attempting to automate, to a large extent, the process of performing quantum mechanical parameterization calculations and developing methods for extracting force field sets from the results. The difficulty in creating transferable parameter sets, however, may indicate a fundamental deficiency in the form of the potential energy function.

VI. Summary

The intent of this assessment has been to describe some of the most significant barriers or gaps that are impeding the application of state of the art molecular simulation techniques to the types of materials systems of most interest for aerospace application. Some of these are practical issues which will be resolved sooner or later, depending on the level of effort and investment of resources devoted to them. Most of the problems described in Section V fall into this category. At least one way (better methods might emerge) of addressing these problems is known and it is largely a matter of time before these cease to be significant issues. The difficulties outlined in Section IV, on the other hand, are more fundamental in nature are unlikely to be completely resolved soon, if at all. That is not to say, however, that they should be dismissed as insuperable problems. Methods such as parallel tempering and generalized ensemble sampling have already shown promise in a number of applications and creative new approaches to increasing sampling efficiency are being reported weekly in the simulation literature. While these problems may never be definitively solved, much can, and is, being done to improve the state of the art. In view of the fact that the field of molecular simulations is now over fifty years old, it continues to be a dynamic and forward-looking endeavor.

VII. References

[1] Feller, D.; Peterson, K. A. "Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than "chemical accuracy"?" *J. Chem. Phys.* **2007**, *126*, 114105.

[2] Helgaker, T.; Ruden, T. A.; Jorgensen, P; Olsen, J.; Klopper, W. "A priori calculation of molecular properties to chemical accuracy" *J. Phys. Org. Chem.* **2004**, *17*, 913.

[3] Schaefer III, H. F. "Methylene: A Paradigm for Computational Quantum Chemistry" *Science* **1986**, *231*, 1100. and references therin.

[4] Martin, R. M. "Electronic Structure: Basic Theory and Practical Methods" Cambridge University Press, 2004.

[5] Duan, Y.; Kollman, PA. "Pathways to a Protein Folding Intermediate Observed in a 1-Microsecond Simulation in Aqueous Solution" *Science* **1988**, *282*, 740.

[6] Lei, H.; Wu, C.; Liu, H.; Duan, Y. "Folding free-energy landscape of villin headpiece subdomain from molecular dynamics simulations" *Proc. Nat. Acad. Sci.* **2007**, *104*, 4925.

[7] Feddolino, P. L.; Liu, F.; Gruebele, M.; Schulten, K. "Ten- Microsecond MD Simulation of a Fast-Folding WW Domain" *Biophys. J.* **2008,** *108*, 131565.

[8] Gee, R. H.; Lacevic, N.; Fried, L. E. "Atomistic Simulations of Spinodal Phase Separation Preceding Polymer Crystallization" *Nature Mat.* **2006**, *5*, 39.

[9] Lacevic, N.; Fried, L. E.; Gee, R. H. "Heterogeneous Directional Mobility in the Early Stages of Polymer Crystallization" *J. Chem. Phys.* **2008**, *128*, 014903.

[10] Auhl, R.; Everaers, R.; Grest, G. S.; Kremer, K.; Plimpton, S. J. "Equilibration of Long Chain Polymer Melts in Computer Simulations" *J. Chem. Phys.* **2003**, *119*, 12718.

[11] Uchida, N.; Grest, G. S.; Everaers, R. Viscoelasticity and Primative Path Analysis of Entangled Polymer Liquids: From F-actin to Polyethylene" *J. Chem. Phys.* **2008**, *128*, 044902.

[12] Knauert, S. T.; Douglas, J. F.; Starr, F. W. "The Effect of Nanoparticle Shape on Polymer-Nanocomposite Rheology and Tensile Strength" *J. Poly. Sci.: B* **2007**, *45*, 1882.

[13] Kadau, K.; Germann, T. C.; Lomdahl, P. S. "Large-Scale Molecular-Dynamics Simulation of 19 Billion Particles" *Int. J. Mod. Phys. C* **2004**, *15*, 193. See also References 14-16.

[14] Kadau, K.; Germann, T. C.; Lomdahl, P. S.; Albers, R. C.; Wark, J. S.; Higginbotham, A.; Holian, B. L. "Shock Waves in Polycrystalline Iron" *Phys. Rev. Lett.* **2007**, *98*, 135701.

[15] Kadau, K.; Germann, T. C.; Lomdahl, P. S. "Molecular Dynamics Comes of Age: 320 Billion Atom Simulation on BlueGene/L" *Int. J. Mod. Phys. C* **2006**, *17*, 1755.

[16] Germann, T. C.; Kadau, K.; Lomdahl, P. S. "25 Tflop/s Multibillion-Atom Molecular Dynamics Simulations and Visualization/Analysis on BlueGene/L" *Supercomputing '05* (SC05, ACM 1-59593-061-2/05/0011), **2005**.

[17] Swendsen, R. H.; Wang, J. S. "Replica Monte Carlo Simulation of Spin Glasses" *Phys. Rev. Lett.* **1986**, *57*, 2607.

[18] Rathore, N.; Chopra, M.; de Pablo, J. J. "Optimal Allocation of Replicas in Parallel Tempering Simulations" *J. Chem. Phys.* **2005**, *122*, 024111.

[19] Trebst, S.; Troyer, M.; Hansmann, U. H. E. "Optimized Parallel Tempering Simulations of Proteins" *J. Chem. Phys.* **2006**, *124*, 174903.

[20] Berg, B. A.; Neuhaus, T. "Multicanonical Algorithms for First Order Phase Transitions" *Phys. Lett. B* **1991**, *267*, 249.

[21] Berg, B. A.; Neuhaus, T. "Multicanonical Ensemble: A New Approach to Simulate First-Order Phase Transitions" *Phys. Rev. Lett.* **1992**, *68*, 9,.

[22] Berg, B. A. "The Multicanonical Ensemble: A New Approach to Computer Simulations" *Int. J. Mod. Phys. C* **1992**, 3, 1083.

[23] Glotzer, S. C. "Molecular and Mesoscale Simulation Methods for Polymer Materials" *Ann. Rev. Mater. Res.* **2002**, *32*, 401.

[24] Yan, Q.; de Pable, J. J. "Hyperparallel Tempering Monte Carlo Simulation of Polymeric Systems" *J. Chem. Phys.* **2000**, *113*, 1276.

[25] Carr, R.; Parrinello, M. "Unified Approach for Molecular Dynamics and Density-Functional Theory" *Phys. Rev. Lett.* **1985**, *55*, 2471.

[26] Tersoff, J. "New Empirical Approach for the Structure and Energy of Covalent Systems" *Phys. Rev. B* **1988**, *37*, 6991.

[27] Brenner, D. W. "Empirical Potential for Hydrocarbons for Use in Simulating Chemical Vapor Deposition of Diamond Films" *Phys. Rev. B* **1990**, *42*, 9458.

[28] Halgren, T. A.; Damm, W. "Polarizable Force Fields" *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

[29] Ponder, J. W.; Case, D. A. "Force Fields for Protein Simulations" *Adv. Prot. Chem.* **2003**, *66*, 27.

REPORT DOCUMENTATION PAGE		*Form Approved* *OMB No. 0704-0188*

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)* 01-05-2008	2. REPORT TYPE Technical Memorandum		3. DATES COVERED *(From - To)*	
4. TITLE AND SUBTITLE State of the Art Assessment of Simulation in Advanced Materials Development			**5a. CONTRACT NUMBER**	
			5b. GRANT NUMBER	
			5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Wise, Kristopher E.			**5d. PROJECT NUMBER**	
			5e. TASK NUMBER	
			5f. WORK UNIT NUMBER 561581.02.08.07.15.02	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NASA Langley Research Center Hampton, VA 23681-2199			**8. PERFORMING ORGANIZATION REPORT NUMBER** L-19457	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001			**10. SPONSOR/MONITOR'S ACRONYM(S)** NASA	
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) NASA/TM-2008-215118	

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified - Unlimited
Subject Category 27
Availability: NASA CASI (301) 621-0390

13. SUPPLEMENTARY NOTES
An electronic version can be found at http://ntrs.nasa.gov

14. ABSTRACT

Advances in both the underlying theory and in the practical implementation of molecular modeling techniques have increased their value in the advanced materials development process. The objective is to accelerate the maturation of emerging materials by tightly integrating modeling with the other critical processes: synthesis, processing, and characterization. The aims of this report are to summarize the state of the art of existing modeling tools and to highlight a number of areas in which additional development is required. In an effort to maintain focus and limit length, this survey is restricted to classical simulation techniques including molecular dynamics and Monte Carlo simulations.

15. SUBJECT TERMS
High performance polymers; Modeling; Molecular dynamics; Nanocomposite; Simulation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON STI Help Desk (email: help@sti.nasa.gov)
a. REPORT	b. ABSTRACT	c. THIS PAGE			
U	U	U	UU	23	19b. TELEPHONE NUMBER *(Include area code)* (301) 621-0390

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18