Contents

1

First Principles Process Planning for Computer-Aided Nano-Manufacturing

Yan Wang

Georgia Institute of Technology, Atlanta, GA 30332, USA

1.1 Introduction

Nanotechnology is the understanding and control of matter at dimensions between 1 and 100 nanometers. Materials at these scales usually exhibit unique characteristics and can provide significant technical and economic advancement with novel applications. Although promising, most nanotechnology research only focuses on dozens of or a few hundred particles or molecules. To realize large-scale devices and commercializable products, massive assembly techniques with high-volume high-rate output are required. This poses a great challenge to the nanomanufacturing research community. That is, how to fabricate nanomaterials and devices in a repeatable and scalable way such that nanotechnology becomes commercially viable. Scaling up nanotechnology from laboratory setup to industry-level production is critical to enable mass-scale impacts of nanotechnology to our daily lives.

Nanomanufacturing is a physical or chemical process that results in new materials and devices with nanoscale structures that exhibit special properties and functions. In the past decade, research efforts have focused on investigating nanomanufacturing techniques that are suitable for large-scale production of materials with nanostructures, such as soft lithography, nanoimprint lithography, and others. However, to enable commercial-scale production, nanomanufacturing techniques are far from mature. Many technical issues related to nanomanufacturing still need to be resolved. For example, in process planning, we need to determine how precursor materials flow through during fabrication or reaction, which paths and what speeds fabrication tools/probes take to minimize wastes, etc. In quality control, we need to know how to adjust controllable parameters of processes such that uncontrollable but important variables are in tolerable ranges. Here, we focus on the issues related to process planning.

In product and process design, computational tools have played important roles. They allow engineers to predict behaviors of products and performance of manufacturing processes. Computer-

aided design (CAD), computer-aided engineering (CAE), and computer-aided manufacturing (CAM) tools are the major enabling technologies and contributors of the success of well-established commercial products such as cars, airplanes, and semiconductor chips. During the so-called *virtual prototyping* formed by CAD/CAM/CAE, CAD tools allow design engineers to build the digital model of a product, then simulate and optimize its behavior by finite-element alike methods in CAE tools. As the next step, CAM tools, especially computer-aided process planning (CAPP) modules in CAM software, enable manufacturing engineers to directly simulate the fabrication processes of the product and evaluate the processes' performance, then design and optimize process parameters (e.g. cutting path and feed speed in machining). The state-of-the-art CAM software can even build digital models of plants and simulate the operation of the complete factory for a new product that is still in the design stage.

Virtual prototyping can reveal and correct problems that may occur within a product's lifecycle (manufacturing, use, and others) and help engineers to understand the impacts of their decisions on product design and process design, before the designs are physically realized. Historically CAD/CAM/CAE has played significant roles in reducing cost and time-to-market while ensuring high-quality and repeatability of mass production in the traditional large-scale manufacturing.

Similar to the traditional manufacturing, future computer-aided nano-design (CAND) tools can be used to build digital models of nanoscale products as the starting point of virtual prototyping. computer-aided nanomanufacturing (CANM) tools are valuable to designing and planning the process of nanomanufacturing. Process planning is the systematic determination of the detailed methods and operations that nano-engineered materials can be manufactured economically and competitively from initial stages to finished stages. The goal of process planning is to ensure the repeatability and scalability of nanomanufacturing processes so that commercial scale nano-engineered products can be mass produced. To be a leading manufacturer, one must implement processes that are well thought through in terms of both physical and software systems. Computing and information technologies are playing an increasingly significant role in the success of new nanomanufacturing systems. Effectiveness and efficiency must be considered early in the design cycle for maximum impact.

We envision that CANM software with CAPP capability would be indispensable in future nanomanufacturing industry for repeatable, low-cost, and high-throughput production, similar to the CAM software in the traditional manufacturing. Our simulation-based process planning approach for CANM described in this chapter is a new paradigm to provide the fundamental understanding of the physical and chemical processes at nanoscale that is necessary to control and optimize those processes to realize scalable nanomanufacturing. Given the complexity, costs, and risks associated with empirical studies by physical experiments, modeling and simulation, especially the first principles approaches, can provide numerous advantages in the research of nanomanufacturing.

The wide spread adoption of nanoscale products, even the most mature ones in semiconductor industry, such as sub-100nm integrated circuit chips, flat panel displays and photovotaics, depend on reproducible and cost effective nanomanufacturing technologies. The challenge of driving cost per function down past critical price points requires understanding and coupled optimization of basic materials, equipment and processes at scales well below 100nm [48]. The full understanding of nanomanufacturing processes is important for process improvement. This understanding provides the basis to design and optimize processes systematically and rationally so that we know why, where, or even how to improve. In return, rational design of processes guides physical experiments and enables efficient implementations. CAPP enables simulation-based process design in traditional manufacturing. Simulation can help predict and correct potential manufacturing problems at very early stages, thus eliminating otherwise induced wastes and costs. Similarly, computational simulation is the most cost-effective way to understand the complex physical and chemical processes at nanoscales.

In this chapter, we present a first principles multiscale simulation framework that simulates both bottom-up and top-down nanomanufacturing processes and enables process planning, improvement,

and optimization. The framework is shown in Figure 1.1. Given an initial set of manufacturing process parameters (e.g., precursor materials, deposition direction, cutting rate, temperature, pressure, external field strength and directions, etc.), a controlled kinetic Monte Carlo simulation (cKMC) model is built. The cKMC generalizes the classical kinetic Monte Carlo (KMC) by introducing controlled events in parallel with self-assembly events so that both top-down and bottom-up fabrication processes can be simulated. The outputs of cKMC simulation are metrics to evaluate and optimize the process parameters. As the inputs for cKMC including states and transition rates, the first principles simulation of density functional theory (DFT) is used to search stable states with minimum energy on a potential energy surface (PES) and transition states that are associated with transition or reaction events, based on some representative small-scale molecular models.

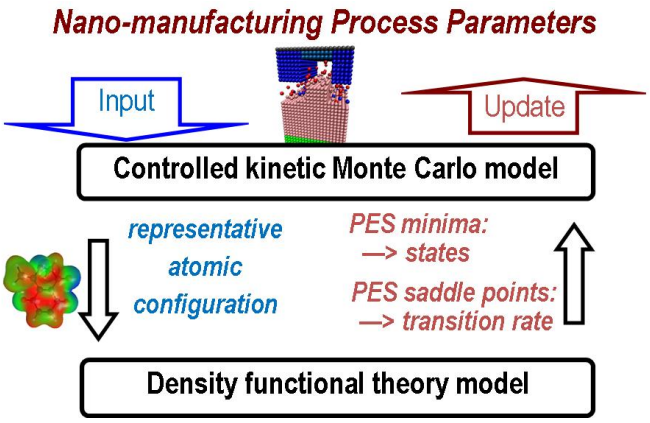


FIGURE 1.1: The proposed first principles multiscale simulation framework to simulate both bottom-up and top-down nanomanufacturing processes and enables process planning.

Scale complexity is inherent in simulating nanomanufacturing processes. At the quantum scale, electrons behave according to the first principles of quantum mechanics. At the atomistic scale, atoms and molecules aggregate or disperse based on the potential energy level as a result of the electron density distribution. At the mesoscale or larger, collective behavior of all atoms and molecules is observed. For length scales, simulating at the atomistic scale is required for nanoscale structural manipulation, yet large-scale production may have thousands of millions of atoms involved simultaneously. Similarly for time scales, electronic and molecular configurations change at the scales of femto (10^{-15}) seconds or shorter, whereas the complete manufacturing processes may last seconds or days. Computational efficiency and effectiveness should be both considered in detailed simulation of large systems. Therefore, a multiscale simulation approach for CANM is necessary to provide fundamental understanding of the processes.

In the rest of the chapter, we first provide a review of the existing nanomanufacturing techniques, including both top-down and bottom-up processes. Then, the most used nanoscale simulation methods are briefly introduced, including quantum Monte Carlo (QMC), DFT, molecular dynamics (MD), and KMC. The main body of the chapter is a proposed generic framework of first principles simulation-based nanomanufacturing process planning. The important elements of stable state and transition state search on first principles PES will be introduced, where new search algorithms to seek a global view of the complex PES are developed. cKMC as a generalization of KMC simulation at the atomistic scale will be presented, where both top-down and bottom-up nanomanufacturing processes can be simulated. The application of the new simulation framework on process planning will be demonstrated.

1.2 Background

1.2.1 Process Planning for Nanomanufacturing

In general, nanomanufacturing techniques are classified as either top-down or bottom-up [68, 9, 61]. In the top-down approaches, materials are removed with low volumes and sizes down to the scale of dozens of nanometers. In the bottom-up approaches, materials are assembled under the guidance of nanoscale templates, either physically or chemically. Some of those techniques are inherited and extended from the traditional semiconductor manufacturing techniques, such as nanoimprint lithography and vapor deposition, since semiconductor is one of the major driving forces of nanomanufacturing.

Among the different nanomanufacturing tools and techniques, some have become mature such as atomic force microscope (AFM), focused ion beam (FIB), chemical vapor deposition (CVD), and physical vapor deposition (PVD), and commercial equipment is available. It is foreseeable that more manufacturing approaches will become commercially available in the coming decade. Some of the key research issues in nanomanufacturing include how to produce and use precursor materials, how to assemble and characterize precursors, how to design and integrate structures into devices and systems, as well as how to develop the corresponding instrumentation and equipment [16]. In addition, we also need to answer questions related to process planning, such as the proper dose of precursor materials, feed rates, injection directions and paths, and so on.

There is very limited research on the issues related to process planning for nanomanufacturing. Collision-free path planning methods in the traditional robotic assembly have been applied to AFM manipulation of nano particles [50, 13, 49, 73] and tip motion control [44, 53]. These methods only apply to one type of top-down fabrication techniques. To improve the throughput of nanomanufacturing systematically, the detailed nanoscale physical and chemical processes, i.e. how atoms are aggregated or dispersed to form nano structures because of self-assembly, should be considered.

1.2.2 Research Gaps in Process Planning for Nanomanufacturing

Despite the rapid progress of fabrication methods, the major challenge for nanomanufacturing is still the issue of how to fabricate high-precision nanostructures with high-throughput rates, which are two conflicting goals. To achieve the level of commercial scale production, the processes themselves have to be improved. Understanding the fundamentals of processes and planning them with the goal of production are important. There are several research gaps related to process planning between current techniques and future needs of commercial-scale mass production, listed as follows.

First, understanding the phenomena and nature of processes is important to predicting and controlling the process outcomes. Lack of fundamental insight on processes prohibits controllable mass production. For instance, although various fabrication methods of CVD and PVD have been widely used to vaporize and grow carbon nanotubes (CNTs), the control of position, orientation, chirality, and length of CNTs is still a daunting challenge to create uniform and isotropic materials batch-to-batch, which is critical for the desired mechanical, electrical and thermal properties. Single-walled CNT (SW-CNT) has superior properties than multi-walled CNT (MW-CNT). However, the purification process for SW-CNT to remove unwanted amorphous carbon and metal catalysts after its growth needs extensive and meticulous procedures. As another example, etching is a commonly-used process for material removal in semiconductors wafer fabrication. However, most of existing research on the process is by empirical studies. They do not provide the basis for understanding the etching rates, etching profiles, selectivity, etc. The large number of independent processing parameters also makes empirical studies time consuming [31]. Future mass production needs repeatable and controllable procedures. The fundamental understandings of the nanomanfuacturing processes

thus are critical. Without them, nanomanufacturing would remain an art that is highly dependent on individuals' skills rather than on systematic engineering.

Second, there is a lack of first principles methodologies that model and study nanomanufacturing techniques to allow for systematic and fundamental understanding. First principles or ab initio approaches rely on the foundational propositions or laws of physics to create models. In contrast, empirical approaches build models from experimental observations. Various empirical or semi-empirical approaches, including the widely used MD simulation, have been developed to model and simulate some processes such as in top-down nanolithography and bottom-up self assembly. However, these empirical models are based on experimental data fitting (e.g. to generate empirical force fields or interatomic potentials in MD) instead of established laws of physics. They can be used to replicate phenomena but do not provide the fundamental understanding. In addition, experimental data at nanoscales are not easy to collect. Insufficient data is the major hurdle to build accurate and precise empirical models. In contrast, first principles modeling and simulation approaches in computational chemistry and physics (such as QMC and DFT) solve Schrödinger's equation to predict the system's energy levels and configurations. Theoretically all physical and chemical properties can be computed and predicted precisely by solving Schrödinger's equation if there were no computational limitation. Ab initio simulation methods have been used to study some of the bottom-up self-assembly processes, such as CVD for crystal growing or diffusion process in doping in semiconductor domains. Yet, generic and integrated first principles methodologies for process planning and controls of both top-down and bottom-up processes are needed.

Third, repeatability and scalability of processes also require that the lifespan of fabrication tools should be predictable. For instance, the physical and chemical property changes of AFM tips during fabrication need to be controllable to ensure quality of products. When mechanical forces or electrical bias are continuously applied in scanning probe, tool wear is an important contributor of defects. There is a need to research on how to design the lithography process so that defects can be minimized and controlled.

Fourth, lack of knowledge of the impact of nanomanufacturing processes on environment, health and safety prohibits setting standards and regulations to prevent hazards. New chemicals are being generated in the processes, which could potentially pose risks to environment and safety. Future manufacturers of nanomaterials in large quantity should ensure that their production does not put their employees and general public in danger. Risks associated with manufacturing processes should be assessed at conceptual stages of process design. To avoid the unnecessary risks of conducting empirical studies, first principles simulation approaches are extremely valuable to assess unknown effects of any new chemicals without physical experiments.

Therefore, to reduce production time, waste, cost and associated risks to achieve realistic commercial-scale production, the fundamental understanding of nanomanufacturing processes via first principles approaches should be greatly enhanced. In this chapter, we will directly address the first two research gaps. We will demonstrate a first principles simulation-based process planning framework for nanomanufacturing. This simulation-based approach provides a first principles view of both top-down and bottom-up approaches starting from the quantum level simulation. Thus the most fine-grained information ever about detailed physical and chemical processes can be captured by simulation and provides the fundamental insights.

1.3 Simulation at Nanoscales

Various simulation methods that are focused on some specific length and time scales have been well developed. As illustrated schematically in Figure 1.2, either discrete or continuous approaches have

been used to simulate at the respective time and length scales. To construct these simulation models, either first principles or empirical methods are used to parameterize the models. For instance, QMC, Hartree-Fock self-consistent field, and DFT are considered to be first principles, whereas commonly used MD is empirical.

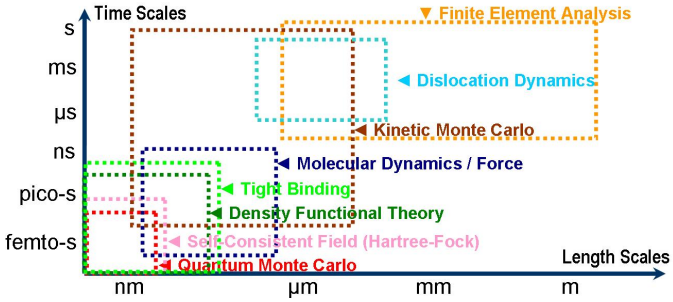


FIGURE 1.2: Simulation methods are developed at respective time and length scales.

A simulation method will become incapable or inefficient in simulating behaviors beyond its major scale. For instance, MD simulation is designed to describe molecules' behavior at time scales from femto seconds to pico seconds. It becomes extremely inefficient in simulating long-term behaviors over micro seconds or longer. Because most computational time is spent on thermal behaviors of atoms at the femto-second scale, the most interesting behaviors are actually much slower and are "rare events", which are hardly observable in simulation. Similarly, for length scales, DFT simulation as a popular first principles approach can be useful for dozens of atoms. But it becomes too expensive to simulate large molecules with thousands of atoms. Therefore, using a simulation model at its major length and time scale is important with the consideration of accuracy and efficiency. In this section, the most widely used QMC, DFT, MD, and KMC simulation approaches are introduced.

1.3.1 Quantum Monte Carlo

Among all computational methods that can calculate electronic structures (thus the atomic system potential energy) with low order (≤ 3) polynomial times and up to a few thousand electrons, QMC is the most accurate one. QMC directly computes the expected energy of electronic structures from the many-body Schrödinger equation by the sampling approach. As a result, QMC can compute system energy at both ground states (most stable with minimal energy) and excited states (with elevated energy level because of spikes of external energy injection).

Two varieties of QMC, variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC), are commonly used. VMC applies the variational principle to approximate the state of a system by optimizing the electron wave functions, whereas DMC employs a Green function to solve the evolution of wave function system toward the actual state with an imaginary time. Typically both approaches are used together. First, VMC is applied. The result of VMC becomes a good initial starting point of DMC for more accurate calculation.

Although the latest implementation of QMC can simulate an atomistic system with up to thousands of electrons in parallel computation, it is still not practical and economical to simulate systems at the typical engineering scale with at least hundreds of atoms.

1.3.2 Density Functional Theory

DFT is a much more efficient alternative of QMC to simulate electronic structure at atomistic scale. Its formulation is based on a theorem of Hohenberg and Kohn, which states that the ground-state properties of a many-electron system may be obtained by minimizing a functional of the electron density. With the electron density expressed in terms of the one-electron wave functions of the non-interacting system, the energy functional includes the components of kinetic energy of the non-interacting system, the interaction between the electron and ions as the potential energy, the Hartree term that represents pair-wise electron-electron Coulomb repulsion, and an exchange-correlation potential that approximates the many-particle interactions. By minimizing the total energy functional, one could find the ground-state wave functions and thus the total energy level. Numerically the Kohn-Sham equations, as the approximation of Schrödinger's equation, are solved iteratively to obtain self-consistent solutions.

Various exchange-correlation potential approximations have been developed. The well-known one is the so-called local-density approximation (LDA), where the exchange-correlation potential depends only on the local electron density, which itself is assumed to be uniformly distributed. This approximation also works surprisingly well when the distribution of electrons is strongly inhomogeneous, such as at surfaces and in molecules. The analogous approximation for spin-polarized systems, known as the local spin-density approximation (LSDA) [6, 62], has also been proved with good accuracy. Higher accuracy can be achieved by the generalized gradient approximation (GGA) [41], which includes no only the local density as in LDA and LSDA, but also the gradients of local density. The accuracy can be even further improved by considering higher-order derivatives. Several nonlocal functionals have also been proposed, such as average-density approximation and weighted density approximation.

In addition to the traditional ground state simulation, various extensions of DFT have been developed to simulate excited state (time-dependent density functional theory), and the influence of magnetic field (magnetic field density functional theory and current density functional theory).

The results obtained by DFT are typically an order of magnitude less accurate than those from QMC. Nevertheless, it is computationally much more efficient than QMC. As a result, it has been widely applied in a variety of fields in physics, geophysics, chemistry, biochemistry, and materials science.

1.3.3 Molecular Dynamics

MD is a simulation approach to model molecular systems with physical interaction between particles by solving the Newton's equations of motion. The interaction is characterized by the interatomic potential energy, which is a function of atomic positions. The physical forces between atoms are computed as the gradient of the potential, which leads to the changes of atoms' momenta and subsequently positions. For a system with N atoms in a 3 dimensional spatial domain with Cartesian coordinates, its $6N$ dimensional phase space consists of $3N$ values of positions and $3N$ values of momenta. The state of the system is a point in the phase space, and the MD simulation is to numerically compute a trajectory in the phase space.

At each time step, the interatomic potential values are updated. With the calculated forces, the accelerations of the atoms are computed too. Based on the numerical finite-difference version of the Newton's equations (e.g. the well-known verlet algorithm), the momenta and positions of particles are updated. To reduce the numerical truncation error for solving the ordinary differential equations, higher-order multi-step algorithms (at the cost of memory) have also been developed.

The conventional MD was formulated for systems of the microcanonical (NVE) ensemble with constant volume and energy, which however does not conform with most of experimental measures. Therefore, MD formulations for other systems have been proposed. For instance, the constrained dynamics [18] with kinetic energy constraints, Nosé-Hoover dynamics [46, 29, 42] based on an

enhanced Hamiltonian, and canonical samplings with velocity scaling [7, 10] were developed for the canonical (NVT) ensemble. The Anderson [1] and Parrinello-Rahman [47] formulations were for the isoenthalpic-isobaric (NPH) ensemble. Additionally, constrained MD with algorithms such as SHAKE [51] and RATTLE [2] were developed to improve the efficiency of simulating large molecules with fixed geometric variables (e.g. bond lengths and angles) to decouple vibrations from translations and rotations.

MD simulation is designed to describe molecules' behaviors at time scales from femto seconds to pico seconds. It is not economically realistic in simulating processes that are over micro seconds or longer. Because most computational time is spent on thermal vibrations of atoms.

1.3.4 Kinetic Monte Carlo

In contrast to MD, the atomic scale KMC [11] is much more efficient in simulating the infrequent transition or reaction processes with longer times than thermal vibrations. KMC does not simulate a system based on its continuous evolution along time as in MD. Instead, it defines a discrete set of states of the system (i.e. all possible configurations). It simulates state transitions between states which are triggered by events (also called processes in chemistry-oriented literature) that cause state changes. In physical processes, each 'snapshot' of the system at a time where atoms are located in space is a state of the KMC model. In chemical reactions, the numbers of different species naturally form the states. Reactions themselves are the events.

During the KMC simulation, events are randomly selected to occur sequentially based on their respective probabilities of occurrence, which follow exponential distributions in general. When an event occurs (or it is usually called the event is fired), the system's state is updated according to the nature of the event. At the same time, the system clock that keeps track of the current time of the system needs to be advanced by a certain period, which is also randomly generated based on probabilistic principles. The system clock provides information of how long the processes last. This discrete-event approach allows us to simulate much larger atomic systems for much longer times than those MD can provide.

The possible states and transitions/reactions with numerical rates are the input parameters of KMC models. Empirically they may be found by physical experiments. From the first principles viewpoint, all of them can be found by computing PES.

1.3.5 MD Applications in Nanomanufacturing

In the domain of nanomanufacturing, MD simulation has been used to study machining in the past two decades [38]. More recently, it was used to simulate nano-lithography [19, 32] and cutting with diamond tools [36, 37, 69, 25]. For other processes, MD has been used to simulate laser ablation of bulk materials [56], laser-based sintering [66], and nano-indentation [54]. However, the major issue of MD is its short time scale that is not compatible with those in nanomanufacturing. Most of the computational time in MD is spent on thermal vibration of atoms, whereas our interested processes are usually longer than microseconds. Those events of interest with longer time scale than thermal behaviors are rare events. MD is very inefficient in simulating these rare events. Unrealistic assumptions have to be made to accommodate the time scale. For instance, unrealistically high cutting speed (e.g. 100 m/second) was used in those MD simulation studies, which only simulate pico seconds of the processes. With such speed and time duration, the accuracy of predictions (e.g. required forces and acceptable defects) is affected.

To simulate the rare events of transitions or reactions, several improvements of MD have been proposed to bridge the gap of time scale and accelerate the simulation speed of rare events, such as by running multiple trajectories [64], introducing bias potentials [63], or increasing temperatures [55]. However, the inherent inefficiency of MD is that computational time is spent on trajectory prediction, which is not important for rare event simulations.

1.3.6 KMC Applications in Nanomanufacturing

In the domain of nanomanufacturing, KMC simulation has been used in some of the bottom-up self-assembly processes. For instance, in PVD processes, it is used to simulate deposition and diffusion events in sputtering (e.g. [70, 30]). In addition, there is a question on how to incorporate extra information of kinetic energy in particles to make correct predictions, such as reflection, resputtering, latent heat, kinetic energy induced diffusion, and others [71, 65, 59]. Approximations based on molecular dynamics and the embedded atom method (EAM) are applied. In CVD processes, KMC is used to simulate on-site chemical reaction events of film growth (e.g. [4, 3, 24, 33]) and etching (e.g. [22, 45, 74]). However, there is no KMC mechanism available to simulate top-down manufacturing processes. In Section 1.5, we will demonstrate a new and generic KMC approach, controlled KMC, for both top-down and bottom-up processes.

The input parameters for KMC simulation are the stable states and the transition rates associated with transitions between stable states. The transition rates are determined by the energy difference between stable states and transition states. The search of stable and transitions states to calculate the inputs for KMC is introduced in the following section.

1.4 Searching Stable and Transition States

The first important element of the proposed first principles simulation framework to predict physical or chemical processes is to search for stable or metastable states, which correspond to the local minima on PES. The stable or metastable states are the possible system configurations of atoms that exist in the physical world. The second important element of the first principles simulation is the transition state search, which correspond to the saddle points on PES. The energy differences between saddle points and local mimima indicate the rates of transitions between stable states.

As illustrated in Figure 1.3, PES is a high-dimensional hyper-surface, or simply speaking a function with system configuration values (e.g. coordinates of all atoms) as the inputs and the system's total potential energy level as the output. The configurations with local minimum energies on PES correspond to the states that are stable or metastable, shown as current state, state i, and state j in Figure 1.3. A transition/reaction rate is determined by the energy barrier between two stable states on PES according to the transition state theory [40]. In Figure 1.3, the activation energy ΔE_j for the transition from current state to state j is the energy difference between the saddle point and current state along the transition path. That is, transition rate k_j from current state to state j is proportional to $\exp(-\Delta E_j/(k_B T))$, i.e. $k_j = \nu_j \exp(-\Delta E_j/(k_B T))$, where k_B is the Boltzmann constant, T is the temperature, and ν_j is the pre-exponential factor that can be calculated from the vibrational frequencies.

Different numerical algorithms for optimization can be applied, including both local and global searching methods, to locate stable states. The most used searching algorithms are gradient-based steepest descent and quasi-Newton methods. Some global searching algorithms have also been applied in searching stable structures on PES. Yet, these methods only search one stable structure with the local minimum energy staring from the initial guess or location. For real-world applications, the potential energy landscape as the PES is very rough and complex where many local minima exist. Therefore the relaxation result sensitively depends on the initial guess. To receive a more

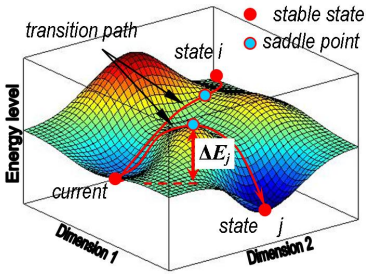


FIGURE 1.3: A illustration of stable and transition states on PES.

comprehensive view of the possible meta-stable structures at nearby locations, it is more desirable to have search algorithms to find multiple local minima from one initial guess. In Section 1.4.1, two new particle swarm optimization search algorithms, which can search multiple local minima simultaneously, will be presented.

Various numerical methods to search transition paths and saddle points have been developed (see a comprehensive list of references in [14]). They can be categorized into two types, transition path search and saddle point search. Transition path search methods generate the minimum energy path (MEP) on the PES, whereas saddle point search methods aim at finding the saddle points only. The transition path search methods include the chain-of-states methods where a small set of discrete states (also called images) form a 'chain' to approximate the MEP. The search is done by adjusting positions of the images iteratively. The nudged elastic band and string methods are the representative chain-of-states methods. Other ones include conjugate peak refinement and accelerated Langevin dynamics. The saddle point search methods include the local search approaches where the search is done by walking through the PES and identify saddles based on curvature information. Some global search approaches also allow for jumps over local stationary points. The existing searching algorithms only search one transition path or one saddle point at a time. In Section 1.4.2, a new transition path search algorithm will be presented, where multiple transition paths can be found simultaneously on a complex PES. The new search algorithm also locate minima and saddle points altogether, which is also different from existing approaches where at least one stable state with local minimum energy should be known in advance.

1.4.1 Two New Particle Swarm Optimization (PSO) Algorithms

The PSO method is inspired by the movement of a group of insects or birds which fly collectively as a group and also learn from each other about potential sources of food. Using this type of behavior, the swarm can search a large area and at the same time relay useful information to all the individuals in the group. An optimization algorithm based on this principle is useful for finding the global minimum or maximum of a numerical function on a continuous domain. A collection of particles are used to search optima simultaneously. They exchange information of optimality during the search and update their respective positions accordingly. Compared to other global optimization techniques, it has a higher degree of time efficiency, nevertheless at the cost of more memory. By adding more individuals to the swarm and increasing the size of the initial search space, the search can be very comprehensive and converge to a global optimum.

In the classical PSO algorithm [35], the velocity of the j^{th} particle at the i^{th} iteration is

$$\mathbf{v}_i^{(j)} = \theta \mathbf{v}_{i-1}^{(j)} + c_1 r_{1,i}(\mathbf{p}^{(j)} - \mathbf{x}_{i-1}^{(j)}) + c_2 r_{2,i}(\mathbf{o} - \mathbf{x}_{i-1}^{(j)}) \tag{1.1}$$

where $j = 1, 2, ..., N$ for a total of N particles, $\mathbf{x}_{i-1}^{(j)}$ is the location of the j^{th} particle at the $(i-1)^{th}$ iteration, $\mathbf{p}^{(j)}$ is the best location (with either minimum or maximum functional value) found so far for the j^{th} particle, $\mathbf{o}$ is the best location found so far among the neighbors, typically the population of all N particles, $r_{1,i}$ and $r_{2,i}$ are random numbers between 0 and 1, c_1 and c_2 are weight coefficients associated with the difference between the current location of the particle and the best it found so far and the difference between the current location and the global best, and similarly θ is the weight assigned to the previous velocity of the particle. The location of the j^{th} particle at the i^{th} iteration is updated by $\mathbf{x}_i^{(j)} = \mathbf{x}_{i-1}^{(j)} + \mathbf{v}_i^{(j)}$. As search continues, the weights could be adjusted such that the global optimum is located.

The two new PSO algorithms [15] presented here are called random group PSO and active group PSO, where particles search in groups and groups are dispersed so that multiple local minima can be found.

Random Group PSO

The random group PSO (RG-PSO) divides the initial population of particles into K groups with the respective sizes of n_k's ($k = 1, \ldots, K$). Groups are always kept away from each other so that the optima found by the groups are different from each other. To keep track of the positions of groups, each group keeps track of its position, as the average position $\overline{\mathbf{x}}_i^{<k>} = \sum_{j_k=1}^{n_k} \mathbf{x}_i^{(j_k)}/n_k$ for group k at the i^{th} iteration. The velocity and position updates proceed similar to the standard PSO, except that only particles from the same group are considered as neighbors. That is, $\mathbf{o}$ in Eq.(1.1) keeps track of the best in the group here.

A Euclidean distance calculation is made to find the distance between groups. A *criticality* value A is assigned to each group, initially assigned as zero. If the average position of particles in any group moves within a threshold distance of the average position of any other group, the criticality value for one of the groups is increased by one. If any group's criticality reaches a predetermined value A_{max} or higher, all of its particles are relocated randomly in the search space. This relocation scheme allows only one group to converge to a particular location in the search space, regardless of whether the location is a local or global minimum. At the same time, given a sufficient number of iterations, the remaining groups have a chance to converge to another minimum. The algorithm increases the probability of finding the global optimum by avoiding premature convergence to local optima, and more importantly provides more information about the energy landscape as different groups converge to various local optima.

The example in Fig.1.4-(a) is used to illustrate RG-PSO. With the initial guess shown in Fig.1.4-(b), the classical PSO algorithm gives the single optimum result in Fig.1.4-(c), whereas RG-PSO gives multiple results of local optima simultaneously as in Fig.1.4-(d).

Active Group PSO

Rather than re-distributing particles randomly when two groups get too close to each other as in RG-PSO, in active group PSO (AG-PSO), each group actively moves away from its m nearest neighboring groups. The velocity is updated by

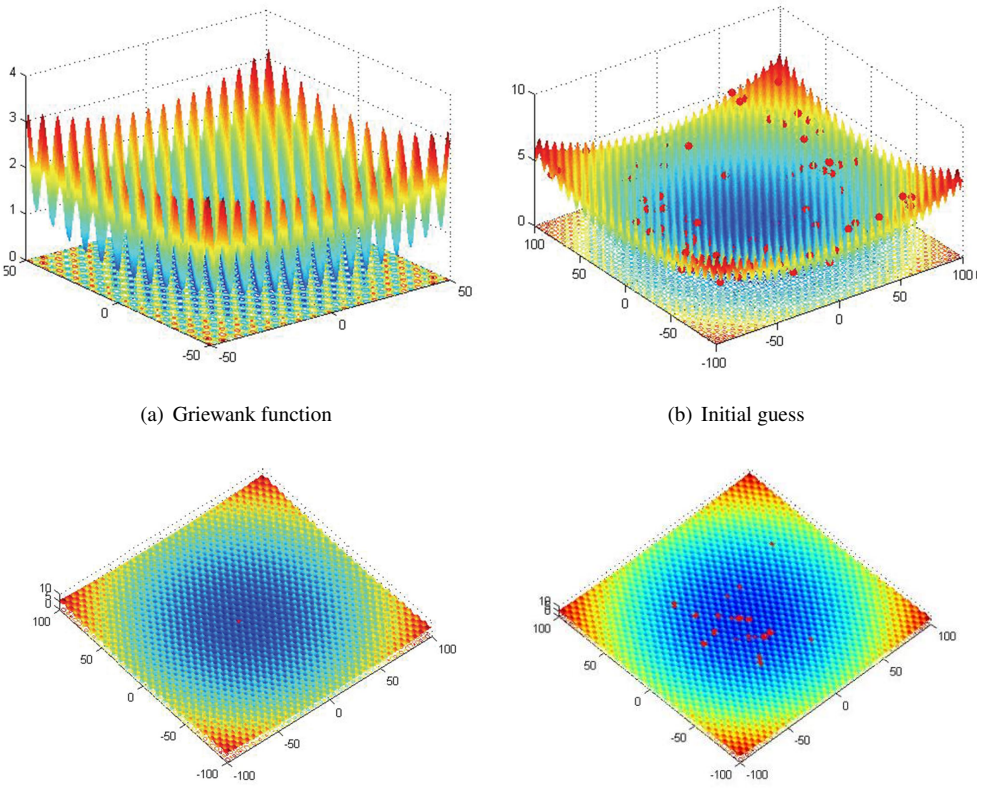
$$\mathbf{v}_i^{(j)} = \theta \mathbf{v}_{i-1}^{(j)} + c_1 r_{1,i}(\mathbf{p}^{(j)} - \mathbf{x}_{i-1}^{(j)}) + c_2 r_{2,i}(\mathbf{o} - \mathbf{x}_{i-1}^{(j)}) - c_3 \sum_{k=1}^{m} r_{k+2,i}(\mathbf{o}^{<k>} - \mathbf{x}_{i-1}^{(j)}) \qquad (1.2)$$

where $\mathbf{o}^{<k>}$ is the best location of the nearest neighboring group k, and $r_{k,i}$'s ($k = 1, ..., m+2$) are random numbers.

Similar to RG-PSO, AG-PSO improves the probability of finding a global optimum when compared to standard PSO and also provides more information about the landscape. The intent of AG-PSO is to avoid wasteful iterations where two or more groups move towards each other before being randomly re-distributed. By actively moving away from each other, the groups are less likely to stay in the same area for multiple iterations. The algorithm should become more efficient as it no longer has to wait to reach the threshold distance and maximum criticality before exploring a different region. By adjusting the value of c_3, AG-PSO can also be more effective at exploring regions that are densely populated with local optima. As two groups approach an area with multiple local optima, they may search within threshold distance of each other if the optima are sufficiently close. Rather than being re-distributed as with RG-PSO, the two groups may converge to different optima within a small region. A similarly thorough search with RG-PSO would require a smaller threshold distance, which makes the algorithm less efficient.

1.4.2 Searching Multiple Transition States

A new concurrent multi-transition path algorithm has been developed [26], which locates both the minimum energy positions and MEP simultaneously. Therefore the precise positions of the stable



(a) Griewank function (b) Initial guess

(c) Result from the classical PSO (d) Result from RG-PSO

FIGURE 1.4: Searching minima of Griewank function $f(x) = (x_1^2 + x_2^2)/4000 - \cos x_1 \cos(x_2/\sqrt{2}) + 1$ by PSO.

states do not need to be known in advance, as in the existing transition path search algorithms. In addition, the new algorithm is able to search multiple transition paths simultaneously, which gives us a comprehensive view of the energy landscape.

Similar to the chain-of-states methods, the transition path is represented by a series of discrete states in the new search algorithm. These discrete states or images are treated as the control points of a Bézier curve, and the MEP is the curve itself. The searching process starts with an initial guess of MEP by a Bézier curve where the two end-point images can be at any states. At the first stage of search, the minimization procedure is applied to the two end-point images respectively so that they move to the nearest stable states. For each of the other intermediate control images, a set of conjugate directions are constructed. Along one of the conjugate directions for one intermediate control image, several steps of line maximization are applied. Then several steps of line minimization along each of the other conjugate directions are applied. This max-min procedure is based on the fact that a first-order saddle point on an n-dimensional PES corresponds to the maximum along one direction and the minimum along the other $n-1$ directions.

In the chain-of-states methods, all discrete states need to be kept apart properly so that they form a chain to approximate the MEP. In the nudged elastic band method [28], a spring is attached between each pair of images. The spring forces thus keep the images apart. In the string method [17], images are discrete evaluations of continuous splines. A reparameterization scheme is used to redefine images so that they are relatively well distributed during the searching process. In our

method, a Bézier curve degree elevation and reduction approach is developed to redistribute the control images very efficiently. Degree elevation for a Bézier curve introduces an extra control point without changing the shape of the curve. Degree reduction removes one control point with the shape of the curve approximated by a new curve with one degree less than the original one.

Once the two end-point images successfully locate two local minima, a curve subdivision scheme is applied to break the curve into two if there is at least an extra local minimum along the curve in addition to the two end-point images. That is, one intermediate control image that is close to the new minimum will now be treated as an end-point image shared by the newly generated two curve segments. The minimization procedure is then applied to this new end-point image to locate the extra minimum position. After the new end-point image is located at the new minimum, each of the two new curve segments will be checked again. This procedure will continue recursively until there is no guarantee that a local minimum will exist between any two end-point images along the respective curve segment. Subdivision rules are defined such that the curve can be effectively divided if there is an extra local minimum along a transition path. More details will be described as follows.

Searching Stable Configurations

At the first stage of the search, the two end-point images move toward the local minima near their respective initial positions. Different local search methods can be used. We use the conjugate gradient (CG) approach to search. The CG approach is efficient for simple objective functions in unconstrained optimization such as those with the quadratic form, where the number of iterations can be predetermined and is the same as the dimension of the search domain.

In our algorithm, the classical Fletcher-Reeves method [21] is employed in calculating the conjugate search directions. That is, for an potential function $f : \mathbb{R}^n \to \mathbb{R}$, the search direction at the starting position $\mathbf{x}_0$ for the first iteration is calculated as $\mathbf{v}_1 = -\nabla f(\mathbf{x}_0)$. Then the position is updated as $\mathbf{x}_1 = \mathbf{x}_0 + \alpha_1 \mathbf{v}_1$. The directions for the rest of $n-1$ iterations are $\mathbf{v}_i = -\nabla f(\mathbf{x}_{i-1}) + (\nabla^T f(\mathbf{x}_{i-1}) \cdot \nabla f(\mathbf{x}_{i-1}))/(\nabla^T f(\mathbf{x}_{i-2}) \cdot \nabla f(\mathbf{x}_{i-2}))\mathbf{v}_{i-1}$ for $i = 2, \ldots, n$. The positions are updated by $\mathbf{x}_i = \mathbf{x}_{i-1} + \alpha_i \mathbf{v}_i$ $(i = 2, \ldots, n)$. In the classical CG approach, the step sizes α_i's are typically determined by the line search so that the function can be minimized along the search direction $\mathbf{v}_i$. Here, we do not find the actual minimum along the line. Because the potential function is not quadratic in general. Most likely, the minimization procedure will go beyond n CG steps for an n-dimensional problem. Therefore, we use a fixed number of mini-steps in the line search. The best step size so far is then chosen as α_i.

After n steps of CG search for minimum for two end-point images, the positions of the intermediate control images will also be updated to approximate the MEP, as described in the next subsection. After both end-point and intermediate control images are updated, this finishes one grand iteration of the searching process. For the next grand iteration, the end-point images will be minimized and intermediate control images will be updated again for the MEP. The grand iterations will finish when the two ends find the minima.

Searching the MEP

The positions of the intermediate control images are also updated by approximated conjugate directions. Different from the end-point images, each intermediate image is maximized along one CG direction and minimized along the other $n-1$ CG directions. The CG directions are estimated similar to Beale's method [5] with the restart direction augmentation.

For the k^{th} intermediate control image out of a total of N images $(k = 2, \ldots, N-1)$, its maximization direction is set to be the finite difference approximation of the tangent direction as $\mathbf{s}_1^{(k)} = \mathbf{x}_0^{(k)} - \mathbf{x}_0^{(k-1)}$ if $k \leq \lceil N/2 \rceil$ and $\mathbf{s}_1^{(k)} = \mathbf{x}_0^{(k)} - \mathbf{x}_0^{(k+1)}$ if $k > \lceil N/2 \rceil$. The position of the k^{th} intermediate control image is then updated as $\mathbf{x}_1^{(k)} = \mathbf{x}_0^{(k)} + \alpha_1 \mathbf{s}_1^{(k)}$, where α_1 is the step size that is determined

by a maximization line search procedure. Similar to the previous subsection of searching stable configurations, we only apply a fixed number of mini-steps instead of finding the actual maximum in order to save some computational time.

The first minimization direction is calculated as

$$\mathbf{s}_2^{(k)} = -\mathbf{g}_1^{(k)} + (\mathbf{g}_1^{(k)T} \cdot (\mathbf{g}_1^{(k)} - \mathbf{g}_0^{(k)}))/(\mathbf{s}_1^{(k)T} \cdot (\mathbf{g}_1^{(k)} - \mathbf{g}_0^{(k)}))\mathbf{s}_1^{(k)} \tag{1.3}$$

where $\mathbf{g}_0^{(k)} = \nabla f((\mathbf{x}_0^{(k-1)} + \mathbf{x}_0^{(k)})/2)$ if $k \leq \lceil N/2 \rceil$ and $\mathbf{g}_0^{(k)} = \nabla f((\mathbf{x}_0^{(k)} + \mathbf{x}_0^{(k+1)})/2)$ if $k > \lceil N/2 \rceil$, and $\mathbf{g}_1^{(k)} = \nabla f(\mathbf{x}_1^{(k)})$. Similarly, the control image is updated as $\mathbf{x}_2^{(k)} = \mathbf{x}_1^{(k)} + \alpha_2 \mathbf{s}_2^{(k)}$ with a fixed number of mini-steps of line minimization to decide the step size α_2.

The rest of the minimization directions are calculated differently with a further augmentation of the direction in the intermediate previous step, as

$$\mathbf{s}_i^{(k)} = -\mathbf{g}_i^{(k)} + (\mathbf{g}_i^{(k)T} \cdot (\mathbf{g}_1^{(k)} - \mathbf{g}_0^{(k)}))/(\mathbf{s}_1^{(k)T} \cdot (\mathbf{g}_1^{(k)} - \mathbf{g}_0^{(k)}))\mathbf{s}_1^{(k)} + (\mathbf{g}_i^{(k)T} \cdot \mathbf{g}_i^{(k)})/(\mathbf{s}_{i-1}^{(k)T} \cdot \mathbf{g}_{i-1}^{(k)})\mathbf{s}_{i-1}^{(k)} \tag{1.4}$$

with $\mathbf{g}_i^{(k)} = \nabla f(\mathbf{x}_i^{(k)})$ and updated position $\mathbf{x}_i^{(k)} = \mathbf{x}_{i-1}^{(k)} + \alpha_i \mathbf{s}_i^{(k)}$ $(i = 3, \ldots, n)$.

The uniqueness of our MEP search method is that the conjugate directions are used in searching saddle points, with specially defined maximization and minimization directions where the location information of neighboring images is utilized to improve the numerical stability.

Constrained Degree Elevation and Reduction

A new approach of degree elevation and reduction is developed to re-distribute the control images during the MEP search. This approach is very efficient in adjusting the positions of control images so that they are evenly distributed and approximate the MEP well. This is done by applying a step of degree elevation, followed by a step of degree reduction. The combination of the two does not change the number of control points. This procedure can also avoid the possible generation of loops or self-intersections of the Bézier curve. Additionally, if we would like to dynamically increase the number of control points to improve the resolution of MEP approximation, single or multiple steps of degree elevation can be applied. Similarly, single or multiple steps of degree reduction can reduce the number of control points. Therefore, the proposed degree elevation and reduction scheme can also provide the dynamic resolution management during the search.

It is well known that the degree elevation of a Bézier curve can replace the existing curve by another curve with one more control point, while the shape of the curve remains unaltered. For a curve with N control points $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(N)}$, one degree elevation step produces a curve with $N+1$ control points $\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(N+1)}$. The new control points are calculated by [20] $\mathbf{y}^{(1)} = \mathbf{x}^{(1)}$, $\mathbf{y}^{(N+1)} = \mathbf{x}^{(N)}$, and $\mathbf{y}^{(k)} = (k/N)\mathbf{x}^{(k-1)} + (1 - k/N)\mathbf{x}^{(k)}$ for $k = 2, \ldots, N$.

Different from degree elevation, there is no degree reduction scheme that can reduce the number of control points while maintaining the same shape of curve. Approximation is inevitable. In our degree reduction scheme, the new control points $\mathbf{x}^{(k)}$'s in the reduced curve are calculated from the original ones $\mathbf{y}^{(k)}$'s by

$$\begin{aligned} &\mathbf{x}^{(1)} = \mathbf{y}^{(1)},\ \mathbf{x}^{(N)} = \mathbf{y}^{(N+1)}, \\ &\mathbf{x}^{(k)} = (1 - \lambda^{(k)})\mathbf{z}_F^{(k)} + \lambda^{(k)}\mathbf{z}_B^{(k)}\ (k = 2, \ldots, N-1) \end{aligned} \tag{1.5}$$

where the forward intermediate points $\mathbf{z}_F^{(k)}$'s are calculated recursively by

$$\begin{aligned} &\mathbf{z}_F^{(1)} = \mathbf{y}^{(1)} \\ &\mathbf{z}_F^{(k)} = \frac{N\mathbf{y}^{(k+2)} + (N-k)\mathbf{y}^{(k+1)} + (N-k)\mathbf{y}^{(k)} - k\mathbf{z}_F^{(k-1)}}{3N - 3k}\ (k = 2, \ldots, N-1) \end{aligned} \tag{1.6}$$

and the backward intermediate points $\mathbf{z}_B^{(k)}$'s are calculated recursively by

$$\begin{aligned}\mathbf{z}_B^{(N)} &= \mathbf{y}^{(N+1)} \\ \mathbf{z}_B^{(k)} &= \frac{N\mathbf{y}^{(k-1)} + (k+1)\mathbf{y}^{(k)} + (k+1)\mathbf{y}^{(k+1)} - (N-k-1)\mathbf{z}_B^{(k+1)}}{3k+3} \; (k = N-1,\ldots,2)\end{aligned} \tag{1.7}$$

The weights are calculated as $\lambda^{(k)} = k/(N-1)$ for $k = 1,\ldots,N-2$. Compared to other Bézier curve degree reduction methods, this formulation simplifies the computation, because we do not require the best approximation of the original curve by the new one. At the same time, the locations of neighboring images are considered in the forward-backward averaging in order to eliminate the undesired loops or self-intersection.

Curve Subdivision for Multiple Pathway Search

A curve subdivision scheme is devised so that the single Bézier curve is divided into two if there is an extra basin of local minimum between the two end-point images along the path. Rules are designed to check whether extra basins exist for the curves with either five or six control points. They are based on the directions of the gradients for those control points, as the indicators of slopes. By comparing the angles between the gradient directions at the locations of control points along the path, we can check if the curve moves up-hill and down-hill on the PES more than once.

If an extra basin of local minimum exists, the control point that is closest to the basin is selected as the breaking point. The breaking point is treated as a new end-point image instead of an intermediate one from now on. The original curve is subdivided into two with the new end point shared by the two new curves. Degree elevation is applied to both curves to introduce extra control points to maintain the necessary resolution. The minimization procedure is applied to the new end point to locate the actual minimum, whereas the intermediate control points for both curves are subject to a new search iteration for their respective MEPs. This break and search process continues until no extra basin of minimum is guaranteed to exist.

Once the MEP search is finished for each one of the curve segments, a climbing procedure is taken for each curve segment, where the intermediate control point that corresponds to the largest potential energy for each curve segment climbs up-hill to locate the actual saddle point.

In summary, with the strategies of local minimum search, constrained degree elevation and reduction, and curve subdivision described as above, the developed concurrent searching algorithm can simultaneously search the multiple local minima and saddle points. An example of search result for a 2-D Rastrigin function $f(x_1,x_2) = 20 + x_1^2 - 10\cos(2\pi x_1) + x_2^2 - 10\cos(2\pi x_2)$ is used to illustrate visually, as shown in Figure 1.5. In general, PES is a hyper-surface in a high-dimensional configuration space.

With the information of a stable state and a saddle point, activation energy, which is the energy difference between the two on the PES, can be used to estimate the transition rate. The transition rate determines how often an event of state transition occurs in KMC simulation. Here, a generalized KMC mechanism is introduced to simulate both bottom-up and top-down nanomanufacturing processes.

1.5 Controlled Kinetic Monte Carlo Simulation

In the bottom-up processes, the simulated events mainly include thermally or chemically induced atomic rearrangements such as diffusion, absorption, desorption, surface reaction and abstraction. These rearrangements occur spontaneously in the form of self-assembly. Those spontaneous events

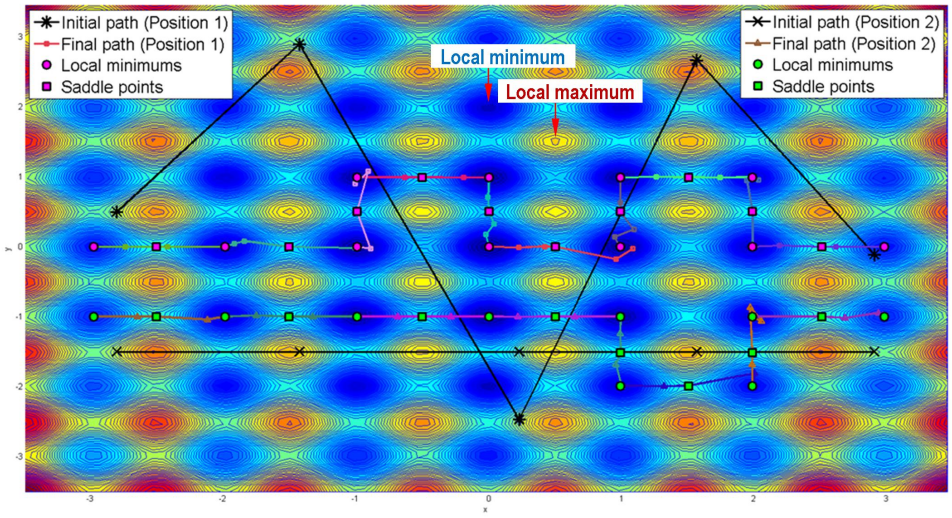


FIGURE 1.5: An example result of the new concurrent searching algorithm that can find multiple local minima and MEPs simultaneously.

are called *self-assembly events*. In the top-down processes, the atomic rearrangement is triggered by some external energy sources such as force or electromagnetic field. The rearrangement is further induced by self-assembly events. These external events could be scanning probe tip interaction with samples, bombardment of high-energy particles, molding materials with attractive and repulsive forces, and others. These external events are scheduled to occur at certain locations, toward certain directions, or at particular times to control the overall process. Therefore, they are called *controlled events*. The major new concept in cKMC is the introduction of controlled events in the algorithm in order to simulate top-down processes in parallel with bottom-up processes.

The main parameters for both self-assembly and controlled events are transition rates, which indicate how often the events occur. Additional parameters are introduced for controlled events, including when and where the events may occur. The time when a controlled event occurs could be deterministic. For example, in the AFM nanolithography, the time when the cutting tip reaches one particular position along the predetermined path is fixed. Furthermore, the spatial location where a controlled event occurs could also be predetermined. For example, in kinetic energy induced diffusion (e.g. ions or electrons in focused beam lithography under electromagnetic field), the preferred directions that particles move toward are fixed. Therefore in the cKMC simulation, there is a time, a direction vector, or both, associated with a controlled event. As the simulation clock is advanced based on self-assembly events, controlled events are checked constantly so that those with the predetermined time earlier than the current time should be fired. If a predetermined direction is associated with a controlled event and the reaction or transition direction is different, then the event is aborted. More details will be described in the following sections.

Notice that the particles in cKMC are not necessarily atoms only. They could be electrons, photons, monomers, or molecules, as long as the internal behavior are not of interest and internal structure is assumed to be at an equilibrium state. It also should be noted that controlled events are not only necessary to simulate top-down processes, but also useful for some bottom-up processes such as controlled growth of carbon nanotube with directed orientations, which so far still largely depends on MD simulation.

1.5.1 An Illustration of cKMC with Scanning Probe Lithography

Here the scanning probe lithography process is used to illustrate the basic idea of cKMC. The details of cKMC can be found in [67]. For the model in Figure 1.6, besides the regular workpiece species, there is a controlled species, even though chemically they could be identical. Controlled species are species where the times of associated events are deterministic. That is, the events will be fired at pre-scheduled times. In this case, it is based on the moving speed of the probe. The locations of controlled species are also pre-determined based on the planned scanning or cutting path. There is also a vaporized workpiece species which escapes the solid body of workpiece in regular diffusion. In the controlled reaction of the controlled species, atoms of the controlled species are converted to activated control species based on the scheduled times sequentially. Here, the probe is moving along the path denoted by the sites of controlled species along the x-axis direction from left to right. The atoms of controlled species on the left have earlier time of conversion than those on the right. The activated control species has a higher kinetic energy than the vaporized workpiece species. Therefore, they are associated with controlled diffusion events. The controlled diffusion events have several restrictions, such as the controlled directions (illustrated by the dotted arrows in Figure 1.6) by which activated atoms can diffuse to neighboring sites or by which they can interact with neighboring atoms of workpiece species. In addition, vacancy exists between occupied spatial lattices. Therefore, a vacancy species can be defined. Finally, there is an absorbent species at the boundary of simulated domain, where the lifecycle of the atoms of vaporized workpiece species and controlled species are terminated.

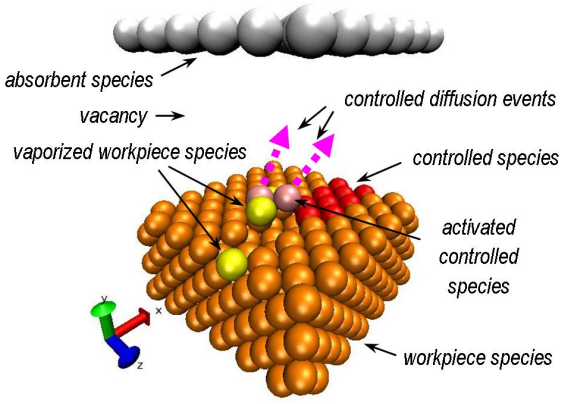


FIGURE 1.6: Illustration of events in scanning probe lithography.

In the simplified scanning probe lithography process in Figure 1.6, several example events are listed in Table 1.1. Events are organized into several categories based on the number of neighboring sites involved. Events are denoted by reaction-like equations. The left of equations are reactants and the right are products. The first reactant is the owner of the event, meaning that the event is associated with the site where the first reactant initially resides. After the reaction, the reactant species are replaced by the product species at the corresponding sites. Among the events in Table 1.1, reaction R1 has only one site involved. It is a controlled event when the controlled species is converted to the activated controlled species based on the scheduled time along the cutting path. R2 to R6 have two neighboring sites involved. R2 is the controlled diffusion event associated with activated controlled species. It models the effect of kinetic energy associated with particles transferred from the moving probe. A particle of activated controlled species can diffuse to a vacant neighboring site if the direction from the current site to the vacant site is aligned with the controlled

TABLE 1.1: Example events in scanning probe lithography

Number of Sites Involved	Reaction/Transition Event
1	R1: *controlled_species* → *activated_controlled_species* (controlled)
2	R2: *activated_controlled_species* + *vacancy* → *vacancy* + *activated_controlled_species* (controlled) R3: *vaporized_workpiece_species* + *vacancy* → *vacancy* + *vaporized_workpiece_species* R4: *workpiece_species* + *vacancy* → *vacancy* + *workpiece_species* R5: *vaporized_workpiece_species* + *absorbent* → *vacancy* + *absorbent* R6: *activated_controlled_species* + *absorbent* → *vacancy* + *absorbent*
3	R7: *workpiece_species* + *workpiece_species* + *vacancy* → *vacancy* + *workpiece_species* + *workpiece_species* R8: *vaporized_workpiece_species* + *workpiece_species* + *workpiece_species* → *workpiece_species* + *workpiece_species* + *workpiece_species*
4	R9: *activated_controlled_species* + *workpiece_species* + *vacancy* + *vacancy* → *vacancy* + *workpiece_species* + *vacancy* + *vaporized_workpiece_species* R10: *activated_controlled_species* + *workpiece_species* + *workpiece_species* + *workpiece_species* → *workpiece_species* + *workpiece_species* + *workpiece_species* + *workpiece_species*

direction. In contrast, R3 and R4 are regular diffusion events without directional constraints. R5 and R6 absorb particles at the boundary of simulation. Among the reactions with three reactants, R7 is a coordinated diffusion event where three sites are involved. It has been revealed by first principles simulation that coordinated diffusion with multiple sites involved could have lower energy barriers than the traditional single-hop diffusion (e.g. in R4) as originally thought [57]. R8 is an adsorption event. For reactions with four reactants, R9 is an example interaction between the activated controlled species and the workpiece species where kinetic energy is transferred. R10 is another absorption event where kinetic energy of activated controlled species is absorbed and atoms settle down. It should be noted that Table 1.1 only lists some examples. More events can be introduced in a similar format, which allows for necessary refinement and extension.

The controlled events can be generally used in various complex top-down fabrication processes. In MD simulation, special treatments of potential energy between atoms of workpiece and tool are needed so that the rapid movement of atoms can be modeled. Here, controlled events are used to model the special movement of atoms under directional constraints.

1.5.2 Detailed cKMC Algorithm

The key components of cKMC are the controlled species and controlled events. In the implementation of cKMC in SPPARKS [39], controlled species can be specified by an input script "***control_species*** *reactant rate product x y z neighbor_lower neighbor_upper init_time*" where ***control_species*** is the command, *reactant* is the name of controlled species, *product* is the resultant species in the controlled reaction, *rate* is the numerical value of reaction rate, and x, y and z specify the direction along which the reaction occur. That is, the reactants of controlled species react and convert to products sequentially along the specified direction at the rate. The next two integer values, *neighbor_lower* and *neighbor_upper* specify the lower and upper bounds of neighboring sites

where the controlled species may react. If the number of its neighbors is not within the bound, the controlled reaction is not initiated. This provides more controls on where controlled reactions can occur. Finally *init_time* specifies when the controlled reaction is started.

Similarly events can be specified by the input script "***event*** $reactant_1$ *(*$reactant_2$*, etc.)* *rate* $product_1$ *(*$product_2$*, etc.)*" where ***event*** is the command, *rate* is the numerical value of reaction rate, and corresponding reactants and products at the particular site are defined. For instance, after the event is fired, $reactant_1$ is replaced by $product_1$, $reactant_2$ by $product_2$, and so on.

Controlled events are specified in two types. The first type is by the direction along which the reaction is going toward, and the second one is by the target location where the reaction is targeted at. The input script is "***control_event*** *index dx dy dz* x_0 y_0 z_0 *theta neighbor_lower neighbor_upper*" where ***control_event*** is the command, *index* refers to the event index previously specified by the command event, *dx*, *dy* and *dz* specify the direction which the first and second reactants should aligned with, and *theta* is an angular allowance such that the direction formed by the first and second reactants can be within the range of $+/-$ *theta* and the reaction still occurs. In other words, the controlled events provide directional selections of events. Controlled events are only fired if the direction formed by the first two reactants is the specified direction or close enough within the $+/-$ *theta* range. Before a controlled event is inserted into the event list, the directional criterion is checked. Those controlled events that do not satisfy the constraints are disregarded. When $dx = dy = dz = 0$, x_0, y_0 and z_0 take effect and they specify the target location by which the direction of reaction is from the current site to the target location. Two cases can be specified. When theta is positive, the reaction direction "converges" toward the target location. When theta is negative, the reaction direction "diverges" away from the target location. Similarly, if the direction formed by the first two reactants is far from the specified, events are discarded. Again, *neighbor_lower* and *neighbor_upper* specify the lower and upper bounds of neighboring sites where the controlled event may occur. Therefore, rather than directly modeling kinetic energy, controlled events provide a directional selection for cKMC.

Before a simulation starts, all species, sites, and events are specified by script commands. Internally species, controlled species, events, controlled events, reaction sites and others are stored in array-type data structures. For each site where a controlled species resides, there is an associated *keyvalue*, which usually is the time when the controlled reaction occurs at site j of controlled species i. It is calculated as $t_{ij} = t_i^{(0)} + (\mathbf{r}_{ij} \cdot \mathbf{s}_i - d_i)/a_i$ where $t_i^{(0)}$ is the reaction start time for controlled species i as previously specified by *init_time*, $\mathbf{r}_{ij}$ is the 3D coordinate of site j of i^{th} controlled species, $\mathbf{s}_i$ is the direction of controlled reaction for controlled species i, d_i is calculated as $d_i = \min_i(\mathbf{r}_{ij} \cdot \mathbf{s}_i)$, and a_i is the reaction rate for controlled species i. In this way, the sites of controlled species are sorted based on the keyvalues corresponding to the reaction direction so that the controlled reactions can be fired deterministically and the time is controllable. Nevertheless, keyvalues can be values other than times if necessary.

1.5.3 Formal Model of cKMC

In this section, a formal model of cKMC will be described. It will be shown that cKMC is a generalization of the conventional KMC. Suppose that a system consisting of N state variables that form the state vector $\mathbf{X}(t) = (X_1(t), \ldots, X_N(t)) \in \mathbb{Z}^+$ at time t where $\mathbb{Z}^+ = \mathbb{N} \cup \{0\}$ is the set of nonnegative integers. There are a total of M events or reaction channels R_j's ($j = 1, \ldots, M$), each of which is characterized by a propensity function $a_j(\mathbf{X}(t) = \mathbf{x})$ or $a_j(\mathbf{x}, t)$ given the current state $\mathbf{x}$ at time t, where $a_j : \mathbb{Z}^+ \to \mathbb{R}^+$, and its state change vector $\mathbf{v}_j = (v_{j1}, \ldots, v_{jN})$. $a_j(\mathbf{X}(t))$ indicates how likely event R_j will fire in the next infinitesimal time interval and transit out of the current state, defined as $p(\mathbf{X}(t+\delta t)|\mathbf{X}(t)) = a_j(\mathbf{X}(t)\delta t + o(\delta t)$ where $\lim_{\delta t \to 0} o(\delta t)/\delta t = 0$, and $o(0) = 0$. From the sum of propensity function $a_0(\mathbf{X}(t) = \sum_{j=1}^{M} a_j(\mathbf{X}(t)))$, the probability of firing R_j at time t is $\Pr(R_j, t) = a_j(\mathbf{X}(t))/a_0(\mathbf{X}(t))$.

cKMC can be modeled as a continuous-time Markov chain and described by the *chemical master equation*

$$\begin{aligned}\frac{d}{dt}p(\mathbf{X}(t)=\mathbf{x}|\mathbf{X}(t_0)=\mathbf{x}_0) = &\sum_{j=1}^{M} a_j(\mathbf{x}-\mathbf{v}_j,t)p(\mathbf{X}(t)=\mathbf{x}-\mathbf{v}_j|\mathbf{X}(t_0)=\mathbf{x}_0) \\ &-[\sum_{j=1}^{M} a_j(\mathbf{x},t)p(\mathbf{X}(t)=\mathbf{x}|\mathbf{X}(t_0)=\mathbf{x}_0)]\end{aligned} \tag{1.8}$$

where $p(\mathbf{X}(t)=\mathbf{x}|\mathbf{X}(t_0)=\mathbf{x}_0)$ is the probability of $\mathbf{X}(t)=\mathbf{x}$ given the initial distribution $\mathbf{X}(t_0)=\mathbf{x}_0$. For a particular set of states $\{\mathbf{x}_d\}$ and times $\{t_k\}$ for which controlled events R_C's are defined, $a_C(\mathbf{x}_d,t_k)=\delta(t-t_k)$ where $\delta(t)$ is the Dirac delta function. That is, when a pre-scheduled deterministic event occurs, the corresponding rate for R_C is $a_C(\mathbf{x}_d,t_k)=\infty$. Eq.(1.8) generally describes the evolution of the probability distribution of the system modeled by cKMC. When $\{\mathbf{x}_d\}=\emptyset$ and $\{t_k\}=\emptyset$, cKMC just becomes the classical KMC.

1.6 Simulation-based Process Planning

Once the simulation models are constructed and validated, they can be used to predict the actual physical and chemical phenomena occurring during the nanomanufacturing processes. Simulation can provide us the analytical relations between process parameters (e.g. pressure, temperature, deposit path, dose, directions, distance to substrate, etc.) and performance (e.g. throughput rate, defect rate, etc.) through design of experiment. Process parameters can then be adjusted and optimized according to the simulation prediction during the process planning.

Here we use an example of nano-imprint lithography to illustrate. Nanoimprint lithography is a promising approach to achieve high-precision high-throughput nanoscale patterning. As shown in Figure 1.7(a), a scanning electron microscopy (SEM) image of a PMMA layer imprinted by a Chromium (Cr) stamp with an array of columns [72]. The corresponding cKMC model is constructed with species illustrated in Figure 1.7(b). The simulated profile of PMMA is shown in Figure 1.7(c). During process planning, different configurations and parameters can be tested by simulation. For instance, a profile with increased temperature is shown in Figure 1.7(d) where the profile is smoother and less edgy compared the original one. A second modification is the stamp geometry change with a smaller size, as shown in Figure 1.7(e), where smaller holes are produced.

The inputs of cKMC model include the species, events, and event rates. The event rates associated with controlled species can be deterministically estimated by controllable parameters in the process, such as moving speed of probe and intensity of electromagnetic field. Those values can be adjusted easily in either top-down or bottom-up manufacturing environments. The event rates associated with reaction or self-assembly events, however, need much effort to select and calibrate. As experiments are not always possible, the first principles approach shows its advantage. Atomistic DFT models with dozens of atoms can reveal both the stable and transition states. They provide the insight of how atoms move physically. This information is used to calibrate cKMC models which may have hundreds of thousands atoms. In addition, DFT simulation may also reveal the new chemical species during the nanomanufacturing processes. Aggregated molecules may move collectively as a clustered particle. In this case, the cKMC models can be built based on the clustered particle instead of at the atomic level. This can also save the computational time of cKMC simulation.

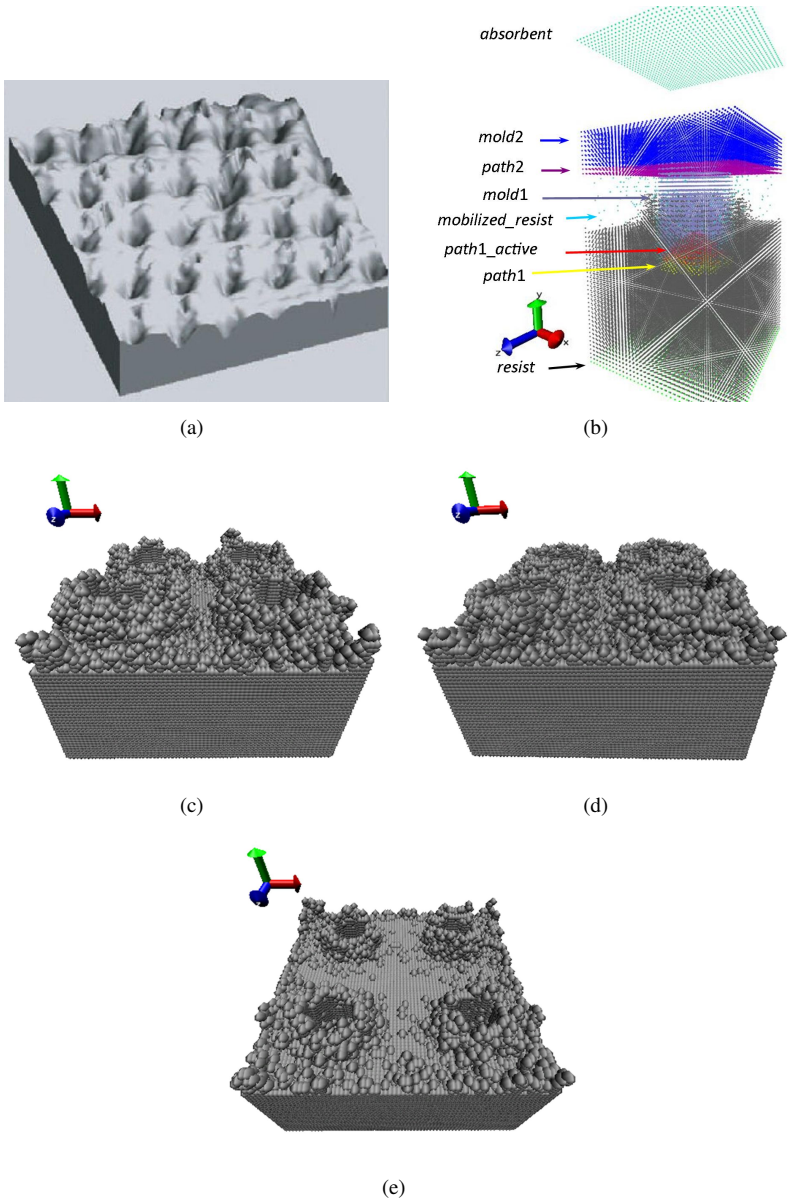


FIGURE 1.7: An illustration of simulation based process planning for nano-imprint lithography. (a) SEM image of PMMA imprinted by a Cr stamp (Zankovych et al.) (b) Species in cKMC model of NIL. (c) cKMC simulation of the PMMA profile. (d) An increased temperature yields a smoother and less edgy profile. (e) A smaller stamp size results in smaller holes.

1.7 Concluding Remarks

In this chapter, a first principles multiscale simulation framework that enables simulation based nanomanufacturing process planning is presented. This general framework can provide manufacturing industry and simulation practitioners a new approach to analyze nanomanufacturing processes

in a virtual environment. This virtual prototyping approach should be an important element to assist rational design of nanomanufacturing processes to improve product quality, increase throughput, and reduce costs. Computational tools helped the success of manufacturing and engineering at the classical bulk scale. They will also significantly help nanotechnology industry to advance to the next stage and bring great benefits to the society.

Although the importance of KMC in rare-event simulation has been recognized, there are still three major challenges related to accuracy, robustness, and efficiency. The first challenge of accuracy for KMC is that ideally all events have to be known a priori and listed in the event catalog so that the dynamics of physical processes can be simulated accurately. That is, we should know all possible events and the associated probabilities of occurrence under all configurations when building KMC models. Fortunately first principles simulation is available for us to explore all possible events. Some self-learning or adaptive multiscale KMC methods have also been proposed to accumulate the knowledge of events and build the event catalog on-the-fly while running KMC [27, 60, 43].

The second challenge of robustness for KMC is that during simulation the occurring probabilities of events are assumed to be accurate and fixed. In the real world, however, uncertainty is always involved in estimating the transition rates thus the probabilities of events. When the rates are estimated by physical experiments, measurement errors are unavoidable. When rates are estimated from first principles calculation, numerical setup and approximations for computability bring unintentional uncertainties, again, such as the exchange-correlation treatment in DFT. Model uncertainty is also inevitable when mathematical models are used to describe physical phenomena. In addition, these rates are dynamically changing over time. For instance, in diffusions, when the material structure is under dynamic mechanical load, the stress will change the diffusion rate. Bulk diffusion or inter-layer mass transport is very sensitive to temperature compared to surface diffusion or intra-layer mass transport. Stress and temperature variation will affect the accuracy of KMC prediction. In biochemical processes, the crowding effect [8, 52], where macromolecules block reaction paths, changes kinetic constants of reactions.

The third challenge of efficiency for KMC is that the time scales of the events vary significantly. Thus it is possible that the frequencies or probabilities of those events are in very different scales. For instance, in simulating CVD of diamond growth, surface diffusion could be several orders faster than adsorption and desorption. Computational time is not optimized to focus more on slower but critical events. Several approaches have been proposed to accelerate the KMC simulation. The tau-leaping method [23] accelerates the simulation by firing multiple events at the same time, and the clock is advanced with a fixed length. Improved methods such as leaping based on binomial [58, 12] have been proposed. The coarse-grained KMC [34] enables simulation of larger length and time scales by grouping lattice sites into coarse cells which act as the basic spatial unit in simulation.

At last, it should be noted that the proposed simulation based process planning framework and the core components are generic. Both top-down and bottom-up physical and chemical processes that occur during nanomanufacturing can be predicted. Yet, in order to apply the framework to a specific process, a domain expert whose working experiences and knowledge about the particular process is critical to build accurate models. Simulation practitioners should work closely with experimentalists on such process planning and optimization.

References

[1] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics*, 72(4):2384, 1980.

[2] H. C. Andersen. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1):24–34, 1983.

[3] C. C. Battaile and D. J. Srolovitz. Kinetic monte carlo simulation of chemical vapor deposition. *Annual Review of Materials Research*, 32(1):297–319, 2002.

[4] C. C. Battaile, D. J. Srolovitz, and J. E. Butler. A kinetic monte carlo method for the atomic-scale simulation of chemical vapor deposition: Application to diamond. *Journal of applied physics*, 82(12):6293–6300, 1997.

[5] E. M. L. Beale. A derivation of conjugate gradients. In F.A. Lootsma, editor, *Numerical methods for nonlinear optimization*, pages 39–43, London, 1972. Academic Press.

[6] A. D. Becke. Correlation energy of an inhomogeneous electron gas: A coordinate-space model. *The Journal of chemical physics*, 88:1053, 1988.

[7] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. R. H. J. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81:3684, 1984.

[8] H. Berry. Monte carlo simulations of enzyme reactions in two dimensions: fractal kinetics and spatial segregation. *Biophysical journal*, 83(4):1891, 2002.

[9] Ahmed Busnaina. *Nanomanufacturing handbook*. CRC, 2006.

[10] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101–014101, 2007.

[11] A. Chatterjee and D. G. Vlachos. An overview of spatial microscopic and accelerated kinetic monte carlo methods. *Journal of Computer-Aided Materials Design*, 14(2):253–308, 2007.

[12] A. Chatterjee, D. G. Vlachos, and M. A. Katsoulakis. Binomial distribution based τ-leap accelerated stochastic simulation. *The Journal of chemical physics*, 122:024112, 2005.

[13] Heping Chen, Ning Xi, Guangyong Li, Jiangbo Zhang, and Mathew Prokos. Planning and control for automated nanorobotic assembly. In *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 169–174. IEEE, 2005.

[14] E. Crnkic, L. He, and Y. Wang. Geometry guided crystal phase transition pathway search. *Computer-Aided Design*, 45(1):53–64, 2013.

[15] Edin Crnkic. Geometry guided phase transition pathway and stable structure search for crystals. Master's thesis, Georgia Institute of Technology, 2012.

[16] H. Doumanidis. The nanomanufacturing programme at the national science foundation. *Nanotechnology*, 13(3):248, 2002.

[17] W. E, W. Ren, and E. Vanden-Eijnden. String method for the study of rare events. *Physical Review B*, 66(5):052301, 2002.

[18] D. J. Evans, W. G. Hoover, B. H. Failor, B. Moran, and A. J. C. Ladd. Nonequilibrium molecular dynamics via gauss's principle of least constraint. *Physical Review A*, 28(2):1016, 1983.

[19] T.-H. Fang, C.-I. Weng, and J.-G. Chang. Molecular dynamics simulation of nano-lithography process using atomic force microscopy. *Surface Science*, 501(1):138–147, 2002.

[20] Gerald E. Farin. *Curves and Surfaces for Computer-Aided Geometric Design: A Practical Code*. Academic Press, Inc., 1996.

[21] R. Fletcher and C. M. Reeves. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

[22] J. Flidr, Y.-C. Huang, T. A. Newton, and M. A. Hines. Extracting site-specific reaction rates from steady state surface morphologies: Kinetic monte carlo simulations of aqueous si (111) etching. *The Journal of chemical physics*, 108:5542, 1998.

[23] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115:1716, 2001.

[24] M. Grujicic and S. G. Lai. Atomistic simulation of chemical vapor deposition of (111)-oriented diamond film using a kinetic monte carlo method. *Journal of materials science*, 34(1):7–20, 1999.

[25] X. Han. Study micromechanism of surface planarization in the polishing technology using numerical simulation method. *Applied surface science*, 253(14):6211–6216, 2007.

[26] Lijuan He. A global search algorithm for phase transition pathways in computer-aided nano-design. Master's thesis, Georgia Institute of Technology, 2013.

[27] G. Henkelman and H. Jónsson. Long time scale kinetic monte carlo simulations without lattice approximation and predefined event table. *The Journal of Chemical Physics*, 115:9657, 2001.

[28] G. Henkelman, B. P. Uberuaga, and H. Jónsson. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics*, 113:9901, 2000.

[29] W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695, 1985.

[30] H. Huang and G. H. Gilmer. An atomistic simulator for thin film deposition in three dimensions. *Journal of Applied Physics*, 84(7):3636–3649, 1998.

[31] W. Jin and H. H. Sawin. Feature profile evolution in high-density plasma etching of silicon with cl_2. *Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films*, 21(4):911–921, 2003.

[32] S. Jun, Y. Lee, S. Y. Kim, and S. Im. Large-scale molecular dynamics simulations of al (111) nanoscratching. *Nanotechnology*, 15(9):1169, 2004.

[33] M. Kalke and D. V. Baxter. A kinetic monte carlo simulation of chemical vapor deposition: non-monotonic variation of surface roughness with growth temperature. *Surface science*, 477(2):95–101, 2001.

[34] M. A. Katsoulakis, A. J. Majda, and D. G. Vlachos. Coarse-grained stochastic processes for microscopic lattice systems. *Proceedings of the National Academy of Sciences*, 100(3):782–787, 2003.

[35] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948. IEEE, 1995.

[36] R. Komanduri, N. Chandrasekaran, and L. M. Raff. Effect of tool geometry in nanometric cutting: a molecular dynamics simulation approach. *Wear*, 219(1):84–97, 1998.

[37] R. Komanduri and L. M. Raff. Molecular dynamics simulation of the nanometric cutting of silicon. *Philosophical Magazine B*, 81(12):1989–2019, 2001.

[38] R. Komanduri and L. M. Raff. A review on the molecular dynamics simulation of machining at the atomic scale. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 215(12):1639–1672, 2001.

[39] Sandia National Laboratory. Spparks. http://spparks.sandia.gov/.

[40] K. J. Laidler and M. C. King. Development of transition-state theory. *The Journal of physical chemistry*, 87(15):2657–2664, 1983.

[41] D. C. Langreth and J. P. Perdew. Theory of nonuniform electronic systems. i. analysis of the gradient approximation and a generalization that works. *Physical Review B*, 21(12):5469, 1980.

[42] G. J. Martyna, M. L. Klein, and M. Tuckerman. Nosé–hoover chains: the canonical ensemble via continuous dynamics. *The Journal of chemical physics*, 97:2635, 1992.

[43] D. Mei, L. Xu, and G. Henkelman. Potential energy surface of methanol decomposition on cu (110). *The Journal of Physical Chemistry C*, 113(11):4522–4537, 2009.

[44] B. Mokaberi and A. A. G. Requicha. Drift compensation for automatic nanomanipulation with scanning probe microscopes. *Automation Science and Engineering, IEEE Transactions on*, 3(3):199–207, 2006.

[45] A. Netto and M. Frenklach. Kinetic monte carlo simulations of cvd diamond growth - interlay among growth, etching, and migration. *Diamond and related materials*, 14(10):1630–1646, 2005.

[46] S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81:511, 1984.

[47] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52:7182, 1981.

[48] M. R. Pinto. Nanomanufacturing technology and opportunities through physically-based simulation. *Simulation of Semiconductor Processes and Devices 2007*, pages 1–8, 2007.

[49] A. A. G. Requicha, D. J. Arbuckle, B. Mokaberi, and J. Yun. Algorithms and software for nanomanipulation with atomic force microscopes. *The International Journal of Robotics Research*, 28(4):512–522, 2009.

[50] A. A. G. Requicha, S. Meltzer, F. P. T. Arce, J. H. Makaliwe, H. Sikén, S. Hsieh, D. Lewis, B. E. Koel, and M. E. Thompson. Manipulation of nanoscale components with the afm: principles and applications. In *Nanotechnology, 2001. IEEE-NANO 2001. Proceedings of the 2001 1st IEEE Conference on*, pages 81–86. IEEE, 2001.

[51] J.-P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of¡ i¿n¡/i¿-alkanes. *Journal of Computational Physics*, 23(3):327–341, 1977.

[52] S. Schnell and T. E. Turner. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. *Progress in biophysics and molecular biology*, 85(2):235–260, 2004.

[53] H. Shakir and W.-J. Kim. Nanoscale path planning and motion control with maglev positioners. *Mechatronics, IEEE/ASME Transactions on*, 11(5):625–633, 2006.

[54] S. Shimada, N. Ikawa, T. Inamura, N. Takezawa, H. Ohmori, and T. Sata. Brittle-ductile transition phenomena in microindentation and micromachining. *CIRP Annals-Manufacturing Technology*, 44(1):523–526, 1995.

[55] M. R. Sørensen and A. F. Voter. Temperature-accelerated dynamics for simulation of infrequent events. *The Journal of Chemical Physics*, 112:9599, 2000.

[56] P. Stavropoulos and G. Chryssolouris. Molecular dynamics simulations of laser ablation: the morse potential function approach. *International Journal of Nanomanufacturing*, 1(6):736–750, 2007.

[57] R. Stumpf and M. Scheffler. Theory of self-diffusion at and growth of al (111). *Physical review letters*, 72(2):254–257, 1994.

[58] T. Tian and K. Burrage. Binomial leap methods for simulating stochastic chemical kinetics. *The Journal of chemical physics*, 121:10356, 2004.

[59] J. Dalla Torre, G. H. Gilmer, D. L. Windt, R. Kalyanaraman, F. H. Baumann, P. L. OSullivan, J. Sapjeta, and M. Djafari Rouhani. Microstructure of thin tantalum films sputtered onto inclined substrates: experiments and atomistic simulations. *Journal of applied physics*, 94(1):263–271, 2003.

[60] O. Trushin, A. Karim, A. Kara, and T. S. Rahman. Self-learning kinetic monte carlo method: Application to cu (111). *Physical Review B*, 72(11):115401, 2005.

[61] Ampere A. Tseng. *Nanofabrication: fundamentals and applications*. World Scientific Publishing Company Incorporated, 2008.

[62] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[63] A. F. Voter. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Physical Review Letters*, 78(20):3908–3911, 1997.

[64] A. F. Voter. Parallel replica method for dynamics of infrequent events. *Physical Review B*, 57(22):13985–13988, 1998.

[65] L. Wang and P. Clancy. Kinetic monte carlo simulation of the growth of polycrystalline cu films. *Surface Science*, 473(1):25–38, 2001.

[66] N. Wang, S. I. Rokhlin, and D. F. Farson. Nanoparticle coalescence and sintering: molecular dynamics simulation. *International Journal of Nanomanufacturing*, 1(6):810–824, 2007.

[67] Y. Wang. Controlled kinetic monte carlo simulation of nanomanufacturing processes. In *Proc. ASME 2011 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*, pages No.DETC2011–48570, Washington DC, 2011. ASME.

[68] Y. Xia, J. A. Rogers, K. E. Paul, and G. M. Whitesides. Unconventional methods for fabricating and patterning nanostructures. *Chemical Reviews-Columbus*, 99(7):1823–1848, 1999.

[69] Y. Yang, S. Chen, K. Cheng, and X. Sun. Diamond turning of microstructured surfaces: modelling and simulation. *International Journal of Nanomanufacturing*, 1(5):627–640, 2007.

[70] Y. G. Yang, R. A. Johnson, and H. N. G. Wadley. A monte carlo simulation of the physical vapor deposition of nickel. *Acta materialia*, 45(4):1455–1468, 1997.

[71] Y. G. Yang, X. W. Zhou, R. A. Johnson, and H. N. G. Wadley. Monte carlo simulation of hyperthermal physical vapor deposition. *Acta materialia*, 49(16):3321–3332, 2001.

[72] S. Zankovych, T. Hoffmann, J. Seekamp, J. U. Bruch, and C. M. Sotomayor Torres. Nanoimprint lithography: challenges and prospects. *Nanotechnology*, 12(2):91, 2001.

[73] W. Zhao, K. Xu, X. Qian, and R. Wang. Tip based nanomanipulation through successive directional push. *Journal of manufacturing science and engineering*, 132:030909–1, 2010.

[74] H. Zhou, J. Fu, and M. Richard. Time-resolved kinetic monte-carlo simulation study on si (111) etching. *The Journal of Physical Chemistry C*, 111(9):3566–3574, 2007.