Chapter 1 Principles of computational materials design

In the context of materials science, the ultimate microscopic constituents of materials are ions and valence electrons; interactions among them at the atomic level determine the behavior



of the material at the macroscopic scale, the latter being the scale of interest for technological applications. To simulate the behaviors of materials, one can readily realize four characteristic length and time scales:

> • The atomic scale $(\sim 10^{-9}/\text{m or a few})$ nanometers), in which

the electrons are the players, and their quantum-mechanical state dictates the interactions among the atoms.

- The microscopic scale ($\sim 10^{-6}$ /m or a few micrometers), where atoms are the players and their interactions can be described by classical interatomic potentials (CIP) that encapsulate the effects of bonding between them, which is mediated by electrons.
- The mesoscopic scale ($\sim 10^{-4}$ /m or hundreds of micrometers), where lattice defects such as dislocations, grain boundaries, and other microstructural elements are the players. Their interactions are usually derived from phenomenological theories that encompass the effects of interactions between the atoms.
- The macroscopic scale ($\sim 10^{-2}$ /m or centimeters and beyond), where a constitutive law governs the behavior of the physical system, which is viewed as a continuous medium. In the macroscale, continuum fields such as density, velocity, temperature, displacement and stress fields, and so forth are the players. The constitutive laws are usually formulated so that they can capture the effects on materials properties from lattice defects and microstructural elements.Phenomena at each length scale typically have a corresponding timescale that, in correspondence to the four length scales mentioned above, ranges roughly from femtoseconds to picoseconds, to nanoseconds, to milliseconds and beyond.

Material Simulations for the 21st Century are likely to be peeked from the following two viewpoints:

(1) The virtual alchemist (linear-response theory),

(2) The nanotechnologist (tight-binding mappings).

However, realistic descriptions of large-scale nanostructures from first-principles are limited

by computation power. One way to circumvent the limitation is to map the electronic structure calculations into model Hamiltonians in terms of tight-binding orbitals (for functionalized nano building blocks) to implement a concept of electronic-structure LEGO bricks. The question is how the material simulation approach is going to make an impact?



The first thing we have to keep in mind is that we must understand what we are doing. Computations should not substitute for lack of knowledge. To illustrate the formal approach of material simulation, let us study the issue of intergranular embrittlement of Fe. Based on observation, we know that: Phosphor impurity/doping in steel will embrittle high strength steel, while boron enhances the intergranular cohesion. To interpret the observation, Rice and Wang proposed a theory: "Embritting tendency of solute depends on difference in segregation energy at grain boundary and free surface". Based on the theory, we can calculate segregation energy for boron and phosphor at free surface and grain boundary and study the brittle and ductile behavior of solids. For the detailed result see R. Wu, A. J. Freeman, and G. B. Olson, Science 265, (1994) 376-380.

When you can not think through the relation between macroscopic behavior and "computable" properties on the atomic scale, we can make an effort to derive relation statistically by using data-mining techniques.



In this regard, some learning methods could be useful to be combined with *ab-initio* calculation.



For example, we can predict stable crystal structures in a binary alloy from knowledge of only the energies $(e_{struc 1}, e_{struc 2})$ of a few compounds following the concept depicting in the

diagram:



Comparing with the method without the data mining, the hybrid approach accelerates the structural prediction scheme by 4 times.



Another example we will discuss in more detail is the development of thermoelectric materials.

A useful figure of merit for thermoelectric is $ZT = \frac{\sigma}{\kappa}S^2T$. To have a high-efficiency

thermoelectric material, we want to have low thermal conductivity κ , which can be calculated but quite tedious. In addition, we want semiconductors with high carrier conductivity and

Seeback coefficient S. Note $\sigma = \frac{e\tau}{3\sigma T} \int d\varepsilon \left(-\frac{\partial f}{\partial \varepsilon}\right) N(\varepsilon) \nu^2(\varepsilon) H(\varepsilon - \varepsilon_F)$, which can be calculated

from band structures. In chapter 6, we will introduce an idea to yield high-efficiency TE materials, which proposes to synthesize complex unit cells to have "ratteling" ions and cause scattering of phonons (skutterudites).

1. Conceptual Derivation of DFT

We learn from quantum mechanics that all information we can have about a given system is contained in the system's wave function Ψ . In a molecular system or solid, the nuclear degrees of freedom appear only in the form of a potential acting on the electrons, so that the wave function depends only on the electronic coordinates. For a single electron moving in a potential $v(\vec{r})$, the wave function satisfies the Schrodinger's equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v(\vec{r})\right]\Psi(\vec{r}) = E\,\Psi(\vec{r}) \quad . \tag{1-1}$$

If there is more than one electron, the equation of motion becomes

$$\left|\sum_{i=1}^{N} \left(-\frac{\hbar^2 \nabla_i^2}{2m} + v(\vec{r}_i) \right) + \sum_{i < j} U(\vec{r}_i, \vec{r}_j) \right| \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = E \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N)$$
(1-2)

where *N* is the number of electrons and $U(\vec{r_i}, \vec{r_j})$ is the electron-electron interaction. For a Coulomb system one has $\hat{U} = \sum_{i < j} U(\vec{r_i}, \vec{r_j}) = \sum_{i < j} q^2 / |\vec{r_i} - \vec{r_j}|$. The kinetic energy operator

 $\hat{T} = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2$ is the same for any system. For a molecule or a solid, the potential on the

electrons is $\hat{V} = \sum_{i} v(\vec{r_i}) = \sum_{i,k} Q_k q / |\vec{r_i} - \vec{R_k}|$, where the sum on k extends over all nuclei in the

system, each with charge $Q_k = Z_k e$ and position \vec{R}_k . Thus, it is only through the term \hat{U} of Eq. (2) that the single-body quantum mechanics of Eq. (1) differs from the extremely complex many-body problem posed by Eq. (2).

The usual quantum-mechanical approach to Schrodinger's equation (SE) can be summarized by the following sequence

$$v(\vec{r}) \xrightarrow{SE} \Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N) \xrightarrow{\langle \Psi | ... | \Psi \rangle} observables \quad . \tag{1-3}$$

One specifies the system by choosing $v(\vec{r})$, plugs it into Schrodinger's equation, solves that equation for the wave function Ψ , and then calculates observables by taking expectation values of operators with this wave function. One among the observables that are calculated in this way is the particle density

$$n(\vec{r}) = N \int d^3 r_2 \int d^3 r_3 \dots \int d^3 r_N \Psi^*(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) \quad . \tag{1-4}$$

DFT explicitly recognizes that Coulomb systems differ only by their potential $v(\vec{r})$, and supplies a prescription for dealing with the universal operators \hat{T} and \hat{U} once and for all. Furthermore, DFT provides a way to systematically map the many body problem with \hat{U} , onto a single-body problem without \hat{U} . All this is done by promoting the particle density $n(\vec{r})$ the status of key variable, on which the calculation of all other observables can be based. The density-functional approach can be summarized by the sequence

$$n(\vec{r}) \Rightarrow \Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N) \Rightarrow v(\vec{r}), \qquad (1-5)$$

i.e., knowledge of $n(\vec{r})$ implies knowledge of the wave function and the potential, and hence of all other observables.

1.1 DFT as a many-body theory

Since according to the above sequence the wave function is determined by the density, we can write it as $\Psi(\vec{r_1}, \vec{r_2}, ..., \vec{r_N}) = \Psi[n(\vec{r})]$, which indicates that Ψ is a function of its Nspatial variables, but a functional of $n(\vec{r})$.

Functionals. A functional F[n] can be defined as a rule for going from a function to a number, such as $N[n] = \int n(\vec{r}) d^3r = N$. Functional depends on the *function* itself, not on its variable, hence we do not need to distinguish $F[n(\vec{r})]$ from $F[n(\vec{r}\,')]$. An important case is the functional depends on a parameter, such as $v_H[n](\vec{r}) = q^2 \int d^3r n(\vec{r}\,')/|\vec{r}-\vec{r}\,'|$, which is a

rule that for any value of the parameter \vec{r} associates a value $v_H[n](\vec{r})$ with the function $n(\vec{r})$.

Functional variation. Given a function of one variable, y = f(x), one can think of two types of variations of *y*, one associated with *x*, the other with *f*. For a fixed point *x*, the functional variation δy measures how the value *y* at this point changes as a result of a variation in the functional form *f*(*x*).

Functional derivative. The functional derivative measures the first-order change in a functional upon a functional variation of its argument δf :

 $F[f(x) + \delta f(x)] = F[f(x)] + \int s(x)\delta f(x) dx + O(\delta f^2)$, where the integral arises because the

variation in the functional F is determined by variations in the function at all points in space. The first-order coefficient s(x) is defined to be the functional derivative $\delta F[f]/\delta f(x)$. A general expression for obtaining functional derivatives with respect to n(x) of a functional $F[n] = \int f(n, n', n'', ...; x) dx$, where primes indicate ordinary derivatives of n(x) with respect to x, is

 $\frac{\delta F[n]}{\delta n(x)} = \frac{\delta F}{\delta n} - \frac{d}{dx} \frac{\partial f}{\partial n'} + \frac{d^2}{dx^2} \frac{\partial f}{\partial n''} - \frac{d^3}{dx^3} \frac{\partial f}{\partial n'''} + \dots$ (1-6).

1.2 The Hohenberg-Kohn theorem

This theorem states that for ground state

 $n(\vec{r}) = N \int d^3 r_2 \int d^3 r_3 \dots \int d^3 r_N \Psi^*(\vec{r_1}, \vec{r_2}, \dots, \vec{r_N}) \Psi(\vec{r_1}, \vec{r_2}, \dots, \vec{r_N}) \text{ can be inverted: Given a ground}$

state density $n_0(\vec{r})$, it is possible to calculate the corresponding ground-state wave

function
$$\Psi_0(\vec{r}_1, \vec{r}_2, ..., \vec{r}_N)$$
.

If Ψ_0 can be calculated from $n_0(\vec{r})$ and vice versa, both functions are equivalent and contain exactly the same information. The crucial fact which makes this possible is that the ground-state wave function Ψ_0 must not only reproduce the ground-state density, but also minimize the energy:

$$E_{v,0} = \min_{\Psi \to n_0} \left\langle \Psi \mid \hat{T} + U + V \mid \Psi \right\rangle, \quad (1-7)$$

where $E_{v,0}$ denotes the ground-state energy in potential $v(\vec{r})$. The total-energy functional $E_v[n(\vec{r})]$ can be written as

$$E_{v}[n] = \min_{\Psi \to n} \left\langle \Psi \mid \hat{T} + \hat{U} \mid \Psi \right\rangle + \int d^{3}r \, n(\vec{r}) \, v(\vec{r}) =: F[n] + V[n], \quad (1-8)$$

where the internal-energy functional $F[n] = \min_{\Psi \to n} \langle \Psi \mid \hat{T} + \hat{U} \mid \Psi \rangle$ is independent of the

potential $v(\vec{r})$, and determined only by the structure of the operators \hat{U} and \hat{T} . This universality of the internal-energy functional allows us to define the ground-state wave function Ψ_0 as that antisymmetric *N*-particle function that delivers the minimum of F[n] and reproduces n_0 . For a Coulomb system, one can read off all information necessary for completely specifying the Hamiltonian (and thus implicitly all its eigenstates) directly from examining the density distribution. This can be achieved by noting: the integral over $n(\mathbf{r})$ yields N; the positions of the cusps of $n(\mathbf{r})$ yield the positions of the nuclei \vec{R}_k ; and the derivative of $n(\mathbf{r})$ at these positions yields Z_k .

The content of the HK theorem consists of four statements:

(1) The nondegenerate ground-state (GS) wave function $\Psi_0(\vec{r_1}, \vec{r_2}, ..., \vec{r_N})$ is a unique functional

of the GS density $n_0(\vec{r})$ by $\Psi_0(\vec{r_1},\vec{r_2},...,\vec{r_N}) = \Psi[n_0(\vec{r})]$.

(2) Perhaps the most important observable is the GS energy. This energy

$$E_{v,0} = E_v[n_0] = \left\langle \Psi[n_0] \mid \hat{H} \mid \Psi[n_0] \right\rangle = \left\langle \Psi[n_0] \mid \hat{T} + U + V \mid \Psi[n_0] \right\rangle \le E_v[n'(\vec{r})].$$

(3) Recalling that the kinetic and interaction energies of a Coulomb system are described by universal operators, we can also write

$$E_{v}[n] = T[n] + U[n] + V[n] = F[n] + V[n], \quad (1-9)$$

where T[n] and U[n] are *universal functionals*, independent of $v(\vec{r})$. (4) Fourth statement of the HK theorem: If $v(\vec{r})$ is not hold fixed, the functional V[n]

becomes universal. The GS density determines not only the GS wave function Ψ_0 , but

also the potential $V = V[n_0]$. Thus, $n_0(\vec{r})$ determines not only the GS wave function Ψ_0 , but the complete Hamiltonian and all excited states too (*i.e.*, the entire spectrum of the many-body Hamiltonian \hat{H}).

1.2 DFT as an effective single-body theory: The Kohn-Sham equations

Density-functional theory can be implemented in many ways. But the most popular way is the Kohn-Sham approach, which looks like a single-particle theory, although many-body effects are still included.

Exchange-correlation energy. A more accurate way to treat the kinetic-energy functional of interacting electrons, T[n], is to decompose it into one part that represents the kinetic energy $T_s[n]$ of noninteracting particles of density n, and one that represents the remainder,

denoted $T_c[n]$: $T[n] = T_s[n] + T_c[n]$. $T_s[n]$ is not known exactly as a functional of *n*, but it is

easily expressed in terms of the single-particle orbitals { $\phi_i(\vec{r})$ } of a noninteracting system

with density *n* as:
$$T_s[n] = -\frac{\hbar^2}{2m} \sum_{i=1}^N \int d^3 r \, \phi_i^*(\vec{r}) \nabla^2 \phi_i(\vec{r})$$
. Since all $\phi_i(\vec{r})$ are functionals of *n*,

this expression for T_s is an implicit density functional $T_s[n] = T_s[\{\phi_i[n]\}]$.

We now rewrite the exact energy functional as

$$E[n] = T[n] + U[n] + V[n] = T_s[\{\phi_i(\vec{r})\}] + U_H[n] + E_{xc}[n] + V[n], \text{ where by definition } E_{xc}$$

contains the differences $T_c = T - T_s$ and $U - U_H$. This definition shows that a significant

part of the correlation energy E_c is due to T_c . The exchange energy can be written explicitly in terms of the single-particle orbitals as

$$E_x[\{\phi_i[n]\}] = -\frac{q^2}{2} \sum_{jk} \int \int d^3r \, d^3r \, d^3r \, \cdot \frac{\phi_j^{\,*}(\vec{r}\,)\phi_k^{\,*}(\vec{r}\,')\phi_j(\vec{r}\,')\phi_k(\vec{r}\,)}{\left|\vec{r}-\vec{r}\,'\right|}$$

For the correlation energy no general explicit expression is known. Recalling the quantum mechanical interpretation of the wave function as a probability amplitude. Let us express two-particle wave function as $\Psi(\vec{r_1}, \vec{r_2}) = \alpha(\vec{r_1})\beta(\vec{r_2})$. But note that mathematically, the probability of a composed event $\Psi(\vec{r_1}, \vec{r_2})$ is equal to the probability of the individual events $\alpha(\vec{r_1})\beta(\vec{r_2})$ only if the individual events are independent (*i.e.*, uncorrelated). Physically, as a consequence of the Coulomb interaction, the electrons will try to avoid each other. The correlation energy is an additional energy lowering due to the mutual avoidance of the interacting electrons.

The xc energy can be written as a Hartree-like interaction

$$E_{xc}[n] = \frac{q^2}{2} \int d^3r \int d^3r' \frac{n(\vec{r})n_{xc}(\vec{r},\vec{r}\,')}{\left|\vec{r}-\vec{r}\,'\right|} \text{ between the charge distribution } n(\mathbf{r}) \text{ and the } xc \text{ hole}$$

 $n_{xc}(\vec{r},\vec{r}\,')$, which can be written as $n_{xc}(\vec{r},\vec{r}\,') = n(\vec{r}\,')(\overline{g}[n](\vec{r},\vec{r}\,') - 1)$. Here $\overline{g}[n](\vec{r},\vec{r}\,')$ is the average of the pair-correlation function. The practical advantage to write $E[n] = T[n] + U[n] + V[n] = T_s[\{\phi_i(\vec{r})\}] + U_H[n] + E_{xc}[n] + V[n]$ is that the unknown functional $E_{xc}[n]$ is typically much smaller than the known terms T_s , U_H and V. One can thus hope that reasonably simple approximations for $E_{xc}[n]$ provide useful results for E[n].

1.4 Static Kohn-Sham equations

First, write the energy minimization as

$$0 = \frac{\delta E[n]}{\delta n(\vec{r})} = \frac{\delta T_s[n]}{\delta n(\vec{r})} + \frac{\delta V[n]}{\delta n(\vec{r})} + \frac{\delta U[n]}{\delta n(\vec{r})} + \frac{\delta E_{xc}[n]}{\delta n(\vec{r})} = \frac{\delta T_s[n]}{\delta n(\vec{r})} + v(\vec{r}) + v_H(\vec{r}) + v_{xc}(\vec{r}) \,. \tag{1-10}$$

For the term $\frac{\delta E_{xc}[n]}{\delta n(\vec{r})} = v_{xc}(\vec{r})$ can only be calculated explicitly once an approximation for

 $E_{xc}[n]$ has been chosen.

Consider now a system of noninteracting particles moving in the potential $v_s(\vec{r})$. For this system, the minimization condition is simply

$$0 = \frac{\delta E_s[n]}{\delta n(\vec{r})} = \frac{\delta T_s[n]}{\delta n(\vec{r})} + \frac{\delta V_s[n]}{\delta n(\vec{r})} = \frac{\delta T_s[n]}{\delta n(\vec{r})} + v_s(\vec{r}), \qquad (1-11)$$



since there are no Hartree and xc terms in the absence of interactions. Comparing Eq. (1-11) with Eq. (10), we find that

both minimizations have the same solution $n_s(\vec{r})=\,n(\vec{r})$, if

$$v_s(\vec{r})$$
 is chosen to be $v_s(\vec{r}) = v(\vec{r}) + v_H(\vec{r}) + v_{xc}(\vec{r})$ (i.e., the

density of the interacting system of interest can be obtained as the density of an auxiliary system of non-interacting particles moving in an effective local single-particle potential, the so-called Kohn Sham potential.) Consequently, one can calculate the density of the interacting (many-body) system in potential $v(\vec{r})$, described by a many-body Schrodinger equation, by solving the equations of a

noninteracting system in potential $v_s(\vec{r})$:

$$-rac{\hbar^2}{2m}
abla^2+v_s(ec r)\left|\phi_i(ec r)=arepsilon_i\,\phi_i(ec r)\,,
ight.$$

which yields orbitals that reproduce the density $n(\vec{r})$ of the original system

$$n(\vec{r}) \equiv n_s(\vec{r}) = \sum_{i=1}^N f_i |\phi_i(\vec{r})|^2$$
 with f_i is the occupation of the *i*'th orbital. Since both $v_H(\vec{r})$

and $v_{xc}(\vec{r})$ depend on $n(\vec{r})$, which depends on the $\{\phi_i(\vec{r})\}$, which in turn depend on $v_s(\vec{r})$, the problem of solving the KS equations is a nonlinear one. The usual way of solving such problems (see the figure on the left) is to start with an initial guess for $n(\vec{r})$, calculate the corresponding $v_s(\vec{r})$, and then solve the KS equation for the $\{\phi_i(\vec{r})\}$. From these one

calculates a new density $n(\vec{r}) = \sum_{i=1}^{N} f_i \left| \phi_i(\vec{r}) \right|^2$ and starts again.

Simulation techniques based on density functional theory are now viewed as an important approach in materials science, with the exchange-correlation energy being treated by the local density approximation or the generalized-gradient approximation. There is also the use of hybrid exchange functionals within DFT. Hybrid functionals are a class of approximations to the exchange-correlation energy functional in DFT that incorporate a portion of exact exchange from Hartree-Fock theory. A widely used implementation of DFT combines a plane-wave basis set with the "pseudopotential" method, in which the pseudopotential replaces the core electrons by an effective potential in which the valence electron states can evolve.

Once one has a converged solution $n_0(\vec{r})$, one can calculate the total energy by

writing
$$V[n]$$
 in $E[n] = T_s[\{\phi_i(\vec{r})\}] + U_H[n] + E_{xc}[n] + V[n]$ as

$$V[n] = \int d^3r \, v(\vec{r})n(\vec{r}) = \int d^3r \, [v_s(\vec{r}) - v_H(\vec{r}) - v_{xc}(\vec{r})]n(\vec{r}) = V_s[n] - \int d^3r \, [v_H(\vec{r}) + v_{xc}(\vec{r})]n(\vec{r}) \, .$$

We identify the energy of the noninteracting (Kohn-Sham) system as $E_s = \sum_{i=1}^N \varepsilon_i = T_s + V_s$.

The total energy can then be calculated by using

$$E_0 = \sum_{i=1}^N \varepsilon_i - \frac{q^2}{2} \int d^3r \int d^3r \cdot \frac{n_0(\vec{r})n_0(\vec{r}\,')}{\left|\vec{r} - \vec{r}\,'\right|} - \int d^3r \, v_{xc}(\vec{r})n_0(\vec{r}) + E_{xc}[n_0],$$

which is not simply the sum of all { ε_i }. In fact, { ε_i } are introduced as artificial objects: they are the eigenvalues of an auxiliary single body equation whose eigenfunctions (orbitals) yield the correct density. It is only this density that has strict physical meaning in the KS equations. The main exception to this rule is the highest occupied KS eigenvalue. Denoting by $\varepsilon_N(M)$ the N'th eigenvalue of a system with M electrons, one can show rigorously that $\varepsilon_N(N) = -I$,

the negative of the first ionization energy of the N-body system, and $\varepsilon_{N+1}(N+1) = -A$, the

negative of the electron affinity of the same N-body system. When calculated with an approximate functional of the LDA or GGA type, the highest eigenvalues usually do not provide good approximations to the experimental I and A. Better results for these observables are obtained by calculating them as total-energy differences, according to

 $I = E_0(N+1) - E_0(N)$ and $A = E_0(N) - E_0(N+1)$, where $E_0(N)$ is the ground-state

energy of the N-body system.

There are basically three distinct types of approximations involved in a DFT calculation. One is conceptual, and concerns the interpretation of KS eigenvalues and orbitals as physical energies and wave functions. The second type of approximation is numerical, and concerns methods for actually solving the KS equation. A main aspect here is the selection of suitable basis functions. The third type of approximation involves constructing an expression for the

unknown xc functional $E_{xc}[n]$, which contains all many-body aspects of the problem.

2. Overview of Time-Dependent Density Functional Theory

2.1 Introduction

In this section, we introduce all the core elements of TDDFT, and how it differs from the ground-state case. The evolution of an interacting many-electron system is governed by the time-dependent Schrodinger equation

$$i\frac{\partial}{\partial t}\Psi(\{\vec{r}\},t) = \hat{H}(\{\vec{r}\},t)\Psi(\{\vec{r}\},t) \quad \Psi(\{\vec{r}\},t_0) \text{ given }, \quad (2-1)$$

where \hat{H} is the Hamiltonian operator of the system and $\{\vec{r}\} = \{\vec{r}_1, \vec{r}_2, ..., \vec{r}_N\}$ are the spatial coordinates of the *N* electrons. The Hamiltonian can be written in the form

$$\hat{H}(\{\vec{r}\},t) = \sum_{i=1}^{N} \left(-\frac{1}{2} \nabla_i^2 + v_{ext}(\vec{r}_i,t) \right) + \sum_{i$$

Here $v_{ext}(\vec{r}_i, t)$ is the time-dependent external potential and $v(\vec{r}_i - \vec{r}_j) = 1/|\vec{r}_i - \vec{r}_j|$ the Coulomb interaction.

The analog of the Hohenberg-Kohn theorem for time-dependent problems is the *Runge-Gross theorem*: Consider *N* electrons, mutually interacting *via* the Coulomb repulsion, in a time-dependent external potential. The densities $n(\vec{r},t)$ and $n'(\vec{r},t)$ evolving from a

common initial state $\Psi_0=\Psi(t=0)~$ under the influence of two external potentials $v_{ext}(\vec{r},t)$

and $v_{ext}^{'}(\vec{r},t)$, which differ by more than a purely time-dependent function:

 $\Delta v_{ext} = v_{ext}(\vec{r},t) - v'_{ext}(\vec{r},t) \neq c(t) \quad \text{(Note: The addition of a purely time-dependent function}$ only changes the phase of the wavefunction but not the density). Thus, there is a one-to-one mapping between *densities* and *potentials*, and the time-dependent potential can be expressed as a functional of the time-dependent density and the initial state $v_{ext}[n, \Psi_0](\vec{r}, t)$.

To prove the theorem, Runge and Gross first derived $\frac{\partial \Delta j(\vec{r},t)}{\partial t}\Big|_{t=0} = -n_0(\vec{r})\nabla[\Delta v_{ext}(\vec{r},0)]$, implying that if the Taylor-expansion about t = 0 of the difference of the two potentials is not spatially uniform for some order, then the Taylor-expansion of the current density difference will be non-zero at a finite order. And from the continuity equation, they further found $\frac{\partial^2 \Delta n(\vec{r},t)}{\partial t^2}\Big|_{t=0} = \nabla \cdot \{n_0(\vec{r})\nabla[\Delta v_{ext}(\vec{r},0)]\}$. Because $\Delta v_{ext}(\vec{r},0)$ is non-uniform, for real systems,

the right-hand side of the equation must be finite, causing the densities to differ in 2nd order in

t. This argument applies to each order and the densities $n(\vec{r},t)$ and $n'(\vec{r},t)$ will become different infinitesimally later than *t*.

Notes:

4 The difference between $n(\vec{r},t)$ and $n'(\vec{r},t)$ is nonvanishing already in first order of

 $\Delta v_{ext}(\vec{r},t)$, ensuring the invertibility of the linear response operators.

- Since the density determines the potential up to a time-dependent constant c(t), the wavefunction is in turn determined up to a time-dependent phase, which cancels out of the expectation value of any operator.
- We write $v_{ext}[n, \Psi_0](\vec{r}, t)$ because it depends on both the history of the density and the

initial wavefunction. This functional is a very complex one, however, knowledge of it implies solution of all time-dependent Coulomb interacting problems.

If we always begin in a non-degenerate ground state, the initial-state dependence can be subsumed by the Hohenberg-Kohn theorem, and then $v_{ext}(\vec{r},t)$ becomes a functional of

$$n(\vec{r},t)$$
 alone: $v_{ext}[n](\vec{r},t)$.

2.2 Time-Dependent Kohn-Sham equations (TDKS). Once we have a prove that the potential is a functional of the time-dependent density $v_{ext}[n](\vec{r},t)$, it is simple to write the TD Kohn-Sham equations as

$$i\frac{\partial\phi_{i\sigma}(\vec{r},t)}{\partial t} = \left(-\frac{\nabla^2}{2} + v_{KS\sigma}[n](\vec{r},t)\right)\phi_{i\sigma}(\vec{r},t) \quad , \quad (2-3)$$

whose potential is uniquely chosen (via the RG theorem) to reproduce the exact spin densities

 $n_{\sigma}(\vec{r},t) = \sum_{i=1}^{N_{\sigma}} \left| \phi_{i\sigma}(\vec{r},t) \right|^2$ of the interacting system. We define the exchange-correlation

potential via $v_{KS\sigma}(\vec{r},t) = v_{ext\sigma}(\vec{r},t) + v_{H\sigma}[n](\vec{r},t) + v_{xc\sigma}(\vec{r},t)$, where the second term is the

Hartree potential $v_{H\sigma}[n](\vec{r},t) = \int d^3r n_{\sigma}(\vec{r}',t) / |\vec{r} - \vec{r}'|.$

Notes:

4 The exchange-correlation potential $v_{xc \sigma}(\vec{r}, t)$ is a functional of the entire history of the

densities $n_{\sigma}(\vec{r},t)$, the initial interacting wavefunction $\Psi(0)$, and the initial Kohn-Sham wavefunction $\Phi(0)$. But if both the KS and interacting initial wavefunctions are non-degenerate ground states, it becomes a simple functional of $n_{\sigma}(\vec{r},t)$ alone.

- By inverting the single doubly-occupied KS equation for a spin-unpolarized two-electron system, it is quite straightforward to find the TDKS potential from an exact time-dependent density.
- In practical calculations, some approximation is used for $v_{xc\sigma}(\vec{r},t)$ as a functional of the

density, and so modifications of traditional TDSE schemes are needed for the propagation.

- Unlike the ground-state case, there is no self-consistency of a density dependent Hamiltonian.
- Contrary to the ground-state, there is no central role played by the ground-state energy $E[n_{\sigma}]$. In fact, an action was written down in the RG paper, but extremizing it was later shown not to yield the TDKS equations.

2.3 Linear-response theory

If the time-dependent external perturbation in $v_{ext}(\vec{r},t) = v_{ext}^{(0)}(\vec{r}) + v_{ext}^{(1)}(\vec{r},t)$ is weak, then linear-response theory can be exploited to describe the dynamics of a system more efficiently than with a full solution of the Kohn–Sham equations. In this case the density is expanded in orders of $v_{ext}^{(1)}(\vec{r},t)$ according to $n(\vec{r},t) = n^{(0)}(\vec{r}) + n^{(1)}(\vec{r},t) + \dots$, where the first-order correction is given by

$$n^{(1)}(\vec{r},t) = \int_{-\infty}^{+\infty} dt' \int d^3r' \,\chi(\vec{r},\vec{r}\,',t-t') v^{(1)}_{ext}(\vec{r}\,',t') \quad , \quad (2-4)$$

in terms of the linear density-response function $\chi(\vec{r},\vec{r}\,',t-t\,') = \frac{\delta n(\vec{r},t)}{\delta v_{ext}(\vec{r}\,',t\,')} \bigg|_{v_{ext}(\vec{r}\,',t\,')=v_{ext}^{(0)}(\vec{r}\,')}$.

Causality requires $\chi(\vec{r}, \vec{r}', t - t') = 0$ t<t. To calculate the linear density-response function, one exploits the fact that the density of the real system is identical to that of the

non-interacting Kohn–Sham electrons. As the latter move in the effective potential $v_{KS}(\vec{r}, t')$,

one starts by applying the chain rule for functional derivatives

$$\chi(\vec{r},\vec{r}\,',t-t\,') = \int_{-\infty}^{+\infty} dt'' \int d^3r'' \,\frac{\delta n(\vec{r},t)}{\delta v_{KS}(\vec{r}\,',t'')} \frac{\delta v_{KS}(\vec{r}\,',t'')}{\delta v_{ext}(\vec{r}\,',t\,')} \quad . \tag{2-5}$$

The first term on the right-hand side corresponds to the linear density-response function $\delta n(\vec{r},t) / \delta v_{KS}(\vec{r}'',t'') = \chi_{KS}(\vec{r},\vec{r}'',t-t'')$ of the non-interacting Kohn–Sham system, since the effective potential plays the role of the 'external potential' of the Kohn–Sham system. The second term in $\chi(\vec{r},\vec{r}',t-t')$ can be derived to be

$$\frac{\delta v_{KS}(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} = \frac{\delta}{\delta v_{ext}(\vec{r}^{'},t^{'})} \{v_{ext}(\vec{r}^{"},t^{"}) + v_{H}[n](\vec{r}^{"},t^{"}) + v_{xc\sigma}[n](\vec{r}^{"},t^{"})\} \\
= \delta(\vec{r}^{"}-\vec{r}^{'})\delta(t^{"}-t^{'}) + \frac{\delta v_{H}(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} + \frac{\delta v_{xc}(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} . (2-6)$$

As both the Hartree potential and the exchange–correlation potential are functionals of the density, one can apply the chain rule once more and rewrite these two contributions as

$$\frac{\delta v_{H}(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} = \int_{-\infty}^{+\infty} dt^{"} \int d^{3}r^{"} \frac{\delta v_{H}(\vec{r}^{"},t^{"})}{\delta n(\vec{r}^{"},t^{"})} \frac{\delta n(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} = \int_{-\infty}^{+\infty} dt^{"} \int d^{3}r^{"} \frac{\delta(t^{"}-t^{"})}{|\vec{r}^{"},\vec{r}^{"}|} \chi(\vec{r}^{"},\vec{r}^{'},t^{"}-t^{'})$$

$$\frac{\delta v_{xc}(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})} = \int_{-\infty}^{+\infty} dt^{"} \int d^{3}r^{"} \frac{\delta v_{xc}(\vec{r}^{"},t^{"})}{\delta n(\vec{r}^{"},t^{"})} \frac{\delta n(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})}, \quad (2-7)$$

$$= \int_{-\infty}^{+\infty} dt^{"} \int d^{3}r^{"} \frac{\delta v_{xc}(\vec{r}^{"},t^{"})}{\delta n(\vec{r}^{"},t^{"})} \frac{\delta n(\vec{r}^{"},t^{"})}{\delta v_{ext}(\vec{r}^{'},t^{'})}, \quad (2-8)$$

$$= \int_{-\infty}^{+\infty} dt^{"} \int d^{3}r^{"} f_{xc}(\vec{r}^{"},\vec{r}^{"},t^{"}-t^{"}) \chi(\vec{r}^{"},\vec{r}^{'},t^{"}-t^{'}), \quad (2-8)$$

which contains the so-called exchange–correlation kernel $f_{xc}(\vec{r}^{"},\vec{r}^{"},t^{"}-t^{"})$. After collecting all terms and performing a Fourier transform to frequency space, whereby convolutions on the time axis turn into simple multiplications, one obtains the final integral equation

$$\chi(\vec{r},\vec{r}',\omega) = \chi_{KS}(\vec{r},\vec{r}',\omega) + \int d^3r'' \int d^3r''' \chi_{KS}(\vec{r},\vec{r}'',\omega) \left(\frac{1}{|\vec{r}''-\vec{r}'''|} + f_{xc}(\vec{r}'',\vec{r}'',\omega)\right) \chi(\vec{r}''',\vec{r}',\omega) .$$
(2-9)

Notes:

- The *xc* kernel is a much simpler quantity than $v_{xc}[n](\vec{r},t)$, since the kernel is a functional of the ground-state density only.
- The kernel is non-local in both space and time. The non-locality in time manifests itself as a frequency dependence in the Fourier transform.
- If $f_{xc}(\vec{r}, \vec{r}', \omega)$ is set to zero in the above equation, it is called the Random Phase Approximation (RPA).
- \blacksquare The Hartree kernel is instantaneous (*i.e.*, local in time, has no memory, given

by an adiabatic approximation, is frequency independent).

- **4** The frequency-dependent kernel makes the solution of an RPA-type equation yield the exact χ .
- The kernel is, in general, complex, with real and imaginary parts related via Kramers-Kronig transformation.

The TDDFT equations in the linear-response regime can be cast in numerous different forms. For solids, in most implementations the integral equation (5) is solved by projecting all quantities onto a suitable set of basis, such as planewave representation within the pseudotential approximation. Equation (5) thus turns into a matrix equation

 $\chi(\omega) = \chi_{KS}(\omega) + \chi_{KS}(\omega)[v + f_{xc}(\omega)]\chi(\omega)$, in reciprocal lattice vectors in the case of periodic

systems, where $\chi = \chi_{GG'}(q)$. To obtain an absorption spectrum at a given momentum transfer

only one component of the matrix $\chi = \chi_{GG'}(\vec{q}, \omega)$ is required. This can be obtained by solving a linear system of equations.

3.4 Excitation energies. In the framework of TDDFT the relevant information about the excited

states is contained in the linear density-response function. The true excitation energies are the poles of $\chi(\vec{r}, \vec{r}', \omega)$. TDDFT has the great advantage that it yields the complete excitation spectrum. In order to see this, one can calculate the density change due to the external potential at first order. The stationary eigenstates of the original unperturbed Hamiltonian are labeled by $\Psi_j(\{\vec{r}\},t) = \Psi_j(\{\vec{r}\})\exp(-iE_jt)$, where E_j denotes the corresponding energy eigenvalues. After the onset of the time-dependent perturbation, the first-order correction (in orders of $v_{ext}(\vec{r},t)$) to the wavefunction $\Psi(\{\vec{r}\},t)$ that evolves from the ground-state

$$\Psi^{(0)}({\vec{r}},t) = \Psi_0({\vec{r}}) \exp(-iE_0 t)$$
 is

$$\begin{split} \Psi^{(1)}(\{\vec{r}\},t) &= -i\sum_{j=0}^{\infty} \Psi_{j}(\{\vec{r}\},t) \int_{-\infty}^{t} dt' \int d^{3}r'_{1} \dots \int d^{3}r'_{N} \ \Psi_{j}^{*}(\{\vec{r}'\},t') \\ &\times \left(\sum_{i=1}^{N} v_{ext}^{(1)}(\vec{r}'_{i},t')\right) \Psi_{0}(\{\vec{r}'\},t') \end{split}$$

$$(2-10)$$

The corresponding change in the density is

$$n^{(1)}(\vec{r},t) = N \int d^3 r_2 \dots \int d^3 r_N \left\{ \left[\Psi^{(1)}(\vec{r},\vec{r}_2,\dots,\vec{r}_N,t) \right]^* \Psi^{(0)}(\vec{r},\vec{r}_2,\dots,\vec{r}_N,t) + \left[\Psi^{(0)}(\vec{r},\vec{r}_2,\dots,\vec{r}_N,t) \right]^* \Psi^{(1)}(\vec{r},\vec{r}_2,\dots,\vec{r}_N,t) \right\}$$
(2-11)

In order to simplify the notation we introduce the overlap functions

$$n_{j}(\vec{r}) = N \int d^{3}r_{2}...\int d^{3}r_{N} \ \Psi_{0}^{*}(\vec{r},\vec{r}_{2},...,\vec{r}_{N})\Psi_{j}(\vec{r},\vec{r}_{2},...,\vec{r}_{N}) \quad , \quad (2-12)$$

We obtain

$$n^{(1)}(\vec{r},t) = \int_{-\infty}^{\infty} dt' \int d^3r' \left[-i \sum_{j=0}^{\infty} \left(n_j(\vec{r}) n_j^*(\vec{r}') e^{-i(E_j - E_0)(t - t')} - n_j(\vec{r}') n_j^*(\vec{r}) e^{-i(E_j - E_0)(t - t')} \right) \Theta(t - t') \right] v_{ext}^{(1)}(\vec{r}',t'), \quad (2-13)$$

the term in square brackets equals the linear density-response function $\chi(\vec{r},\vec{r}',t-t')$. After a

Fourier transform to frequency space and using $\Theta(t) = i/2\pi \lim_{\eta \to 0+} \int_{-\infty}^{\infty} d\omega \frac{1}{\omega + i\eta} e^{it\omega}$, one

arrives at the Lehmann representation of the density-response function:

$$\chi(\vec{r},\vec{r}',\omega) = \lim_{\eta \to 0+} \sum_{j=1}^{\infty} \left(\frac{n_j(\vec{r})n_j^*(\vec{r}')}{\omega - E_j + E_0 + i\eta} - \frac{n_j(\vec{r}')n_j^*(\vec{r})}{\omega + E_j - E_0 + i\eta} \right).$$
(2-14)

It is evident that the poles of $\chi(\vec{r}, \vec{r}', \omega)$, ω) correspond to the exact excitation energies $E_j - E_0$.

Furthermore, all quantities on the right-hand side depend only on the Hamiltonian of the unperturbed stationary system. By virtue of the Hohenberg–Kohn theorem the linear density-response function is hence a functional of the static ground-state density $n_{GS}(\vec{r})$.

3. Accelerated MD

3.1 The time-scale problem of the Molecular Dynamics

The premiere tool in simulating the dynamical evolution of systems of atoms is molecular dynamics (MD), in which one propagates the classical equations of motion forward in time. This requires choosing an interatomic potential for the atoms and a set of boundary conditions. Integrating the classical equations of motion forward in time, the behavior of the system emerges naturally. A serious limitation of MD is that accurate integration requires time steps about 10^{-15} s to resolve the atomic vibrations. Considering a reactant-to-product reaction, the

rate constant can be estimated with $k_{RP} = A \cdot \exp[- E_b^* \big/ k_B T].$ By using $A = 10^{13}~{\rm sec}$,

$$k_B T(300^0 K) = 25 meV$$
 and $E_b^* = 1 eV$, it is found $k_{RP} = 10^{-4} \text{ sec}^{-1}$, implying that to

simulate the chemical kinetics it requires at least 10^{19} MD steps. For processes such as surface diffusion and chemical kinetics, simulations to much longer time scales are needed to reveal the detailed kinetics. This is the "time-scale problem". Kinetic Monte Carlo attempts to overcome this limitation by exploiting the fact that the long-time dynamics of this kind of system typically consists of diffusive jumps from state to state.

3.2 Different Approaches



Several time acceleration concepts had been proposed for molecular dynamical simulations to circumvent the so-call time-scale problem. These are illustrated in Fig. x, which shows schematically the behavior of a single atom moving in a series of simple one-dimensional potential energy surfaces (PES) denoted by solid black lines. In the hyperdynamics

scheme conceived by Voter, shown in Fig. xa, the original PES, V(x), is augmented by a bias potential ΔV , which is zero at the dividing surface between the two energy minima, and acts to increase the frequency of barrier crossing by 'filling in' the areas of low energy. In the regions where the bias potential is non-zero, the effective simulation time passes more quickly by a factor of $\exp[\Delta V/k_BT]$. The ratio of accumulated hypertime to the standard MD clock

time is known as the 'boost', and can be as large as 106 if an appropriate form of biasing potential is chosen. It follows that the boost drops off considerably as rare events become more frequent. Unfortunately, it is not easy to find a general method of specifying the bias potential, and this area still is a topic of ongoing research. However, particularly convenient forms have been devised, such as the 'bond-boost

method' of Miron and Fichthorn, and also the bias potential dynamics (BPD) method by Hamelberg et al., the latter of which is implemented in the DL POLY code (version 2.20), available under academic license from Daresbury Laboratory in the UK.

A technique related to hyperdynamics is the metadynamics method of Laio and Parinello, in which a series of Gaussian functions are added to the PES in order to flatten it and force the atom to explore other regions of phase space, as shown in Fig. xb. Metadynamics enables the rapid exploration of free energy surfaces in some chosen set of coordinates, but there is no direct connection to a timescale and so any dynamics is largely fictitious. However, the metadynamics does have an advantage that it requires no a priori knowledge of the bias potential, and that the sum of Gaussians deposited up to a particular time provides an unbiased estimate of the free energy in the region explored during the simulation.

Two other accelerated dynamics methods also developed by Voter and co-workers that do not rely on biasing the PES are parallel replica dynamics (PRD) and temperature accelerated dynamics (TAD). Parallel replica dynamics is illustrated schematically in Fig. xc, where the canonical dynamics of a single system is replicated on a number (say, M) of processors running in parallel. Having allowed the replicas to become locally uncorrelated in the original basin, the dynamics of all M systems are then monitored until a transition occurs in a single one. The simulation clock is then advanced by the elapsed time summed over all M replicas, and the replication process is continued from the replica which made the transition. TAD, as illustrated in Fig. xd, is based on the concept of raising the temperature of the system to enable rare events to occur more frequently, while at the same time preventing the system from evolving along diffusion pathways. This can be done as follows: As the high temperature system is arrested just after making a transition, harmonic TST can be used to extrapolate the low temperature rate for the transition by finding the appropriate barrier height. By confining the system to its local basin, the TAD method essentially tries to find all possible escape routes at high temperatures, and then selects the one with the shortest time to occur at low temperatures. Compared to other accelerated dynamics methods, TAD is the most approximate, relying heavily on the assumption of harmonic TST, whereas PRD is the most accurate. On the other hand, TAD can yield large boost factors when energy barriers are high, whereas PRD is dependent on the number of available processors to achieve a given boost.

4. Kinetic Monte Carlo (kMC) Method

Monte Carlo refers to a broad class of algorithms that solve problems through the use of

random numbers. The most famous of the Monte Carlo methods is the Metropolis algorithm, which offers an elegant and powerful way to generate a sampling of geometries appropriate for a desired physical ensemble for the simulations of equilibrium properties of physical systems.

It was highly desirable to develop a different kind of Monte Carlo algorithm for evolving systems dynamically from state to state. This type of MC methodology is termed as kinetic Monte Carlo method. An appealing property of kMC is that it can, in principle, give the exact dynamical evolution of a system.

4.1 Rare-event systems and the kMC concept

A rare-event system is one in which the dynamics is characterized by occasional transitions from one state to another, with long periods of relative inactivity between these transitions. Each state corresponds to a single energy basin, and the long time between transitions arises because the system must surmount an energy barrier to get from one basin to another, schematically indicated in Fig. 1.

The key property of a rare-event system caught in a particular basin is that because it stays there for a much longer time than that of one vibrational period, it forgets how it got there. Then, for each possible escape pathway from the basin *i* to an adjacent basin *j*, there is a

rate constant k_{ii} that characterizes the escaping probability per unit time, and these rate

constants are independent of what state preceded state *i* (*i.e.*, a Markov chain process).

In general, there are two factors involved in k_{ij} : The first factor describes the inverse of the

time it takes for a particle with thermal velocity to traverse the barrier region. The second factor accounts for the small probability that a sufficient amount of energy to overcome the barrier is present in the reaction coordinate. Various portions of energy usually scattered among the many degrees of freedom of the heat bath happen by chance to be collected for a short moment in the motion along the reaction coordinate.

4.2 The rate constant of first-order processes

Because the system loses its memory of how it entered state *i* on a time scale that is short



Molecular Dynamics: the whole trajectory



Kinetic Monte Carlo: coarse-grained hops

compared to the time it takes to escape. Thus, during each short increment of time, it has the same probability of finding an escape path as it had in the previous increment of time. This gives rise to a first-order process with exponential decay statistics. The probability the system has not yet escaped from state *i* is given by $p_{survival}(t) = \exp(-k_{ij}t)$. We are particularly interested in the waiting-time probability distribution function $p_w(t)$ for the time of first



escape from the state.

Integrating $p_w(t)$ from 0 to t' gives the probability that the system has escaped by time t', which must equate

$$\operatorname{to} 1 - p_{survival}(t') = 1 - \exp(-k_{ij}t').$$

Thus, taking the negative of the time derivative of $p_{succent}(t')$ gives

the waiting-time probability distribution function as $p_w(t) = k_{ij} \exp(-k_{ij}t)$. The average waiting time τ before escape is just the first moment of this

distribution $\tau = \int t p_w(t) dt = 1/k_{ij}$. Because escape can occur along any of a number of pathways, we can make the same statement as above about each of these pathways. Each of these pathways has its own rate constant k_{ij} , and the total escape rate must be the sum of these rates: $k_{tot} = \sum_j k_{ij}$.

4.3 The kMC procedure

To determine a waiting time t_w with the distribution $p_w(t) = k_{tot} \exp(-k_{tot}t)$, we first draw a random number ρ on the interval (0, 1), and then form $t_w = -\ln \rho / k_{tot}$. A time drawn in this way is an appropriate realization for the time of first escape for a first-order process with rate constant k_{tot} .

An efficient rejection-free algorithm was implemented in most kMC simulation. Figure 2 compiles a flowchart of this algorithm. It starts with the determination of all N possible pathways, out of the present state. The corresponding N different rate constants are then summed to yield the total rate k_{tot} . Executed is the process q, which fulfills the condition

constant $\sum_{j=1}^{q-1} k_{ij} \le \rho k_{tot} \le \sum_{j=1}^{q} k_{ij}$. In order to understand the idea behind this condition imagine

for each process a block of height k_{ij} . If we then stack all of these blocks on top of each other as illustrated in Fig. 2, we arrive at a stack of total height k_{tot} . Aligning a random height $0 < \rho k_{tot} \leq k_{tot}$ along this stack will point to one of the blocks and this is the process that is selected. Obviously, a process with a large rate constant, *i.e.* a large block height, has a higher chance of being chosen in this way, and this probability weighted selection is precisely what the partial sums achieve.

The kMC procedure is summarized as follows:

(a) At time *t*, calculate $k_{tot} = \sum_{i} k_{ij}$ with the catalog of rate constants.

(b) Generate a random numbers ρ uniformly distributed in (0, 1) and compute the waiting time for escape $t_w = -\ln \rho / k_{tot}$.

(c) Make an array of partial sums as
$$s(m) = \sum_{j=1}^m k_{ij}$$
 . Find q from $s(q-1) \le \rho k_{tot} \le s(q)$ to

determine the chosen pathway for the escape at time $t + t_{\rm w}$.

(d) Advance the system clock and the entire cycle starts anew from the new system state.

4.4 Determining the rates

Assuming we know about the possible pathways, we can use transition state theory (TST) to compute the rate constant for each pathway. Transition state theory makes the following assumptions:



i) Reaction system passes the barrier only once (*i.e.*, no re-crossings);

ii) Energy distribution of reactant degree of freedom (DOF) is Boltzmann-like (many collisions compared to reaction events yield equilibrium between activated complex and reactant IS, except with respect to the reaction coordinate);

iii) Passage over barrier is the motion of only oneDOF, the reaction coordinate, which is independent of

all other motions of the activated complex (no concerted motions);

iv) Passage over barrier is a classical event (no tunneling).

The rate constant for escape from state i to state j is taken to be the equilibrium flux through a dividing surface separating the two states. It is possible that the trajectory may recross the dividing surface one or more times before either falling into state j or falling back into state i occur. The correct rate constant is to minimize the equilibrium flux by placing the dividing surface at the ridgetop.

The harmonic approximation to TST further simplifies the calculations of KMC rate constants. In hTST, we assume that the potential energy near the basin minimum to be well described with a second-order energy expansion and that the same is true for the modes perpendicular to the reaction coordinate at the saddle point. The dividing surface is taken to be the saddle plane and for a system with *N* moving atoms gives the simple form of

$$k^{hTST} = \frac{\prod_{i=1}^{3N} \nu_i^{\min}}{\prod_{i=1}^{3N-1} \nu_i^{sad}} \exp[-E_b / k_B T] \quad .$$
(4-1)

Here E_b is the static barrier height (energy difference between the saddle point and the minimum). In the pre-exponential factor, { ν_i^{\min} } are the 3N normal mode frequencies at the minimum and { ν_i^{sad} } are the 3N-1 nonimaginary normal mode frequencies at the saddle. The computation of k^{HTST} thus requires information only about the minimum and the saddle point for a given pathway.

4.5 The lattice assumption and the rate catalog

Typically in KMC simulations, the atoms in the system are mapped onto a lattice. An event may move one atom or many atoms, but in the final state, each atom will again map onto a unique lattice point.



Lattice mapping also makes it easy to exploit locality in determining rates. We assume that only the atoms near a defect affect the rate constant as shown in Fig. 4.1 for a pathway corresponding to the jump of a lattice vacancy. The number of possible rates, ignoring symmetry, is

 $n_{rate} = (n_{type} + 1)^{n_{site}}$, where n_{site} is the number of sites explicitly considered and n_{type} is the number of possible atom types that can be at each of those sites. The set of rates computed in this way comprise a "rate catalog", which we can then use to look up the rates we need for every state the system visits. The work can be reduced somewhat by splitting the neighborhood into two sets of sites, one set of sites that most influence the active atom in the initial state (e.g., sites 1, 2, 3, 5, 7, 8, 9 in Fig. 4) and another set of sites that most influence the active atom at the saddle point (e.g., sites 2, 3, 8, 9 in Fig. 4). Two catalogs are then generated, one for the minima and one for the saddle points. Each catalog entry gives the energy required to remove the active atom from the system. Subtracting these special vacancy formation energies for a given minimum-saddle pair gives the energy barrier for that process. It is important to make sure that all types of pathway are included, as missing pathways often cause larger errors in the final kMC dynamics.

Within the HTST framework, an easy simplification is to assume that the barrier height can be approximated by additive interactions. For example, in Fig. 4.2, the neighboring atoms



can be categorized as class m1 (nearest neighbors 2, 5, 8 when the system is at the minimum), class m2 (second nearest neighbors 1, 3, 7, 9 when the system is at the minimum), class s1 (first neighbors 2, 3, 8, 9 when the system is at the saddle point), and so forth. The barrier energy is then

approximated by $E_b = E_{sat} - E_{min}$ with $E_{min} = E_{min}^0 + n_{m1}E_{m1} + n_{m2}E_{m2}$ and

 $E_{sad} = E_{sad}^0 + n_{s1}E_{s1} + n_{s2}E_{s2}$. In this way, the rate catalog is replaced by a small number of additive interaction energies, which can be simply specified or adjusted to give simulation results that match experiment.

For every pair of connected states i and j in steady state, the number of transitions per unit time from i to j must equal the number of transitions per unit time from j to i as

 $p_i\,k_{ij}\,=\,p_j\,k_{ji}\,$ (the detailed balance criterion), where $\,p_i\,(p_j)\,$ is the fractional population of

state *i* (*j*). At equilibrium, the fractional population of state *i* is proportional to $\exp[-G_i / k_B T]$.

Thus, $k_{ij}/k_{ji} = \left\langle e^{-[G_j(T) - G_i(T)/k_BT} \right\rangle$ becomes a Boltzmann weighted free-energy difference of these two states. The detailed balance equation places requirements on the rate constants. If a

rate catalog is constructed that violates detailed balance, then the dynamical evolution will not correspond to a physical system.

We observe that on present-day computers, one can take roughly 10¹⁰ steps in a few hours of computer time. Assuming that for every state there is one fast escape pathway with a

fixed lowest barrier E_b and a prefactor of 10^{13} , then we can achieve a simulation time of

 $10^{10}/[10^{13} \exp(-E_b/k_B T)]$. For $E_b = 0.5$ eV, this gives a total simulation time of 2.5×10^5 at

T=300K, and 16 s at T=600K. For a system with very low barrier, properly depicting the low barrier process requires much shorter time step. To cover a similar time span, the number of simulation steps will be increased significantly. How to properly simulate the kinetics of a system with low barrier prevails as one of the long-standing challenges to kMC simulations.

5. Nudged Elastic Band Method

Calculation of transition states and reaction pathways is an important task for theoretical chemistry and in condensed matter physics. When studying chemical reactions where bond breaking is involved, QM methods are required. However, the QM calculation of reaction rates from transition state theory is difficult, since due to the exponential factor a chemical accuracy ≤ 1 kcal/mole is necessary to get meaningful rates. This requires high level calculations.

A popular approach to retrieve the transition state geometry, reaction pathways, and



activation energies is the nudged elastic band (NEB) method. NEB is based on the "chain-of-states" idea, where a number of images of the system are connected together between the endpoint configurations to trace out a transition pathway. The NEB approach initially makes an estimate of the minimum energy path (MEP) with a linear interpolation. Several images are constructed along the initial estimate of the MEP and forms a band of N+1 images (typically N=3-10), denoting by [$\mathbf{R}_0, \mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_N$] where \mathbf{R}_i defines the co-ordinates of image *i*. \mathbf{R}_0 and \mathbf{R}_N are the end points

and remain fixed throughout the calculation. The basic idea is to optimize the energy of all N+1 systems simultaneously, but to keep the N-1 non-minimum structures along the path, we put springs between them and optimize the energy:

$$E = \sum_{i=0}^{N} V(\mathbf{R}_{i}) + \sum_{i=0}^{N-1} \frac{1}{2} k(\mathbf{R}_{i+1} - \mathbf{R}_{i})^{2}.$$
 (5-1)

However, two problems can arise:

■ If the spring constants are too small, the images at steep potentials will slide down the potential energy hill, and all images will be in the valley, none be at or close to the transition state.

■ If the spring constants are too large, the images will move away from the MEP, the corners will be cut.

To get this clearer, let us consider the forces parallel and perpendicular to the NEB path as given by the sequence of images *i*=1,..., N-1. By calculating the tangent vector $\vec{\tau}_i$ on the path, the forces, which make the images slide down the hill, will be $\nabla V(\mathbf{R}_i) \cdot \vec{\tau}_i$. We can project those out to get the contributions of the potential perpendicular to the NEB path: $\nabla V(\mathbf{R}_i)|_{\perp} = \nabla V(\mathbf{R}_i) - \nabla V(\mathbf{R}_i) \cdot \vec{\tau}_i$. The forces leading to the corner cutting come from the contributions of the springs perpendicular to the NEB path. Therefore, to remove the effect, we consider only those parallel to the path

$$\vec{F}_i^s \Big|_{\mathbb{H}} = \left[k(\mathbf{R}_{i+1} - \mathbf{R}_i) - k(\mathbf{R}_i - \mathbf{R}_{i-1}) \right] \cdot \vec{\tau}_i.$$
(5-2)

If we optimize the whole system using only the forces $\nabla V(\mathbf{R}_i)|_{\perp}$ and $\vec{F}_i^s|_{\parallel}$ to eliminate

interferences between the spring forces and the true forces, the images will evenly space along the NEB path and perfectly follow the MEP.

The tangent estimation can be refined as

$$\vec{\tau}_{i} = \begin{cases} \vec{\tau}_{i}^{+} & \text{if } E_{i+1} > E_{i} > E_{i-1} \\ \vec{\tau}_{i}^{-} & \text{if } E_{i+1} < E_{i} < E_{i-1} \end{cases} ,$$
(5-3)

where $\vec{\tau}_i^+ = \mathbf{R}_{i+1} - \mathbf{R}_i$ and $\vec{\tau}_i^- = \mathbf{R}_i - \mathbf{R}_{i-1}$. If image *i* is at a minimum $(V_{i+1} > V_i < V_{i-1})$ or at a maximum $(V_{i+1} < V_i > V_{i-1})$, the tangent estimate becomes



$$\vec{\tau}_{i} = \begin{cases} \vec{\tau}_{i}^{+} \Delta V_{i}^{\max} + \vec{\tau}_{i}^{-} \Delta V_{i}^{\min} & \text{if } V_{i+1} > V_{i-1} \\ \vec{\tau}_{i}^{+} \Delta V_{i}^{\min} + \vec{\tau}_{i}^{-} \Delta V_{i}^{\max} & \text{if } V_{i+1} < V_{i-1} \end{cases}, \quad (5-4)$$

where $\Delta V_i^{\max} = \max(|V_{i+1} - V_i|, |V_{i-1} - V_i|)$ and

 $\Delta V_i^{\min} = \min(|V_{i+1} - V_i|, |V_{i-1} - V_i|)$. The tangents must then be normalized.

The implementation of the NEB is quite simple: First, set up an initial guess with a linear interpolation between the two end points. The energy and its gradient

are evaluated for each image in the elastic band, using some description of the energetics (*ab initio* or empirical force fields). For each image, the local tangent to the path is estimated and

the force is evaluated for an initial guess of the path. The subsequent minimization for the magnitude of the forces $\vec{F}_i = \vec{F}_i^s |_{\parallel} - \nabla V(\mathbf{R}_i)|_{\perp}$ with respect to the coordinates of the intermediate images can be carried out with the velocity Verlet algorithm

$$\mathbf{R}(t+\delta t) = \mathbf{R}(t) + \vec{v}(t)\delta t + \frac{1}{2}\vec{a}(t)\delta t^2$$
(5-5)

$$\vec{v}(t+\frac{\delta t}{2}) = \vec{v}(t) + \frac{1}{2}\vec{a}(t)\delta t$$
(5-6)

$$\vec{v}(t+\delta t) = \vec{v}(t+\frac{\delta t}{2}) + \frac{1}{2}\vec{a}(t+\delta t)\delta t$$
(5-7)

where $\vec{a}(t) = -\vec{F}(t)/m$ is the acceleration at time *t*. $\vec{v}(t)$ is the

velocity at time t and m is the mass of a given atom. The time step is a constant. If this is set too small the convergence to an MEP will be very slow, if it is set too large then the system may oscillate or become unstable. The velocity is initially zero and is

allowed to increase after each time step in the direction of the current force. To obtain a minimization using this method, it is necessary to damp the kinetic energy, this is achieved by keeping only the velocity component which is parallel to the force at the current step. It has been suggested that the most efficient place to quench the velocities is after equation (5-7) as

$$\vec{v}(t + \frac{\delta t}{2}) = \vec{v}(t + \frac{\delta t}{2}) \frac{\vec{v}(t + \frac{\delta t}{2}) \cdot \vec{a}(t)}{\left|\vec{a}(t)\right|}$$
. Further improvement with higher resolution around the

<u>.</u>,

saddle point can be done by using variable elastic constants. Spring constants can be chosen so that springs are stiffer where the potential energy is higher. Once the climbing direction is specified, a given image (x_i) can be made "to climb" up-hill the PES. This can be used to achieve an improved localization of saddle points.



6. Multiscale Simulations of Materials

6.1 Overview of Multiscale Simulation

Many important phenomena in material science and other research fields evolve over multiple length and time scales which cannot be captured both accurately and comprehensively in a single model. For example, materials models can be ordered as levels in a multiscale hierarchy based on their length and time scales as well as on a hierarchy of relations between them. Considering the ordering in length and time scales, *ab initio* models typically are in the range 0.1–10 nm length scales and time scales of 0.1–100 ps. However, in the study of dislocation dynamics, it often requires modeling to cover ranges of about 10 nm–10 µm and time scales of about 1 ns–1000 ns. Continuum mechanics can be successfully applied at this and even longer scales.

The successful modeling of the phenomena places some fundamental requirements, which include detailed and therefore complex models of material interactions on the one hand and relatively large and complex systems to be modelled on the other hand. The emerging point of view was to look at the different models as members of hierarchies ordered by increasing length and time scales. Coupling each model to the next would in principle allow the construction of both accurate and comprehensive models of material phenomena. The current concepts had been implemented along two main strands: (i) sequential multiscale modelling to construct largerscale material models from first principles and (ii) hybrid concurrent multiscale modeling for the description of unitary systems.

6.2 Sequential Multiscale Approache and A Case Study

In a sequential multiscale scheme, the models in the multiscale hierarchy can be coupled by determining the material parameters in each model from the outcome of calculations in a lower level, more detailed model. These material parameters are used to represent the interactions of the elements of the model and are generally obtained in the multiscale approach by considering these elements as systems in the lower level model.



The sequential multiscale approach offers the possibility of systematic construction from

ab initio principles of detailed, realistic models of material behavior at longer scales which are relevant to experimentally interesting materials science issues. This approach places the material models at longer length and time scales on a theoretically more secure basis. The sequential model has been used in different combinations for a wide variety of systems and phenomena.

[A case study of *ab initio* kMC on heterogeneous catalysis].

6.3 Concurrent Multiscale Approaches and A Case Study

A more formal approach is to conceive of the hierarchy of models as being generated from a single Hamiltonian by averaging out over selected degrees of freedom and their elimination, *i.e.* coarse graining (CG). The philosophy of the approach is that the level of detail required to model them is not homogeneous, so that uniform precision is not required. Detailed modelling is only required in relatively small regions during limited time intervals, when and where some 'significant' dynamical evolution of the system occurs. Most of the time and in most of the system detailed modelling is unnecessary. This idea can be implemented by concurrently employing a detailed model where this is required, and a higher level less detailed model for the rest of the system's trajectory. This approach may be successful even in the strongly coupled multiscale systems, as the long range interactions may be still fully taken into account through a (classical) representation which provides the necessary dynamical coupling between the high precision and the low-precision ('embedding') region.

In an atomistic concurrent hybrid scheme, a QM model and a classical model can be



coupled together in a single calculation, with elastic coupling occurring through the boundaries between the two representation regions and long range electrostatic interaction accounted for by the classical Coulomb potentials. The classical model is then used to evolve the environment in which the QM calculation is embedded. Essentially, the higher level calculation provides the dynamical boundary conditions appropriate for the lower level. Note further that typical length scales of 10–100 nm are generally sufficient for

the interactions to be insignificant compared with the thermal noise. This size scale indicates that the main case of hybrid multiscale modeling will have to couple *ab initio* type calculations on the nanometer scale with classical atomistic calculations which can range into tens of nanometers.

[see concurrent multiscale approach for metals, PHYSICAL REVIEW B 73, 024108 (2006)]

Final Remark on This Chapter

Apparently, the future of material simulation depends on computation power. But what does

more computing help? Based on the record of the past 40 years, computation power doubles every two years. This translates to 10^6 increase in computation performance in every 40 years.

For a simulation system with a length L, we note $N = L^3$. From the scaling of computation Molecular Dynamics with potentials: O(N)

 $\begin{array}{ll} DFT (LDA, GGA) & : O(N^3 \mbox{ or } N^2 log(n)) \\ Hartree \mbox{ Fock } & : O(N^4) \end{array}$

Thus, we can forecast the future trend of computation as

Method	Today (atoms)	+40 years
MD (potentials)	10 ⁸ atoms	10 ¹⁴ atoms
LDA (N ³)	1000	100,000
LDA(N)	1000	109
HF +CI(N ⁶)	10	100

Thus, we conclude that computational modeling is very powerful, but we must be very smart in using the power.