Kinetic Monte Carlo modelling of semiconductor growth

Peter Kratzer
Faculty of Physics, University Duisburg-Essen, Germany
Time and length scales

- morphology
- 2D islands
- surface reconstruction

![Graph showing time (s) vs. length (m) with different scales and regions labeled as DFT, molecular dynamics, and kinetic Monte Carlo.](image)
Outline of this talk

• Models in Statistical Physics, Monte Carlo (MC) methods, thermodynamic equilibrium
• How kinetic Monte Carlo exceeds over MC
• From molecular dynamics to kMC, how to make contact to DFT calculations

• Applications
• Summary
Methods of Statistical Physics
Discrete models in Statistical Physics

- Ising model (magnetism)

\[ H(s) = -J_q \sum_i \sum_{j \in n(i)} s_i s_j - \mu_B B \sum_i s_i \]

- Lattice-gas interpretation
  \[ c_1 = 0, 1 \quad s_i = 2c_i - 1 \]

\[ H = -4J_q \sum_i \sum_{j \in n(i)} c_i c_j + 2(qJ_q - \mu_B B) \sum_i c_i - N(qJ_q - \mu_B B) \]

- Goal:
  Calculation of thermal averages
A discrete model for epitaxy: solid-on-solid (SOS) model

- Atoms are symbolized by little cubes placed on a lattice.
- The growth surface has no voids, no “overhangs”.
- Atoms move by discrete hops with rate $\Gamma = \exp(-E/kT)$.
- The binding energy is determined by the # of neighbors $n$
  \[ E = E_D + n E_B \]
Stochastic sampling

• Calculating thermal averages in many-particles systems requires evaluation of high-dimensional integrals.
• Choosing the sampling points in an (almost) random way is a good strategy, in particular in high dimensions!
• Even better: importance sampling -- density of sampling points proportional to local value of the integrand
• Idea: create a stochastic process that achieves importance sampling.

\[ \pi/4 = 0.78 \approx 20/25 = 0.8 \]
Metropolis Sampling

**Solution:** *Importance Sampling* with

\[ w(q) = \frac{\exp(-V(q)/(k_BT))}{Z'} \]

- Generate random support points, distributed according to \( w(q) \), i.e., out of total \( K \) points, \( k_i = K w(q_i) \) in the unit volume around \( q_i \).

- The expectation value of an observable is calculated as

\[ \langle A \rangle \approx \frac{1}{K} \sum_{i=1}^{K} k_i A(q_i) \]

- The Metropolis algorithm generates, starting from \( q_0 \), successively a sequence of \( K \) configurations \( q_i \), distributed according to \( w(q) \).

- Even though we don’t know \( Z' \), this is possible, because it is just the correct relative probabilities that matter:

  - accept new config. \( q_{i+1} \), if \( \exp \left( -\frac{V(q_{i+1}) - V(q_i)}{k_BT} \right) > \text{rnd} \)
  - else reject.

- This assures that

\[ \frac{w(q_{i+1})}{w(q_i)} = \exp \left( -\frac{V(q_{i+1}) - V(q_i)}{k_BT} \right) \]
Metropolis algorithm

arbitrary start configuration

config. is modified \rightarrow test config.

\text{E}_{\text{test}} < E_i? 

\text{yes} \quad \rightarrow \quad \text{i} \rightarrow \text{i+1} \quad \text{yes}

\text{no} \quad \rightarrow \quad \text{draw random x} \in [0, 1] \text{[}
\exp\left(-\frac{(E_{\text{test}} - E_i)}{k_B T}\right) > x?

After “warm up” (i > 1000), add the value of observable A in config. i+1 to the cumulated average

\text{no}
From MC to kMC: the $N$-fold way
Classification of spins according to their neighborhood

<table>
<thead>
<tr>
<th>class</th>
<th>central spin</th>
<th>neighbors</th>
<th>class members $n_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>↑</td>
<td>↑, ↑</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>↑</td>
<td>↑, ↓</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>↑</td>
<td>↓, ↓</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>↓</td>
<td>↓, ↓</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>↓</td>
<td>↑, ↓</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>↓</td>
<td>↑, ↑</td>
<td>3</td>
</tr>
</tbody>
</table>
The \( N \)-fold way algorithm in MC

- processes are chosen with a probability proportional to their rates
- no discarded attempts (in contrast to Metropolis)

<table>
<thead>
<tr>
<th>class</th>
<th>central spin</th>
<th>neighbors</th>
<th>class members ( n_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>↑</td>
<td>↑,↑</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>↑</td>
<td>↑,↓</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>↑</td>
<td>↓,↓</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>↓</td>
<td>↓,↓</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>↓</td>
<td>↑,↓</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>↓</td>
<td>↑,↑</td>
<td>3</td>
</tr>
</tbody>
</table>

pointer steered by random number
Simulations of non-equilibrium processes: kinetic MC

• While being aware of all processes possible at an instant of time, we need a way of (randomly) selecting one process with the appropriate relative probability.
• An internal clock keeps track of the advancement of physical time.
  – If the processes are clearly separated in time, i.e. processes are uncorrelated on the time scale during which the processes takes place, the waiting time for each individual process has Poissonian distribution.
• We need to update the list of all possible processes according to the new situation after the move.

Specific algorithms:
  • process-type list algorithm
  • binary-tree algorithm
  • time-ordered-list algorithm
Application to a lattice-gas model

- example: lattice \( L_x \times L_y \)

- fool’s algorithm: first select one particle, then select one move of that particle

- the correct solution: cumulated partial rates
  \[ r_k = \sum_{i=1}^{k} \Gamma_i \], normalized to the total rate \( R = r_N \)

- selection process: draw a random number \( \rho \) and compare it to all the \( r_k/R \) sequentially; as soon as \( \rho \) exceeds \( r_k/R \), execute process \( k \)

- problem: we need to compare \( \rho \) to many (in the worst case all) of the \( r_k/R \)

- note: Selecting a process with the right probability requires that we can enumerate all \( N \) processes.
idea:
for \( p \) process types, we need to compare only to the \( p \) numbers \( N(k) \Gamma(k) \), \( k=1,p \), rather than to all \( r_k/R \) (which are much more numerous)
flow chart for a kMC algorithm

determine all possible processes for a given configuration of your system and build a list

calculate total rate \( R = \sum_k N^{(k)} \Gamma^{(k)} \)

\( \rho_1, \rho_2, \rho_3 \) random numbers \( \in [0,1] \)

find class \# k such that

\[
\sum_{j=0}^{k-1} N^{(j)} \Gamma^{(j)} \geq \rho_1 R > \sum_{j=0}^{k-1} N^{(j)} \Gamma^{(j)}
\]

execute process number \( \text{int}(\rho_2 N^{(k)}) \) from class \# k

update clock \( t \rightarrow t - \ln(\rho_3)/R \)

delete now obsolete processes from the process list

START

END
Time-ordered list algorithm

1. assign a random waiting time $t_i$ to each individual process
2. sort all processes according to ascending waiting time (requires only log(N) comparisons, if done in a way similar to the binary tree)
3. always select the first process and execute it
4. advance the clock by $t \rightarrow t + t_i$
5. Update the list and return to 1.

- This algorithm requires many exponentially distributed random numbers; thus it’s advisable to use specially a designed random number generator.

Self-learning kMC

- **Idea:** build up a database of rates on the fly
- If a certain environment/certain process is missing in the database, spawn a calculation of the barrier for this process.
- All environments **on a lattice** can be classified by the occupancy of neighbor shells.

superbasin algorithm

• If “fast” hops occur, consolidate them into a superbasin
• several exits with analytically calculated partial probabilities
• various models for exit time distributions available
• superbasins can be created or dismantled "on the fly"

example:
1D potential

conventional kMC


with superbasins
From molecular dynamics to kinetic Monte Carlo
From molecular dynamics to kinetic Monte Carlo

Conceptually, the system must be divided into the motion along the reaction coordinate and a “heat bath”. 
Counter-example: liquid-solid epitaxy

- Molecular dynamics (MD) may be unavoidable in cases when the atoms **not** are sitting on lattice sites
- possibly use some **accelerated** MD
- here: only the solid phase is treated atomistically

Transition State Theory (1-dim)

- Kramer's rate theory

\[ \Gamma = \frac{\lambda}{\omega_b} \left( \frac{\omega_0}{2\pi} \exp\left( - \frac{E_b}{kT} \right) \right) \]

\[ \lambda = \left( \gamma^2/4 + \omega_b^2 \right)^{1/2} - \frac{\gamma}{2} \]

\( \gamma \): friction due to coupling to the heat bath

- high-friction limit

\[ \Gamma = \frac{\omega_0 \omega_b}{2\pi \gamma} \exp\left( - \frac{E_b}{kT} \right) \]

- ‘medium’ friction → transition state theory

\[ \Gamma = \frac{\omega_0}{2\pi} \exp\left( - \frac{E_b}{kT} \right) \]

P. Hänggi, P. Talkner & M. Borkovec, Rev. Mod. Phys. 62, 251 (1990)
From the PES to rate constants $\Gamma$ (multi-dimensional)

$\Gamma = \frac{kT}{h} \frac{Z_{TS}}{Z_i} = \prod_N \nu_{k,i} / \prod_{N-1} \nu_{k,TS} \exp(-\Delta E/kT)$

Idea:
associate minima with the nodes, hops with the interconnects in a network.

Hopping rates derived from the PES:
$E(x_i,y_i) = \min \ E_{\text{tot}}(x_i, y_i, z_i, c_{\alpha})$
How accurate is Transition State Theory?

Three levels of approximation:

1. **direct molecular dynamics**

2. TST with **thermodynamic integration** of partition functions from restricted molecular dynamics at the ‘ridge’ (‘blue-moon-ensemble’)

3. TST within harmonic approximation

Cu/Cu(100): good agreement between method 1) and 2)

<table>
<thead>
<tr>
<th></th>
<th>( \ln \Gamma_0 ) [THz]</th>
<th>( \Delta E ) [eV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TI</td>
<td>MD</td>
</tr>
<tr>
<td>hop</td>
<td>2.9±0.2</td>
<td>3.0±0.2</td>
</tr>
<tr>
<td>exchange</td>
<td>6.5±0.6</td>
<td>6.1±0.7</td>
</tr>
</tbody>
</table>

Application I:

Molecular beam epitaxy on GaAs(001) $\beta_2(2\times4)$
Molecular beam epitaxy of III-V semiconductors

1) adsorption of As₂
2) dissociation of As₂
3) diffusion of As
4) desorption of As₂

5) adsorption of Ga
6) diffusion of Ga
7) desorption of Ga
8) island nucleation
9) growth

Processes:

What is the interplay of these processes for a given temperature and flux?
Rates from first-principles calculations

\[ \Gamma^{(k)} = W(f,i) = \Gamma^{(fi)}_0 \exp\left(-\frac{(E^{(fi)}_{TS} - E_i)}{kT}\right) \]
Surface diffusion on GaAs(001): mapping of PES to network graph

PES from DFT calculations $\rightarrow$ network of hops

barriers minima
kMC with explicit list of process types

Voter’s lattice kMC:

- simulation on a lattice
- group possible transitions $\Gamma(f,i)$ from $i$ to $f$ into classes, each class is characterized by a rate
- classification of initial and final state by ‘atomic neighborhoods’ e.g., the number and relative position of neighbors define a process type

DFT-based kMC:

possible hops in the trench... due to neighbors.
kinetic Monte Carlo simulations for GaAs epitaxy

- 32 microscopically different Ga diffusion processes, and \( \text{As}_2 \) adsorption/desorption are included explicitly
- Computational challenge: widely different time scales (\(10^{-12} \text{ sec to 10 sec}\))
- Simulation cell 160 x 320 sites (64 nm x 128 nm)
kinetics of island nucleation and growth

1/60 of the full simulation cell
As$_2$ pressure = 0.85 x $10^{-8}$ bar
Ga deposition rate = 0.1 ML/s
$T = 700$ K
island density

deposition rate
0.1 ML Ga per second, III/V ratio 1:1000, T=700K
scaling with temperature?

\[ N_{is} = \eta \left( \frac{R}{D} \right) \frac{i^*}{(i^*+2)} \]

- \( N_{is} \): island density
- \( D \): diffusion constant
- \( R \): deposition flux
- \( \eta \): numerical constant
- \( i^* \): critical nucleus


Application II:
kinetics of sintering
Sintering in materials synthesis

- For thermodynamics reasons, some materials (e.g. alloys) cannot be grown from solution
- Polycrystalline samples may be obtained by synthesising small particles and compaction, followed by a temperature and/or heat treatment
- Large crystals grow on the expense of smaller ones and may enforce re-orientation of neighbouring crystallites

Carbonyl iron powder (electron microscopy image)
Hybrid simulation

- particles treated as rigid bodies, using molecular dynamics with few collective variables
- contact dynamics for touching particles
- surface diffusion and growth treated by self-learning kMC

L. Brendel & D.E. Wolf, University Duisburg-Essen
Summary: Bridging the time-scale gap

- molecular dynamics (Car-Parrinello method)
- accelerated molecular dynamics
  - using a boost potential (Voter, Fichthorn, …)
  - temperature-accelerated MD (Montalenti et al. PRL 87, 126101 (2001))
- kinetic Monte Carlo with transition state search on the fly (avoids both lattice approximation and pre-defined rate table)
- lattice kinetic Monte Carlo, N-fold way (Voter PRB 34, 6819 (1986))
“Keep things as simple as possible, but not simpler ..”

Thank you for your attention!

Summary: arXiv:0904.2556
Parallelization of kMC

semi-rigorous synchronous sublattice algorithm
[Y. Shim and J.G. Amar, PRB 71, 115436 (2005)]

A-B-C-D-A-B-C-D- ...

one processor
Temperature-accelerated dynamics (TAD)

Event is observed at $T_{\text{high}}$, but its rate is extrapolated to $T_{\text{low}}$ (using the TST rate law).

“Speculative" TAD

A way to use computational parallelism in kinetic simulations.

If you have many processor cores available,
- spawn a new TAD sub-simulation as soon as a transition is seen
- use the retro-diction from $T_{\text{high}}$ to $T_{\text{low}}$ to assign a time when to expect this event
- branching continues until/unless it has become clear that this transition is not the one to be accepted (at $T_{\text{low}}$).

Example: Vapor-phase epitaxy of Cu on Ag(100)

It took ~1 year to grow 1.5 ML with serial TAD.