#### Recent Advances in Accelerated Molecular Dynamics Methods

Arthur F. Voter Theoretical Division Los Alamos National Laboratory

Quantum and Atomistic Modeling of Materials Defects Institute for Pure and Applied Mathematics (IPAM) UCLA October 2, 2012

# Outline

• Very brief introduction to accelerated molecular dynamics methods

Then I will focus on one recent advance:

• Local formulation of hyperdynamics for large systems

# Acknowledgments

DOE Office of Basic Energy Sciences Los Alamos LDRD ASCR (DOE) SCIDAC (DOE)

Local Hyperdynamics: Soo Young Kim (LANL postdoc) Danny Perez (LANL staff) Shiyu Du (LANL postdoc) Tim Germann (LANL staff)

# The time-scale problem



We have some material system (e.g., adatoms on a surface or a strained region of solid).

On some time scale, activated events will take place.

Using molecular dynamics (MD), we can run about 1 microsecond -- might not even see first event.

We need some other way to predict the long-time evolution.

LA-UR-12-26437

#### Infrequent event system



The system vibrates in a 3N-dimensional basin many times before finding an escape path. For each possible escape path, there is an associated section of dividing surface and a rate constant for escape through that dividing surface.

# Transition State Theory (TST)

TST escape rate = equilibrium flux through dividing surface at x=q

$$k_{A \to B}^{TST} = \langle \delta(x - q) | \dot{x} | \rangle \quad \text{(exact flux)}$$
$$k_{A \to B}^{HTST} = v_0 e^{-\Delta E/k_B T} \quad \text{(harmonic approx.)}$$

- classically exact rate if no recrossings or correlated events
- no dynamics required
- excellent approximation for materials diffusion

# The Problem

Often we don't know what the possible escape paths are.

#### **Accelerated Molecular Dynamics**



The system vibrates in 3N-dimensional basin many times before finding an escape path. The trajectory finds an appropriate way out (i.e., proportional to the rate constant) without knowing about any of the escape paths except the one it first sees. *In the AMD methods, we exploit this property, letting the trajectory find an appropriate escape path, but we carefully trick it into doing this more quickly.* 

LA-UR-12-26437

# AMD methods

We coax the trajectory into finding its own way out of the state, and then we repeat that procedure for the next state, and so on.

This approach (if it works) allows us to follow the state-to-state dynamics out to much longer times than we can reach with direct MD.

# Accelerated Molecular Dynamics Methods

#### Hyperdynamics



#### Parallel Replica Dynamics



- Design bias potential that fills basins.
- MD on biased surface evolves correctly from state to state.
- Accelerated time is statistical quantity. (AFV, J. Chem. Phys., 1997)
- Parallelizes time.
- Very general -- any exponential process.
- Gives exact dynamics if careful.
- Boost requires multiple processors (AFV, Phys. Rev. B, 1998)

#### **Temperature Accelerated Dynamics**



- Raise temperature of MD in this basin.
- Intercept and block every attempted escape.
- Accept event that would have occurred first at the low temperature.
- More approximate; good boost.

(M.R. Sorensen and AFV, J. Chem. Phys., 2000)

LA-UR-12-26437

#### New AMD-like method: Kappa Dynamics

Lu, Makarov and Henkelman, J. Chem. Phys. 133, 201101 (2010).



No TST assumption Exponential boost Very parallelizable Requires accurate sampling

FIG. 1. κ-dynamics trajectories are sampled from a TS surface that divides initial state i from possible final states j and l. Trajectories are followed until one is found which starts in i and goes LAdigect 2.43 product state without recrossing the TS.

#### New AMD-like method: Kappa Dynamics

Lu, Makarov and Henkelman, J. Chem. Phys. 133, 201101 (2010).



No TST assumption Exponential boost Very parallelizable Requires accurate sampling

FIG. 1. κ-dynamics trajectories are sampled from a TS surface that divides initial state i from possible final states j and l. Trajectories are followed until one is found which starts in i and goes LAdigect 2.43 product state without recrossing the TS.

# AMD methods

- Significant speedup over standard MD when barriers are high relative to temperature
- Often encounter unexpected behavior (e.g., surprising reaction mechanisms)
- The low-barrier problem is our biggest challenge for future
- Computational scaling with system size is also an important issue
  - Spatially parallel TAD
    - [Shim, Amar, Uberuaga, and Voter, Phys. Rev. B 76, 205439 (2007)]
  - Today's talk (local hyperdynamics)
- Recent introduction/review of the AMD methods: Perez et al, Acc. Comp. Chem. Res., 5, 79 (2009).

### Hyperdynamics

Builds on umbrella-sampling techniques (e.g., Valleau 1970's) Assumptions:

- infrequent events
- transition state theory (no recrossings)



AFV, J. Chem. Phys. 106, 4665 (1997)

### Hyperdynamics

Builds on umbrella-sampling techniques (e.g., Valleau 1970's) Assumptions:

- infrequent events
- transition state theory (no recrossings)



Procedure:

- design bias potential  $\Delta V$  (zero at dividing surfaces; causes no recrossings)
- run thermostatted trajectory on the biased surface (V+ $\Delta$ V)
- accumulate hypertime as

 $t_{hyper} = \Sigma \Delta t_{MD} exp[\Delta V(R(t))/k_{B}T]$ 

LA-UR-12-26437

AFV, J. Chem. Phys. 106, 4665 (1997)

# Hyperdynamics

Builds on umbrella-sampling techniques (e.g., Valleau 1970's) Assumptions:

- infrequent events
- transition state theory (no recrossings)



Procedure:

- design bias potential  $\Delta V$  (zero at dividing surfaces; causes no recrossings)
- run thermostatted trajectory on the biased surface (V+ $\Delta$ V)
- accumulate hypertime as

 $t_{hyper} = \Sigma \Delta t_{MD} exp[\Delta V(R(t))/k_BT]$ 

Result:

- state-to-state sequence correct
- time converges on correct value in long-time limit (vanishing relative error)

LA-UR-12-26437

AFV, J. Chem. Phys. 106, 4665 (1997)

#### Hyperdynamics – the bias potential

Designing valid and effective bias potential is a key challenge.

Bias potential can be a function of

- the shape of the energy surface (AFV, 1997)
- the energy (Steiner, Genilloud and Wilkins, 1998)
  - (Hamelberg, Mongan & McCammon, 2004)
- the geometry
  - bond lengths, Miron and Fichthorn, 2003, 2005
  - local strain, Hara and Li, 2010

Must be careful that bias is zero on all dividing surfaces or dynamics will be wrong.

When barriers are high relative to T, boost can be many orders of magnitude.

#### Hyperdynamics bias potential

An extremely simple form: flat bias potential



Steiner, Genilloud, and Wilkins, Phys. Rev. B 57, 10236 (1998).

- no more expensive than normal MD (negative overhead(!))
- very effective for low-dimensional systems
- diminishing boost factor for more than a few atoms.

#### Bond-boost bias potential

R.A. Miron and K.A. Fichthorn J. Chem. Phys. **119**, 6210 (2003)

Assumes any transition will signal itself by significant changes in bond lengths

Bias potential is turned on near the minimum in the potential basin, but turns off when any bond is stretched beyond a threshold value

Features:

- fairly general
- very low overhead
- purely geometric behaves better than earlier bias potentials based on slope and curvature of potential

# Co/Cu(001) growth using bond-boost hyperdynamics

Miron and Fichthorn Phys. Rev. B 72, 035415 (2005)

Simulation of growth at 1 ML/s



T=250K

T=310K

#### Simplified bond-boost bias potential

Based on Miron-Fichthorn 2003 form, but simplified.

Bias potential turns off when the relative distortion  $[\epsilon_{ij} = (r_{ij}-r_{ij}^{min})/r_{ij}^{min}]$  of any bond exceeds a threshold value q(as in Miron-Fichthorn 2003).

Simplification:  $\Delta V$  depends purely on coordinate ( $\epsilon_{max}$ ) of mostdistorted bond. Only one bond at a time has any bias force.



LA-UR-12-26437

#### Simple bond-boost bias example

Simple bond-boost bias potential

Cu adatom on Cu(100) surface

Hop barrier = 0.53 eV

<u>T(K)</u>	hop time	boost factor (S=0.4 eV)
350 K	1.36 μs *	1.1x10 <sup>3</sup>
300 K	<b>27</b> μs	3.1x10 <sup>4</sup>
200 K	0.8 s	1.1x10 <sup>8</sup>



\*At T=350K, the hyperdynamics rate matches full harmonic TST ("Vineyard") rate within 5% error bars (743 events). LA-UR-12-26437

#### Hyperdynamics on large systems



Whenever system is near a dividing surface,  $\Delta V$  must be zero.

For a 4x larger system, the trajectory is near a dividing surface ~4x more often, causing a lower overall boost factor.

For very large systems, the boost decays to unity - i.e., there is no speedup, no matter what form of bias potential is used.



e.q., Miron and Fichthorn saw boost~N<sup>-0.9</sup> and Hara and Li saw boost~N<sup>-1</sup>

#### Local Hyperdynamics

S.Y. Kim, D. Perez, and AFV (to be submitted).

Modified formulation of hyperdynamics that gives *constant boost for arbitrarily large systems.* 



<u>Key concept:</u> Most systems we are interested in are intrinsically local in their behavior. A transition, or near-transition, in one region of system should not have any significant effect on atoms that are far away.

LA-UR-12-26437

#### Local hyperdynamics – key components

- define a *local* bias energy and local bias force
- relax the requirement of conservative dynamics
- this bias force can be nonzero in one region even if there is a transition-causing distortion in another region, provided it is far away

The method is probably not exact, but we have some understanding of the error terms, and why they should largely cancel.

Tests on various systems show the method is very accurate.

#### Local hyperdynamics - procedure



Each bond (i) has its own local domain (I).

LA-UR-12-26437

#### Local hyperdynamics - procedure



Each bond (i) has its own local domain (I).

LA-UR-12-26437



Each bond (i) has its own local domain (I).

#### Local hyperdynamics - procedure



Each bond (i) has its own local domain (I).

#### Local hyperdynamics - procedure



Each bond (i) has its own local domain (I).

Each domain has its own bias energy  $\Delta V_I = C_I \Delta V_{\text{bias}}(R_I(t))$  and boost factor  $B_I$  based on the geometry  $R_I$  within the domain (exactly like a global hyperdynamics in that domain).

The bias energy in domain I is used to determine the force only on bond i. E.g., with a simple bond-boost bias there is a force on a bond if and only if it is the most distorted bond in its own domain.



Each bond (i) has its own local domain (I).

Each domain has its own bias energy  $\Delta V_I = C_I \Delta V_{bias}(R_I(t))$  and boost factor  $B_I$  based on the geometry  $R_I$  within the domain (exactly like a global hyperdynamics in that domain).

The bias energy in domain I is used to determine the force only on bond i. E.g., with a simple bond-boost bias there is a force on a bond if and only if it is the most distorted bond in its own domain.



Each bond (i) has its own local domain (I).

Each domain has its own bias energy  $\Delta V_I = C_I \Delta V_{\text{bias}}(R_I(t))$  and boost factor  $B_I$  based on the geometry  $R_I$  within the domain (exactly like a global hyperdynamics in that domain).

The bias energy in domain I is used to determine the force only on bond i. E.g., with a simple bond-boost bias there is a force on a bond if and only if it is the most distorted bond in its own domain.

# Local hyperdynamics - procedure

Each bond (i) has its own local domain (I).

Each domain has its own bias energy  $\Delta V_I = C_I \Delta V_{bias}(R_I(t))$  and boost factor  $B_I$  based on the geometry  $R_I$  within the domain (exactly like a global hyperdynamics in that domain).

The bias energy in domain I is used to determine the force only on bond i. E.g., with a simple bond-boost bias there is a force on a bond if and only if it is the most distorted bond in its own domain.

# Local hyperdynamics - procedure

Each bond (i) has its own local domain (I).

Each domain has its own bias energy  $\Delta V_I = C_I \Delta V_{\text{bias}}(R_I(t))$  and boost factor  $B_I$  based on the geometry  $R_I$  within the domain (exactly like a global hyperdynamics in that domain).

The bias energy in domain I is used to determine the force only on bond i. E.g., with a simple bond-boost bias there is a force on a bond if and only if it is the most distorted bond in its own domain.

A domain-bias multiplier (C<sub>I</sub>) is adjusted for each domain to make its average boost <B<sub>I</sub>> match the *target boost factor* B<sub>target</sub>. LA-UR-12-26437

#### Why it is not conservative



The force on bond i (the center of domain I) is taken to be

 $f_i = -\delta \Delta V_I / \delta x_i$  .

However, the force on bond j is given by

 $f_j{=}{-}\delta\Delta V_J/\delta x_j$  ,

so this is not conservative dynamics.

#### Why it is not conservative



The force on bond i (the center of domain I) is taken to be ("PowerPoint tex":  $\delta$  means partial derivative)  $f_i = -\delta \Delta V_I / \delta x_i$ .

However, the force on bond j is given by

$$f_j = -\delta \Delta V_J / \delta x_j$$
 ,

so this is not conservative dynamics.

#### Why it should work

Atoms near the center of domain I experience forces that closely approximate the forces in a global hyperdynamics on domain I with the same bias potential.



Average force error on some bond  $j = \langle f_j(R_I) - f_j^{global}(R_I) \rangle$ 

 $\Delta V_{I}^{\text{global}} = C_{I} \Delta V_{\text{bias}}(R_{I}(t))$ 



The force on bond j, which is based on domain J

 $f_i = -\delta \Delta V_J / \delta x_i$ 

is different than the force that bond i (domain I) *thinks* that bond j will have

 $f_j^{(I)} = -\delta \Delta V_I / \delta x_j.$ 

The size of this force error depends on the "I crescent" and the "J crescent".

LA-UR-12-26437



The force on bond j, which is based on domain J

 $f_j = -\delta \Delta V_J / \delta x_j$  (= -C<sub>J</sub> 2( $\epsilon_j / q^2$ )  $\delta \epsilon_j / \delta x_j$  if  $\epsilon_j$  is largest in J, else 0)

is different than the force that bond i (domain I) *thinks* that bond j will have

 $f_j^{(I)} = -\delta \Delta V_I / \delta x_j$ . (= -C<sub>I</sub> 2( $\varepsilon_j / q^2$ )  $\delta \varepsilon_j / \delta x_j$  if  $\varepsilon_j$  is largest in I, else 0)

The size of this force error depends on the "I crescent" and the "J crescent".

LA-UR-12-26437

# The domain crescent cases when $C_I = C_J$

(for simple bond boost bias)



Bond j is the most distorted bond in domain I+J  $\rightarrow$  exact ( $f_{j}^{(I)}=f_{j}$ )



Some other bond k in I  $\land$  J is the most distorted in domain I+J  $\rightarrow$  exact (f<sub>j</sub><sup>(I)</sup>=f<sub>j</sub>=0)



Bond j is the most distorted in domain I, but there is an even more distorted bond in the J crescent (that j sees but i does not see)  $\rightarrow$  error (f<sub>i</sub><sup>(I)</sup>≠0, f<sub>i</sub>=0)



Bond j is the most distorted in domain J, but there is an even more distorted one in the I crescent (that i sees but j does not see)  $\rightarrow$  error (f<sub>j</sub><sup>(I)</sup>=0, f<sub>j</sub>≠0)



The most distorted bond in domain I is in the I crescent and most distorted bond in domain J is in the J crescent  $\rightarrow$  exact ( $f_i^{(I)}=f_i=0$ )

LA-UR-12-26437

# The domain crescent cases when $C_I = C_J$

(for simple bond boost bias)



Bond j is the most distorted bond in domain I+J  $\rightarrow$  exact ( $f_j^{(I)}=f_j$ )



Some other bond k in I  $\land$  J is the most distorted in domain I+J  $\rightarrow$  exact (f<sub>j</sub><sup>(I)</sup>=f<sub>j</sub>=0)



Bond j is the most distorted in domain I, but there is an even more distorted bond in the J crescent (that j sees but i does not see)  $\rightarrow$  error (f<sub>i</sub><sup>(I)</sup> $\neq$ 0, f<sub>i</sub>=0)



Bond j is the most distorted in domain J, but there is an even more distorted one in the I crescent (that i sees but j does not see)  $\rightarrow$  error (f<sub>j</sub><sup>(I)</sup>=0, f<sub>j</sub>≠0)



The most distorted bond in domain I is in the I crescent and most distorted bond in domain J is in the J crescent

 $\rightarrow$  exact (f<sub>j</sub><sup>(I)</sup>=f<sub>j</sub>=0)

If I-crescent and J-crescent are statistically equivalent,

LA-UR-12-26437

-26437 then  $C_I = C_J$ , and error 3 and error 4 cancel in a time average. Los Alamos

# The domain crescent cases when $C_I = C_J$

(for simple bond boost bias)



Bond j is the most distorted bond in domain I+J  $\rightarrow$  exact ( $f_j^{(I)}=f_j$ )



Some other bond k in I  $\land$  J is the most distorted in domain I+J  $\rightarrow$  exact (f<sub>j</sub><sup>(I)</sup>=f<sub>j</sub>=0)



Bond j is the most distorted in domain I, but there is an even more distorted bond in the J crescent (that j sees but i does not see)  $\rightarrow$  error (f<sub>i</sub><sup>(I)</sup>≠0, f<sub>i</sub>=0)



Bond j is the most distorted in domain J, but there is an even more distorted one in the I crescent (that i sees but j does not see)  $\rightarrow$  error (f<sub>j</sub><sup>(I)</sup>=0, f<sub>j</sub>≠0)



The most distorted bond in domain I is in the I crescent and most distorted bond in domain J is in the J crescent

 $\rightarrow$  exact (f<sub>i</sub><sup>(I)</sup>=f<sub>i</sub>=0)

If I-crescent and J-crescent are NOT statistically equivalent, error 3 and error 4 LA-URd 2n264 are the compensate this.

#### The boostostat

Instead of pre-adjusting the  $\{C_I\}$  values for the current state, we have found that we can simply apply a "boostostat" to gently but constantly push on these  $\{C_I\}$  values during the simulation to move the boost for each domain towards the correct target boost.

For each domain I at each MD step:

 $C_{I}(t+\Delta t) = C_{I}(t) - \alpha_{B}\Delta t_{MD}[B_{I}(t) - B_{target}]/B_{target}$ 

#### where

 $\alpha_{\rm B}$  = boostostat coupling strength 0.9 'lowcitest1-4.coefvst' using 0:1 'lowcitest1-4.coefvst' using 0:2 lowcitest1-4.coefvst' using 0:3 0.89 0.88 0.87 running boost 0.86 coef 0.85 0.84  $\alpha_{\rm B}$ =2.5X10<sup>9</sup> s<sup>-1</sup> 0.83 0.82 0.81 5000 10000 15000 20000 LA-UR-12-26437 10 ns time (half-ps blocks)

n ( $\sim 10^9 - 10^{10} \, \mathrm{s}^{-1}$ )



#### The boostostat

Instead of pre-adjusting the  $\{C_I\}$  values for the current state, we have found that we can simply apply a "boostostat" to gently but constantly push on these  $\{C_I\}$  values during the simulation to move the boost for each domain towards the correct target boost.

For each domain I at each MD step:

 $C_{I}(t+\Delta t) = C_{I}(t) - \alpha_{B}\Delta t_{MD}[B_{I}(t) - B_{target}]/B_{target}$ 





Narrow strip of Ag(100), 8 adatoms, top layer and adatoms free to move (72 moving atoms), periodic in x, EAM potential.

Small enough to run fast, but large enough and complex enough to test method.

Both hops and exchanges can occur.

Transitions observed and then rejected so the rates stay constant.





Local range set to 10 Å for most tests:







Hop pathway: Ea=0.504 eV,  $v_0$ =3.53x10<sup>12</sup> Hz

Exchange pathway: Ea=0.651 eV,  $v_0$ =6.48x10<sup>12</sup> Hz

Side-hop pathway: Ea=0.743 eV,  $v_0$ =3.02x10<sup>12</sup> Hz

#### Local hyperdynamics – rate tests

T=500K Target boost = 100 Average event time = 2.2 ns (hypertime), 22 ps (MD time)



#### Local hyperdynamics – rate tests



T=500K,  $B_{target}$ =100,  $\alpha_B$ =5X10<sup>9</sup>, range=10 Å, full bond boost

Event type	Vineyard+Kramers	<u>Direct MD</u>	Local Hyperdynamics
Нор	88.7%	92%	90.9% (431/474)
Exchange	11%	8%	8.9% (42/474)
Side-hop	0.3%	~0.3%	0.2% (1/474)
Escape time	e 2.18 ns	2.16 ± .04 ns	$2.21 \pm 0.1$ ns (full BB) $2.25 \pm 0.1$ ns (simple BB)
Kinetic Temperature <boost></boost>			500.1±0.25 K 99.8±0.2



Tests inequivalent crescents.

Main unique events:





T=500K,  $B_{target}$ =100,  $\alpha_B$ =5X10<sup>9</sup>, range=10 Å, simple bond boost



LA-UR-12-26437

Final Energy (eV)

#### Dealing with transitions

When a transition occurs, in the region of the transition:

- the bond definitions must be updated
- the local-domain coefficients must be readjusted to maintain target boost

We have found we can simply let the boostostat make the necessary corrections to fix the  $\{C_i\}$  values after each transition. (Requires a careful choice for  $\alpha_B$  and good starting guesses for  $C_i$ 's for new bonds.)

For a brief period after each transition, the boost values are wrong, but overall the rates are still fairly accurate.

# Boostostatting through a transition Ag(100) test system

- Defective Ag surface: monomer, dimer, vacancy
- 434 atoms, 218 moving
- T=325K, γ=5x10<sup>11</sup> s<sup>-1</sup>, locality radius = 10 A
- On-the-fly boostostatting,  $\alpha_{\rm B}$ =5x10<sup>9</sup> s<sup>-1</sup> B<sub>target</sub>=100
- Coefficients reset locally to 0.2 eV after each transition
- About 25 thermally relevant transitions. Rates vary over four orders of magnitude.



# **Observed processes**



# Ag(100) test results

Ag defective surface, T=325K



LUS Alamos

# Conclusions from Ag(100) tests

- Agreement with Hamonic TST is excellent -- Local hyperdynamics rates are only a few percent high.
- Readjustment of {C<sub>I</sub>} following transition can be done on the fly without corrupting rates too much (although this will be case dependent).
- This is one key piece in maintaining true N scaling to arbitrarily large systems. (The other one is redefining the bonds on the fly when a transition occurs.)

#### Ag(100) tests – Impact of the locality radius



Too-short locality range (D) increases the rates

LA-UR-12-26437

#### Ag(100) tests – Impact of the locality radius



Too-short locality range (D) increases the rates

Transition Rate (1/s)

LA-UR-12-26437

### Local hyperdynamics Massively parallel implementation

10 million atoms for  $\sim 10$  microseconds

- Array of 200 bulk vacancies in EAM Cu
- Target boost =20,000
- T=350K
- $\sim 1$  day on 256 processors

(For this run, the code did not truly scale, because entire system was quenched each time a transition was detected.)





#### Conclusions

A new, local formulation of hyperdynamics makes an advance on the size-scale problem in accelerated molecular dynamics methods.

Boost may be modest (perhaps ten to a few powers of ten).

Should be scalable, with constant boost, to arbitrarily large systems, provided the lowest barrier in system does not keep dropping with increasing system size.