Local Hyperdynamics

Arthur F. Voter
Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM, USA

Stochastic Sampling and Accelerated Time
Dynamics on Multidimensional Surfaces
IPAM/UCLA
October 16-20, 2017

Acknowledgments:
Danny Perez (LANL, T-1), Soo Young Kim (former LANL postdoc)
Steve Plimpton (SNL), Dipanjan Ray (LANL postdoc), Tim Germann (LANL, T-1)
DOE/BES, DOE Exascale Computing Project, Los Alamos LDRD

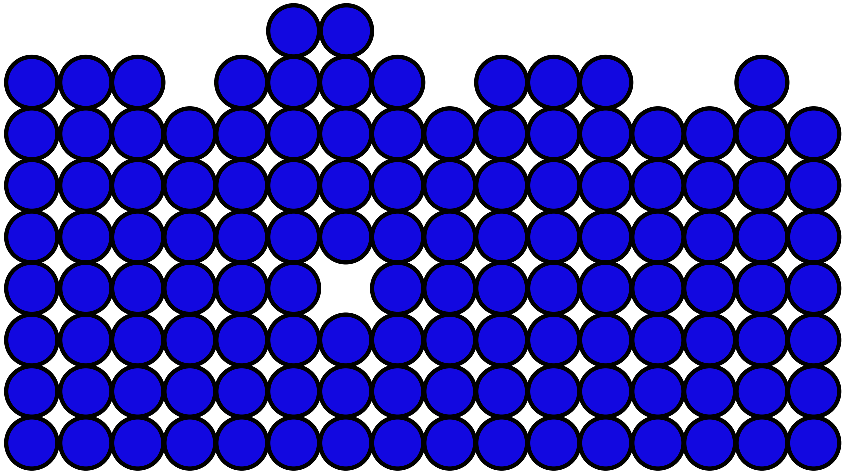
Local Hyperdynamics

Arthur F. Voter
Theoretical Division
Los Alamos National Laboratory
Los Alamos, NM, USA

Stochastic Sampling and Accelerated Time
Dynamics on Multidimensional Surfaces
IPAM/UCLA
October 16-20, 2017

Acknowledgments:
Danny Perez (LANL, T-1), Soo Young Kim (former LANL postdoc)
Steve Plimpton (SNL), Dipanjan Ray (LANL postdoc), Tim Germann (LANL, T-1)
DOE/BES, DOE Exascale Computing Project, Los Alamos LDRD

The time-scale challenge



We have some atomistic system.

Using molecular dynamics (MD), we can run a few microseconds.

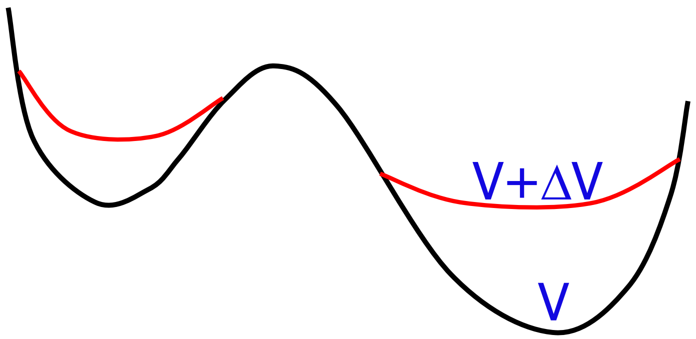
How do we accurately predict the evolution on longer time scales?

Usually, this long-time evolution involves infrequent events.

Hyperdynamics

Assumptions:

- infrequent events
- transition state theory (no recrossings)



Procedure:

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

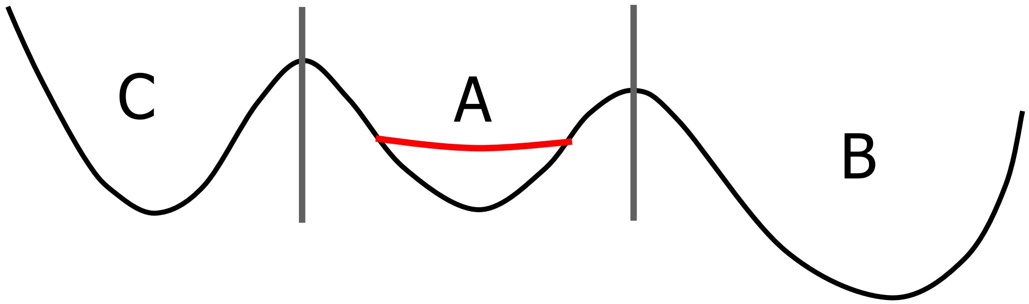
$$t_{hyper} = \Sigma \Delta t_{MD} \exp[\Delta V(R(t))/k_B T]$$

Result:

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

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

The boost factor



The boost factor (the hypertime divided by the MD time) is the average value of exp[+βΔV] on the biased potential:

$$\mathrm{Boost}_A = \left\langle e^{+\beta\Delta V(x)} \right\rangle_{Aboosted}$$

Bond boost bias potential

Miron and Fichthorn, J. Chem. Phys. 119, 6210 (2003)

Bias is based on bond distortions; shuts off completely when the relative distortion ε of any “bond” exceeds a pre-chosen critical value q (e.g., $q=0.3$).

Simple and inexpensive to evaluate.

Probably the best existing bias potential.

Bond boost bias potential

Miron and Fichthorn, J. Chem. Phys. 119, 6210 (2003)

Bias is based on bond distortions; shuts off completely when the relative distortion ε of any "bond" exceeds a pre-chosen critical value *q* (e.g., *q*=0.3).

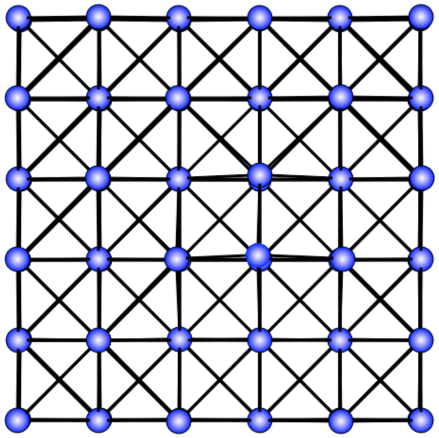
Simple and inexpensive to evaluate.

Probably the best existing bias potential.

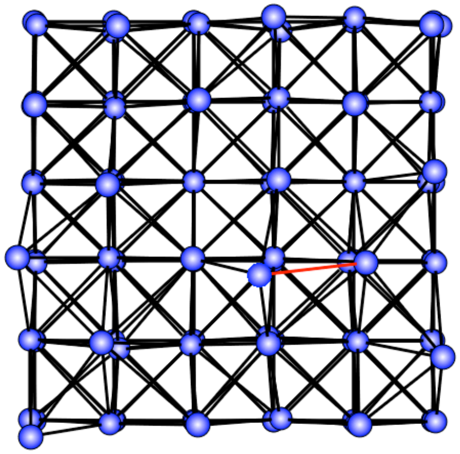
Simplified version ("simple bond boost"): ΔV depends purely on the coordinate (ε_{max}) of the most-distorted bond (Perez et al 2009).

Simplified bond-boost bias potential

ΔV depends purely on coordinate (ε_{max}) of most-distorted bond.
At most, one bond at a time has any bias force.



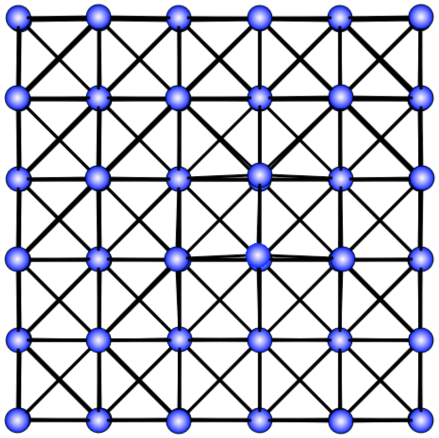
current minimum
defines $\{r_{ij}^{min}\}$ values



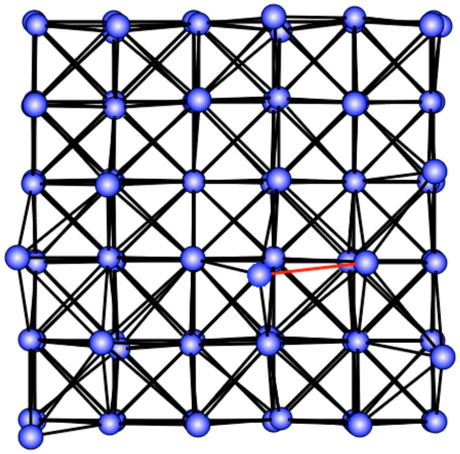
trajectory point
$\varepsilon_{ij} = (r_{ij}-r_{ij}^{min})/r_{ij}^{min}$
(relative distortions)

Simplified bond-boost bias potential

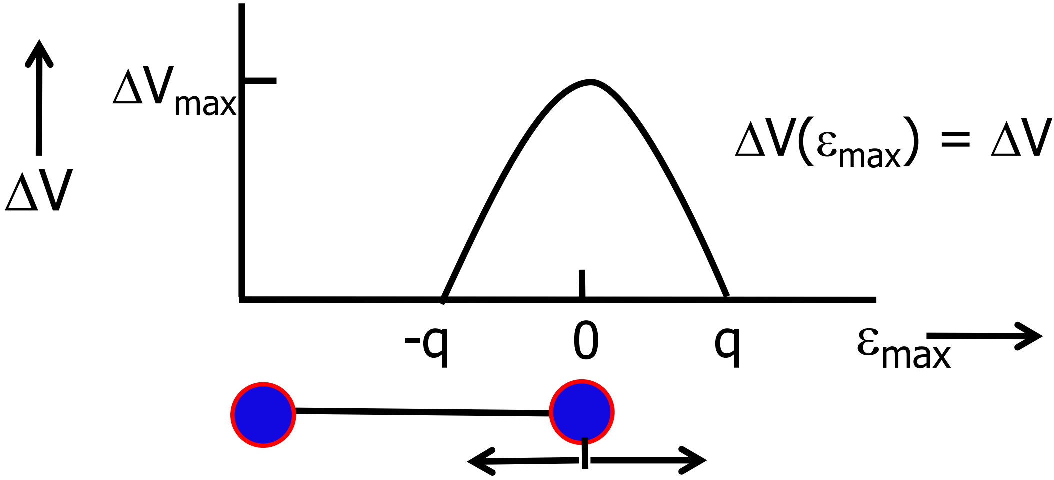
ΔV depends purely on coordinate (ε_{max}) of most-distorted bond.
At most, one bond at a time has any bias force.



current minimum
defines $\{r_{ij}^{min}\}$ values



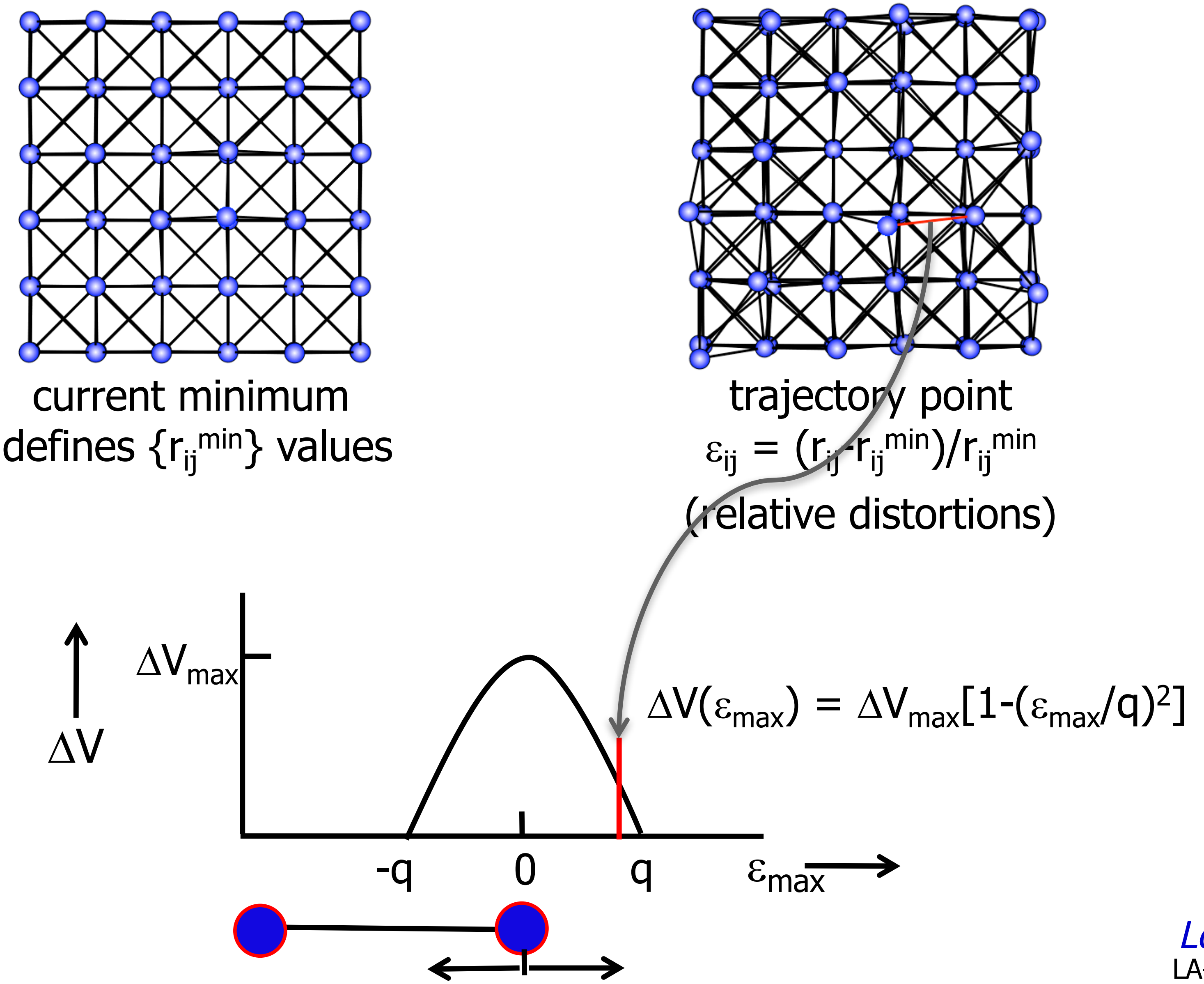
trajectory point
$\varepsilon_{ij} = (r_{ij}-r_{ij}^{min})/r_{ij}^{min}$
(relative distortions)



$$\Delta V(\varepsilon_{max}) = \Delta V_{max}[1-(\varepsilon_{max}/q)^2]$$

Simplified bond-boost bias potential

ΔV depends purely on coordinate (ε_{max}) of most-distorted bond.
At most, one bond at a time has any bias force.



current minimum
defines $\{r_{ij}^{min}\}$ values

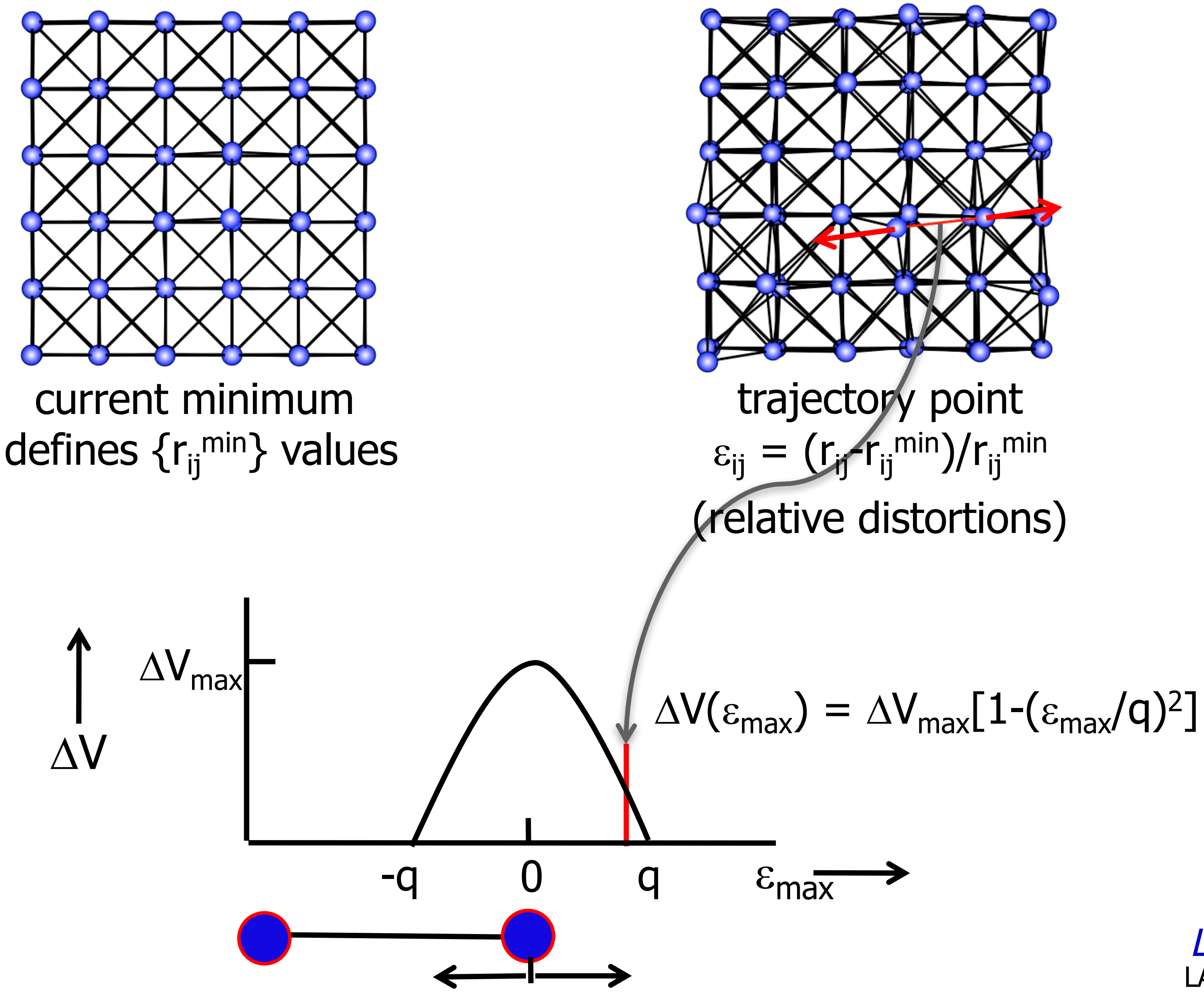
trajectory point
$\varepsilon_{ij} = (r_{ij} - r_{ij}^{min})/r_{ij}^{min}$

(relative distortions)

$$\Delta V(\varepsilon_{max}) = \Delta V_{max}[1-(\varepsilon_{max}/q)^2]$$

Simplified bond-boost bias potential

ΔV depends purely on coordinate (ε_{max}) of most-distorted bond.
At most, one bond at a time has any bias force.



current minimum
defines $\{r_{ij}^{min}\}$ values

trajectory point
$\varepsilon_{ij} = (r_{ij}-r_{ij}^{min})/r_{ij}^{min}$
(relative distortions)

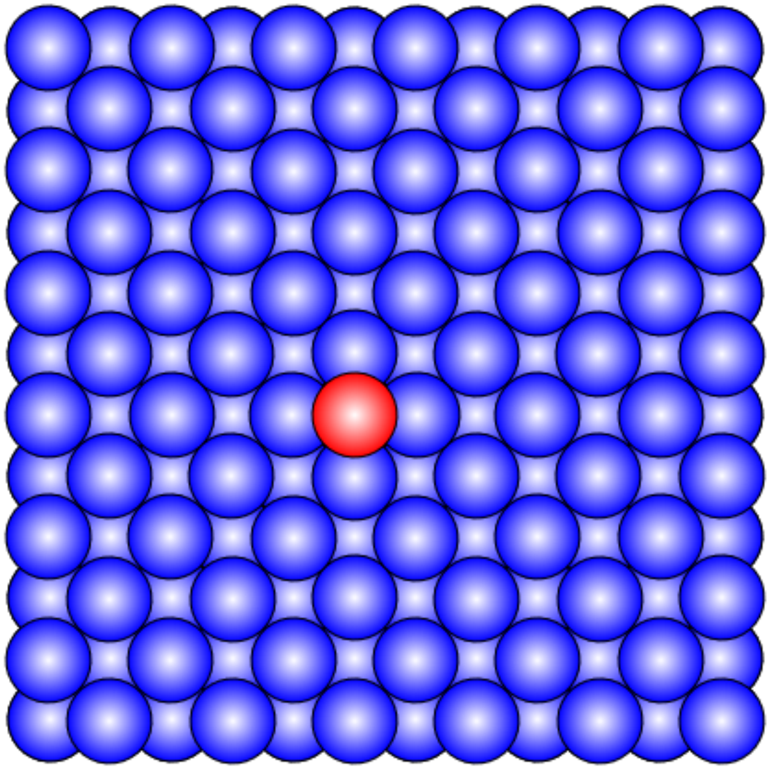
$$\Delta V(\varepsilon_{max}) = \Delta V_{max}[1-(\varepsilon_{max}/q)^2]$$

Simple bond-boost bias example

Cu adatom on Cu(100) surface

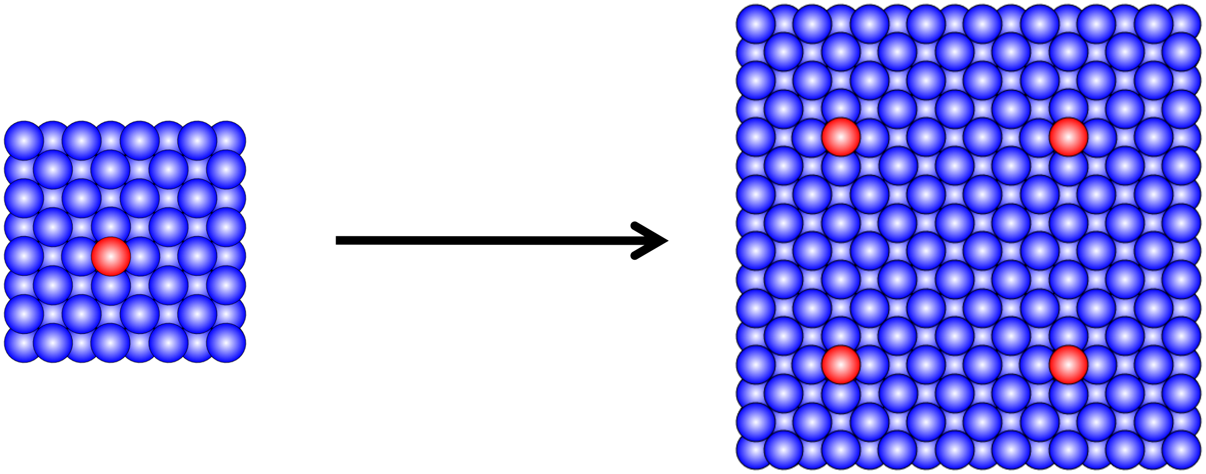
Hop barrier = 0.53 eV
ΔV_{max} = 0.4 eV

T(K)	hop time	boost factor
300 K	27 μs	3.1×10^4
200 K	0.8 s	1.1×10^8



Local hyperdynamics for large systems

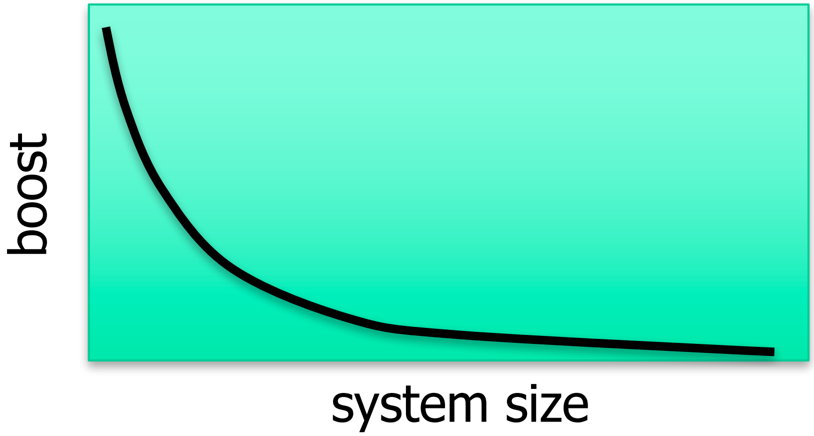
Hyperdynamics on large systems



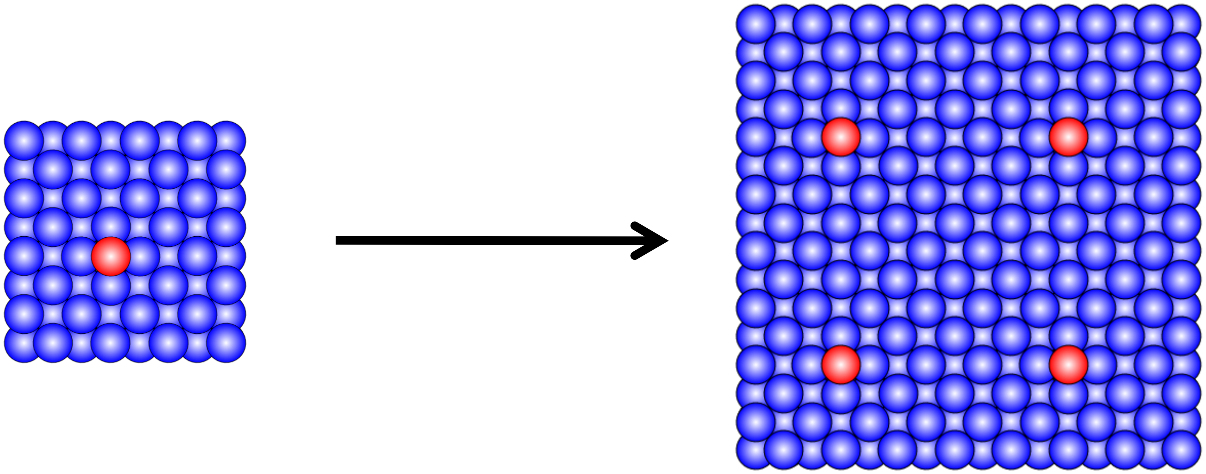
Whenever system is near a dividing surface, ΔV must be zero.

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*.



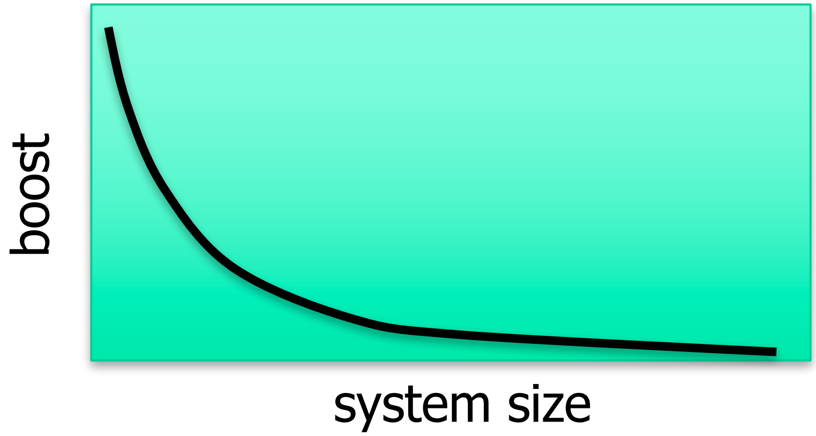
Hyperdynamics on large systems



Whenever system is near a dividing surface, ΔV 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*.

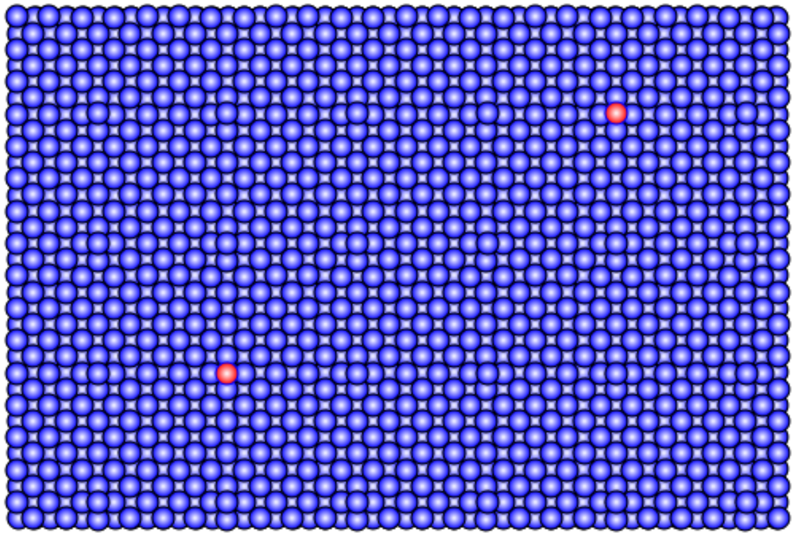


All the AMD methods in their primitive form show poor scaling with system size

Local Hyperdynamics

S.Y. Kim, D. Perez, and AFV, J. Chem. Phys. 139, 144110 (2013).

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

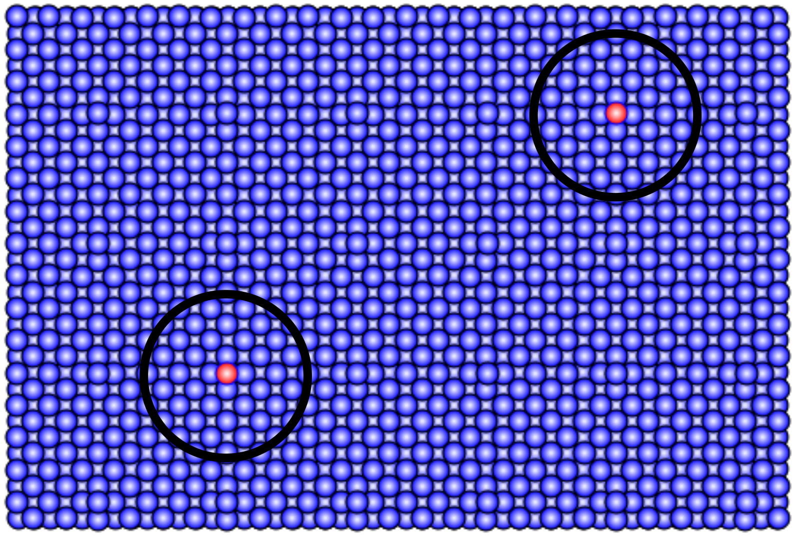


Key concept: Most systems we are interested in are intrinsically local in their behavior. A transition, or near-transition, in one region of system should not have any significant effect on atoms that are far away.

Local Hyperdynamics

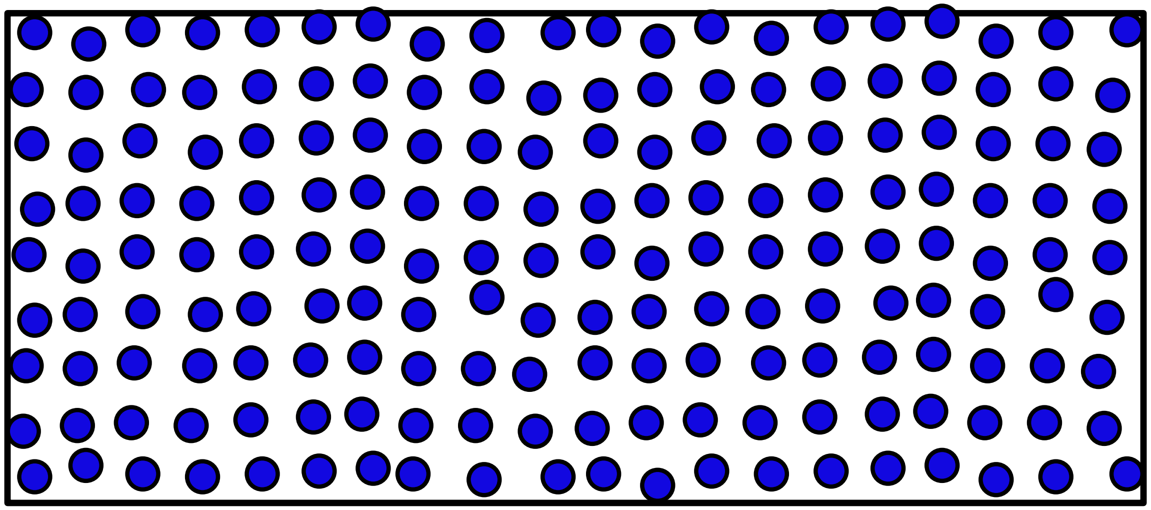
S.Y. Kim, D. Perez, and AFV, J. Chem. Phys. 139, 144110 (2013).

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



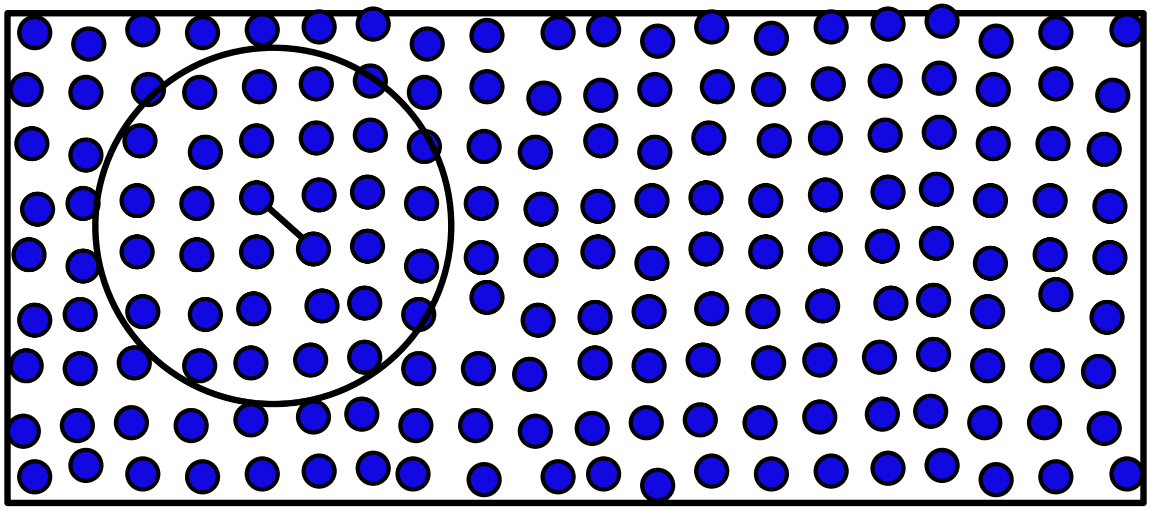
Key concept: Most systems we are interested in are intrinsically local in their behavior. A transition, or near-transition, in one region of system should not have any significant effect on atoms that are far away.

Local hyperdynamics - procedure



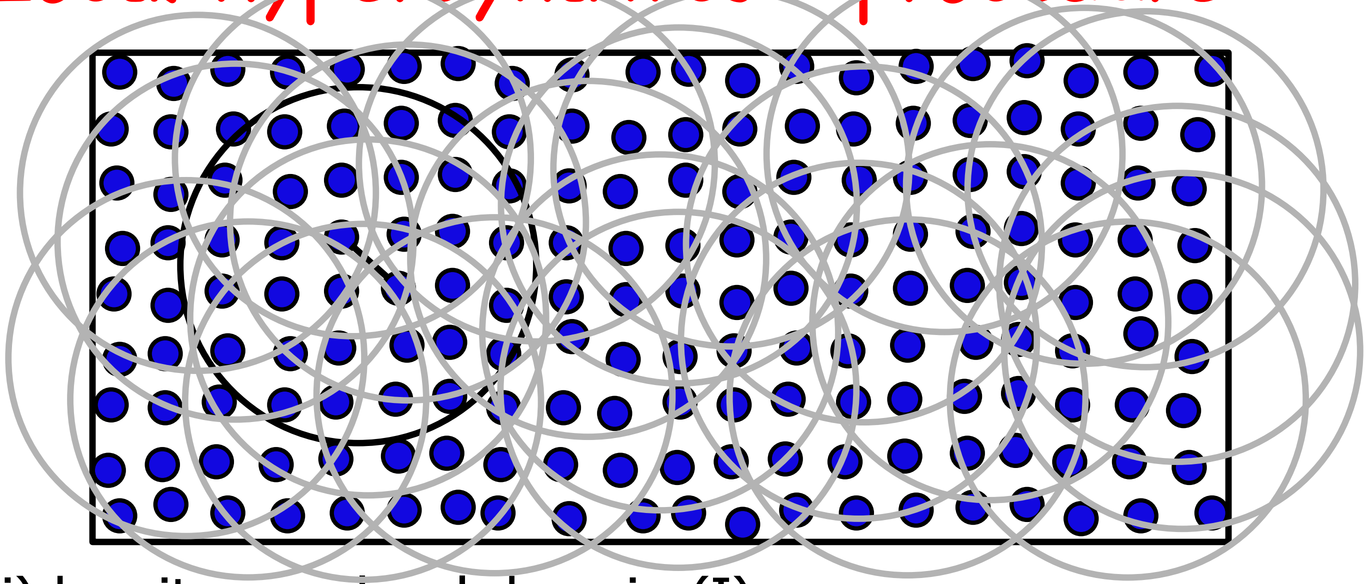
Each bond (i) has its own local domain (I).

Local hyperdynamics - procedure



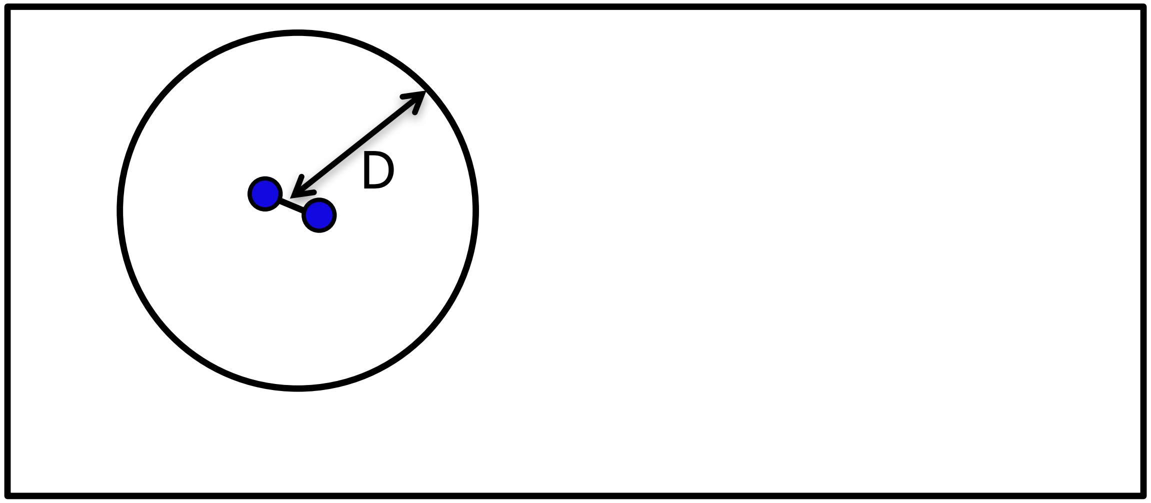
Each bond (i) has its own local domain (I).

Local hyperdynamics - procedure



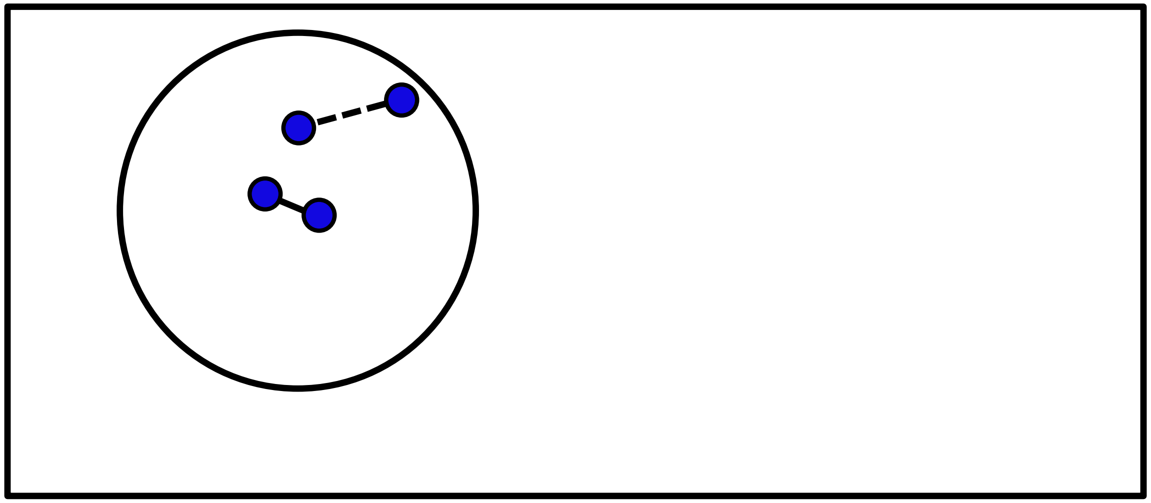
Each bond (i) has its own local domain (I).

Local hyperdynamics - procedure



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

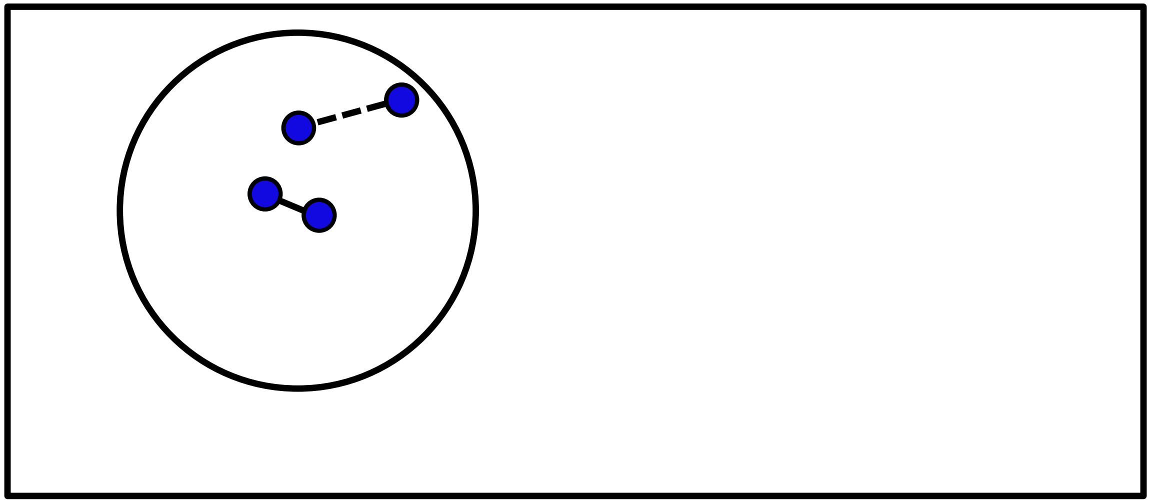
Local hyperdynamics - procedure



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

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

Local hyperdynamics - procedure

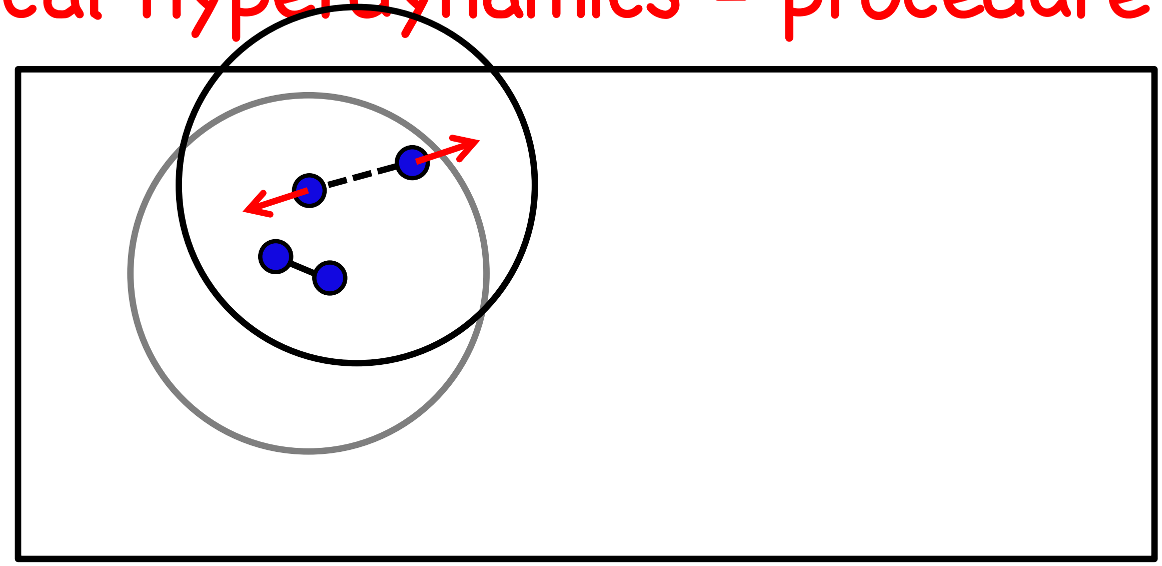


Each bond (i) has its own local domain (I).

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.

Local hyperdynamics - procedure

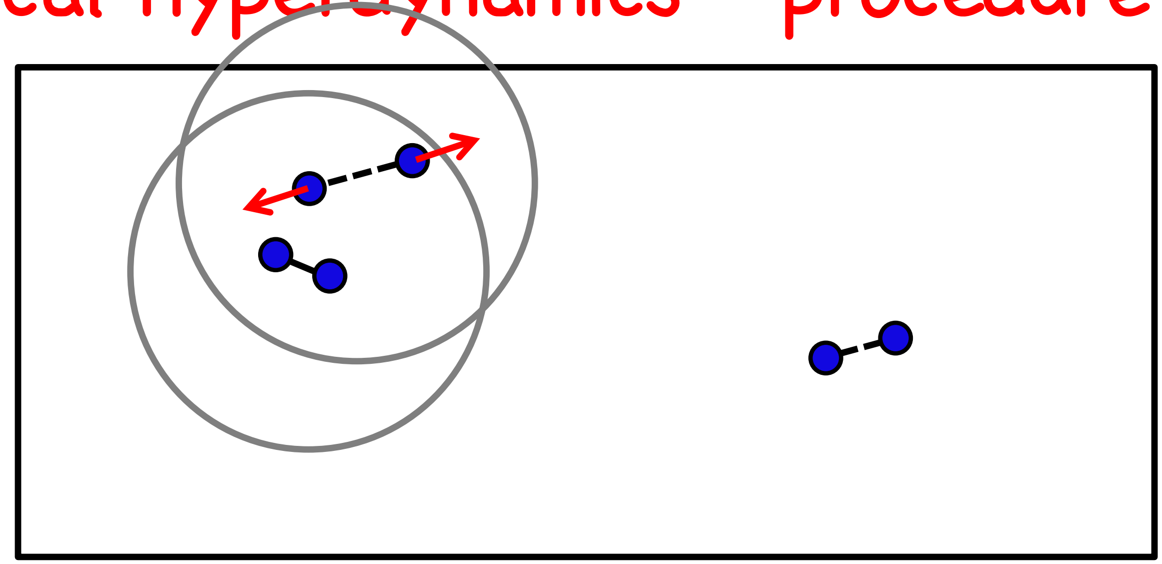


Each bond (i) has its own local domain (I).

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

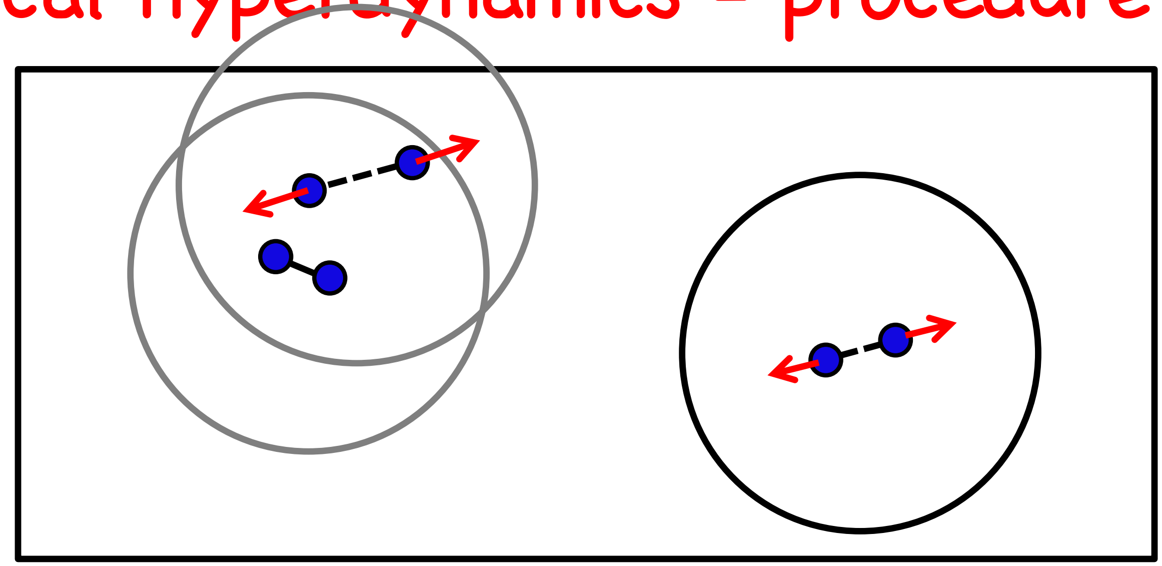


Each bond (i) has its own local domain (I).

Each domain has its own bias energy $\Delta V_I = C_I \Delta V_{bias}(R_I(t))$ 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

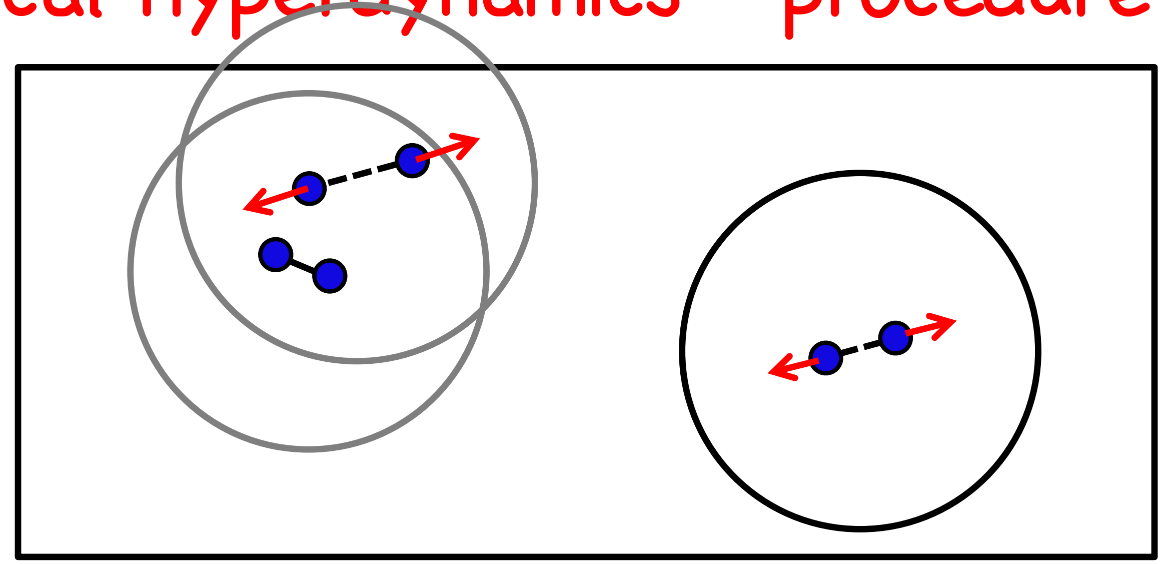


Each bond (i) has its own local domain (I).

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 $\Delta V_I = C_I \Delta V_{bias}(R_I(t))$ 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_I) is adjusted for each domain to make its average boost $\langle B_I \rangle$ match the *target boost factor* B_{target}.

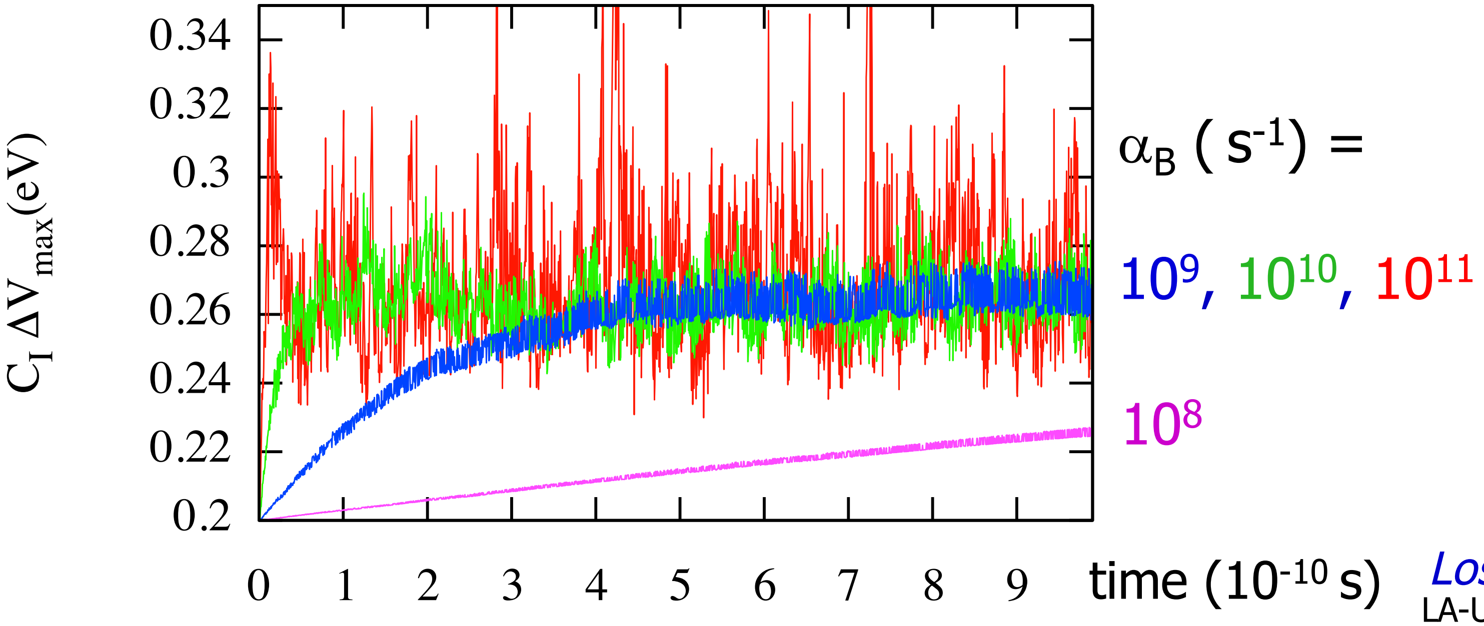
The boostostat

Instead of pre-adjusting the $\{C_I\}$ values for the current state, we 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.

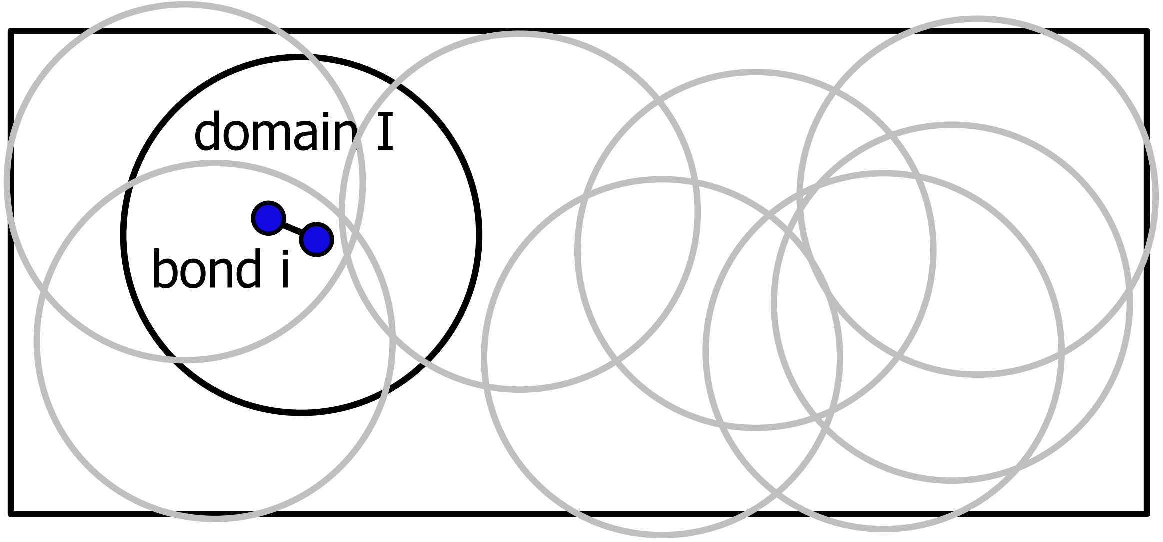
For each domain I at each MD step:

$$C_I(t+\Delta t) = C_I(t) - \alpha_B \Delta t_{MD}[B_I(t) - B_{target}]/B_{target}$$

where α_B = boostostat coupling strength $(\sim 10^9 - 10^{10}\ s^{-1})$



Note – this is not conservative dynamics



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

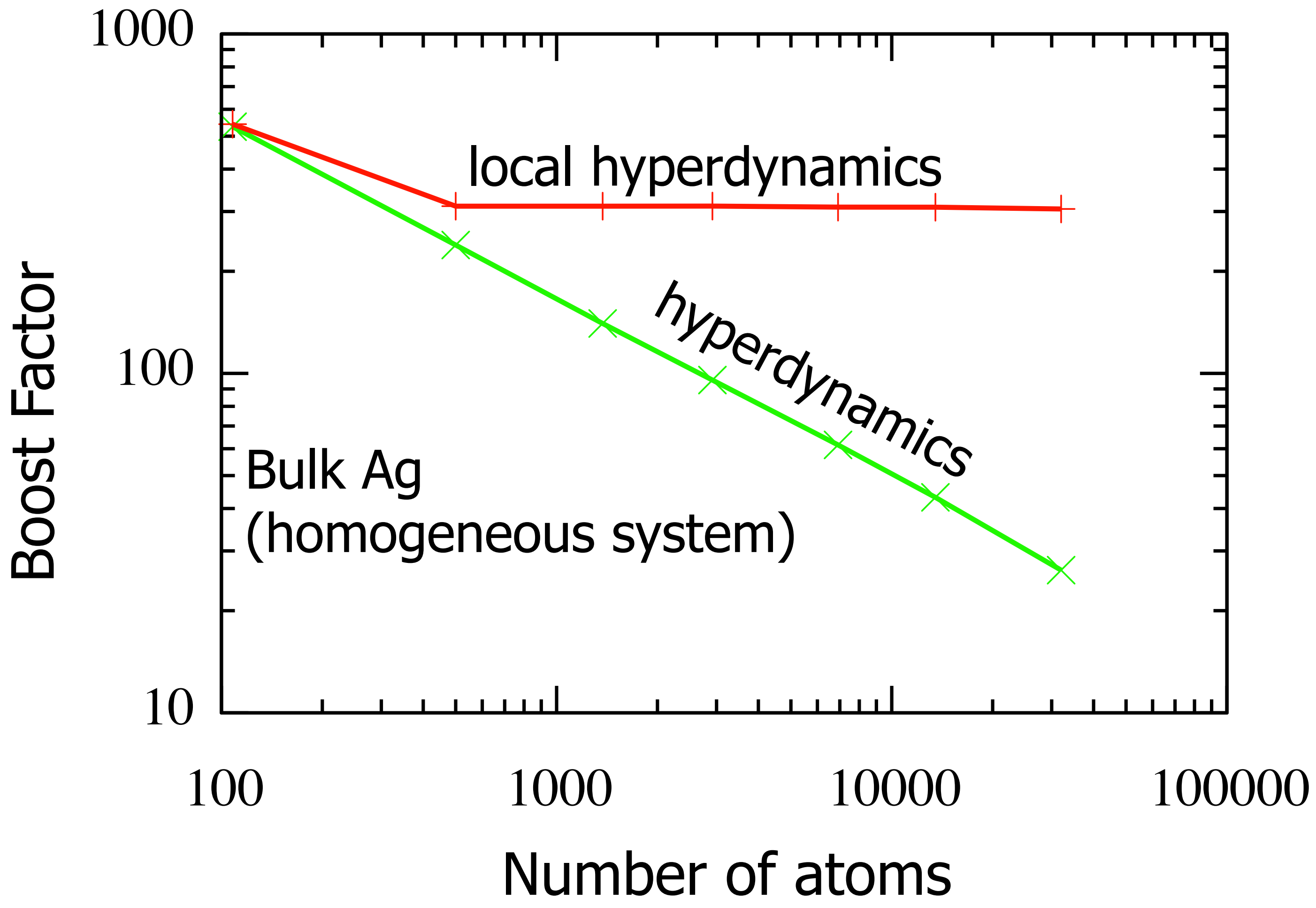
$$f_i = -\partial \Delta V_I / \partial x_i \ .$$

However, the force on nearby bond j is given by

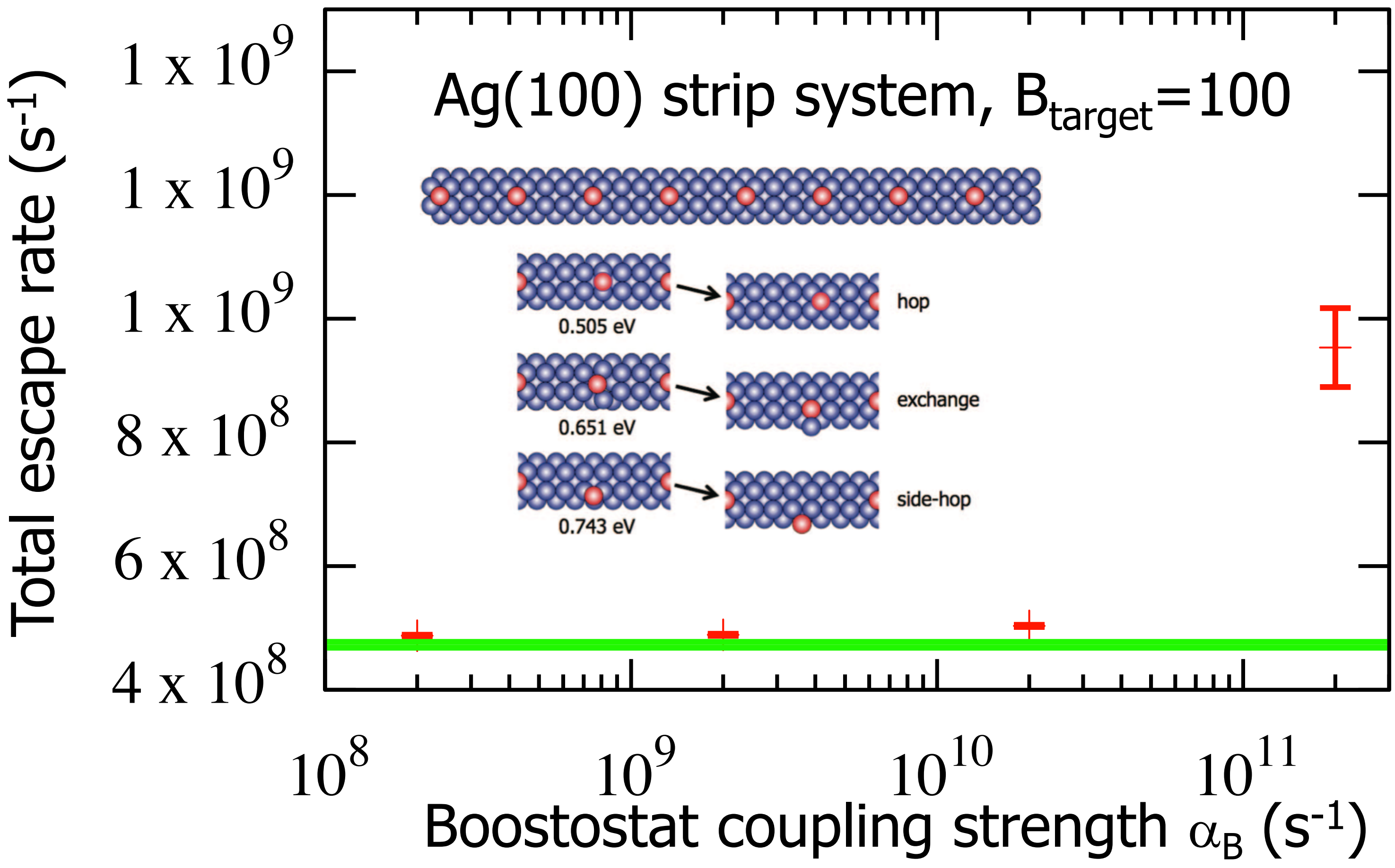
$$f_j = -\partial \Delta V_J / \partial x_j \ ,$$

so this is not conservative dynamics.

Local hyperdynamics – boost scaling



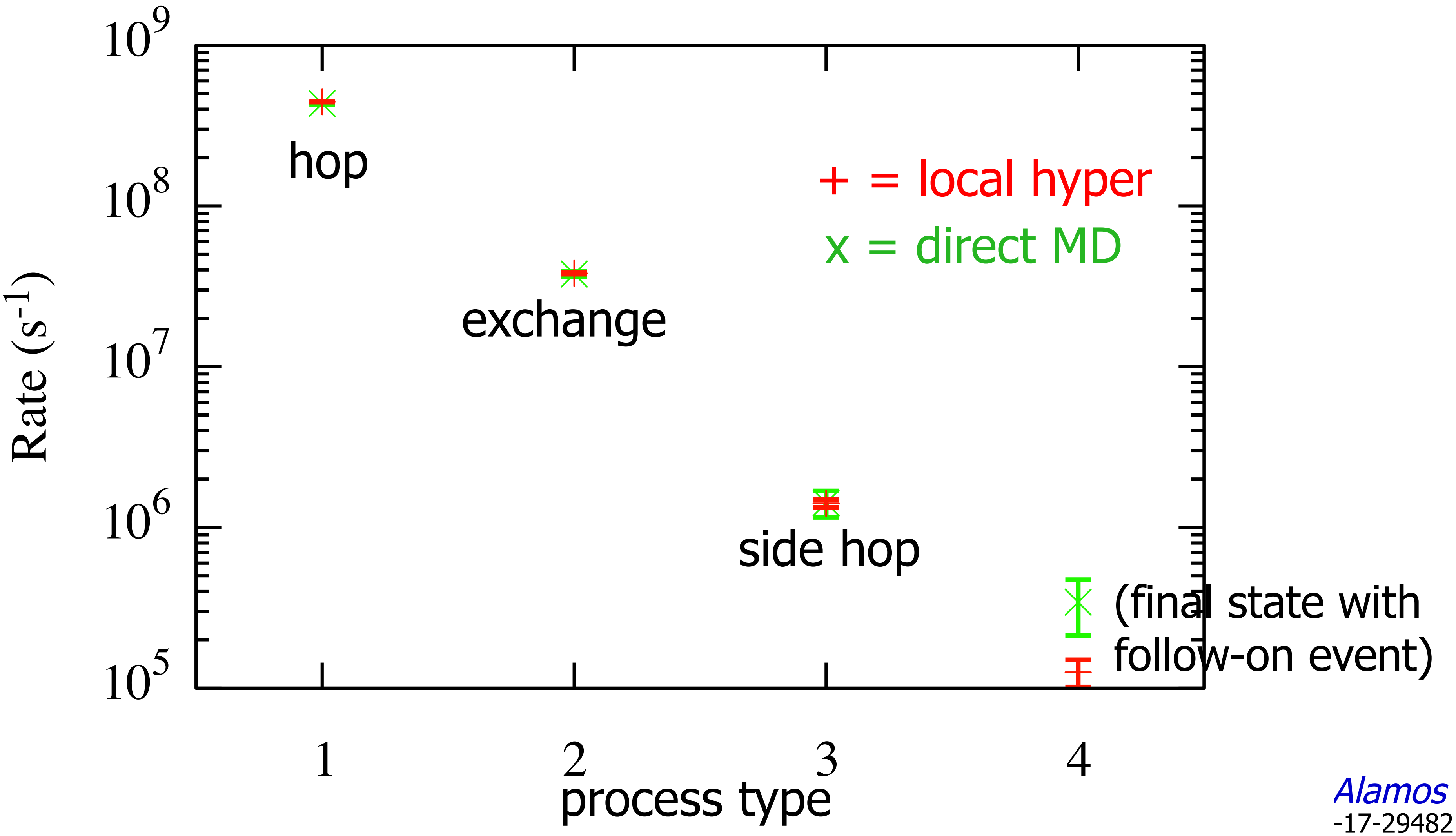
Accuracy of the rates for a test system



Individual rates also extremely accurate.

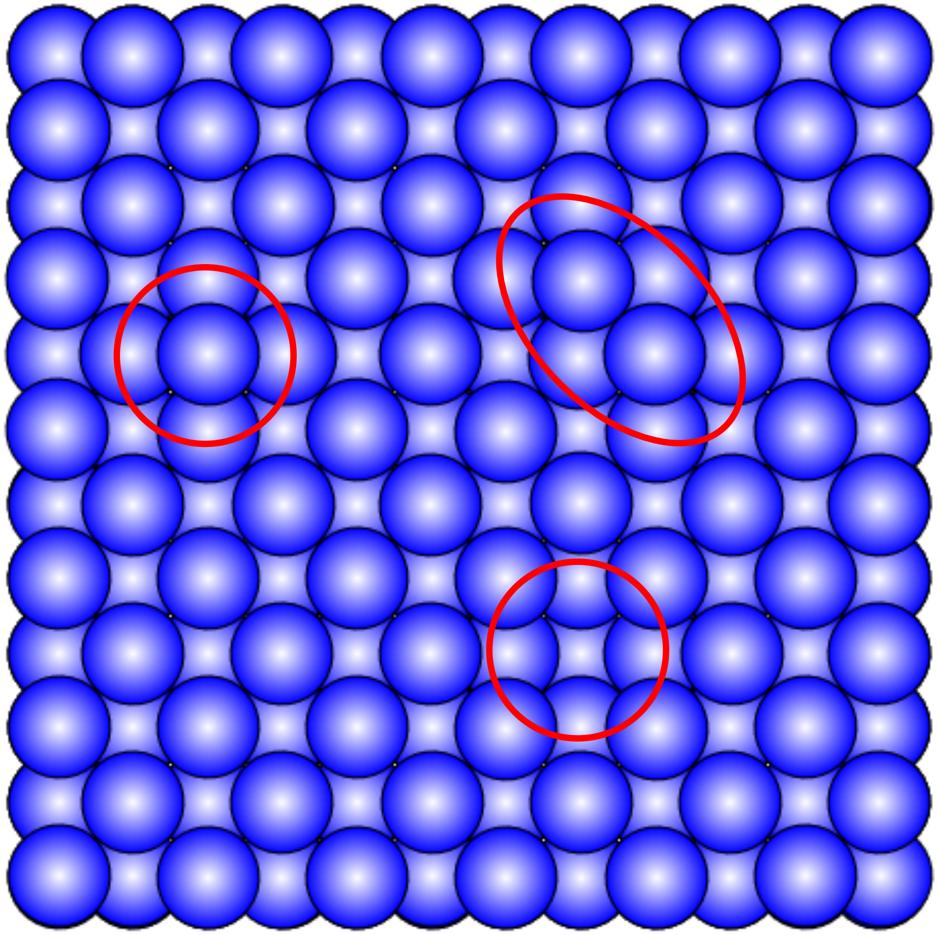
Local hyperdynamics test on strip system

T=500K, B_{target}=100, α_B=2X10^{10}, range D=10 Å

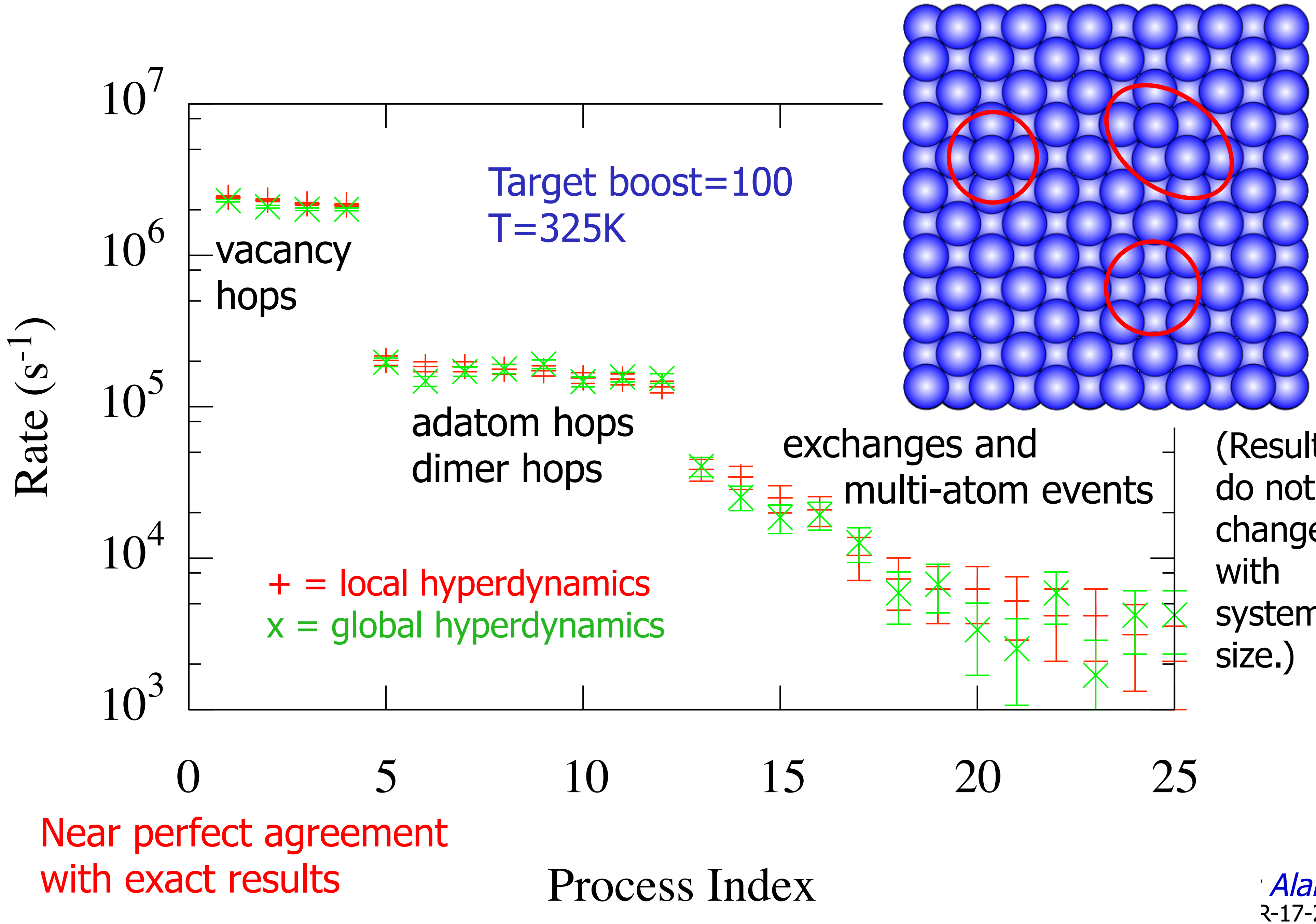


Test on more complex Ag/Ag(100) system

- Defective Ag surface: adatom, adatom dimer, vacancy
- 434 atoms, 218 moving
- T=325K, $\gamma_{\text{Langevin}} = 5\text{x}10^{11}\ \text{s}^{-1}$
- Locality radius D=10 Å
- On-the-fly boostostatting, $\alpha_B=2\text{x}10^{10}\ \text{s}^{-1}$
- $B_{\text{target}}=100$
- Coefficients $\{C_I\}$ reset locally to 0.2 eV after each transition, and system replaced in initial state (as if a transition brought it to this state). Boostostatting and MD never paused.
- About 25 thermally relevant transitions. Rates vary over three orders of magnitude.



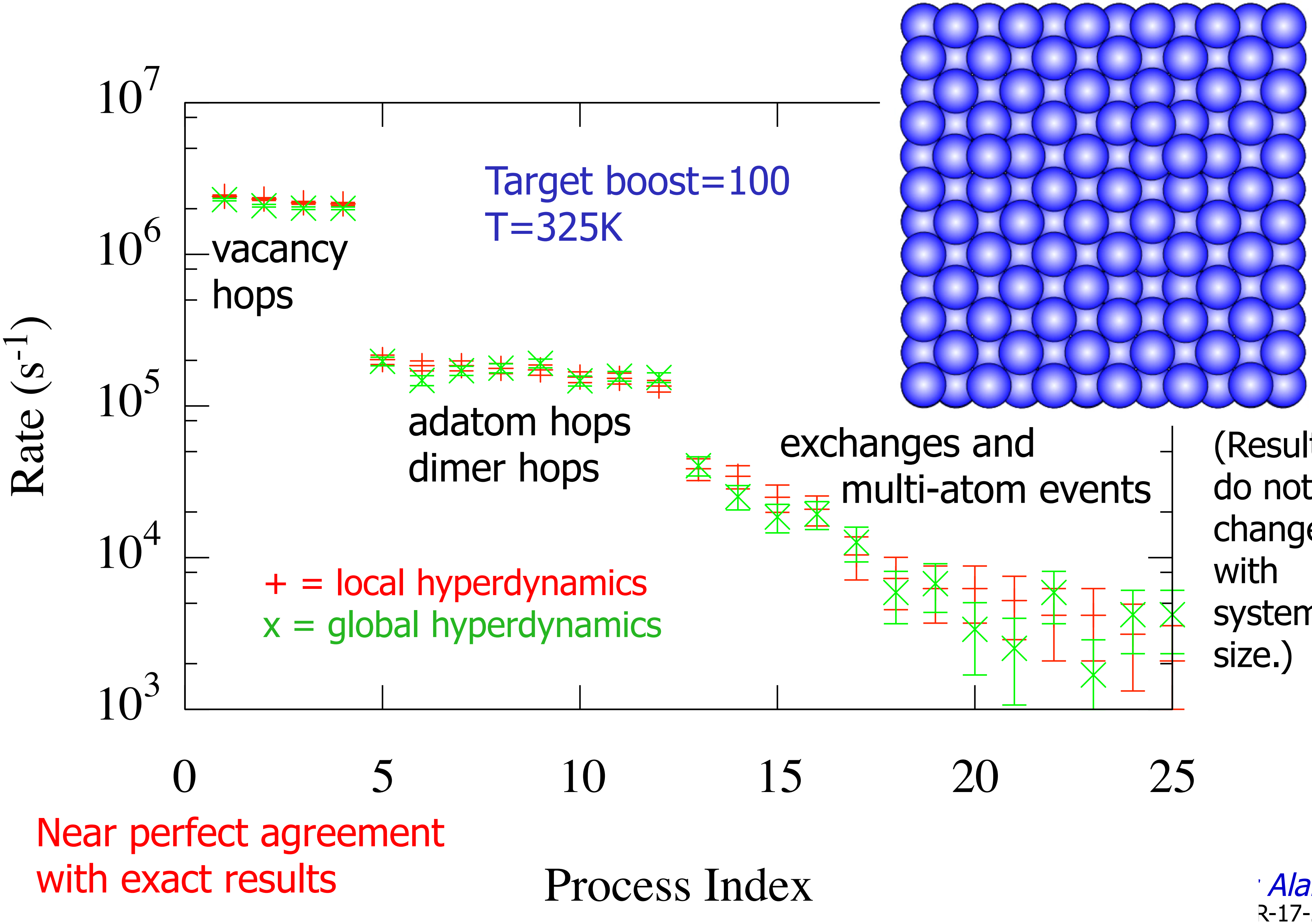
Local Hyperdynamics - Ag(100) tests



(Results
do not
change
with
system
size.)

Near perfect agreement
with exact results

Local Hyperdynamics - Ag(100) tests



(Results do not change with system size.)

Near perfect agreement with exact results

Parallel implementation

Local hyperdynamics coded into LAMMPS by Steve Plimpton

Interrupt dynamics regularly, perform full minimization, check for transitions, redefine bonds and domains as necessary.

Details:

Run 100 time steps (4 fs) of MD with local HD bias
  - bias all bonds which are max strain within D_{cut} region
  - apply boostostat to maintain target boost

Extra operations for HD:
  - 2nd neighbor list out to D_{cut}
  - double loops over atoms and big neighbor list
  - communicate to acquire bond-strain info for ghost atoms

Quench to check if event has occurred
  - if yes: re-form bond list, preserving bond C_I's

For cheap EAM, about 2x more expensive than regular MD
  - half is quench & event search every 100 steps
  - half is comp/comm to find max-strain bonds
  - this 2x shrinks for more expensive potentials

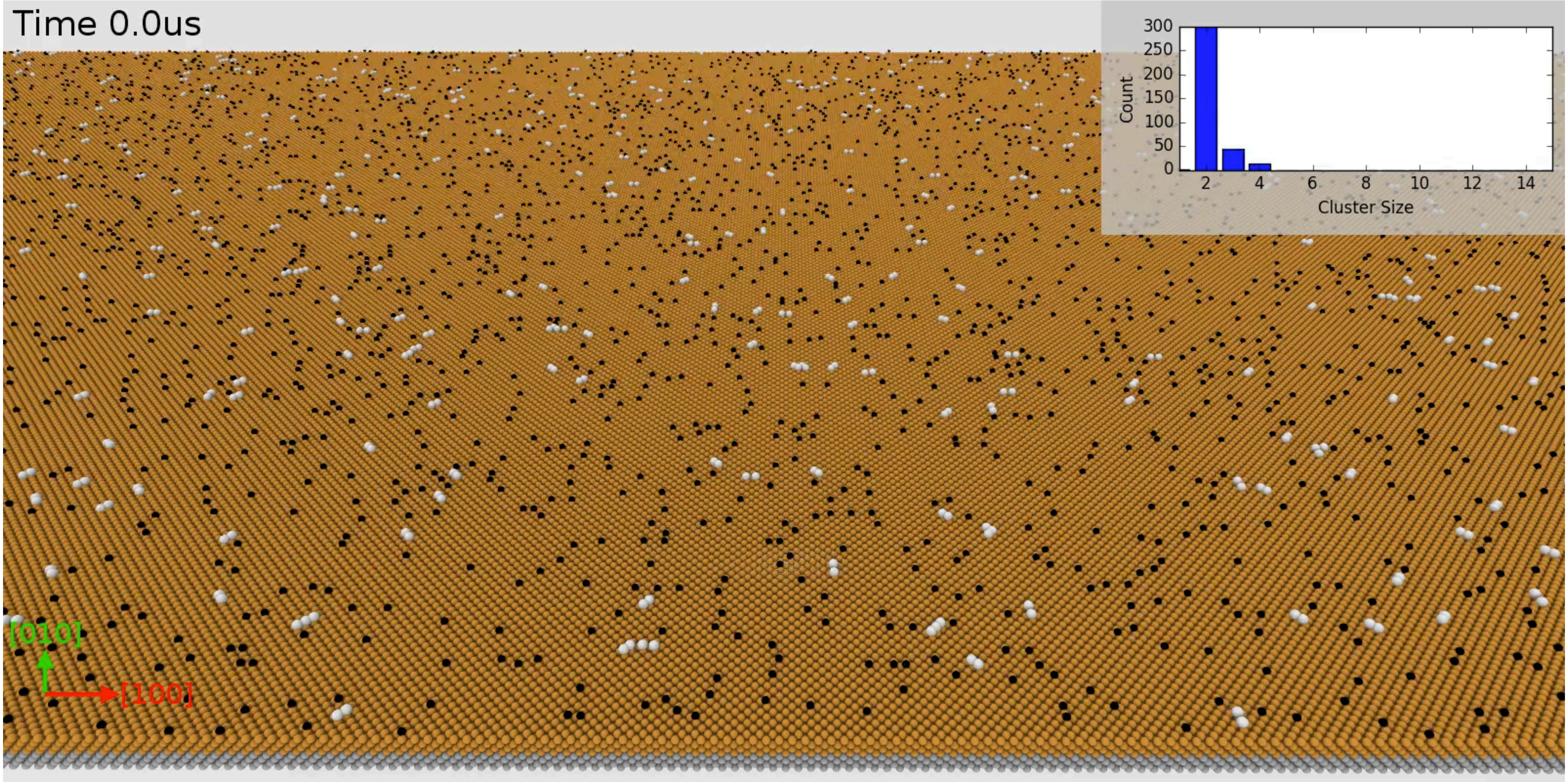
Pt/Pt(100) – 1 million atoms for 1 millisecond

4% adatom coverage, random initial placement

T=400K, 1 million atoms total, top 4 layers free

Target boost = 4000 (gives $<C_I V_{max}>$= ~0.4 eV)

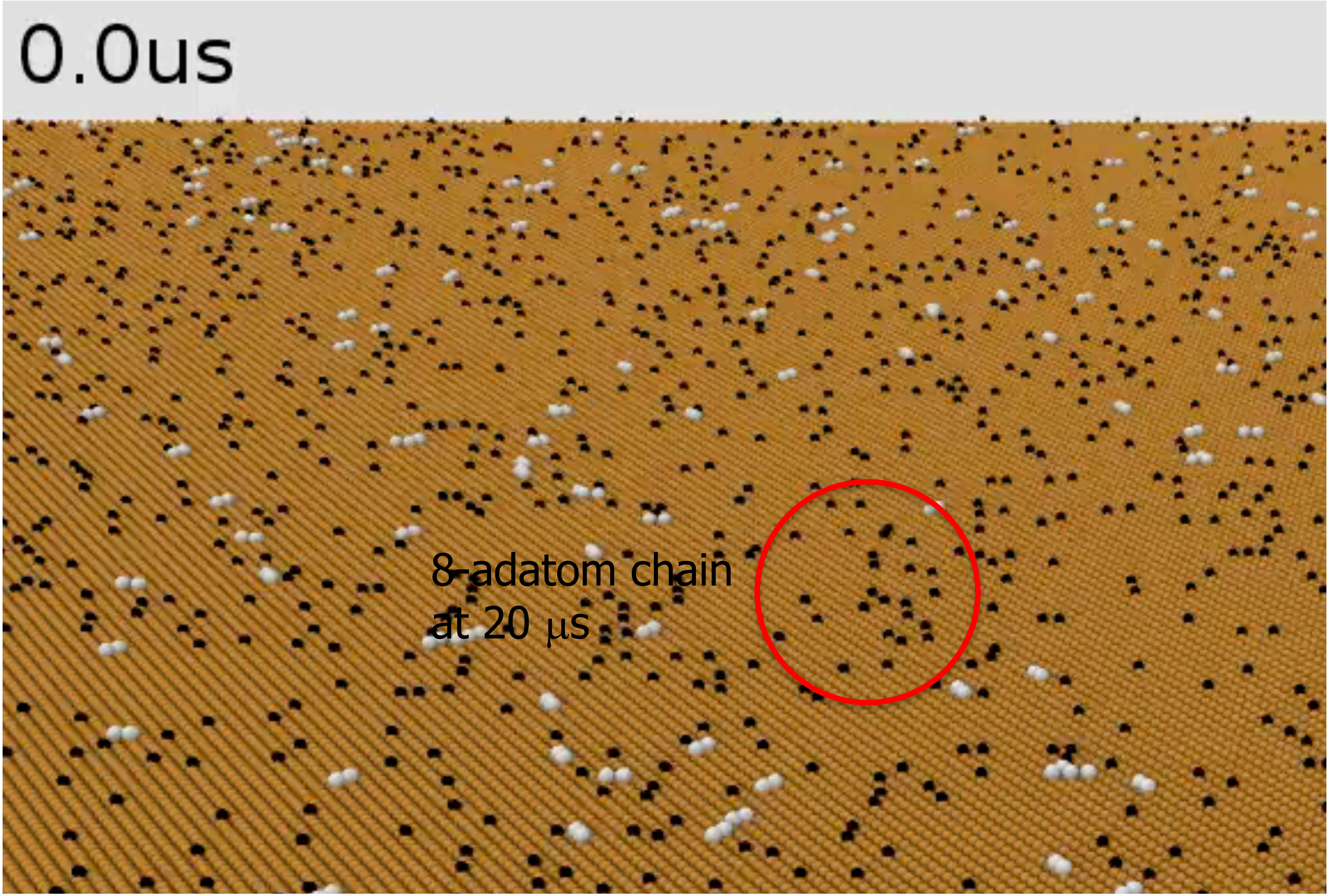
24 hrs on 128 Broadwell nodes (4096 cores, 300 atoms/core)



Steve Plimpton

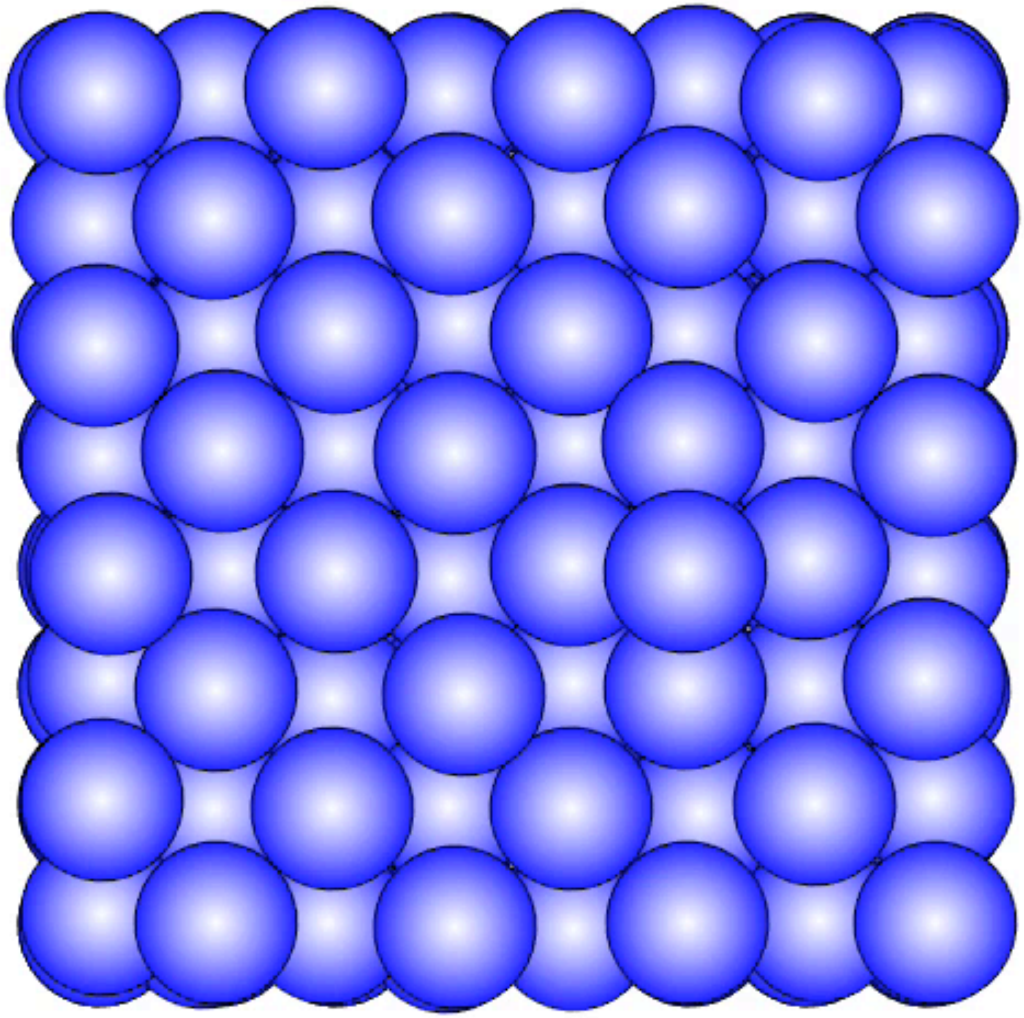
Los Alamos
LA-UR-17-29482

Pt/Pt(100)



8-adatom chain
at 20 μs

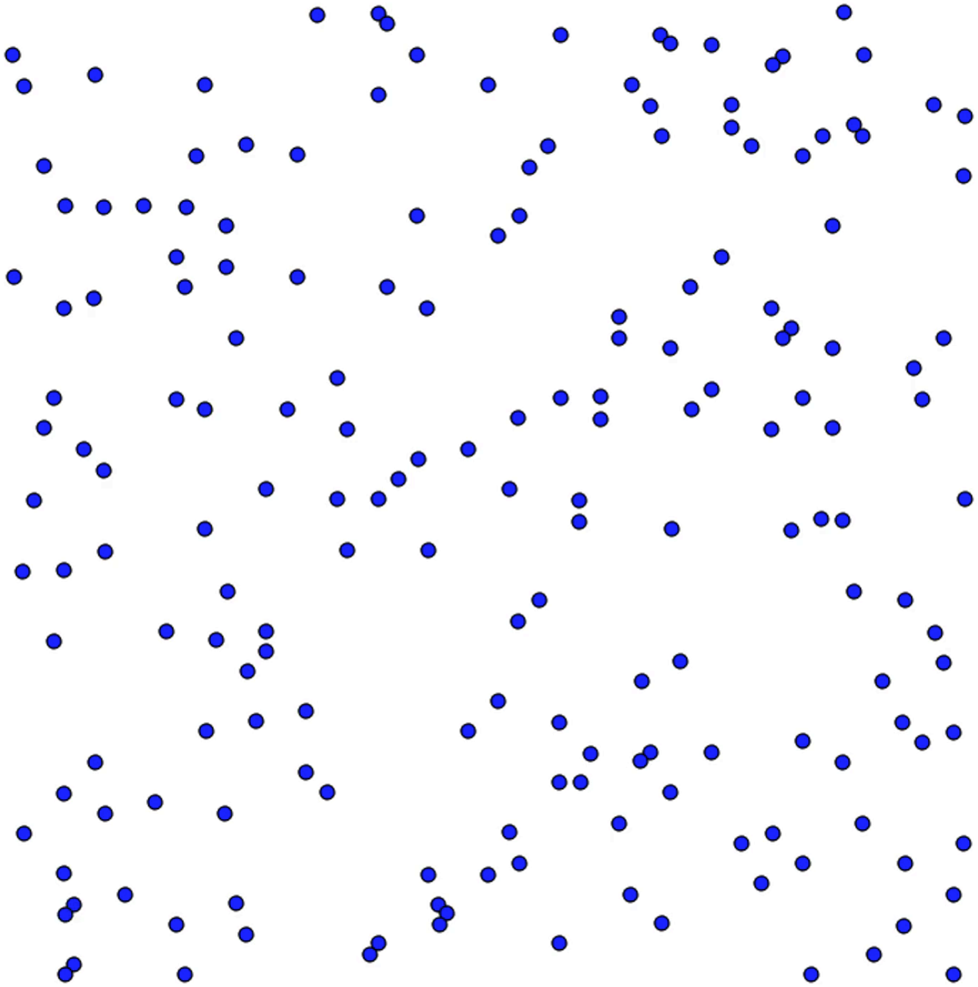
Pt/Pt(100) favors exchange over hop



hop 1.25 eV
exchange 0.64 eV

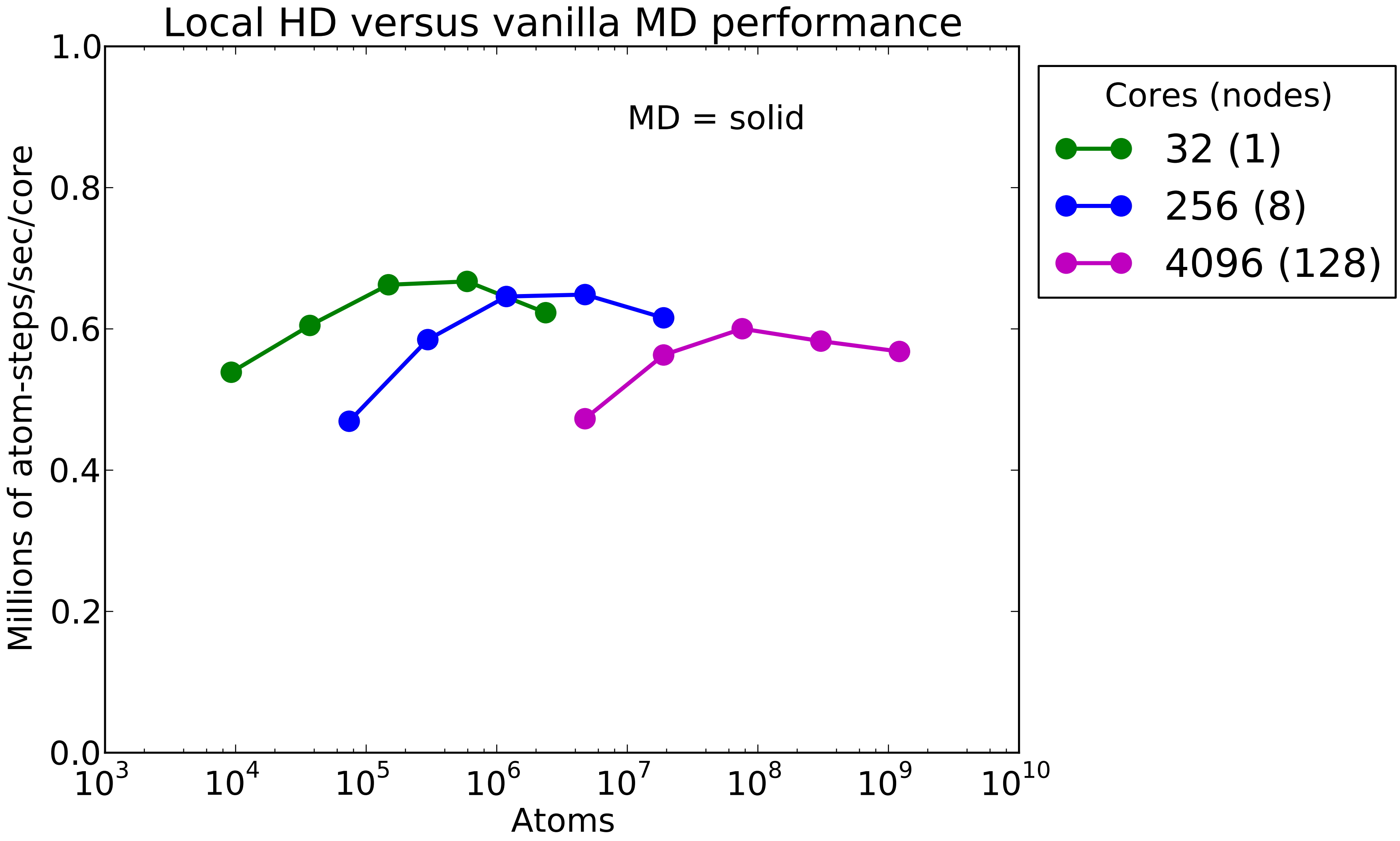
Pt/Pt(100)

individual events in first microsecond or so



Pt/Pt(100) system – computational scaling

4% adatom coverage, random initial placement
T=400K, B_{target} = 4000

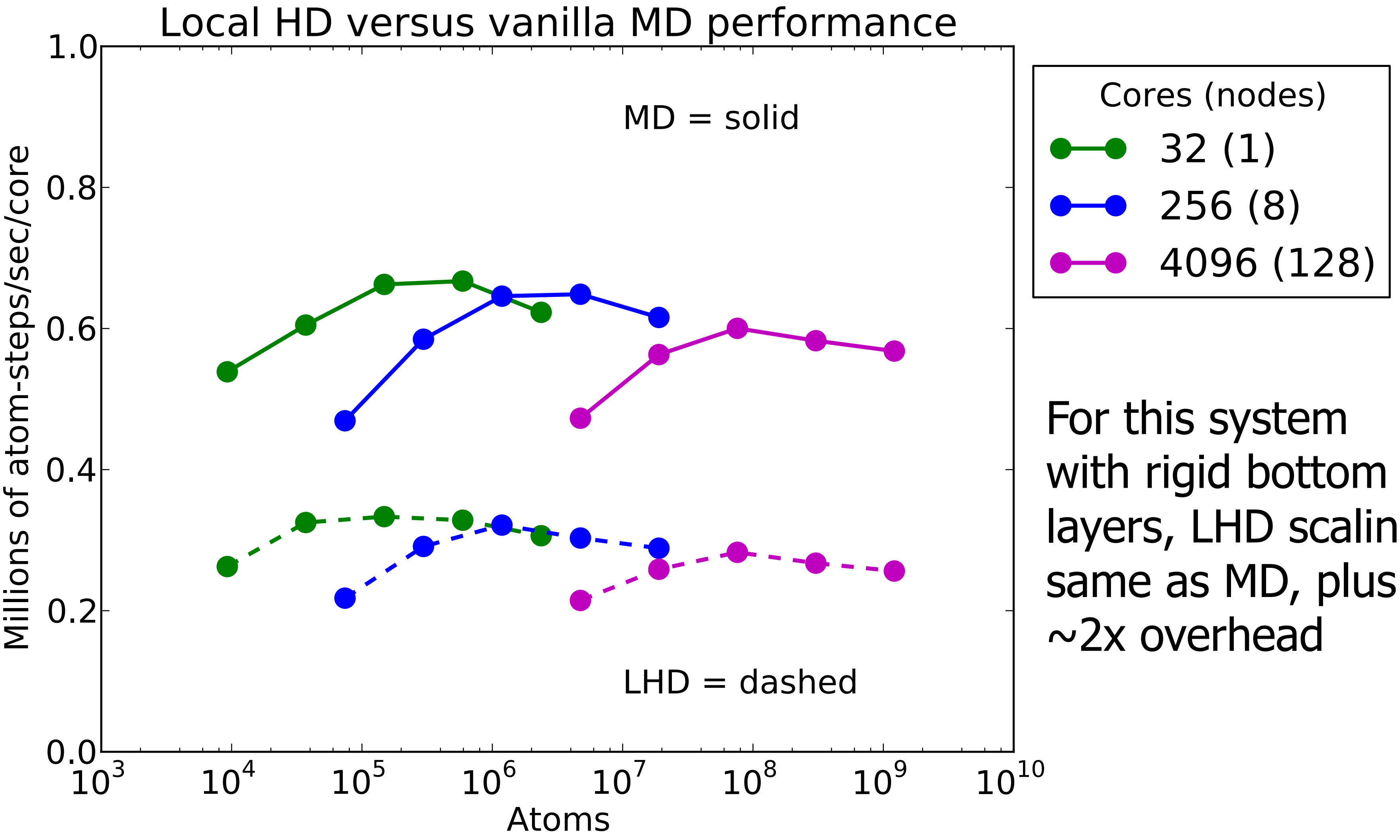


Steve Plimpton

Los Alamos
LA-UR-17-29482

Pt/Pt(100) system – computational scaling

4% adatom coverage, random initial placement
T=400K, B_{target} = 4000



For this system with rigid bottom layers, LHD scaling same as MD, plus ~2x overhead

Steve Plimpton

Los Alamos
LA-UR-17-29482

Local hyperdynamics – could it be exact?

We can show that local hyperdynamics should give increasingly accurate results as the local bias range D is increased, for any proper form of bias potential.

Local hyperdynamics – could it be exact?

We can show that local hyperdynamics should give increasingly accurate results as the local bias range D is increased, for any proper form of bias potential.

For the simple bond-boost bias potential:

We can show that for a homogeneous system (all bonds the same) the force "errors" arising from the non-conservative dynamics "cancel" (as discussed next).

For a non-homogeneous system, it is not clear whether there could be exact cancellation.

Local hyperdynamics – could it be exact?

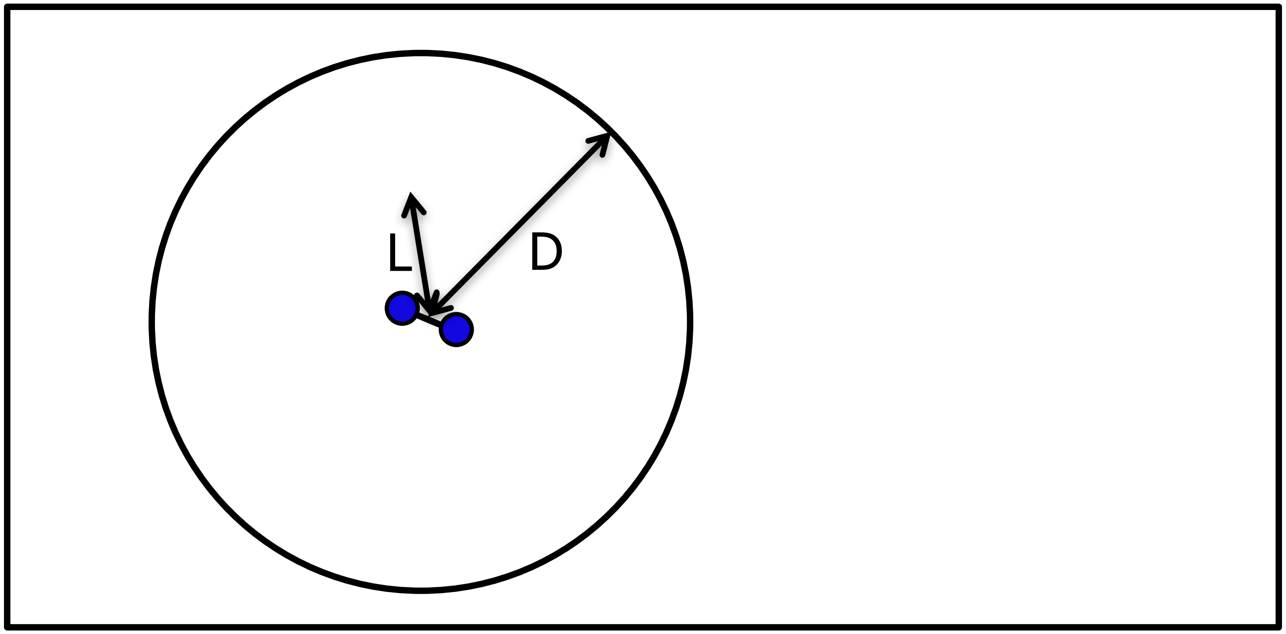
We can show that local hyperdynamics should give increasingly accurate results as the local bias range D is increased, for any proper form of bias potential.

For the simple bond-boost bias potential:

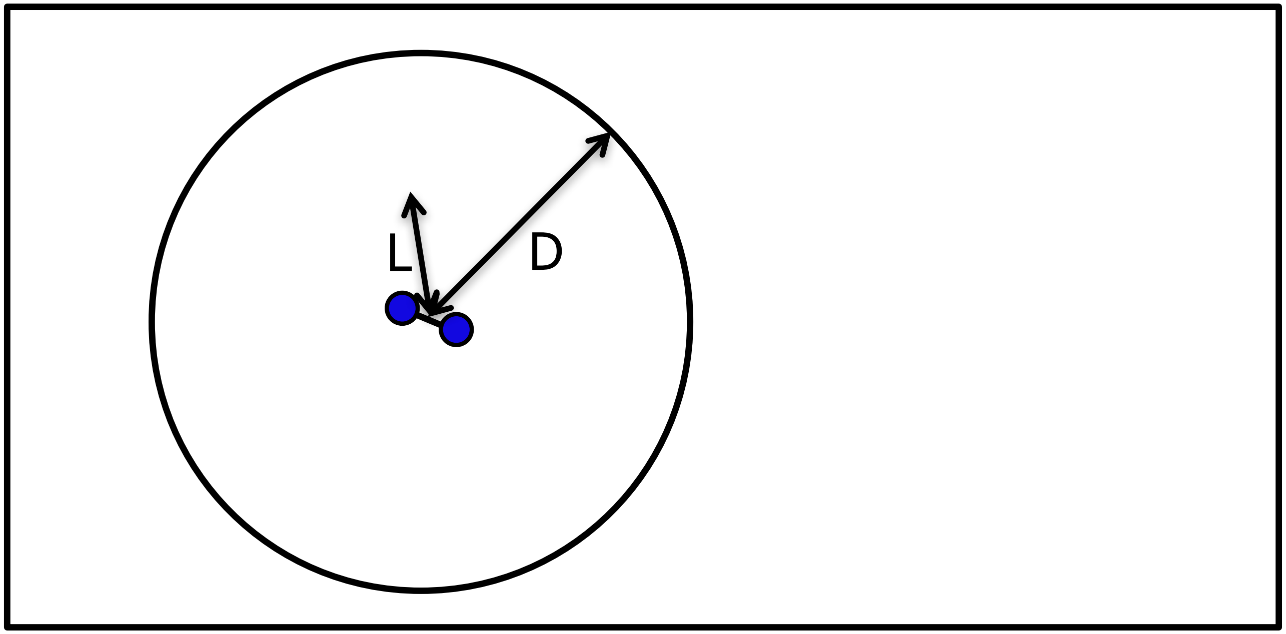
We can show that for a homogeneous system (all bonds the same) the force "errors" arising from the non-conservative dynamics "cancel" (as discussed next).

For a non-homogeneous system, it is not clear whether there could be exact cancellation.

Yet, every test we have done shows very high accuracy, making us suspect there may be something more general that can be derived about this kind of dynamics.



Assume D>2L, where L = the intrinsic locality range.

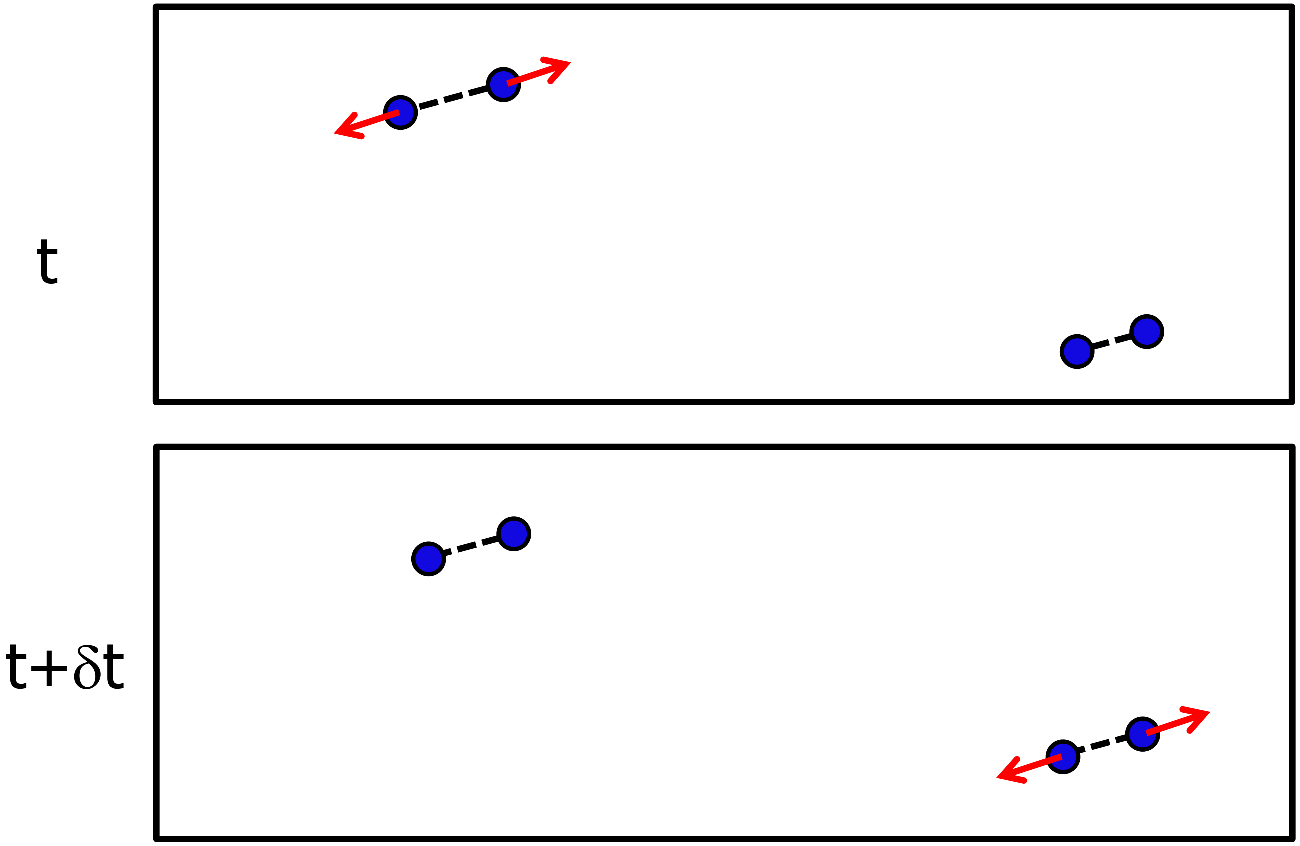


Assume D>2L, where L = the intrinsic locality range.

L is defined by the decay of correlation between bond distortions:

$$\langle \varepsilon_i \varepsilon_j \rangle = \sim 0 \quad \text{if} \quad r_{ij} > L$$

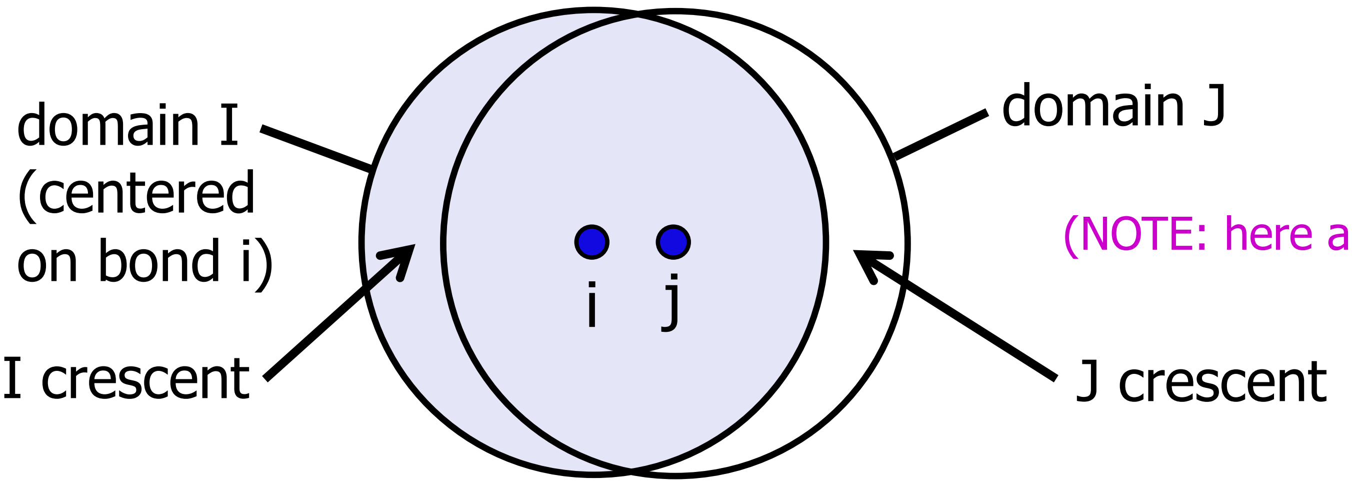
Normal (global) hyperdynamics, SBB bias



There is a discontinuity in the bias force whenever the identity of the most-distorted bond changes. Moreover, it can turn on (or off) abruptly due to the change in length of a bond that is *far away*.

In normal hyperdynamics, we find this has essentially no effect on the quality of the dynamics (the observed kinetic temperature, the boost, the rate,...)

The force mismatch terms



(NOTE: here a dot is a bond)

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

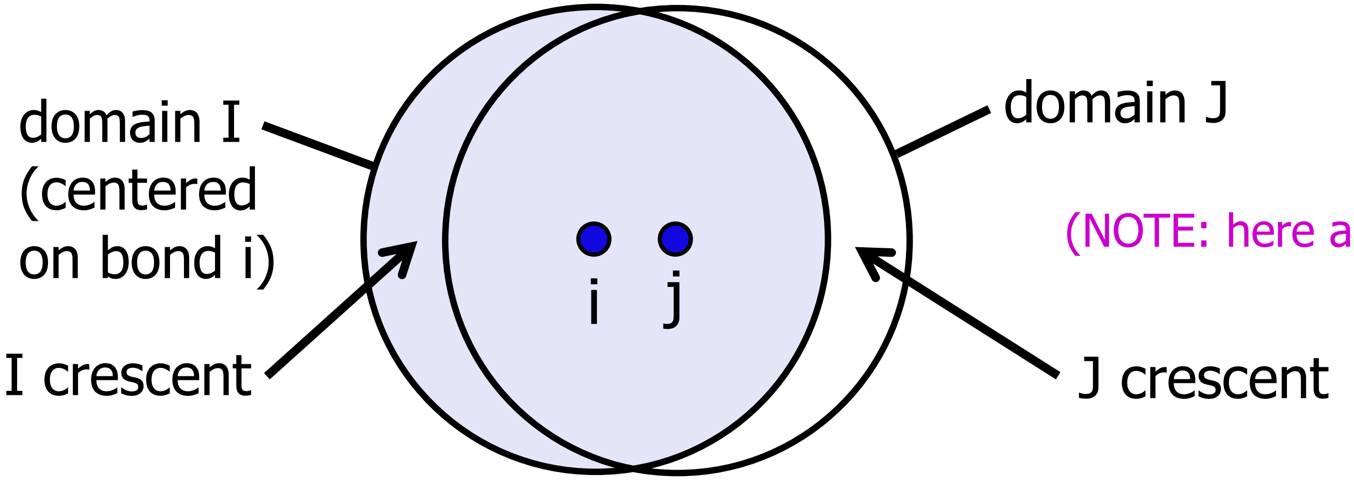
$$f_j = -\partial \Delta V_J / \partial x_j$$

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

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

The nature of this force mismatch ("error") depends on the "I crescent" and the "J crescent".

The force mismatch terms



(NOTE: here a dot is a bond)

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

$$f_j = -\partial \Delta V_J / \partial x_j$$

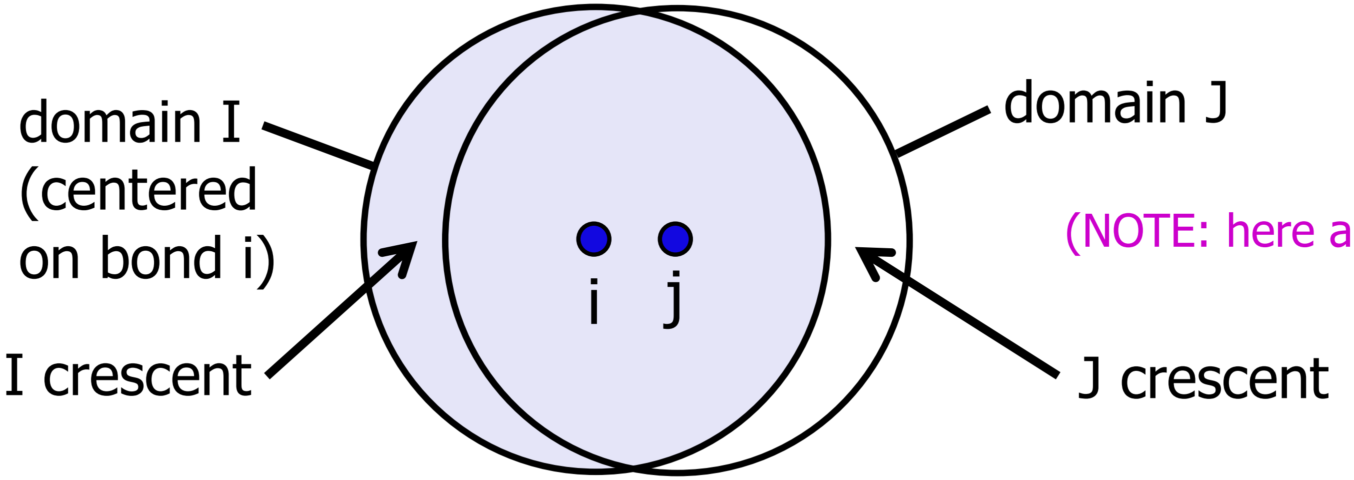
is different than the force that bond i (domain I) *thinks* that bond j will have

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

The nature of this force mismatch (“error”) depends on the “I crescent” and the “J crescent”.

If $r_{ij} < L$, the bonds in the crescents are beyond the interaction range of bond i and bond j. If all bonds in system are equivalent, these force mismatches from the domain edges come and go in a way that is simply a reordering in time relative to what they would be if domain I was the same as domain J.

The force mismatch terms


domain I (centered on bond i)

I crescent

i

j

domain J

J crescent

(NOTE: here a dot is a bond)

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

$$f_j = -\partial \Delta V_J / \partial x_j$$

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

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

The nature of this force mismatch (“error”) depends on the “I crescent” and the “J crescent”.

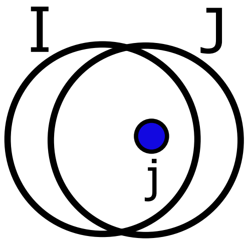
If $r_{ij}<L$, the bonds in the crescents are beyond the interaction range of bond i and bond j. If all bonds in system are equivalent, these force mismatches from the domain edges come and go in a way that is simply a reordering in time relative to what they would be if domain I was the same as domain J.

The domain crescent cases when $C_I=C_J$

(for simple bond boost bias)

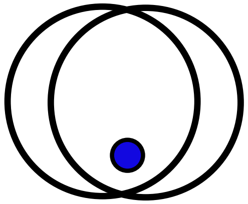
case

1



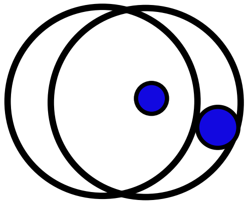
Bond j is the most distorted bond in domain I+J
→ exact $(f_j^{(I)}=f_j) = -C_J\partial\Delta V_{bias}(R_J(t))/\partial x_j$

2



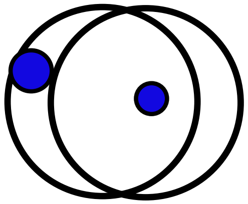
Some other bond k in $I\wedge J$ is the most distorted in domain I+J
→ exact $(f_j^{(I)}=f_j=0)$

3



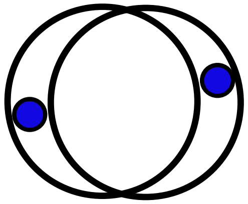
Bond j is the most distorted in domain I, but there is an even more distorted bond in the J crescent (that j sees but i does not see)
→ "error" $(f_j^{(I)}\neq 0,\ f_j=0)$

4



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)
→ "error" $(f_j^{(I)}=0,\ f_j\neq 0)$

5



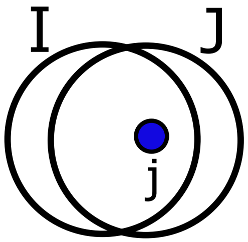
The most distorted bond in domain I is in the I crescent and most distorted bond in domain J is in the J crescent
→ exact $(f_j^{(I)}=f_j=0)$

The domain crescent cases when $C_I=C_J$

case

(for simple bond boost bias)

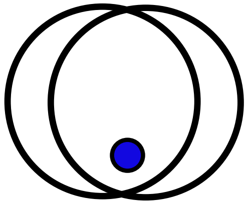
1



Bond j is the most distorted bond in domain I+J

→ exact ($f_j^{(I)}=f_j$) = $-C_J\partial\Delta V_{bias}(R_J(t))/\partial x_j$

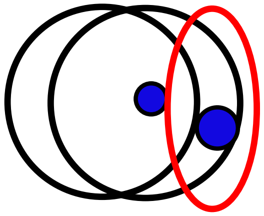
2



Some other bond k in $I\wedge J$ is the most distorted in domain I+J

→ exact ($f_j^{(I)}=f_j=0$)

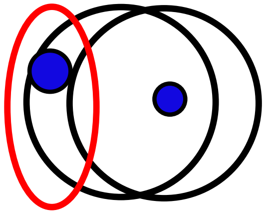
3



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)

→ "error" ($f_j^{(I)}\neq 0$, $f_j=0$)

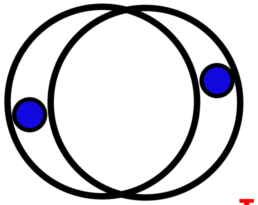
4



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)

→ "error" ($f_j^{(I)}=0$, $f_j\neq 0$)

5



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

→ exact ($f_j^{(I)}=f_j=0$)

If I-crescent and J-crescent are statistically equivalent, then C_I = C_J, and "error" 3 and "error" 4 balance in a time average

The core argument – time reordering

Assume D > 2L and r_{ij}<L. Then all bonds in both crescents are more than a distance L away from both bond *i* and bond *j*. Thus, crescent bonds do not interact directly with *i* or *j*.

Large bond distortions come and go in the I crescent, dictating, from a large distance away, what bond *i* thinks the force on bond *j* should be. Meanwhile, the *actual force* on bond *j* is dictated by the (distant) bonds in crescent J.

Because the bonds in crescent J are identical to the bonds in crescent I (assumed homogeneous system), the bond-length fluctuations in crescent J that can shut down bond *j* are identical to the bond-length fluctuations in crescent I, *other than a random reordering in time*.

The core argument – time reordering

Assume D > 2L and r_{ij}<L. Then all bonds in both crescents are more than a distance L away from both bond *i* and bond *j*. Thus, crescent bonds do not interact directly with *i* or *j*.

Large bond distortions come and go in the I crescent, dictating, from a large distance away, what bond *i* thinks the force on bond *j* should be. Meanwhile, the *actual force* on bond *j* is dictated by the (distant) bonds in crescent J.

Because the bonds in crescent J are identical to the bonds in crescent I (assumed homogeneous system), the bond-length fluctuations in crescent J that can shut down bond *j* are identical to the bond-length fluctuations in crescent I, *other than a random reordering in time*.

A force-expectation discrepancy can arise only when there is no force on bond *i*. Thus, the abrupt turning-on and turning-off of the force on bond j, which also happens in regular hyperdynamics (though less often), will still give appropriate Langevin evolution of bond *j* from the point of view of bond *i,* whether or not the sequence is reordered in time. Bond *i* has no way of telling the difference. *Locally*, then, for all *j* bonds within a range D/2 of bond *i,* the evolution is equivalent to true Langevin evolution on V+ΔV, and should give the same boost factor for that region as a global bias on domain I would give.

Probing for accuracy problems in local hyper

We have tested various inhomogeneous systems and find the results are still very accurate.

The errors we do see are so small that they might be caused by setting the range D too small, or setting the target boost too high, or the integrator time step too long. It is hard to rule out the possibility that LHD is "exact" (for D>2L), or accurate to high order.

Probing for accuracy problems in local hyper

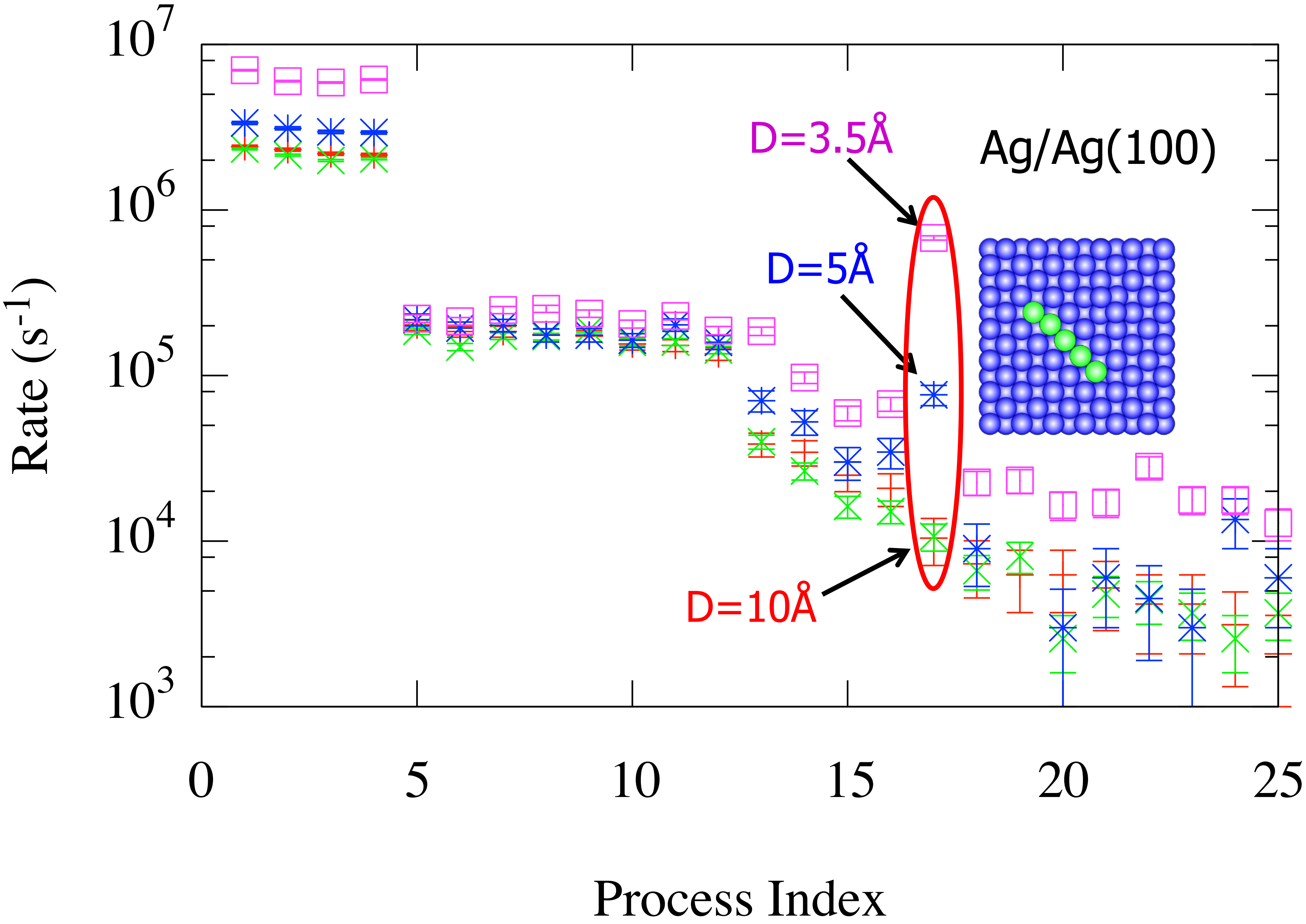
We have tested various inhomogeneous systems and find the results are still very accurate.

The errors we do see are so small that they might be caused by setting the range D too small, or setting the target boost too high, or the integrator time step too long. It is hard to rule out the possibility that LHD is “exact” (for D>2L), or accurate to high order.

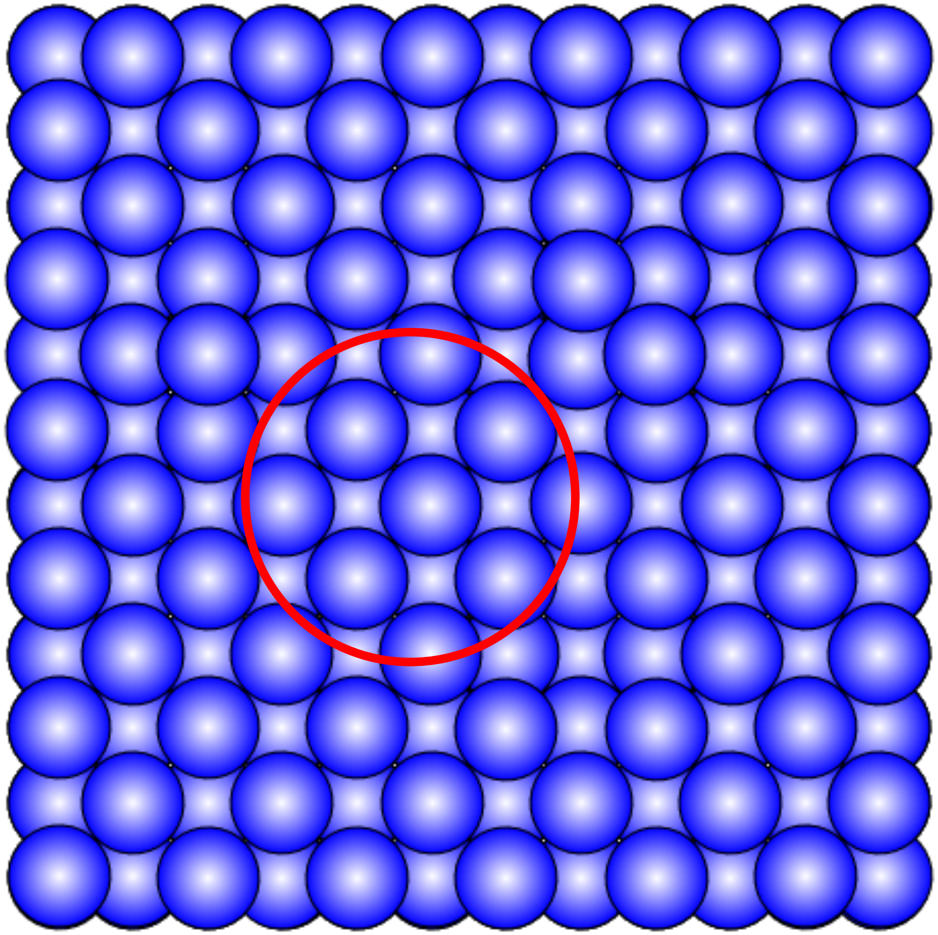
We have found that it is difficult to even create a test case that can clearly distinguish right from wrong (e.g., to use in developing a more accurate variation on local hyperdynamics).

Behavior when locality range is underestimated

Long-range mechanisms too large for D go *faster*



Safely "self correcting" if D is too small

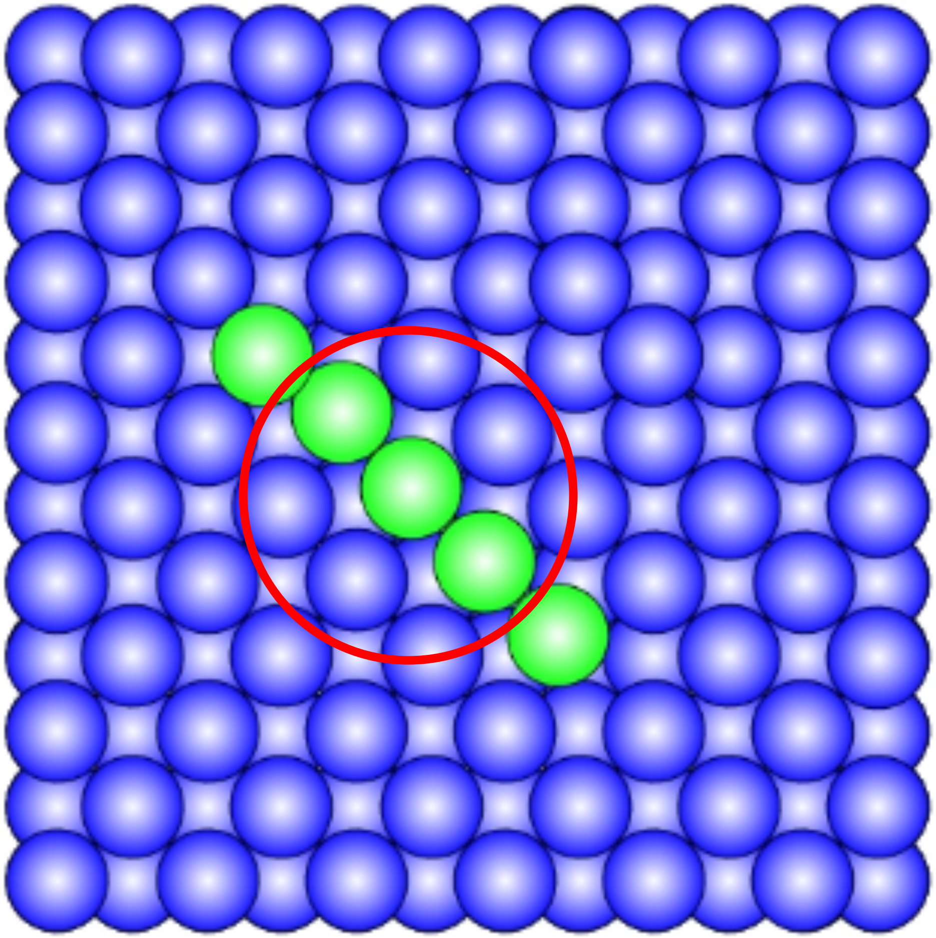


LHD run using D=5Å

D = 5 Å

Safely "self correcting" if D is too small

D = 5 Å

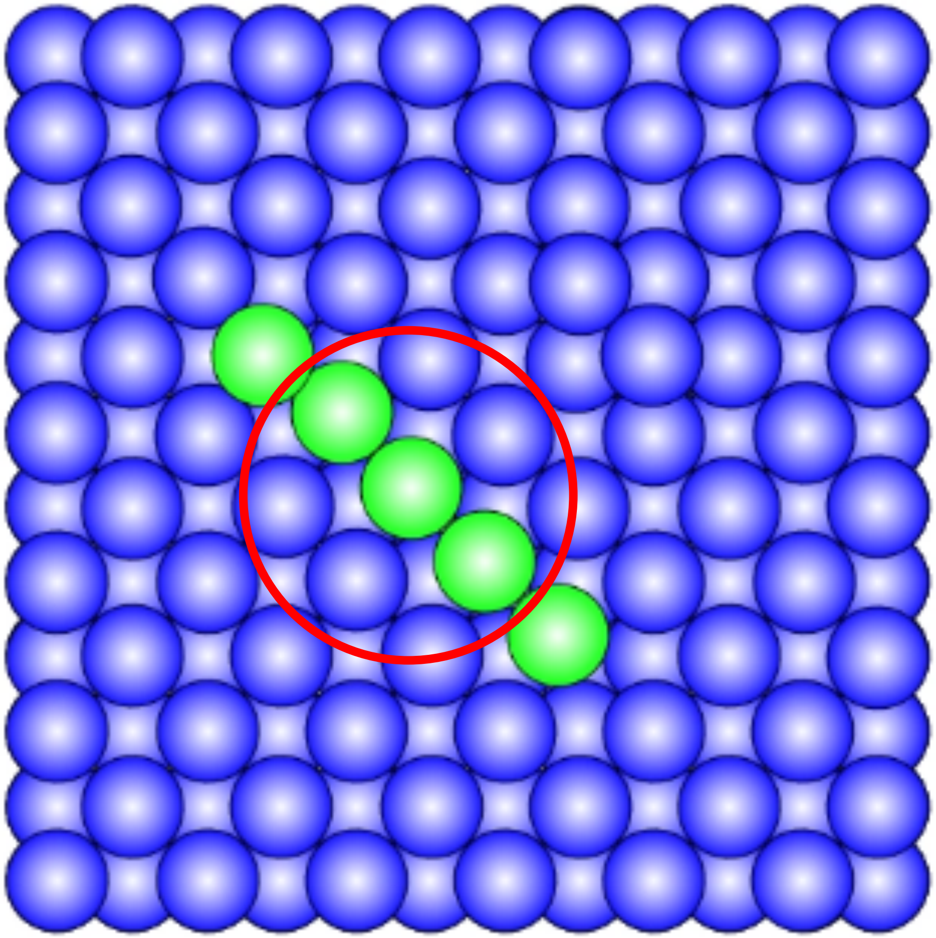


LHD run using D=5Å

This event shows up *sooner* than it should have.

Safely "self correcting" if D is too small

D = 5 Å

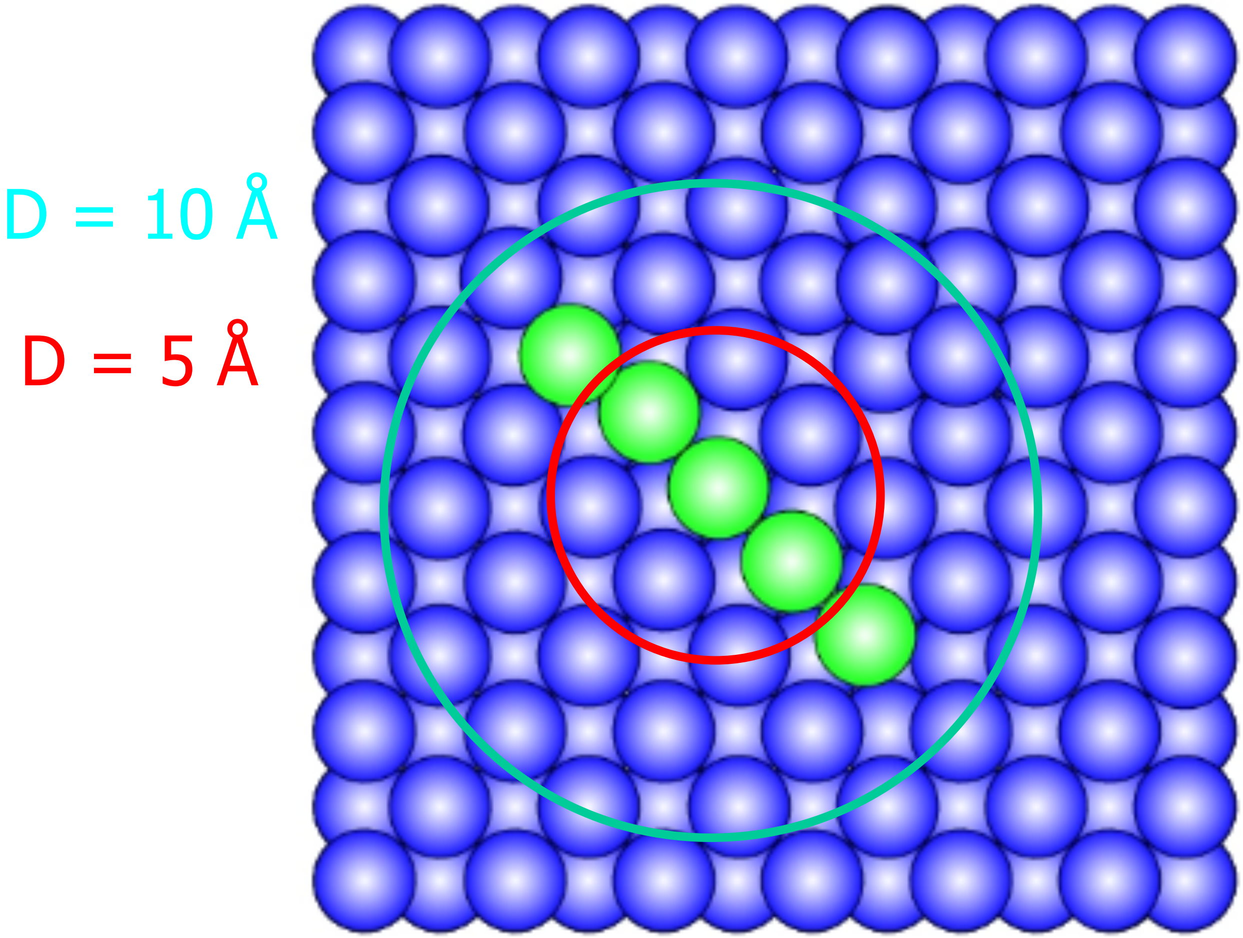


LHD run using D=5Å

This event shows up *sooner* than it should have.

But evolution so far (not counting this event) is correct.

Safely "self correcting" if D is too small



LHD run using D=5Å

This event shows up *sooner* than it should have.

But evolution so far (not counting this event) is correct.

Just discard this event, increase D, and continue running.

Conclusions – Local Hyperdynamics

A local formulation of hyperdynamics is scalable, with constant boost, to arbitrarily large systems, provided the lowest barrier in system does not decrease with increasing system size.

Parallel implementation (now in LAMMPS) looks very promising.

We understand why should be correct for homogeneous systems.

We are surprised how accurate it is for *every* system we have tried, homogeneous or not. Perhaps the inhomogeneous case can be shown to be correct as well…

A remaining issue: adjusting B_{target} on fly if lower barriers show up.

S.Y. Kim, D. Perez, and AFV, J. Chem. Phys. 139, 144110 (2013).