Lifted Markov chains: from solvable models to applications in chemical physics

Werner Krauth
Department of Physics, Ecole normale supérieure Paris, France
Rudolf Peierls Centre for Theoretical Physics, University of Oxford, UK
Simons Center for Computational Physical Chemistry, New York University, USA

10 May 2024
Center for Computational Mathematics
Flatiron Institute, New York (NY) USA

This work was supported by a grant from the Simons Foundation

This work was supported by the Alexander von Humboldt Foundation

Département de Physique
École normale supérieure

Markov-chain Monte Carlo

URNAL OF CHEMICAL PHYSICS VOLUME 21, NUMBER 6 JUN

Equation of State Calculations by Fast Computing Machines

NICHOLAS METROPOLIS, ARIANNA W. ROSENBLUTH, MARSHALL N. ROSENBLUTH, AND AUGUSTA H. TELLER,
Los Alamos Scientific Laboratory, Los Alamos, New Mexico

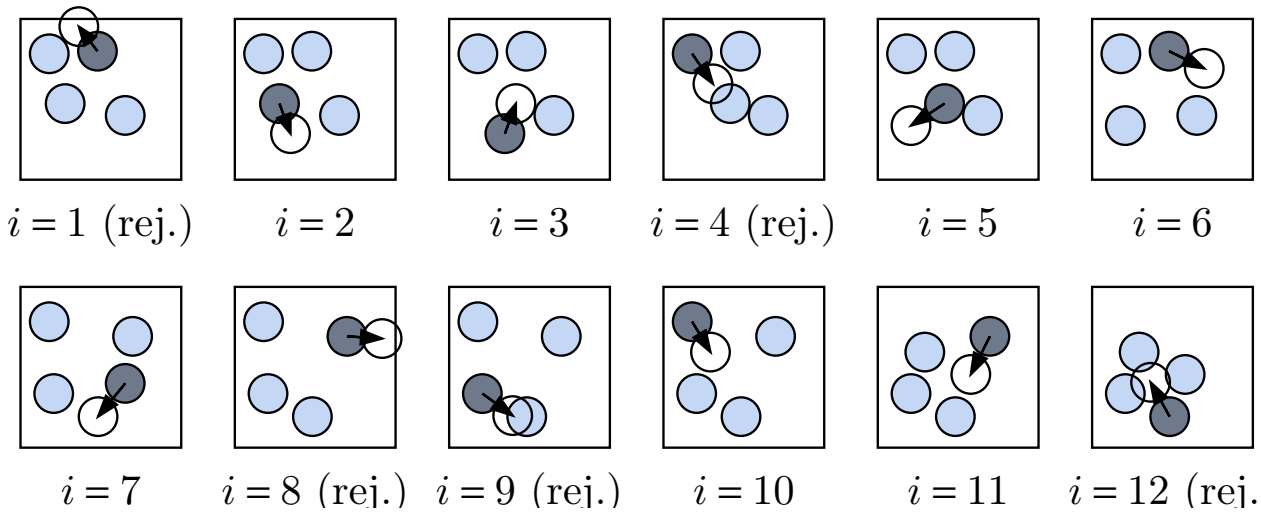
AND

EDWARD TELLER,* *Department of Physics, University of Chicago, Chicago, Illinois*

(Received March 6, 1953)

A general method, suitable for fast computing machines, for investigating such properties as equations of state for substances consisting of interacting individual molecules is described. The method consists of a modified Monte Carlo integration over configuration space. Results for the two-dimensional rigid-sphere system have been obtained on the Los Alamos MANIAC and are presented here. These results are compared to the free volume equation of state and to a four-term virial coefficient expansion.

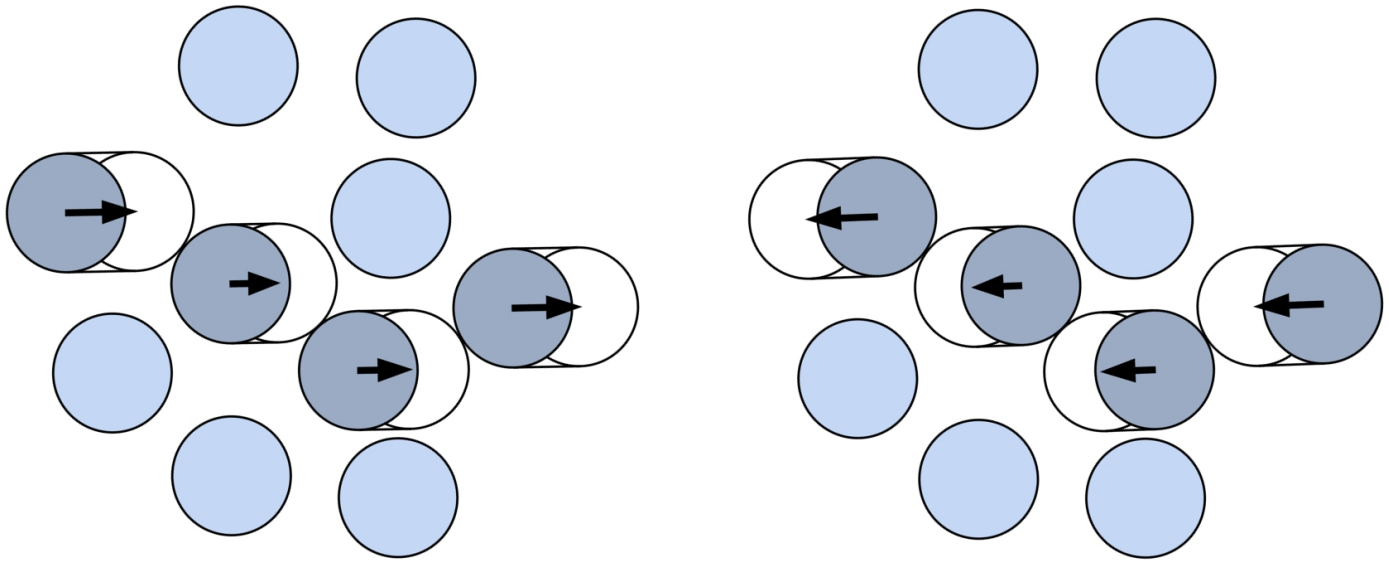
Metropolis algorithm for hard disks



- Metropolis algorithm (1953, mod.)

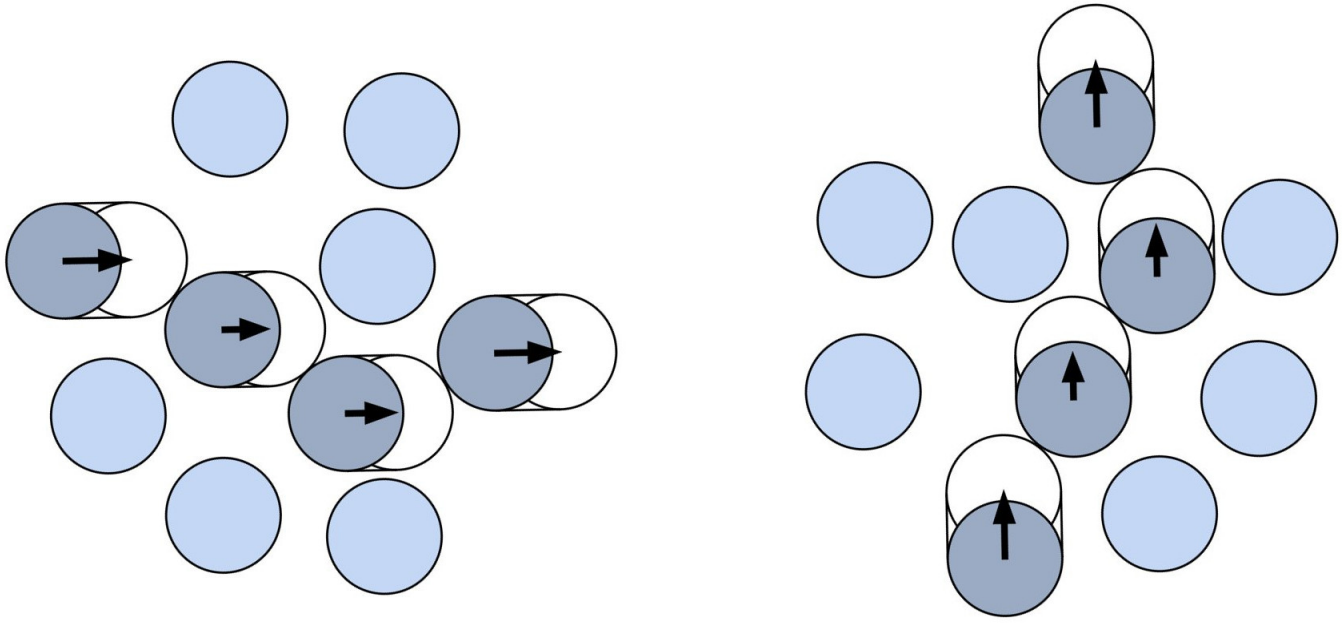
Département de Physique
École normale supérieure

Event-chain Monte Carlo (reversible)



- total length of arrows $= \ell$, fixed (sic!)

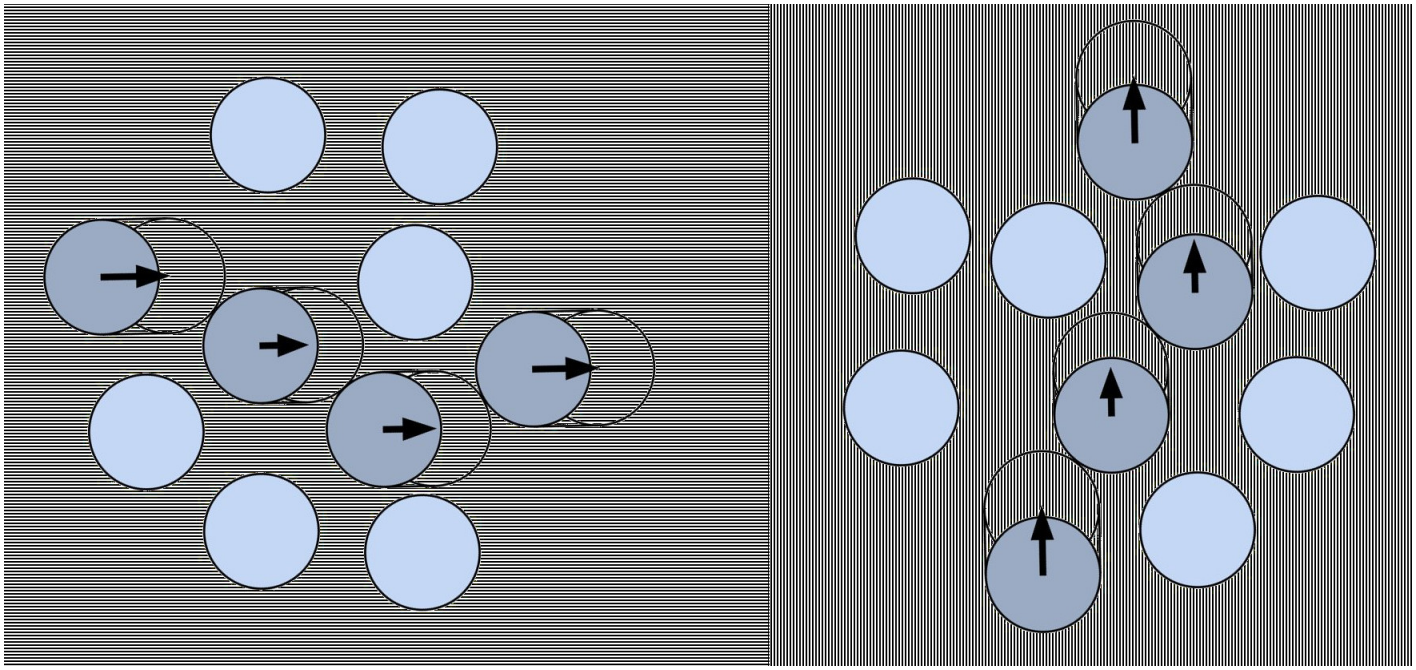
Event-chain Monte Carlo (non-reversible)



- Box of size $L \times L$, periodic boundary conditions, N disks.
- Event-chain Monte Carlo: Bernard, Krauth, Wilson (2009).

Département de Physique
École normale supérieure

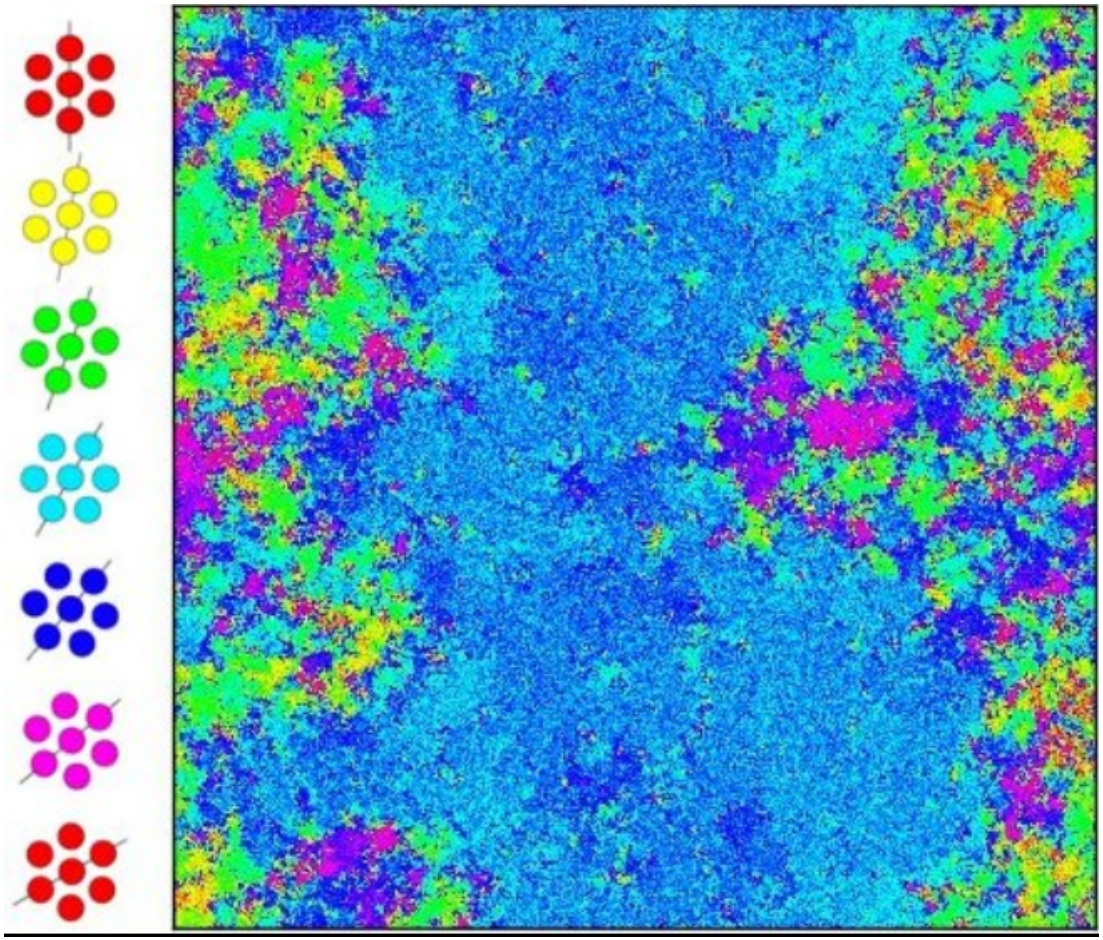
Event-chain Monte Carlo and molecular dynamics



- Box of size $L \times L$, periodic boundary conditions, N disks.
- Event-chain Monte Carlo: Bernard, Krauth, Wilson (2009).
- Molecular dynamics 'on a grating'.
- Hamiltonian Monte Carlo 'on a grating'.

Département de Physique
École normale supérieure

Equilibrium hard-disk configuration



- First equilibrated sample of 10^6 disks ever obtained (Bernard & Krauth 2011, see also Li et al. 2022)
  liquid–hexatic coexistence

Département de Physique
École normale supérieure

Markov chains

- Sample space Ω (e.g. hard disks, water molecules, quarks, ...)
- Markov chain $\leftarrow$ Sequence of random variables
  $(X_0 \sim \pi^{\{0\}}, X_1 \sim \pi^{\{1\}}, X_2 \sim \pi^{\{2\}} \dots)$
  X_{t+1} depends only on X_t, t is a 'time'
- Transition matrix P:
  - P_{ij}: conditional probability to move from sample i to sample j.
  - $\pi^{\{t+1\}} = \pi^{\{t\}} P$: Evolve probability distribution at time t to probability distribution at time $t+1$ (with $\pi^{\{t\}}, t > 0$ often non-explicit,even for $t \to \infty$).
- Move set $\mathcal{L}$: ... from which moves are sampled.
- Equilibrium distribution π: Satisfies global balance:

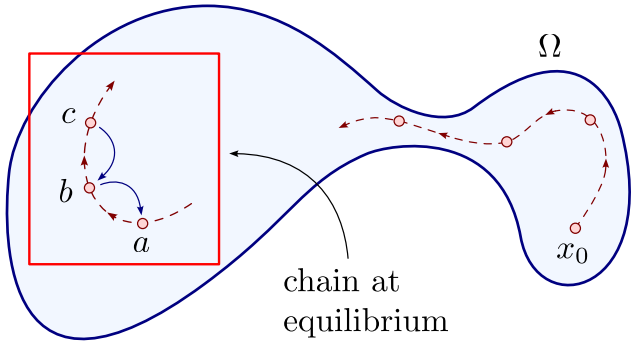
$$\pi_i = \sum_{j \in \Omega} \pi_j P_{ji} \quad \forall i \in \Omega.$$

  NB: P irreducible $\implies$ π unique.
- Aperiodicity: Absence of cycles. P irreducible and aperiodic:

$$\pi^{\{t\}} \to \pi \quad \text{for } t \to \infty$$

Département de Physique
École normale supérieure

Detailed balance, global balance



- Reversible transition matrices P satisfy the 'detailed-balance' condition:

$$\pi_a P_{ab} = \pi_b P_{ba}$$

- Non-reversible transition matrices P only satisfy 'global balance':

$$\pi_a = \sum_{b \in \Omega} \pi_b P_{ba}$$

Conductance (bottleneck ratio)

$$\Phi \equiv \min_{S \subset \Omega, \pi_S \leq \frac{1}{2}} \frac{\mathcal{F}_{S \to \overline{S}}}{\pi_S} = \min_{S \subset \Omega, \pi_S \leq \frac{1}{2}} \frac{\sum_{i \in S, j \notin S} \pi_i P_{ij}}{\pi_S}.$$

- Reversible Markov chains:

$$\frac{1}{\Phi} \leq \tau_{\mathrm{rel}} \leq \frac{8}{\Phi^2}$$

  ('$\leq$': Sinclair & Jerrum (1986), Lemma (3.3))

- Arbitrary Markov chain (see Chen et al. (1999)):

$$\frac{1}{4\Phi} \leq \mathcal{A} \leq \frac{20}{\Phi^2},$$

  ($\mathcal{A}$: set time: Expectation of $\max_S (t_S \times \pi_S)$ from equilibrium)

Département de Physique
École normale supérieure

Lifting (Chen et al. (1999))

- Markov chain Π $(\Omega, P, \pi) \Leftrightarrow$ Lifted Markov chain $\hat{\Pi}$ $(\hat{\Omega}, \hat{P}, \hat{\pi})$
- Mapping f from $\hat{\Omega}$ to Ω.
- **Condition 1**: π is preserved

$$\pi_v = \hat{\pi}\left[f^{-1}(v)\right] = \sum_{i \in f^{-1}(v)} \hat{\pi}_i,$$

- **Condition 2**: flows are preserved

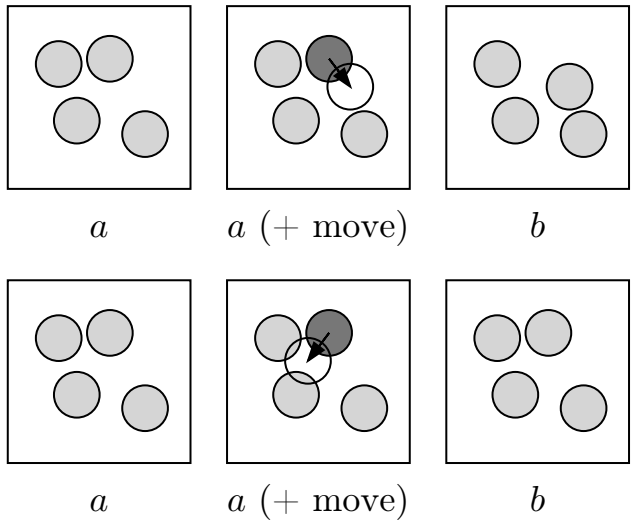
$$\underbrace{\pi_v P_{vu}}_{\text{collapsed flow}} = \sum_{i \in f^{-1}(v), j \in f^{-1}(u)} \overbrace{\hat{\pi}_i \hat{P}_{ij}}^{\text{lifted flow}} .$$

- $\hat{P}$ cannot have larger (better) conductance than P.
- Our work: $\hat{\Omega} = \Omega \times \overline{\mathcal{L}}$ with $\overline{\mathcal{L}}$ subset of move set $\mathcal{L}$.
- Speedup requires non-reversibility (Chen et al. 1999).

Département de Physique
École normale supérieure

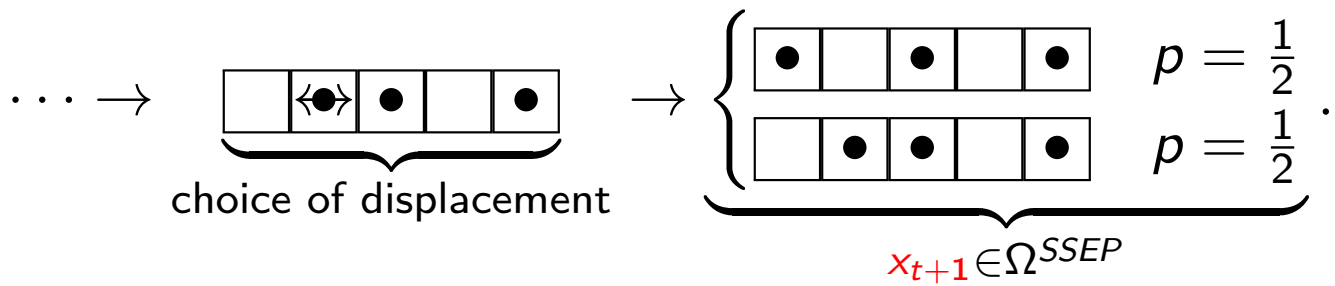
The Metropolis algorithm and the SSEP

1. Metropolis (1953) algorithm:



The Metropolis algorithm is reversible.

2. SSEP:

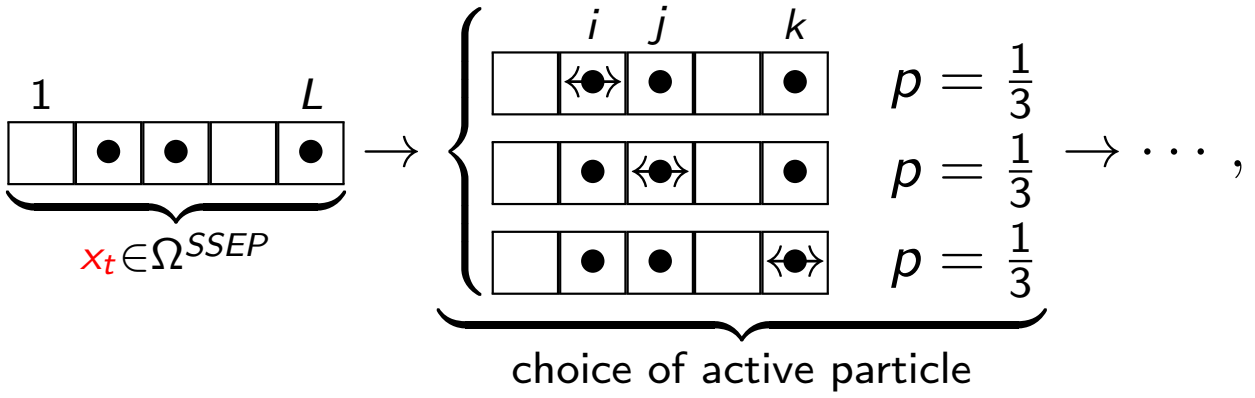


$1D$ lattice reduction of hard-sphere Metropolis algorithm.

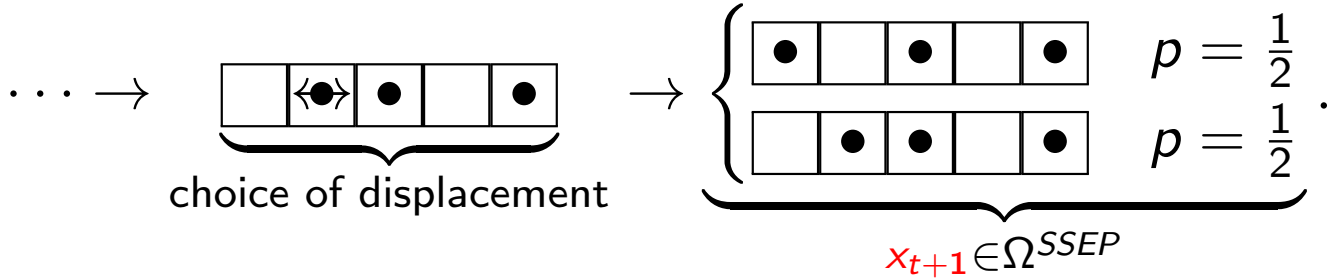
Département de Physique
École normale supérieure

Symmetric simple exclusion process (SSEP)

- Move (first part ...)



- Move (... second part)

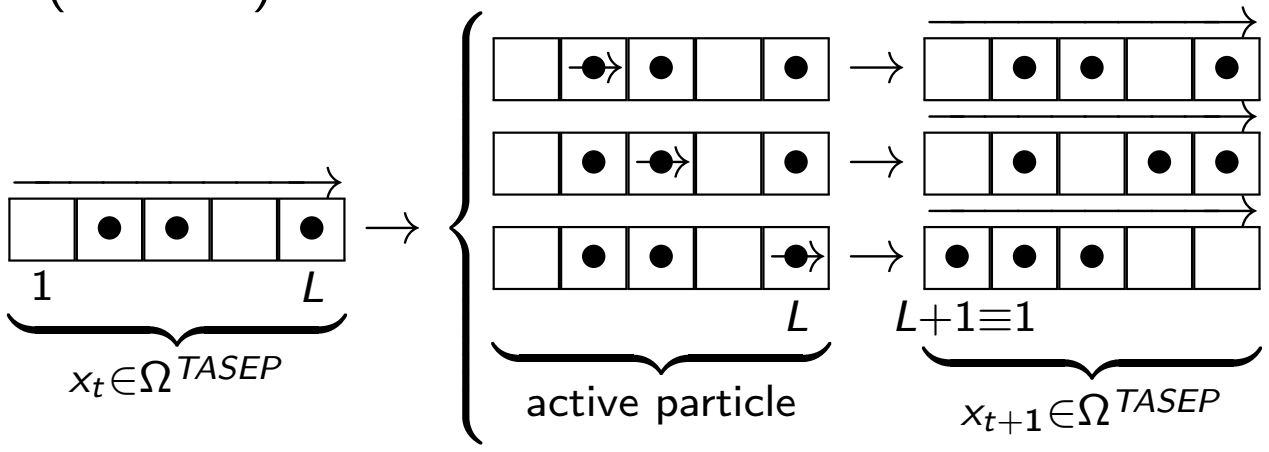


- The SSEP is 'our' collapsed (reversible) Markov chain.
- $\pi = \text{const.}$

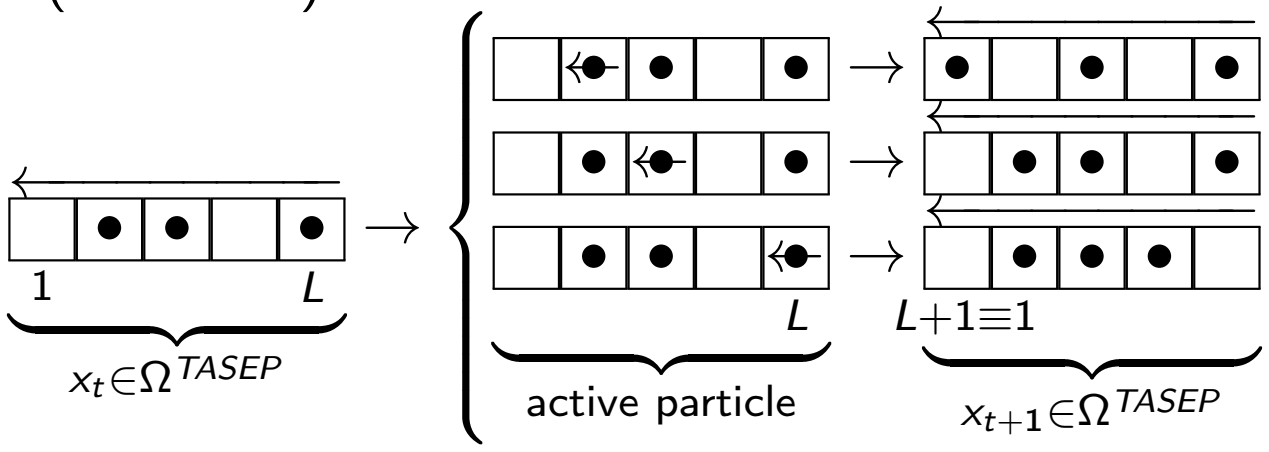
Département de Physique
École normale supérieure

Totally asymmetric simple exclusion process (TASEP)

- Move (forward)



- Move (backward)



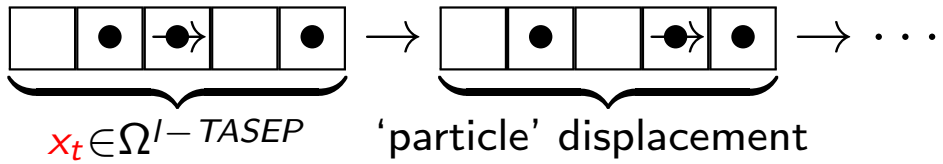
- forward–backward coupling (ad-hoc, or boundary conditions).

NB: Non-reversible, i.e. non-equilibrium, but samples equilibrium Boltzmann distribution.

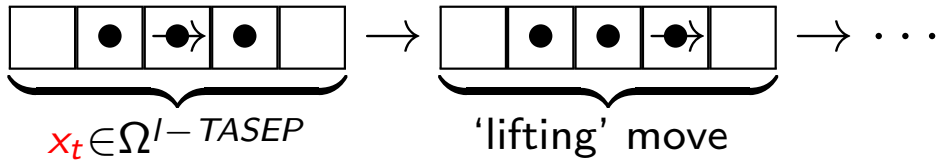
Département de Physique
École normale supérieure

Lifted TASEP (definition)

- $\Omega^{l-TASEP} = \Omega^{SSEP} \times \{-1, +1\} \times \{1, \ldots, N\}$, $\mathcal{L} = \emptyset$
- Move (first part ...)

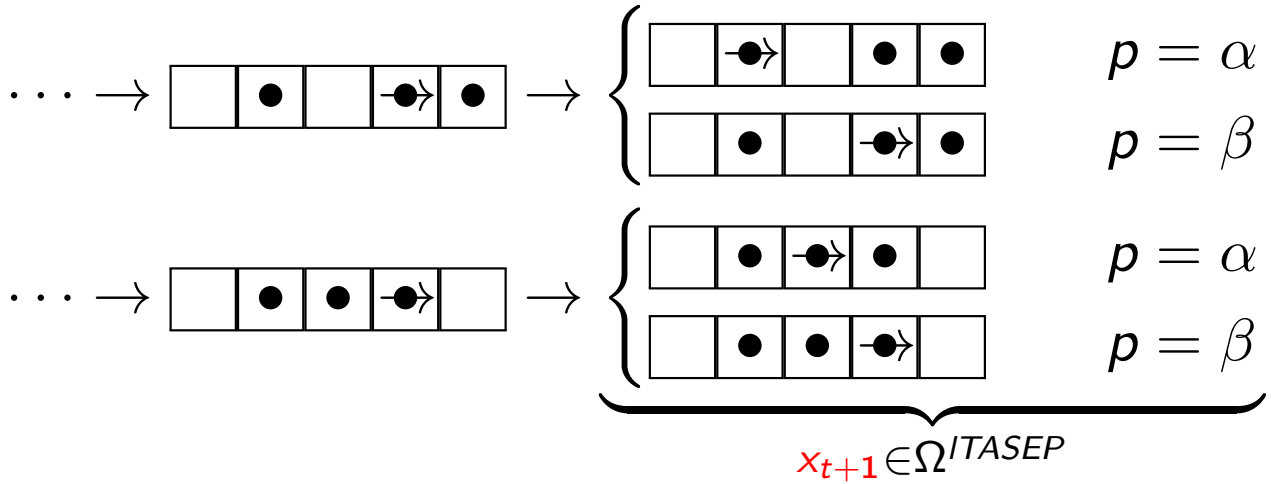


(1)



(2)

- Move (second part ...)

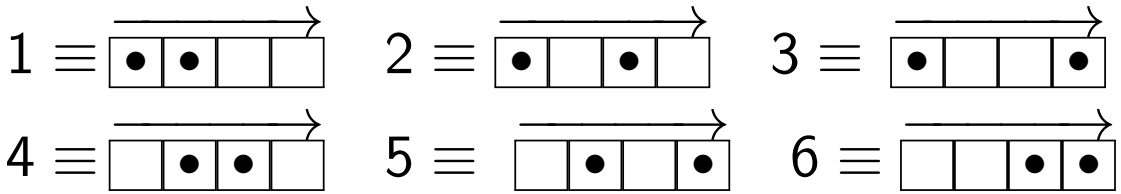


(2b)

(2b)

TASEP (example)

NB: Consider only the forward-moving sector (pbc):

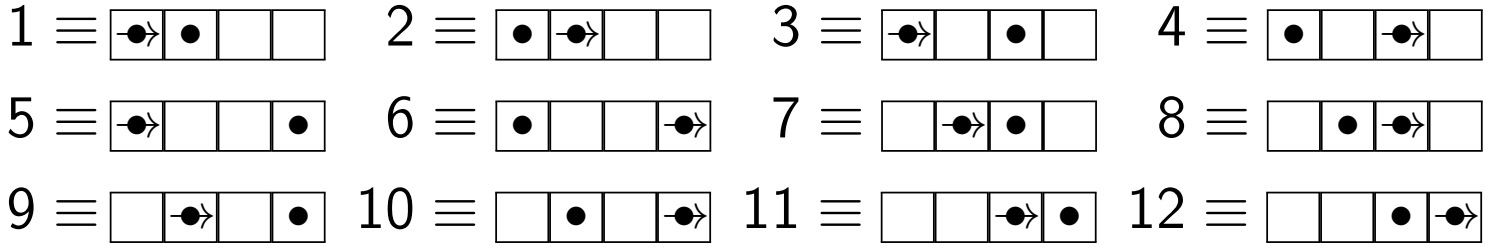


$$P^{\mathrm{TASEP}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & 1 & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & 1 \end{bmatrix} .$$

Département de Physique
École normale supérieure

Lifted TASEP (example)

NB: Consider only the forward-moving sector (pbc):



$$P = \begin{bmatrix}
\alpha & \beta & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \alpha & \beta & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \beta & \alpha & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \alpha & \beta & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \beta & \alpha & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \beta & \alpha & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \alpha & \beta & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \alpha & \beta & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \beta & \alpha \\
\beta & \alpha & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \alpha & \beta \\
\cdot & \cdot & \beta & \alpha & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
\end{bmatrix}$$

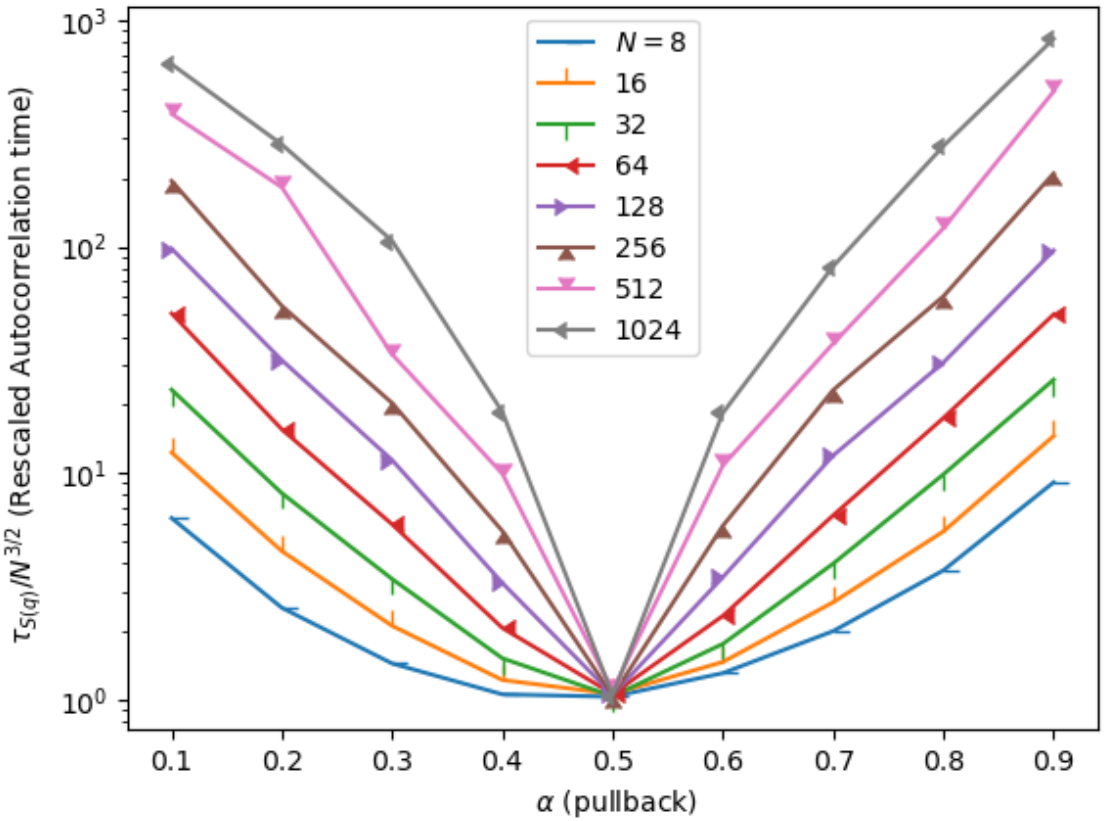
Synopsis of mixing and relaxation times

Algorithm	mixing	inverse gap	Comment
SSEP	$N^3 \log N$	N^3	Edwards–Wilkinson
TASEP	$N^{5/2}$	$N^{5/2}$	KPZ
Lifted TASEP	N^2	N^2 ($N^{3/2}$)	at α_{crit} only, otherwise KPZ.

- continuous-space versions available (Kapfer & Krauth (2017)).
- see Essler & Krauth (2023), results from Bethe ansatz.

Département de Physique
École normale supérieure

Lifted TASEP benchmark

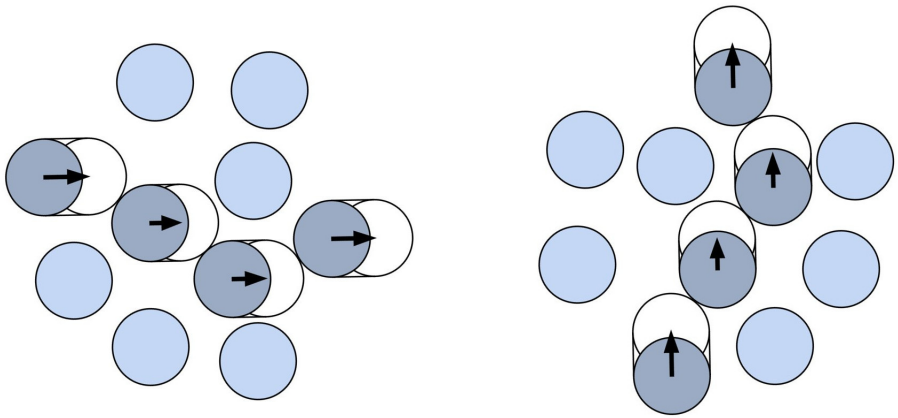


- Autocorrelations of structure factor (small-q density fluctuations) of the lifted TASEP.
- $\alpha = \alpha_{\text{crit}}$: vanishing drift velocity of activity.

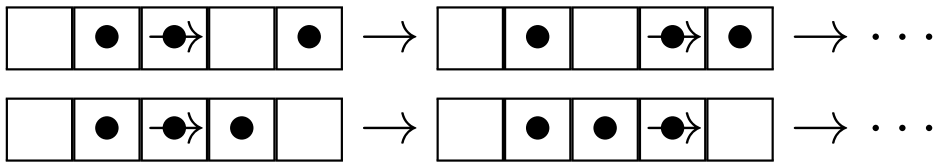
Département de Physique
École normale supérieure

The lifted TASEP and event-chain Monte Carlo

1. Event-chain Monte Carlo algorithm (Bernard, Krauth, Wilson 2009):



2. Lifted TASEP:



One-dimensional lattice reduction of event-chain Monte Carlo.

Département de Physique
École normale supérieure

Factorized Metropolis algorithm (pair potential)

- Metropolis filter:

$$p^{\mathrm{Met}}(a \to b) = \min\left[1, \exp\left(-\beta \sum_{i<j} \Delta U_{i,j}\right)\right]$$

$$p^{\mathrm{Met}}(a \to b) = \min\left[1, \prod_{i<j} \exp\left(-\beta \Delta U_{i,j}\right)\right]$$

- Factorized Metropolis filter (Michel, Kapfer, Krauth 2014):

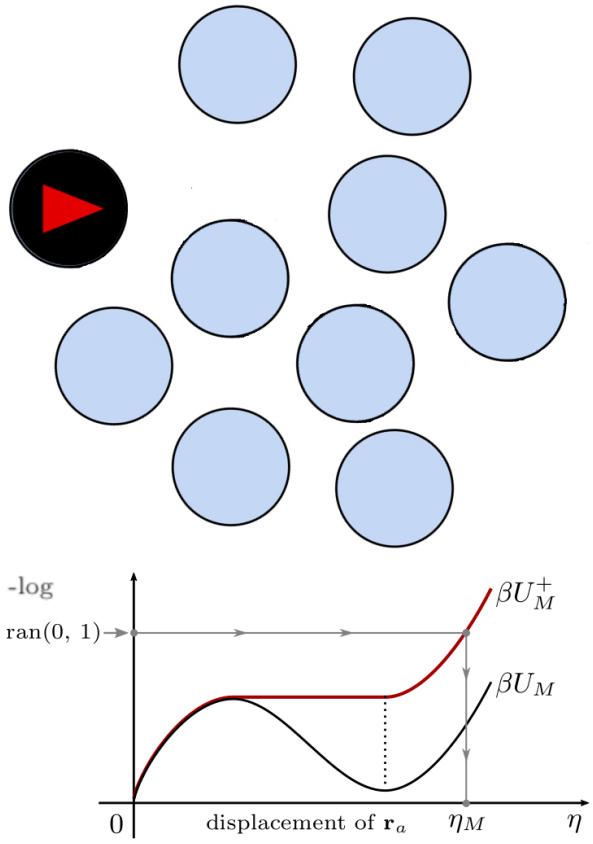
$$p^{\mathrm{Fact}}(a \to b) = \prod_{i<j} \min\left[1, \exp\left(-\beta \Delta U_{i,j}\right)\right].$$

$$X^{\mathrm{Fact}}(a \to b) = X_{1,2} \wedge X_{1,3} \wedge \cdots \wedge X_{N-1,N}$$

- i.e.: Accept move if all pairs (i,j) accept it.
- Consensus-based, not gradient-based.

Département de Physique
École normale supérieure

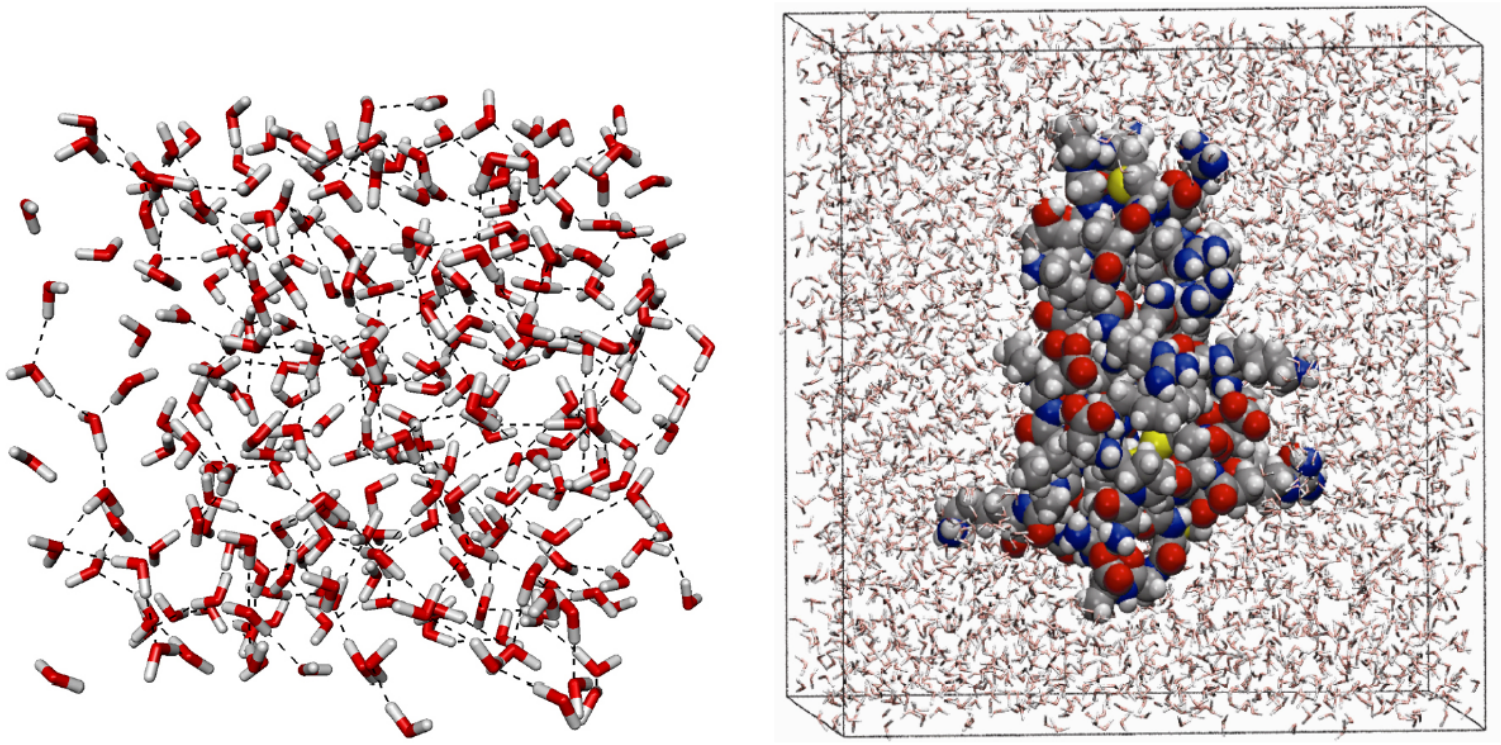
Factorized Metropolis algorithm (pair potential)



- Complexity $\mathcal{O}(1)$ per event for $N \to \infty$ (Kapfer–Krauth 2016)
- Zoo of choice about what to do after the event.
- Piecewise deterministic Markov process, samples $\exp(-\beta U)$ at all continuous times.

Département de Physique
École normale supérieure

All-atom simulation with non-reversible Markov chains



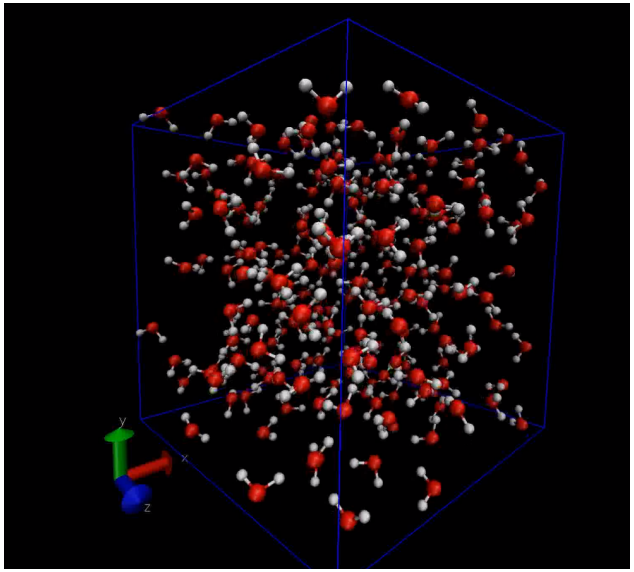
(a) (b)

(a): Höllmer, Maggs, Krauth (2023) (see also Faulkner et al. (2018), Höllmer et al. (2020)).

(b): Yet to be done...

Département de Physique
École normale supérieure

All-Atom Coulomb problem



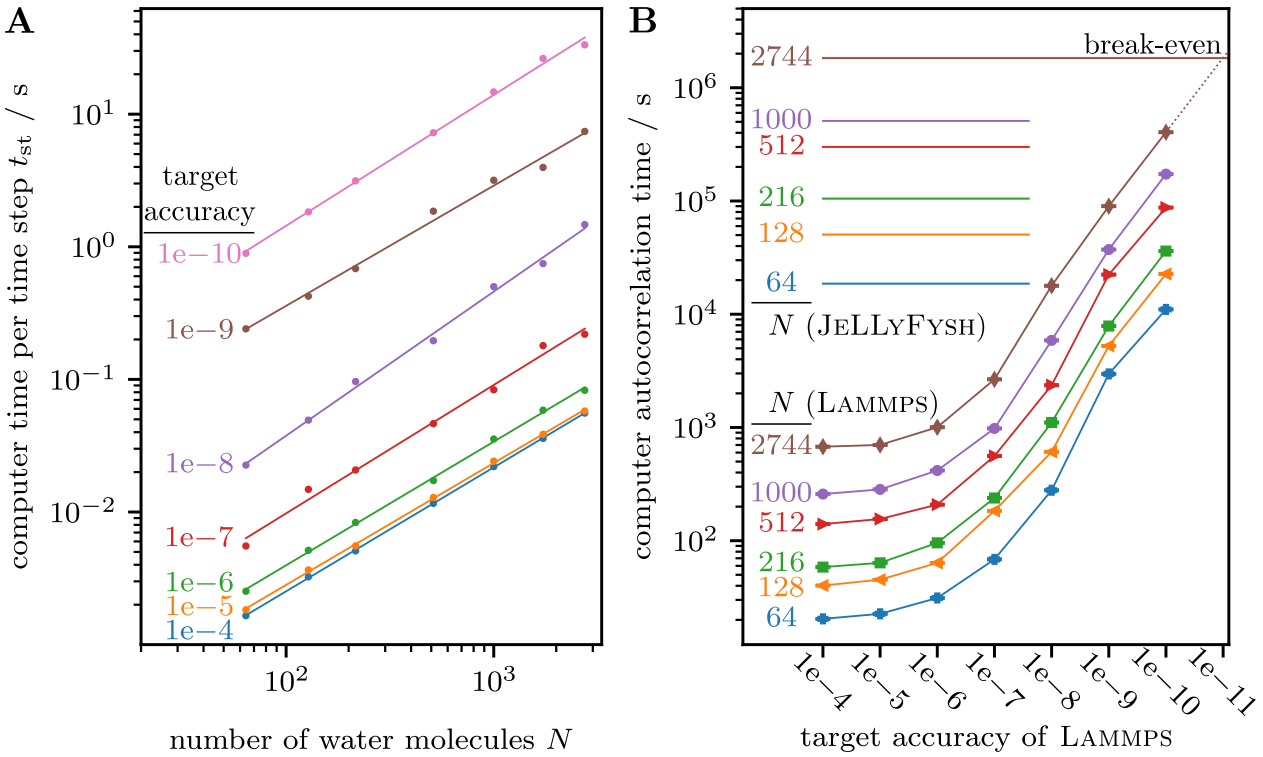
- Factors M: 'bond', 'bending', 'Lennard-Jones', 'Coulomb' (SPC/Fw water model).

$$X^{\text{Fact}}(c \to c') = \bigwedge_{M \in \mathcal{M}} X_M(c_M \to c'_M).$$

- i.e.: Move until a single of $\mathcal{O}(N)$ factors vetoes acceptance.
- Complexity $\mathcal{O}(1)$ per event (Kapfer, Krauth 2016).
- Event-driven, approximation-free, canonic, without thermostat.

Département de Physique
École normale supérieure

Water simulation—benchmark JeLLyFysh—LAMMPS



- See Höllmer, Maggs, Krauth (2023)
- JeLLyFysh (non-reversible Markov chain) is approximation-free: Evaluates neither forces nor total potentials.
- LAMMPS (molecular dynamics) has time-step error and lattice error for computation of electrostatic interaction.

Département de Physique
École normale supérieure

Conclusion & Outlook

Lifted Markov chains:

- From the beginnings to real-world applications.
- Event-driven, rejection-free, non-reversible, fast.
- Consensus-driven, not gradient-driven.
- Non-reversibility at all times appears essential.

Outlook:

- Sampling $\exp(-\beta U)$ without evaluating U.
- 'Natively cutoff-free' MCMC (Coulomb, LJ) in $\mathcal{O}(1)$.
- Applications in chemical physics.

Département de Physique
École normale supérieure