Superposition Enhanced Nested Sampling

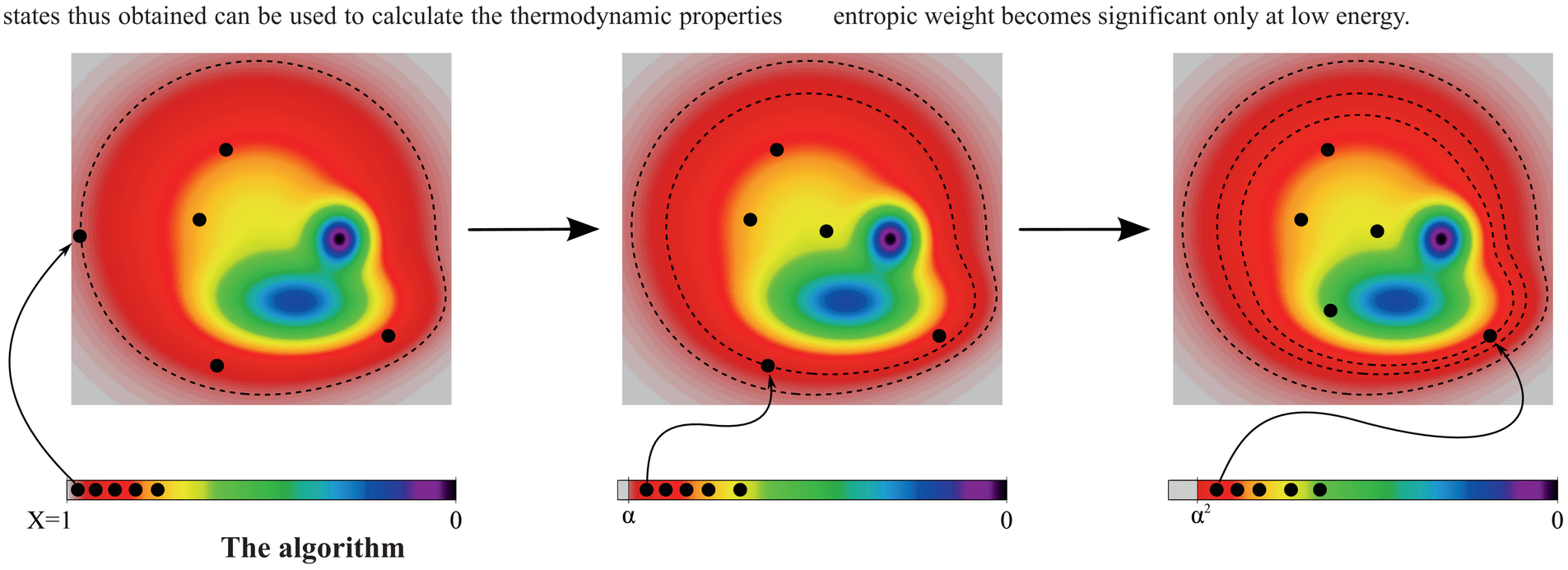
Stefano Martiniani, Jacob Stevenson, David Wales, Daan Frenkel
Department of Chemistry, University of Cambridge
Lensfield Road, Cambridge, CB2 1EW

Funded by the ERC

Nested Sampling

https://github.com/js850/nested_sampling

Nested Sampling (NS) is a Bayesian Monte Carlo algorithm that uses multiple replicas to explore geometrically shrinking regions of phase space by placing a hard constraint in energy at each step. The statistical weight of each fraction of phase space can be estimated statistically by mapping the problem to that of placing multiple replicas on a unit interval. The density of states thus obtained can be used to calculate the thermodynamic properties of the system. NS is self-adapting, reducing phase space volume by a constant fraction at each iterations thus making the method particularly powerful when dealing with pseudo-discontinuous phase transitions. Furthermore NS is trivially parallelisable. The drawback of NS is that the descending hard constraint in energy may lock out basins whose relative entropic weight becomes significant only at low energy.



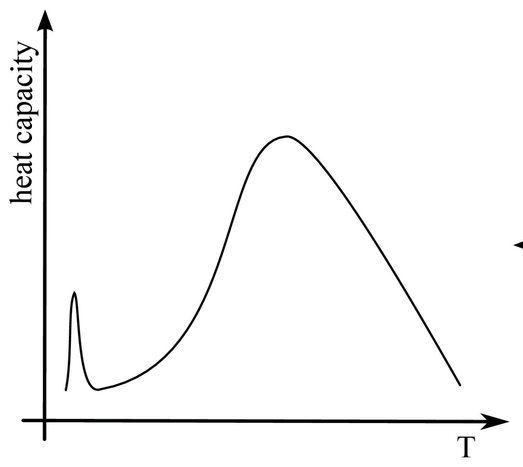
At each iteration, phase space is compressed by a constant factor

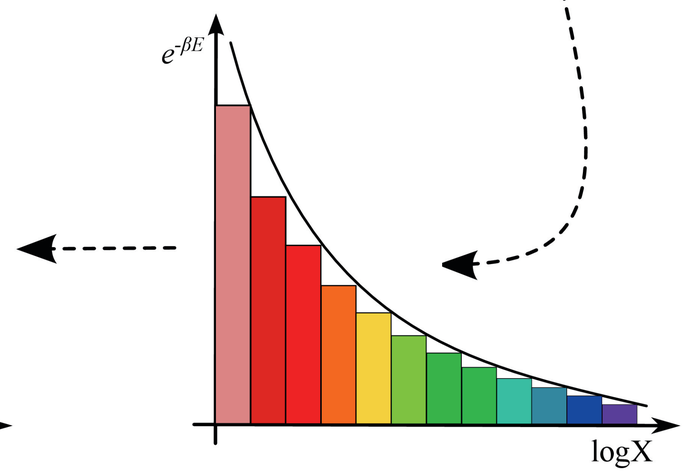
$\alpha = K / (K+1)$

where K is the number of replicas

The algorithm

1. Sample multiple configurations of the system (replicas) at random and store them;
2. Remove the replica with highest energy and store its energy E_i;
3. Obtain a new replica sampled uniformly from configuration space with energy $E \leq E_i$ by starting a Monte Carlo Markov Chain from an exisisting replica chosen at random;
4. repeat 2-4 until a termination condition is met.
5. Estimate statistically the fraction of configurational space, $X_i=\alpha^i$, with energy below E_i.
6. Use the thus obtained density of states $g_i=X_{i-1}-X_i$ to calculate the partition function by numerical integration and the thermodynamic porperties of the system.





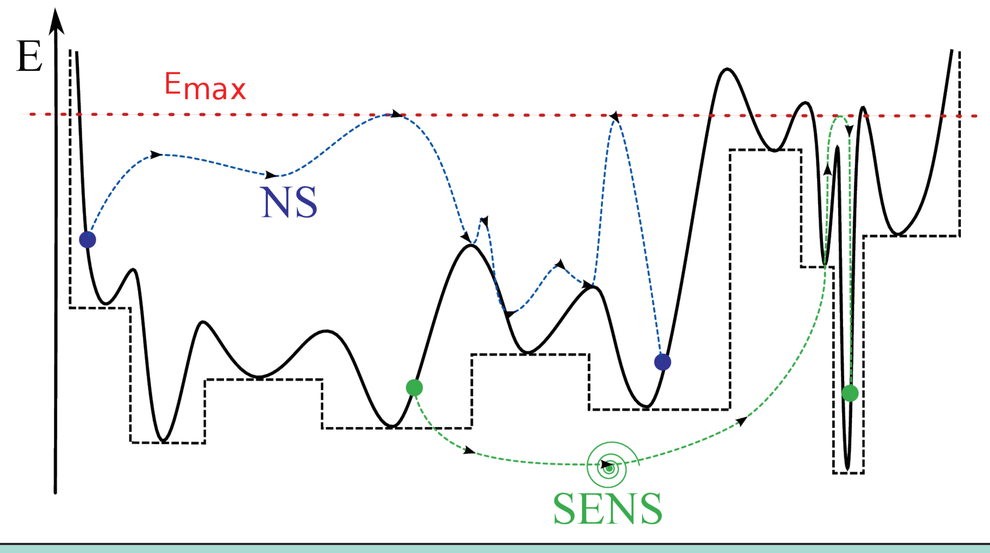
SENS

https://github.com/smcantab/sens

Use knowledge about the minima of the energy landscape obtained using global optimization techniques to improve sampling

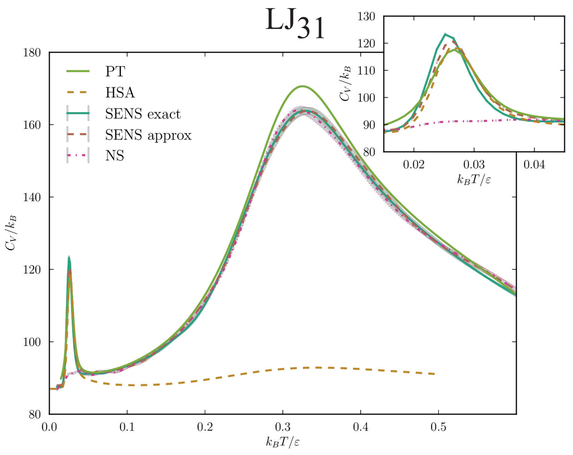
- **Global optimization**: Fast and powerful methods for exploring phase space. Can easily build up a database of local minima. The reason it can be so fast is that you abandon Bolzmann sampling (e.g. no detailed balance). This can be partially recovered using the harmonic superposition approximation (HSA), but this is valid only at low temperatures.
- **Exact SENS**: Use Hamiltonian replica exchange to attempt swaps with the database of minima during the MCMC.
  1. Generate a configuration X_{HSA} randomly (and exactly) from the HSA
  2. Accept a swap if $E(X_{HSA}) < E_{MAX}$ and $E_{HSA}(X) < E_{MAX}$
- **Approximate SENS**: Always accept the swap

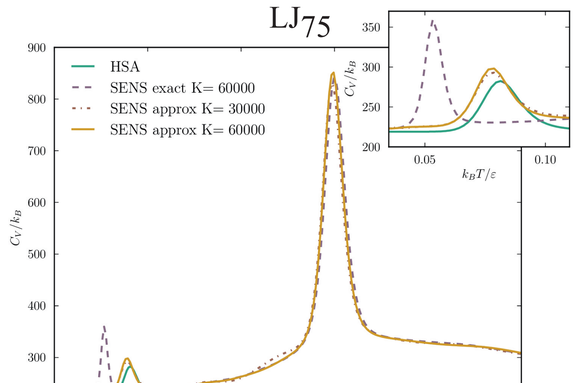
Nested Sampling reproduces the high temperature behaviour of the system and assigns the correct statistical weight to each basin. The database of minima guarantees that the basins associated with the lowest minima are visited even if the energy constraint is below the minimum energy path connecting two basins.

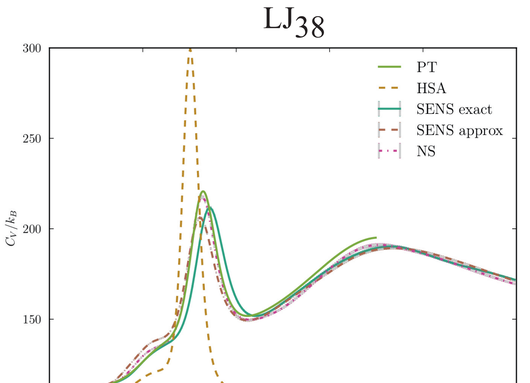


Results

Benchmarking of SENS for Lennard-Jones clusters showed a significant improvement in efficiency over raw Nested Sampling and Parallel Tempering: for LJ_{31} SENS requires 10^{10} energy evaluations while 10^{12} energy evaluations of raw Nested Sampling (or parallel tempering) were necessary to succesfully identify the low temperature solid-solid transition.







References

J. Skilling, Bayesian Analysis, 1, 833–859, 2006.

L. B. Pártay, A. P. Bartók, and G. Csányi, The journal of physical chemistry. B, 114, 10502–12, 2010.

Superposition Enhanced Nested Sampling (this work) arXiv:1402.6306