Abstract

We investigated the

scaling behaviour of the computational effort

of global minimization algorithms for a simple model of protein-folding with respect to the system size N.

We studied a "demixing" model of two species of Lennard-Jones-particles with a Monte-Carlo and a stochastic tunneling method on the lattice and in the continuum and compared these results to a multi-start algorithm.

Motivation

We investigated the scaling behaviour of the computational effort (N_{CPU}) of global minimization algorithms with respect to the system size (N)

$$N_{\mathrm{CPU}} = A \cdot N^{\beta}$$

to obtain

(a) an estimate for realistic applications in the context of protein-structure-prediction and

(b) a criterion to fundamentally distinguish different algorithms

Model

Two types (yellow,red) of particles interacting with a Lennard-Jones-Potential in 2D

$$V_{LJ}(r) = A \cdot \left(\frac{1}{r^{12}} - \frac{2}{r^{6}} \right)$$

$$A = \begin{cases} 1: & \text{different color} \\ 2: & \text{same color} \end{cases}$$

(a) **Local Minima $\equiv$ lattice configurations on Δ lattice**

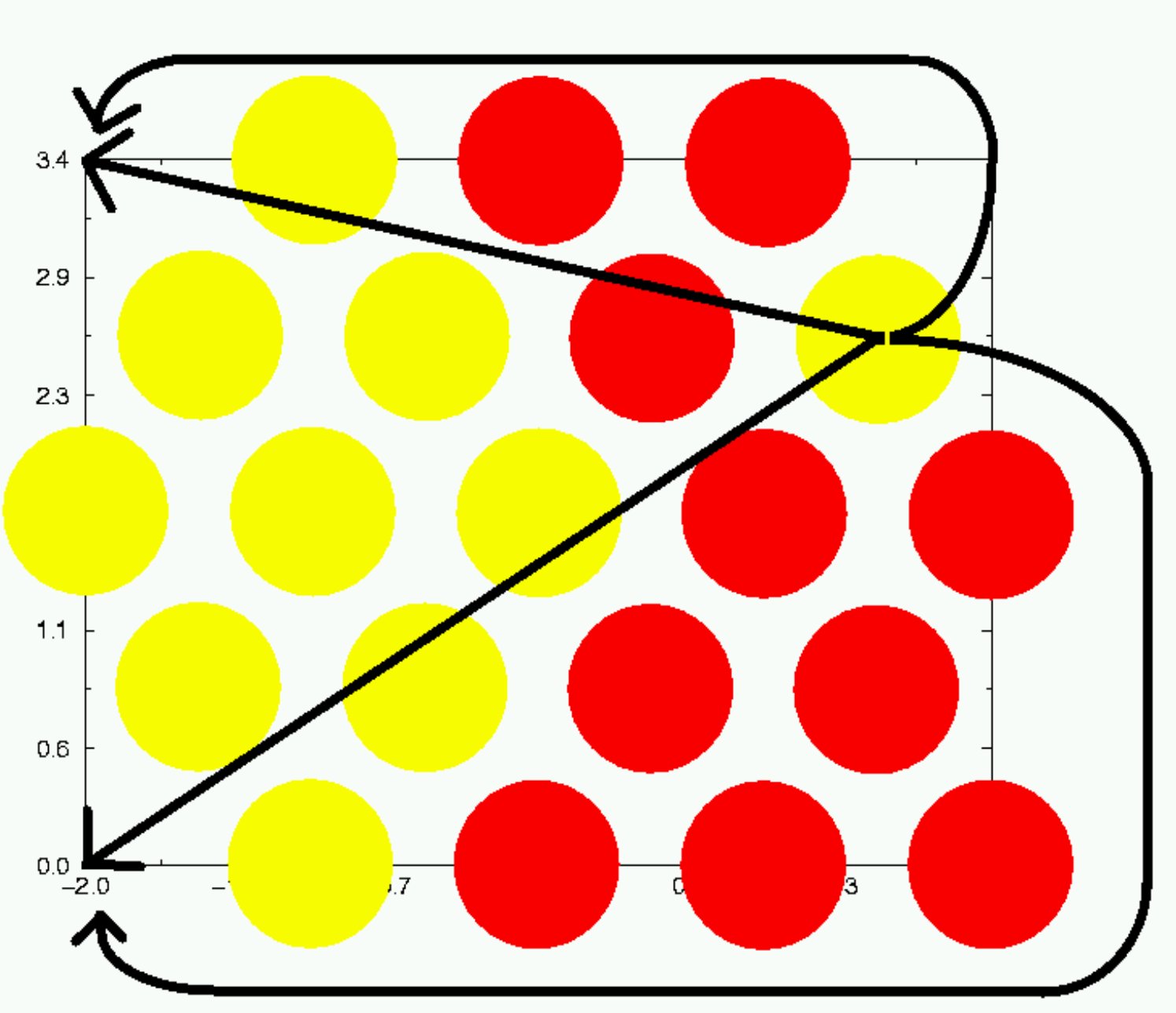
(b) **NP-hard**

(c) **Global Minimization $\equiv$ Demixing**

average pathlength $\sim \sqrt{N}$

Demixing makes the model more complex than it would be for identical particles.

Typical paths to escape from local minima :



Methods for the continuum model

MCM (Scheraga et al.):

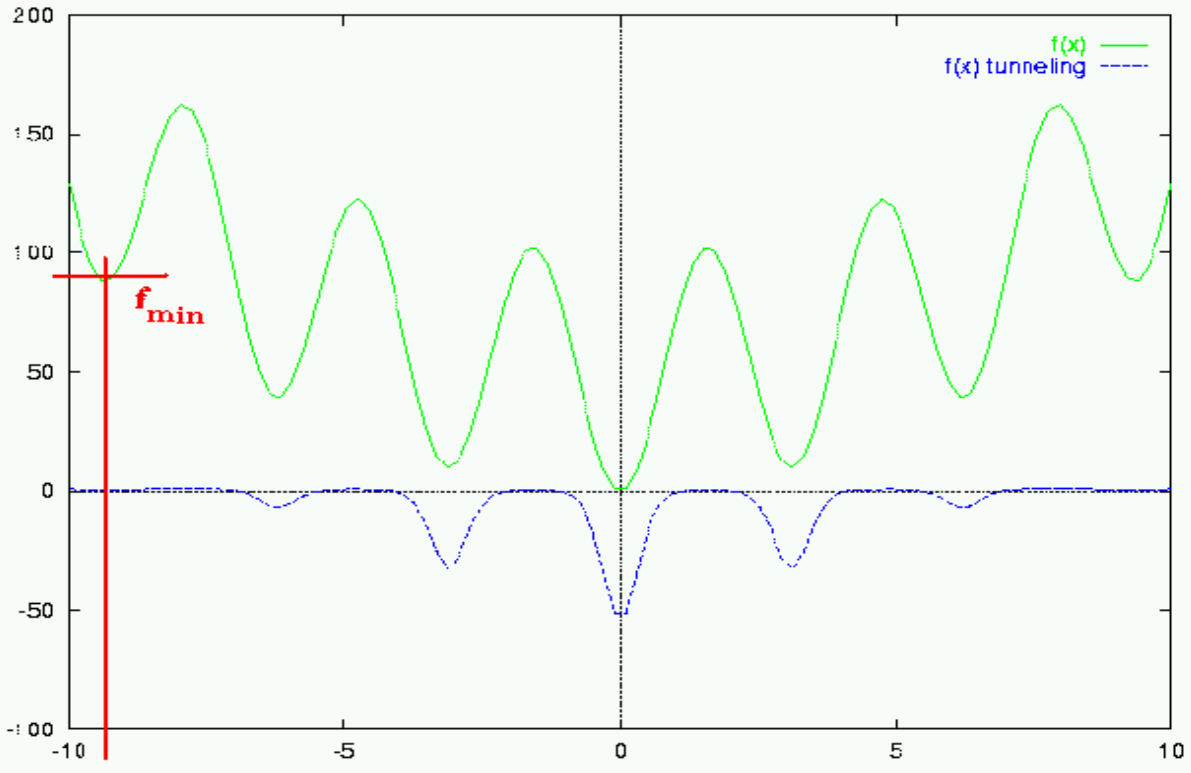
Monte-Carlo on local minima

Ensemble $\sim \exp\left(-\frac{\Delta E}{kT}\right)$

Tunneling (Levy et al.) :

Monte-Carlo on local minima with a Tunneling approach

$$\tilde{f}(\vec{x}) = 1 - \exp\left(-\beta_2 \left[f(\vec{x}) - f_{\min}\right]\right)$$



Multi-Start :

Choose random initial configurations and perform local minimization.

Methods for the lattice Model

Multi-Start :

Choose random initial configurations and swap nearest neighbours as long as this lowers the energy.

MC on lattice $\equiv$ MCM in the continuum :

Ensemble $\sim \exp\left(-\frac{\Delta E}{kT}\right)$ **sampling**

Conclusions

- **Thermodynamic minimization algorithms yield power-law scaling**

$$N^{\text{cont}}_{\text{MCM}} \sim N^{6.5\pm1}$$
$$N^{\text{cont}}_{\text{TUN}} \sim N^{7.2\pm1}$$
$$N^{\text{latt}}_{\text{MC}} \sim N^{4.4\pm1}$$

  whereas the Multi-Start algorithm shows an exponential behaviour.

- **Extrapolation from met-enkephalin**

  $N = 5$ $\qquad$ $t = 1000\,\text{s}$

  to

  α-hemoglobin

  $N = 146$ $\qquad$ $t = 1.6 \times 10^7\,\text{years}$

NEED ALGORITHM WITH LOWER SCALING EXPONENTS

References

[1] A. V. Levy and A. Montalvo, *The tunneling algorithm for the global minimization of functions*, SIAM J. Sci. Stat. Comput. **6** (1985), no. 1, 15–29.

[2] Akbar Nayeem, Jorge Vila, and Harold A. Scheraga, *A comparative study of the simulated-annealing and monte carlo-with- minimization approaches to the minimum-energy structures of polypeptides: [met]-enkephalin*, J. Comp. Chem. **12** (1991), no. 5, 594–605.

[3] David Shalloway Panos M. Pardalos and G. Xue (eds.), *Global minimization of nonconvex energy functions: Molecular conformation and protein folding*, DIMACS – Series in Discrete Mathematics and Theoretical Computer Science, vol. 23, 1995, DIMACS workshop, March 20-21, 1995.

[4] Harold A. Scheraga, *Calculations of stable conformations of polypeptides, proteins, and protein complexes*, Chemica Scripta **29A** (1989), 3–13.

[5] Jeffrey Skolnick and Andrzej Kolinski, *Computer simulations of globular protein folding and tertiary structure*, Annu. Rev. Phys. Chem. **40** (1989), 207–235.