Full Sampling of Atomic Configurational Spaces

Gábor Csányi

Engineering Laboratory
Cambridge

with: Lívia Bartók-Pártay and Albert Bartók-Pártay

Goals

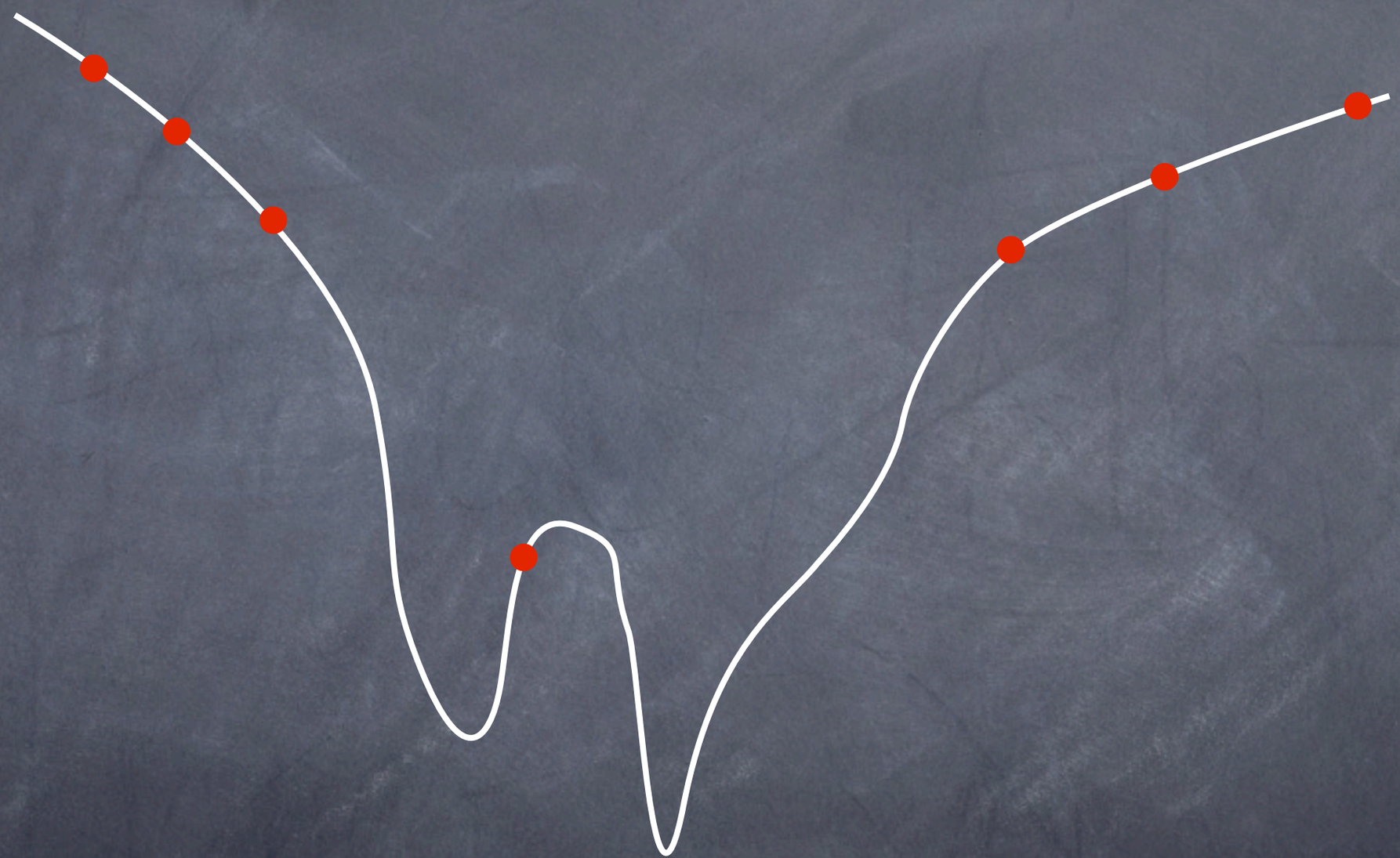
- Sample Potential Energy Landscapes (PES)
- Find low energy configurations
- Evaluate the partition function and similar integrals
- Have error estimates for all measurements

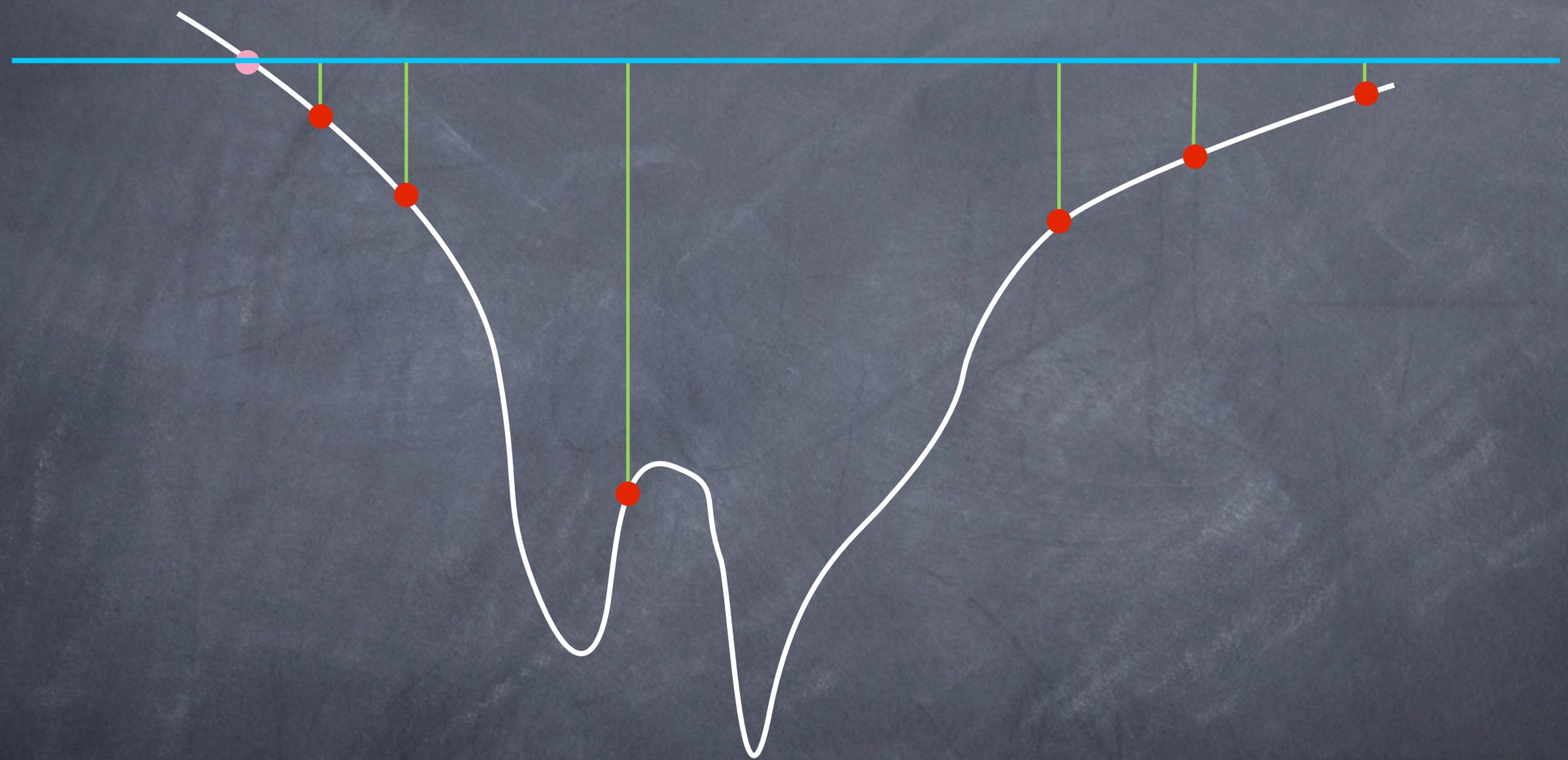
Exploring the PES

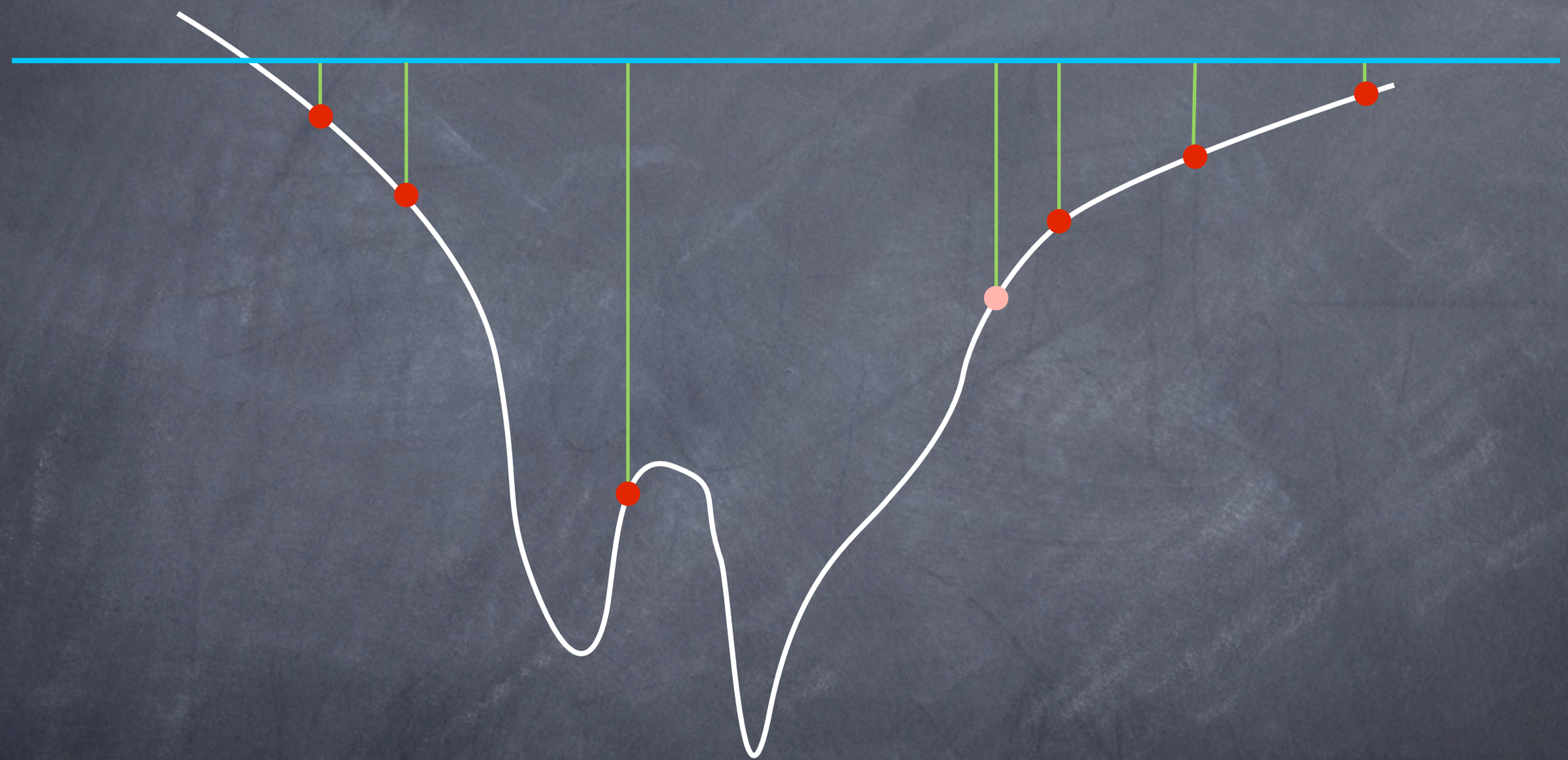
- Search
  - Minimisation
  - Simulated Annealing
  - Genetic Algorithm
  - Basin-hopping
  - Minima-hopping
  - Metadynamics
- Thermodynamics
  - Molecular Dynamics & MCMC
  - Temp. Accelerated Dynamics
  - Parallel Tempering
  - Wang-Landau
  - Clausius-Clapeyron
  - Thermodynamic Integration

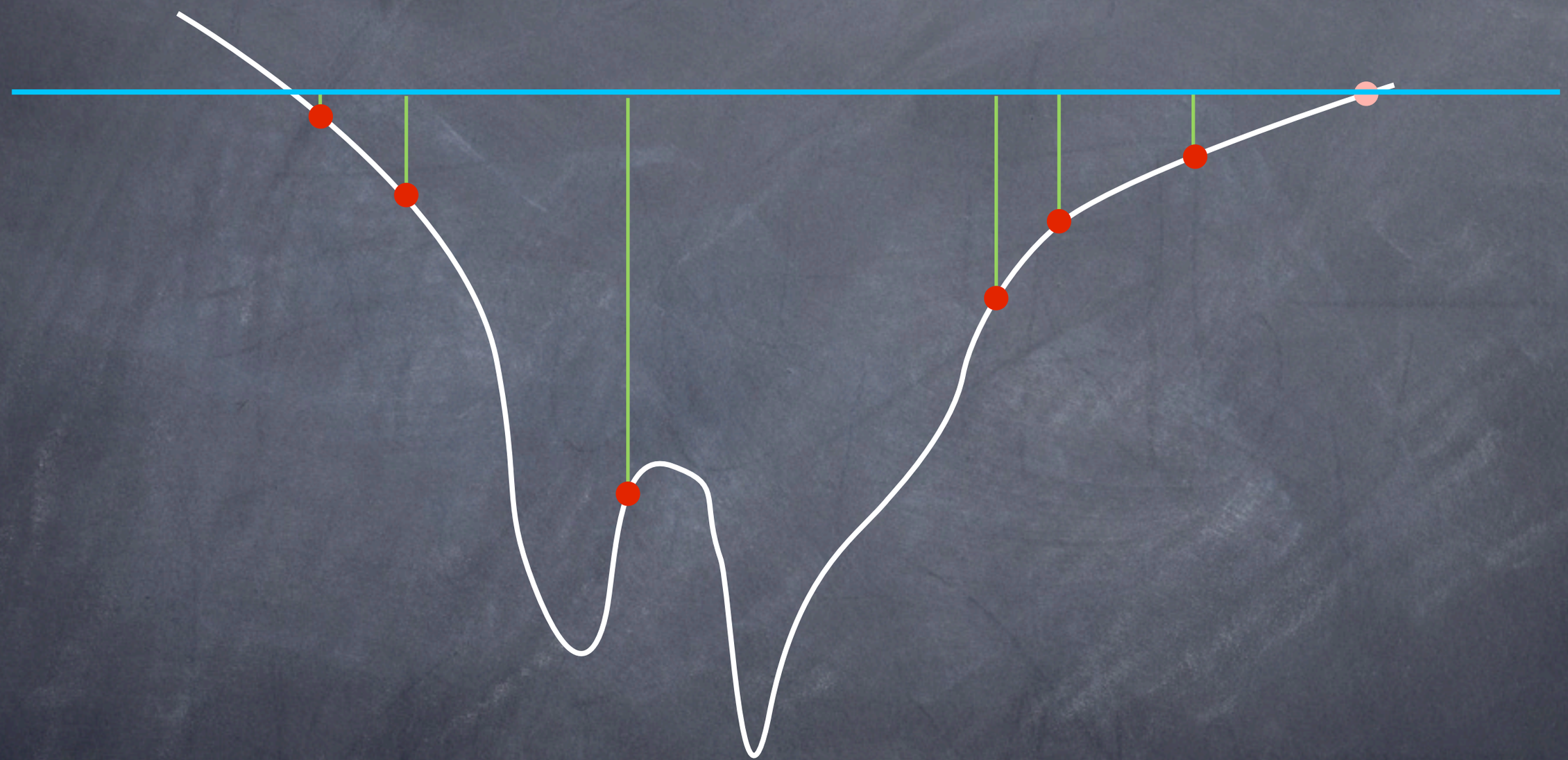
Nested sampling

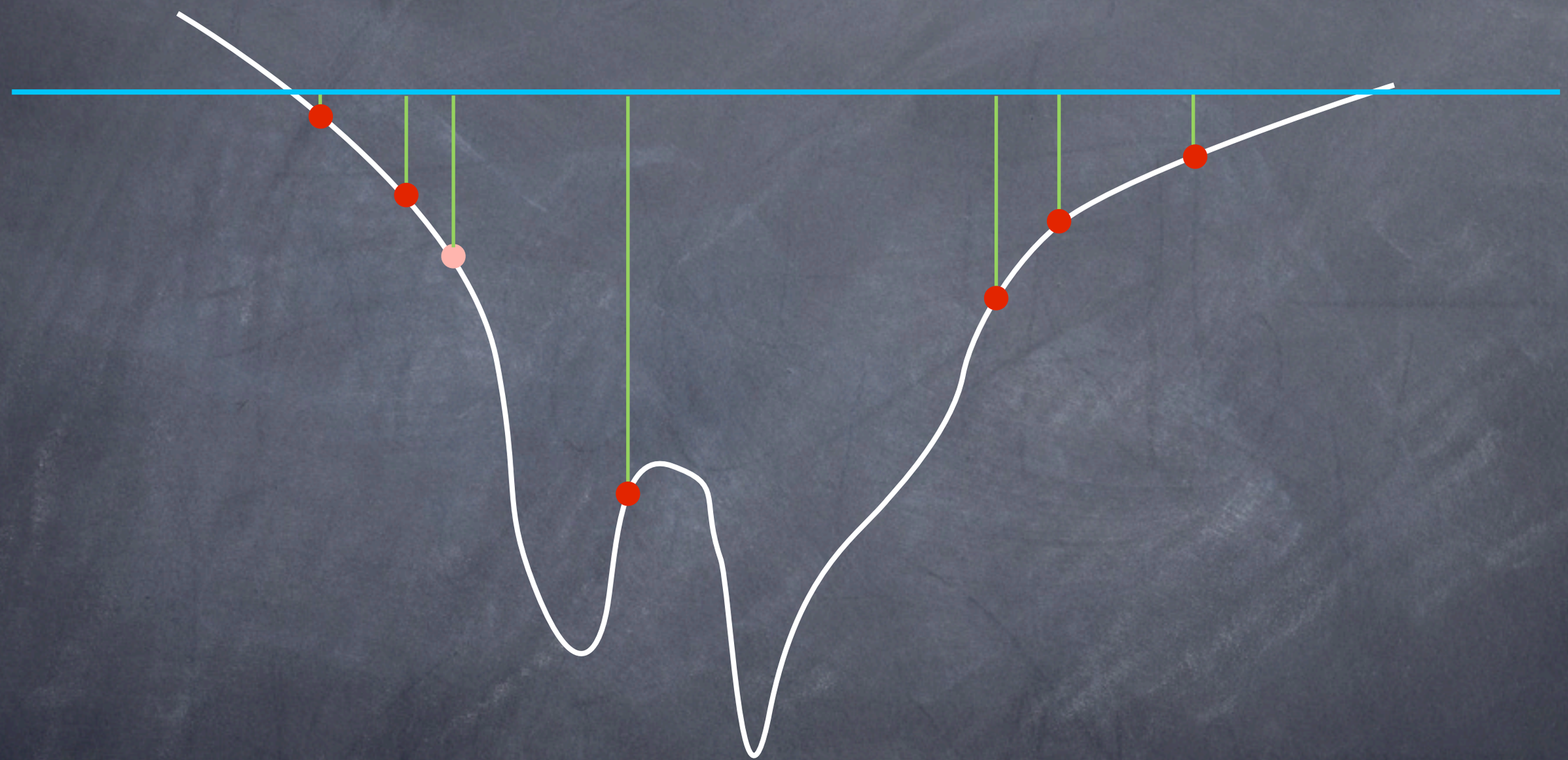
- A very simple algorithm to sample $E(x)$:
  1. Choose N points randomly: $E(x_k)$
  2. Remove one with the highest energy E_i
  3. Replace with a random point, $E(x) < E_i$
  4. $i=i+1$, goto 2.
- At the end $\{E_i\}$ forms a good mesh for integrating things like $\exp(-E(x)/kT)$

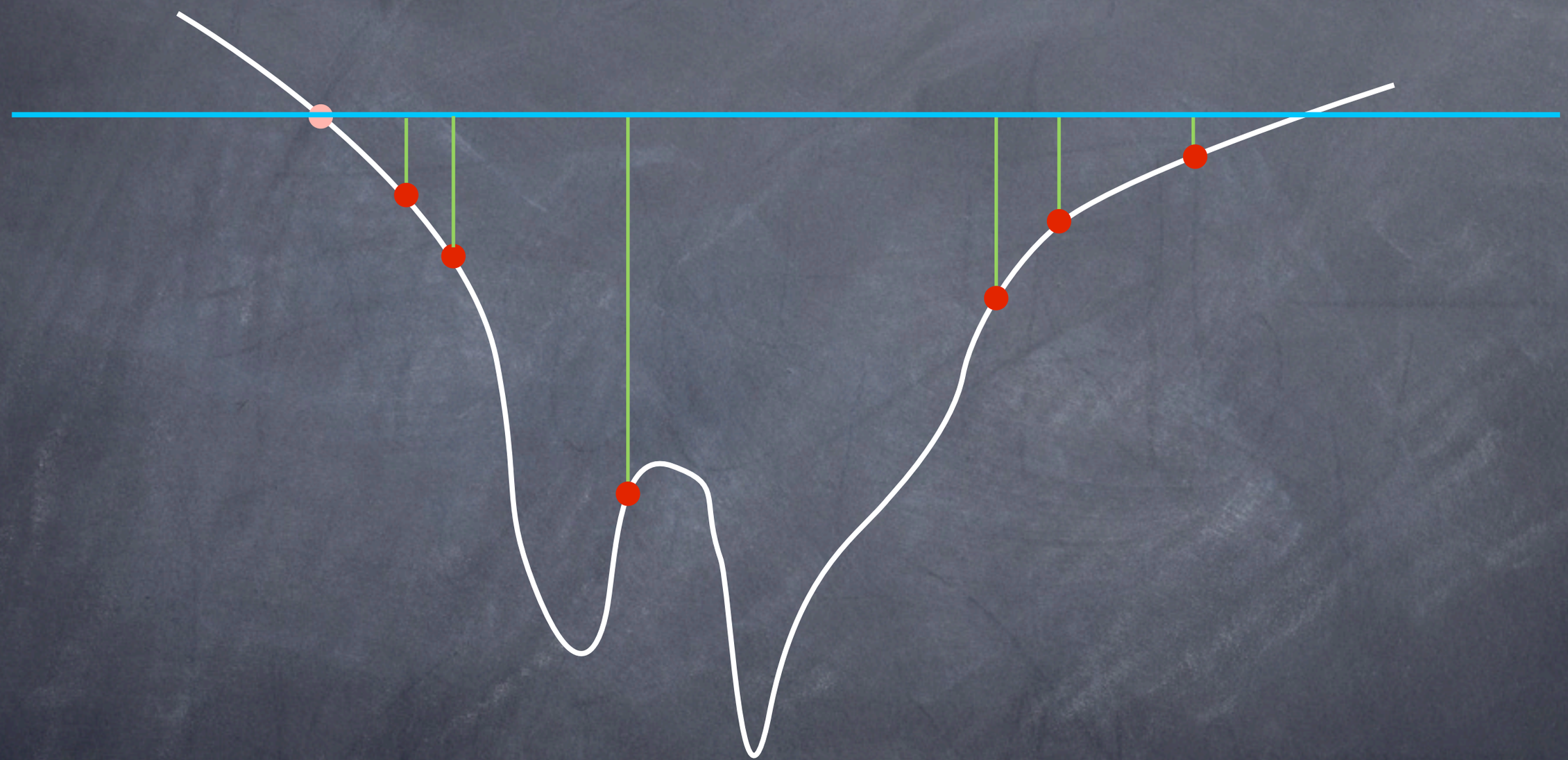


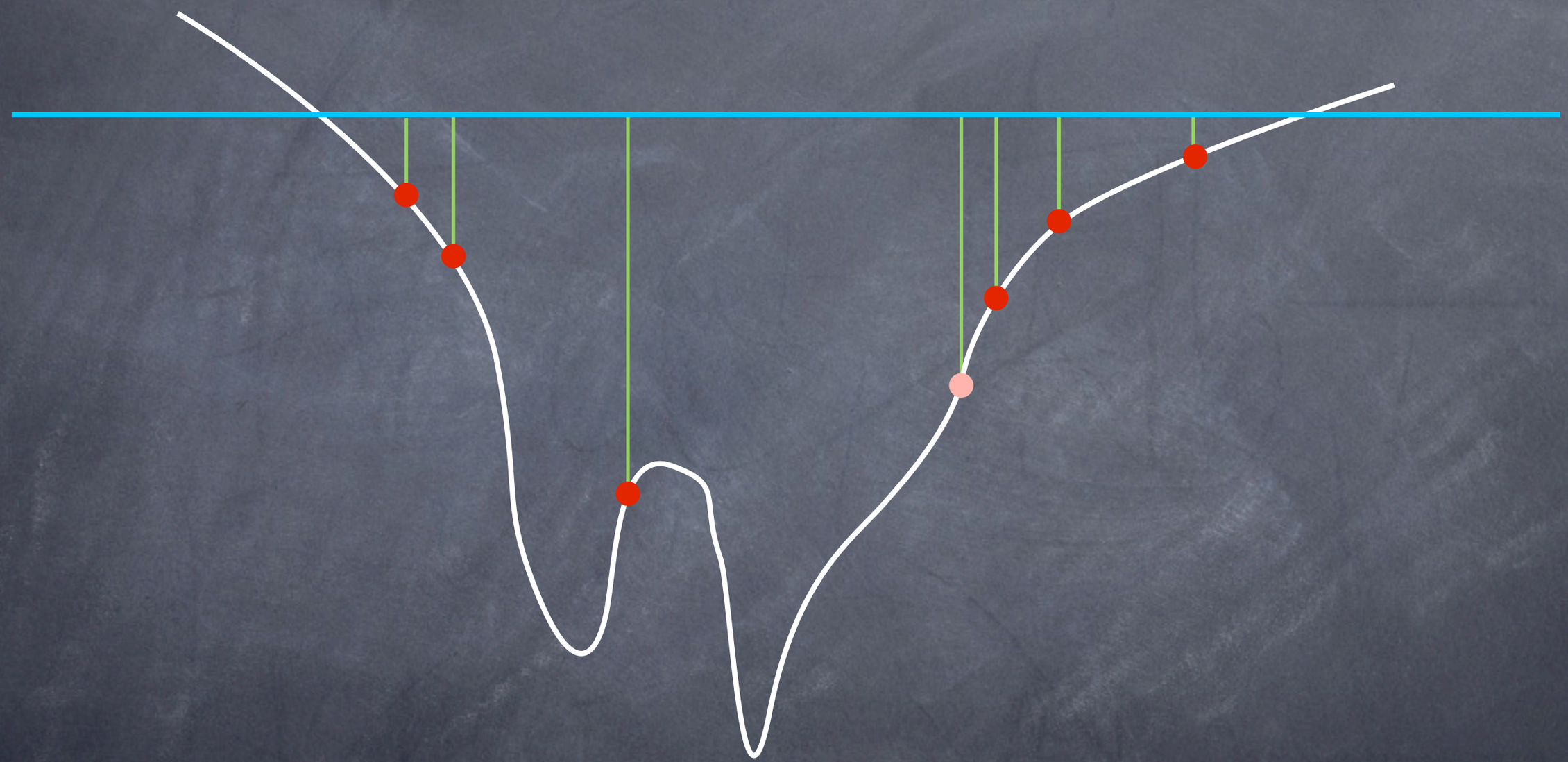


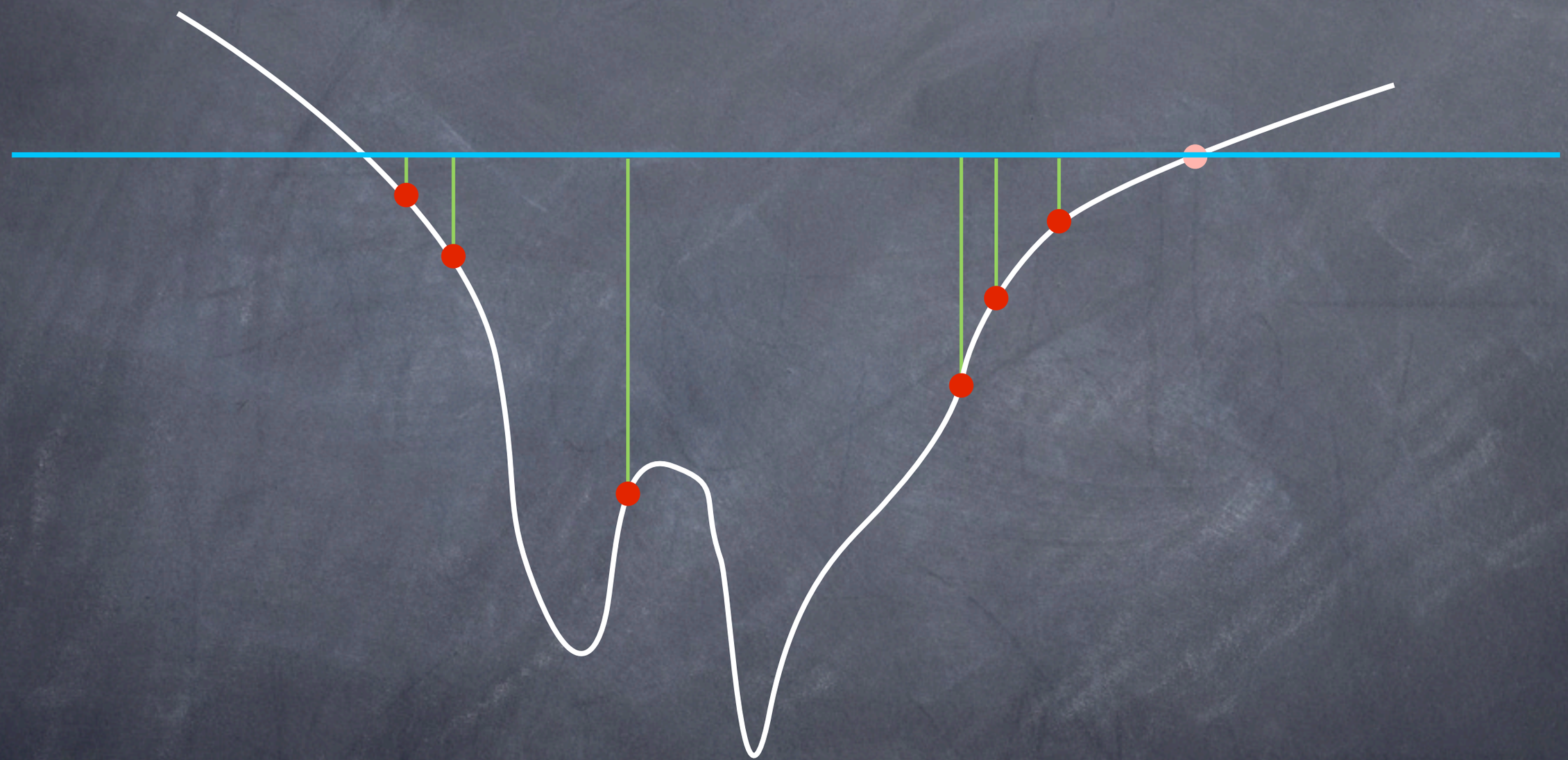


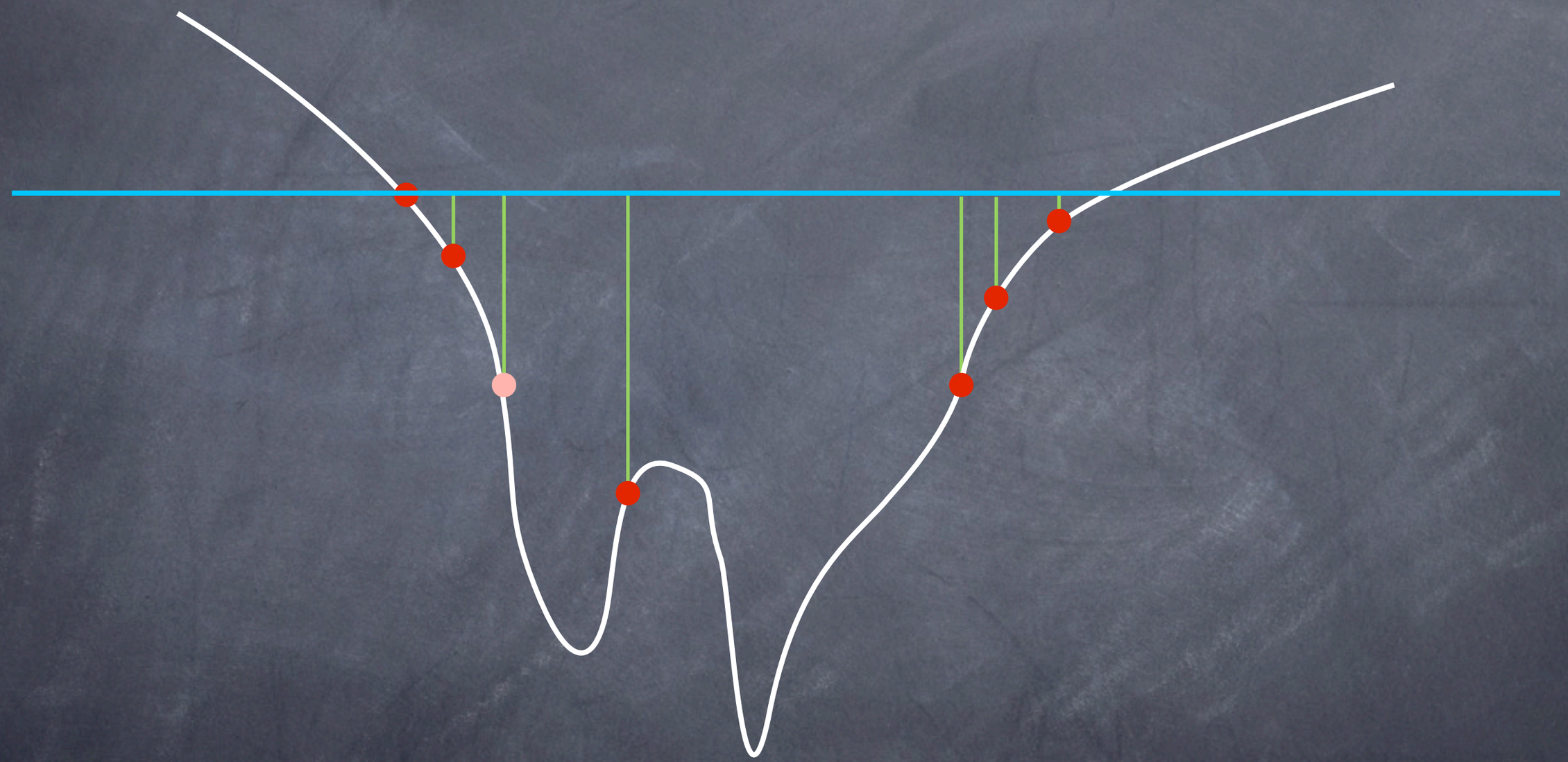


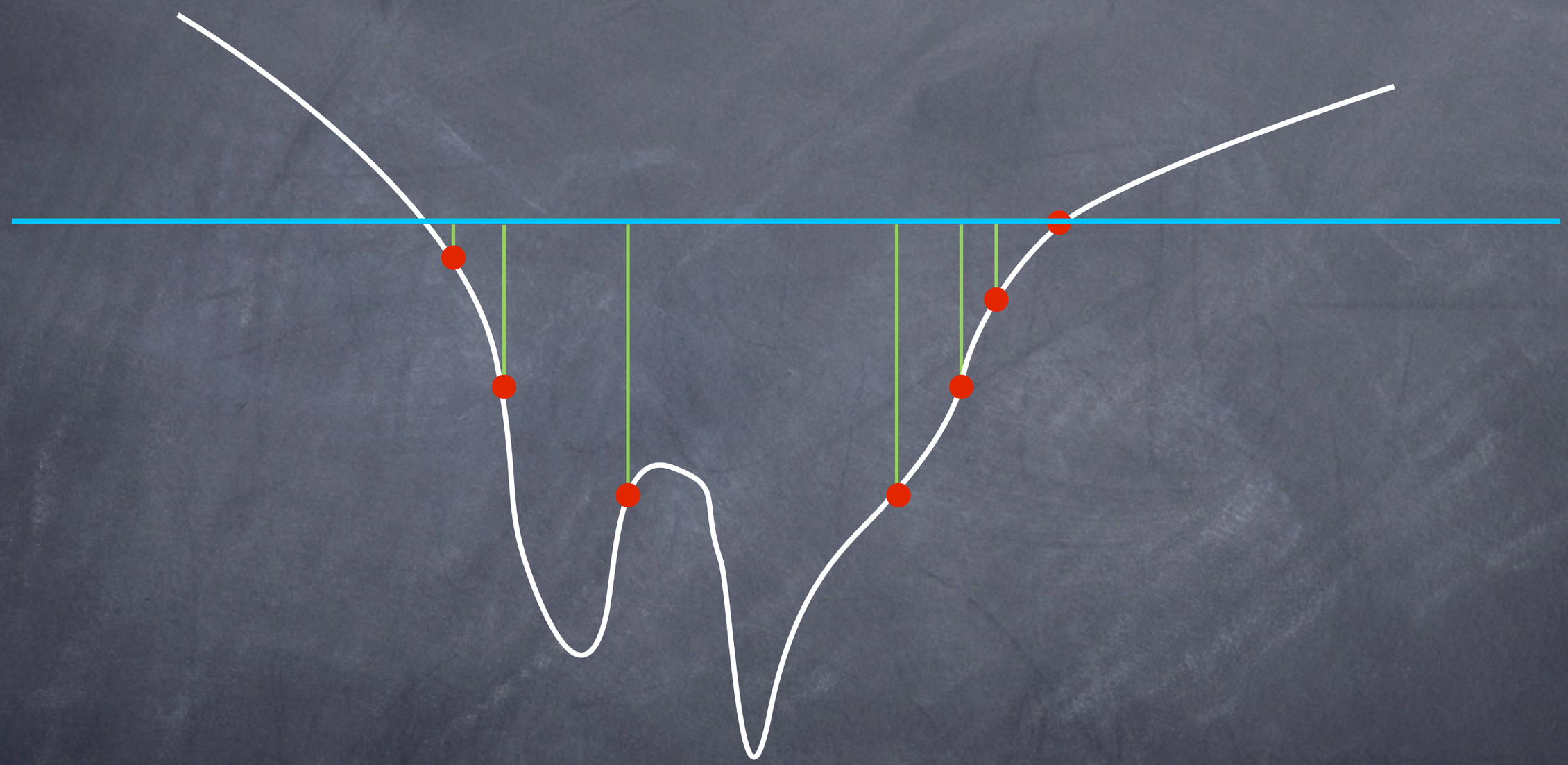


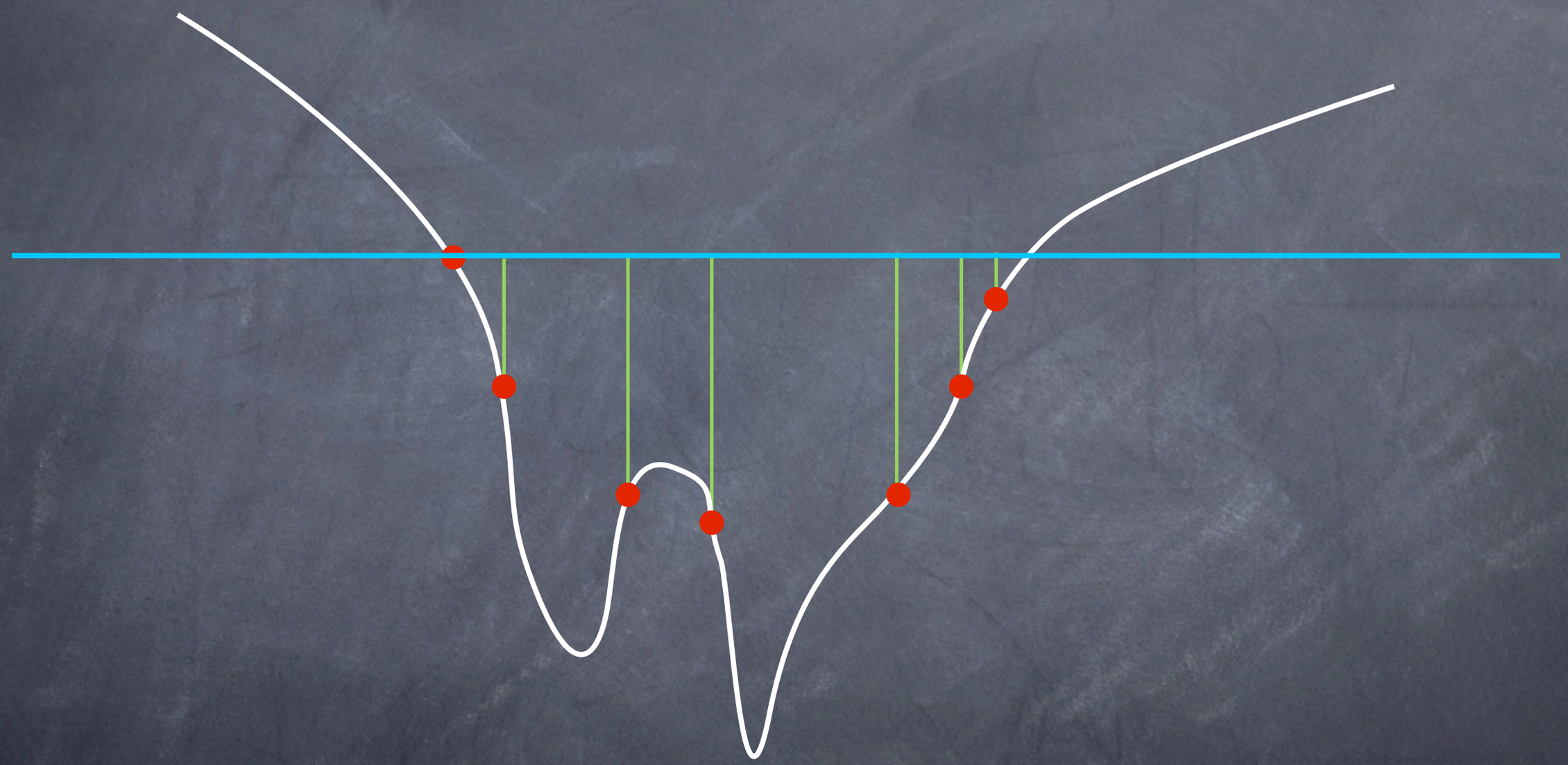


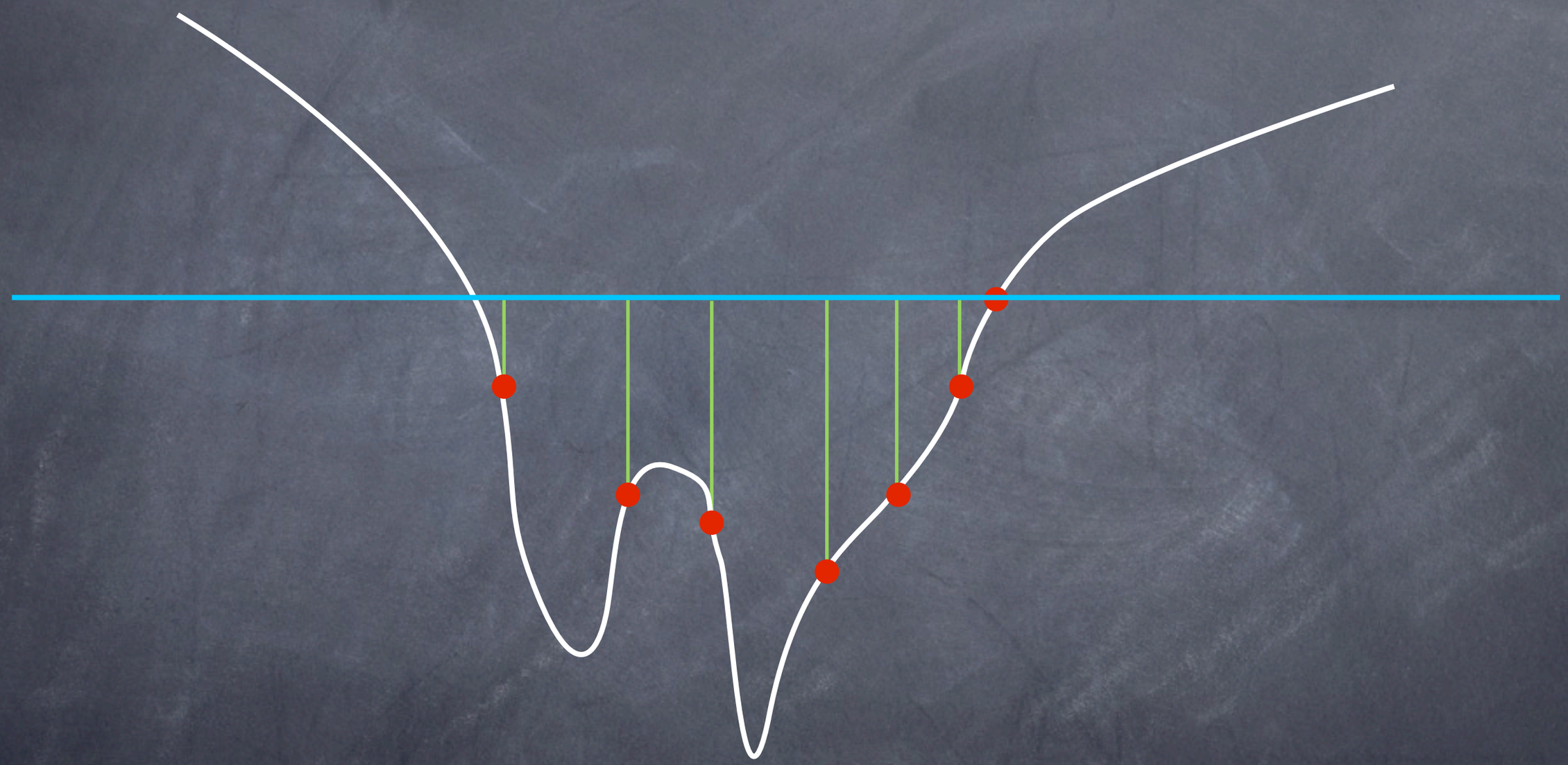


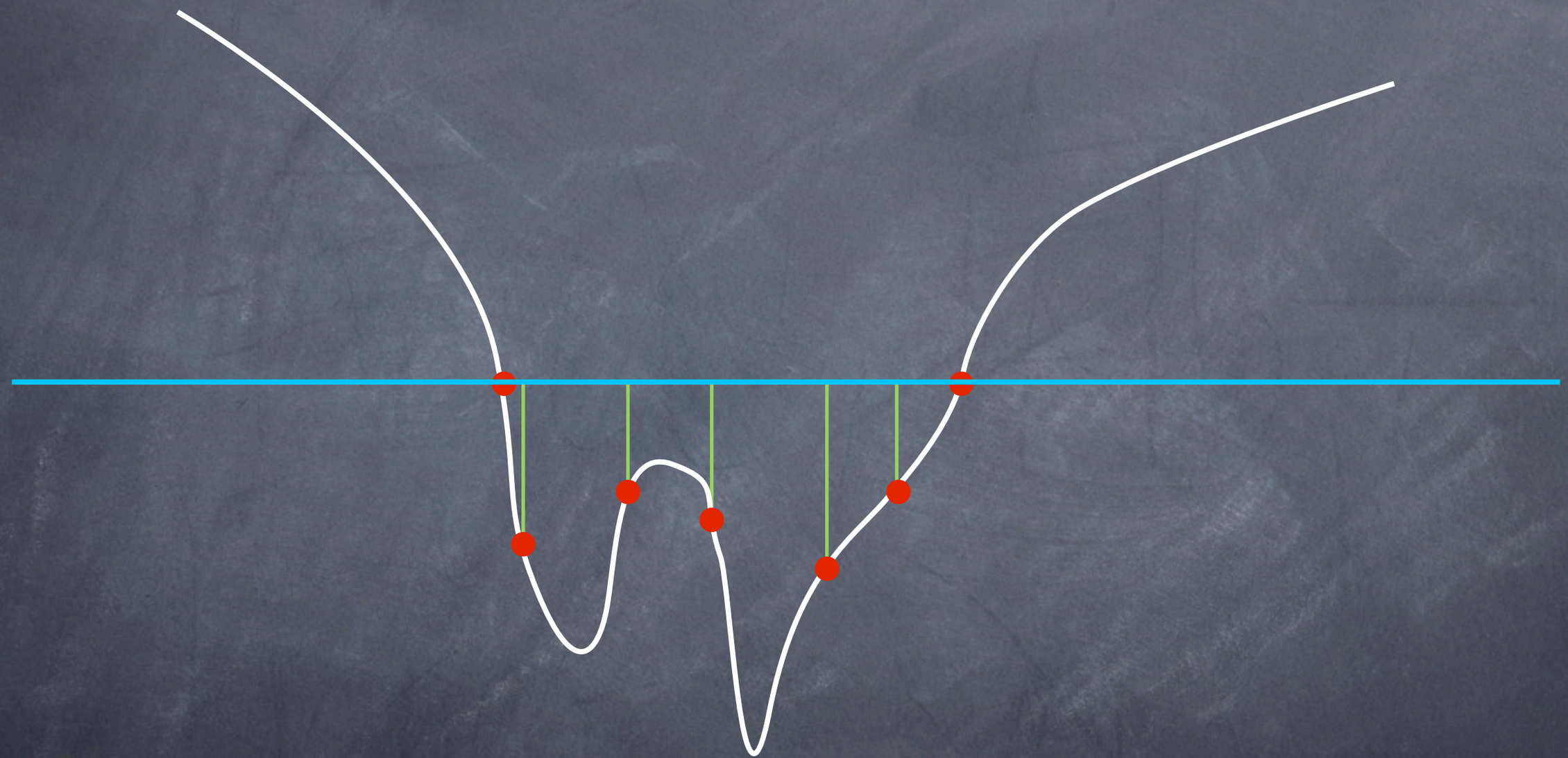


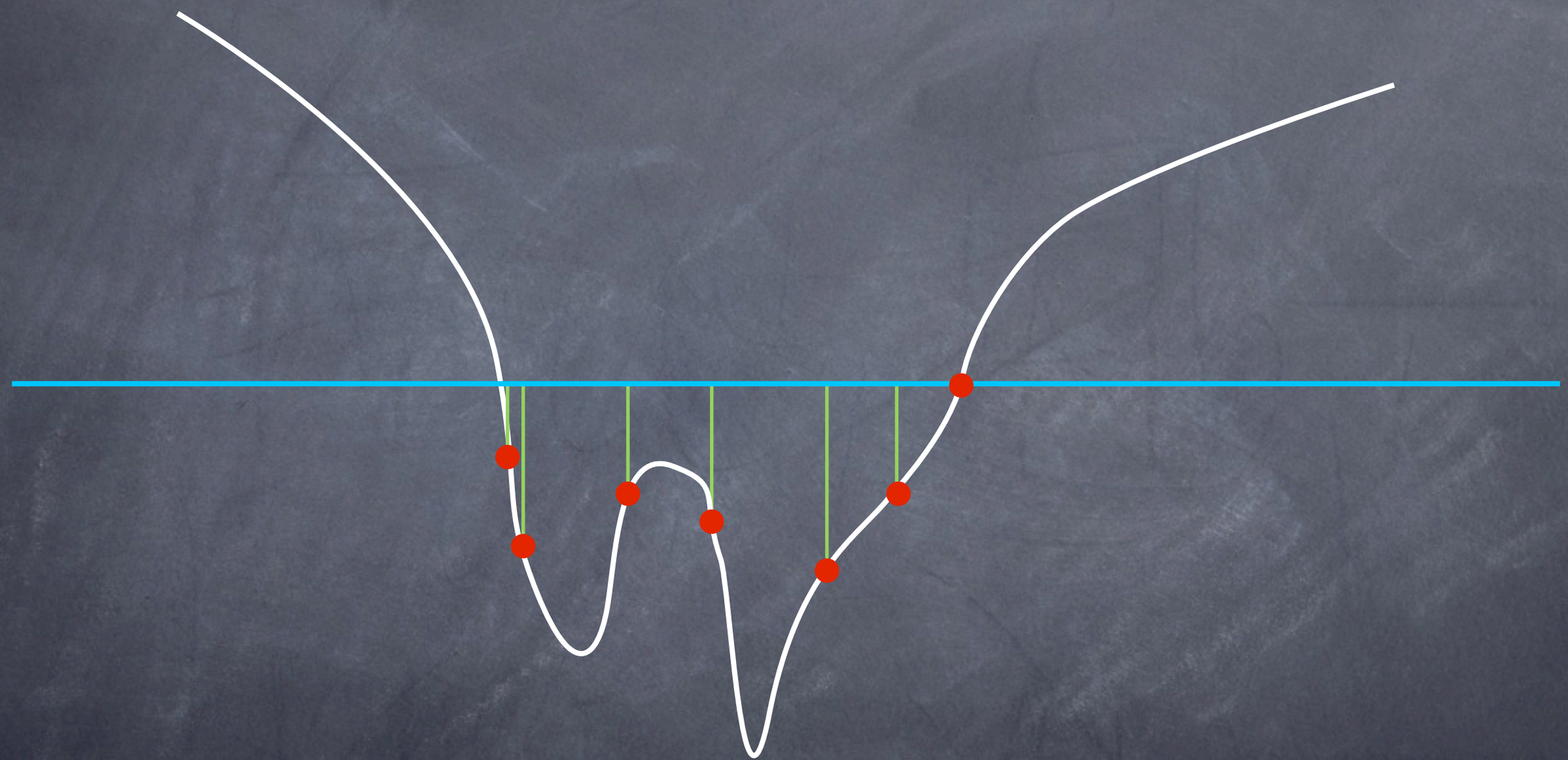


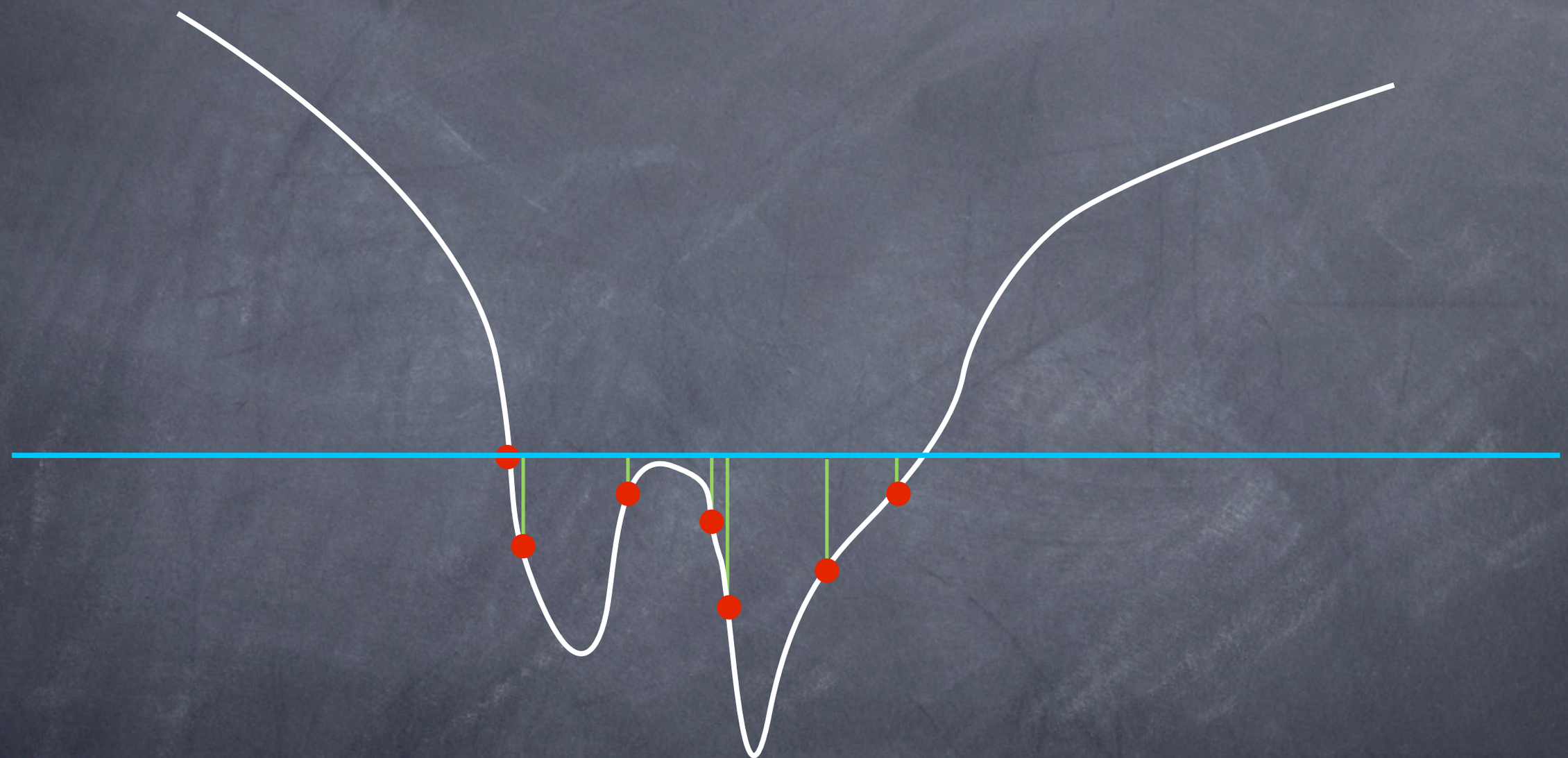


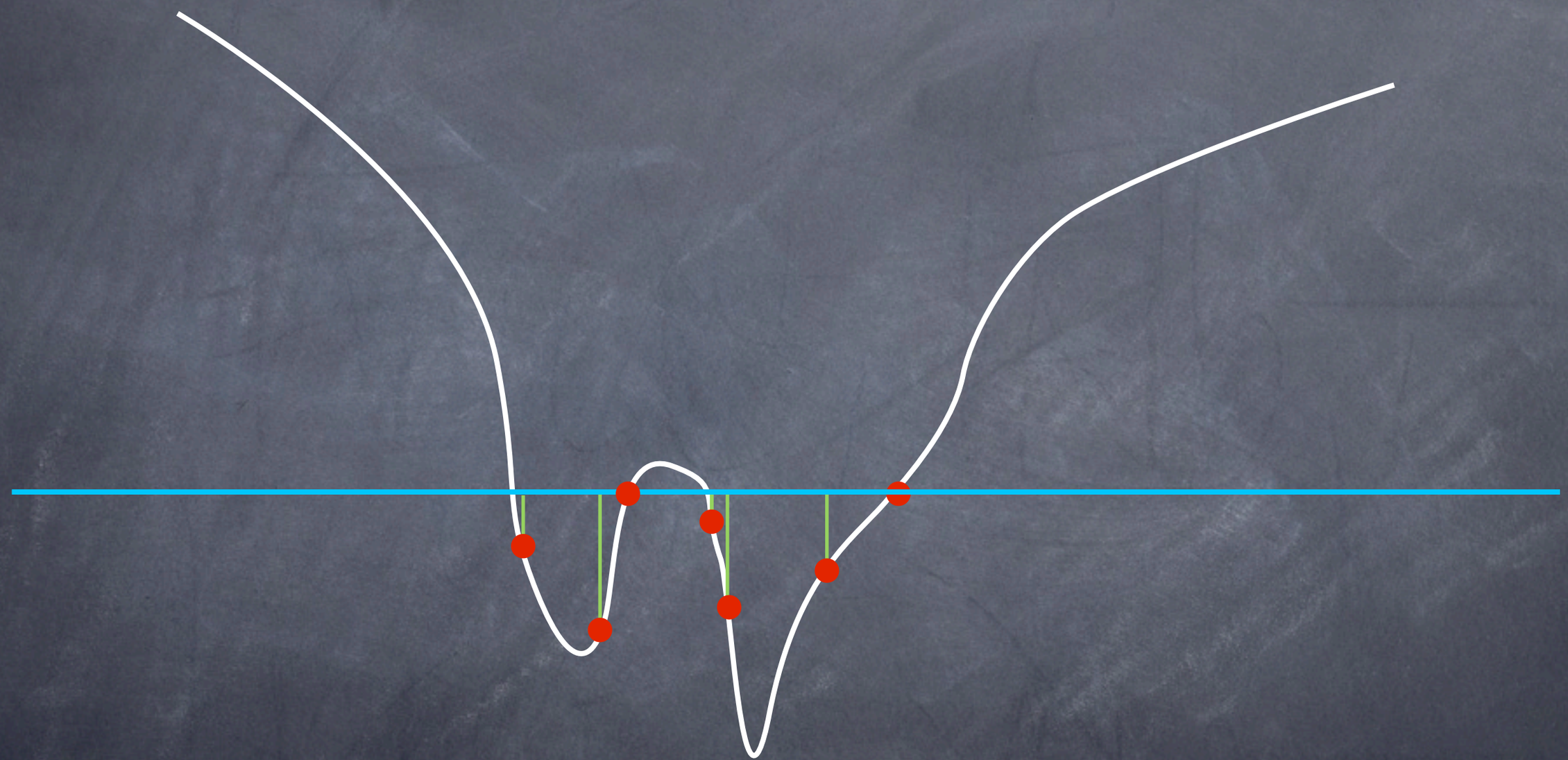


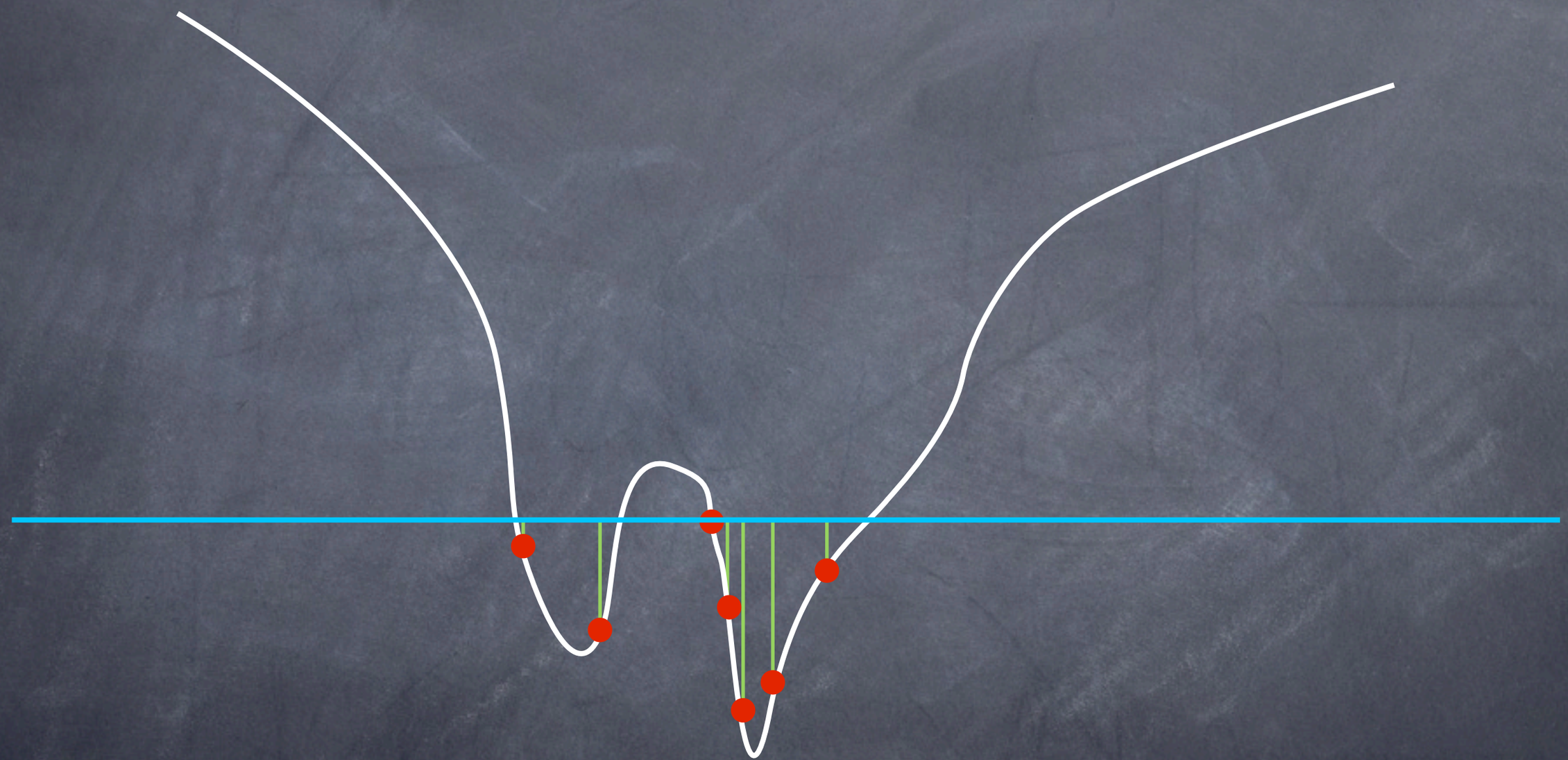


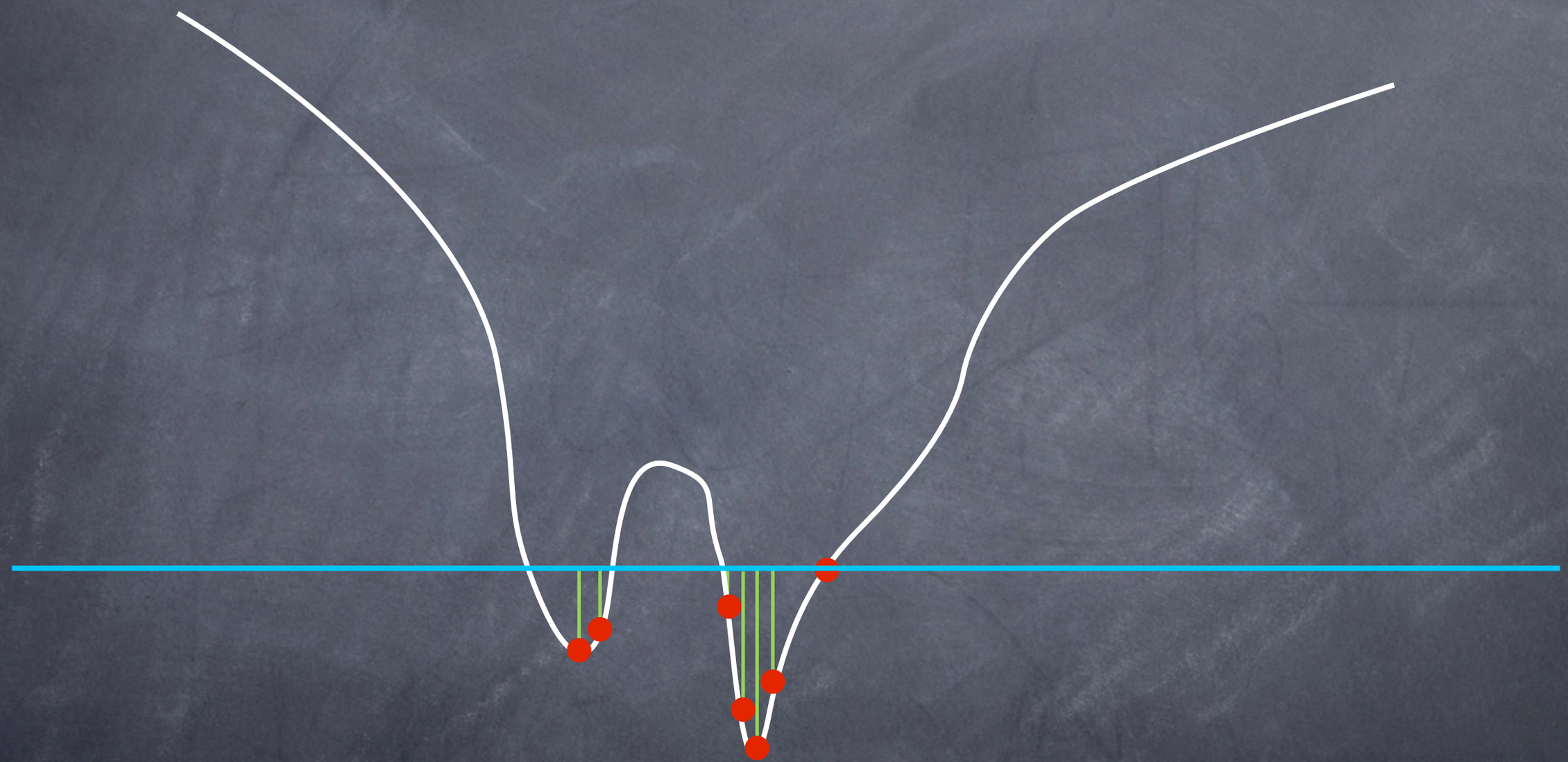


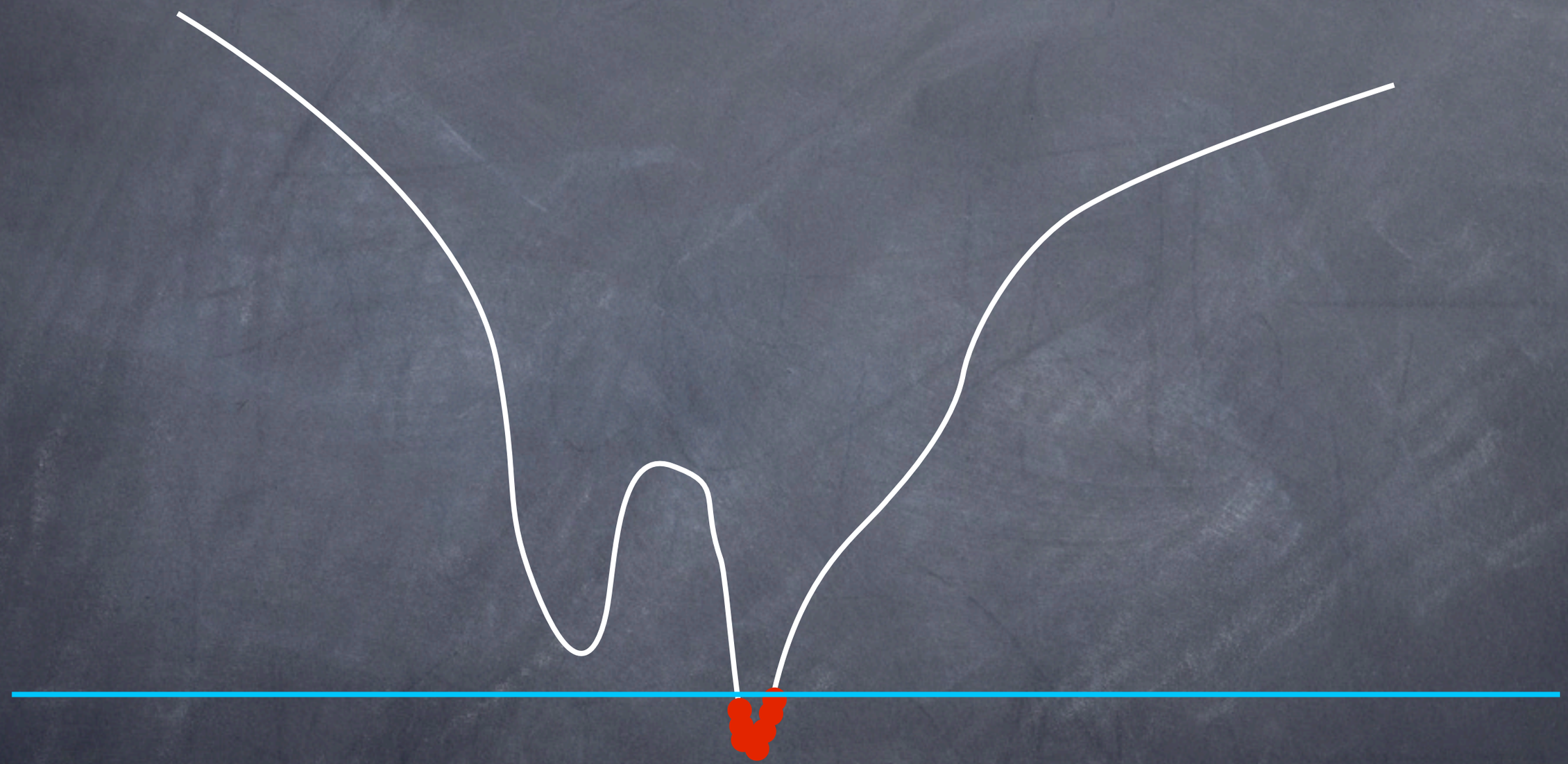




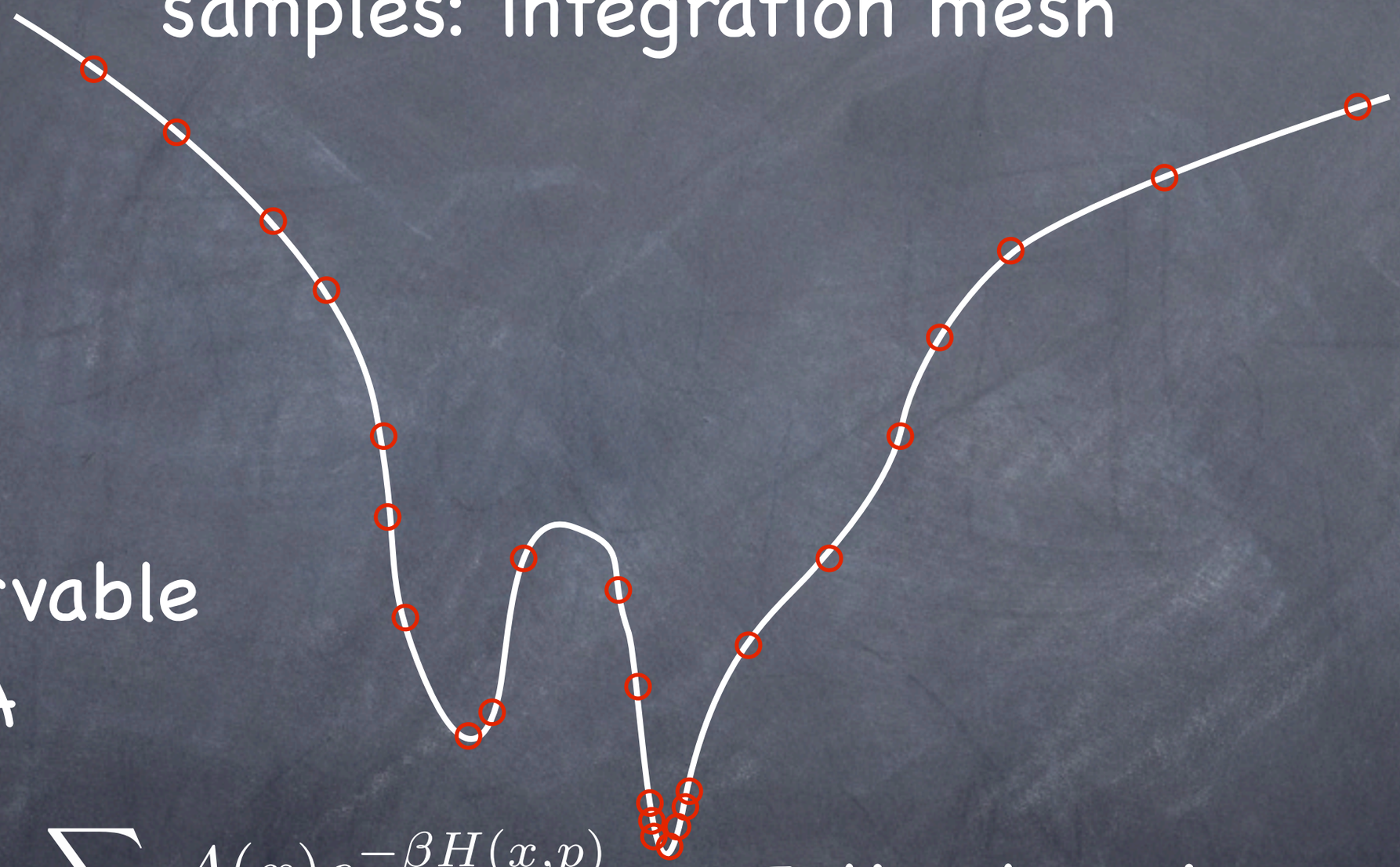








samples: integration mesh



Observable

A

$$\langle A \rangle = \frac{1}{Z} \sum_{\{x,p\}} A(x) e^{-\beta H(x,p)}$$

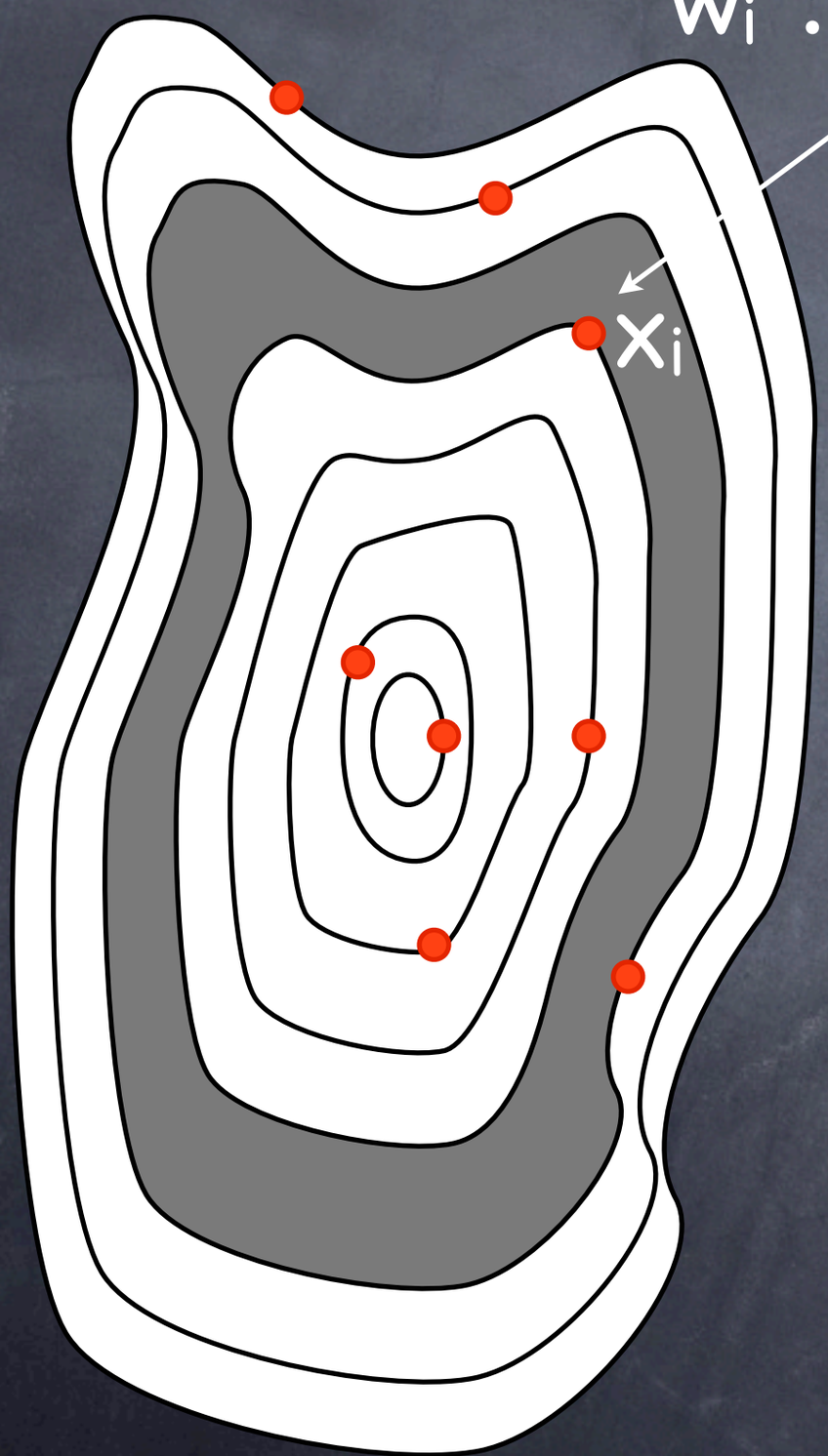
$$Z = \sum_{\{x,p\}} e^{-\beta H(x,p)}$$

Estimate using samples:

$$\langle A \rangle_{\mathrm{est}} = \frac{Z_p}{Z} \sum_i w_i A(x_i) e^{-\beta E(x_i)}$$

Energy contours

w_i : volume of "shell", random variable



Sequence of volumes: Γ_i

$w_i = \Gamma_i - \Gamma_{i+1}$ $\quad t = \Gamma_{i+1} / \Gamma_i$

$$P(t) = N t^{N-1}$$

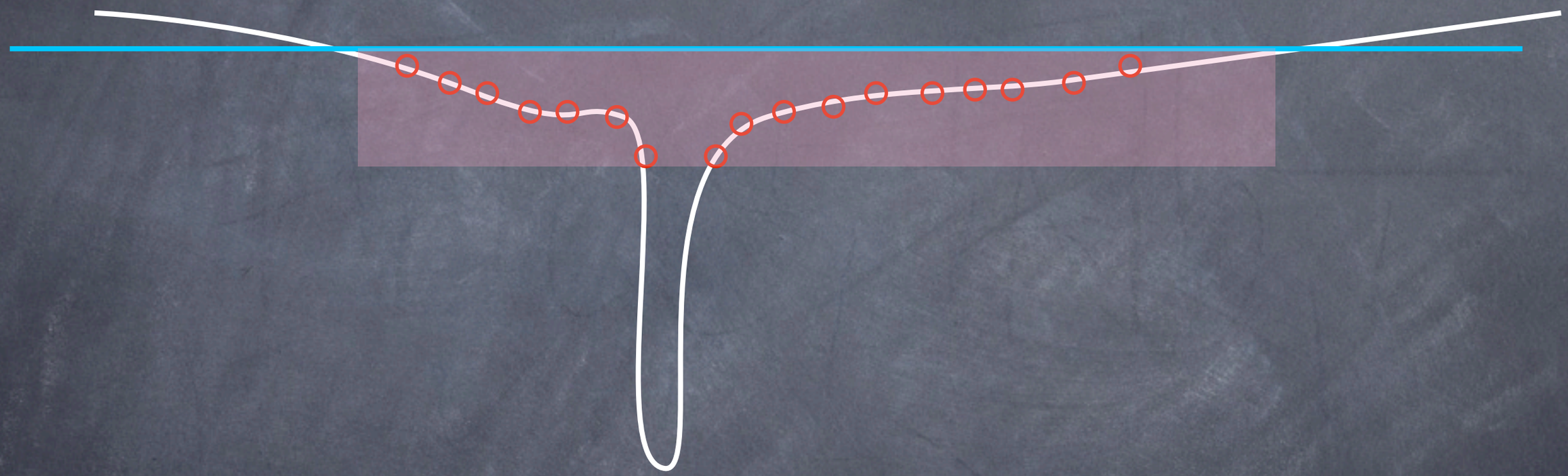
$$\langle \ln \Gamma_i - \ln \Gamma_{i+1} \rangle = \langle \ln \frac{\Gamma_i}{\Gamma_{i+1}} \rangle = -\langle \ln t \rangle$$

$$\langle \ln \Gamma_i \rangle - \langle \ln \Gamma_{i+1} \rangle = -\int_0^1 \ln(t) N t^{N-1} = 1/N$$

$$\langle \Gamma_i \rangle = e^{-i/N}$$

$$\boxed{\langle w_i \rangle = e^{-i/N} - e^{-(i+1)/N}}$$

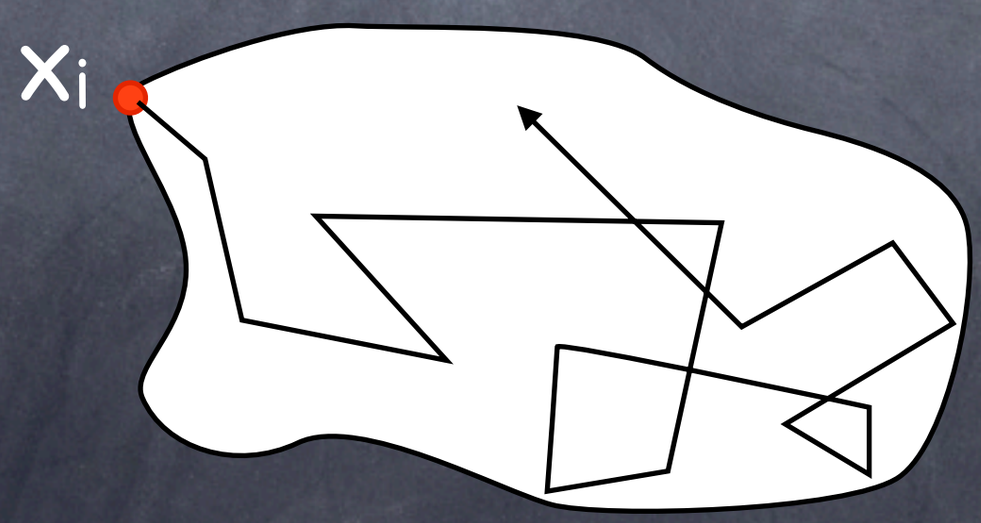
High dimensionality



- Exponential growth of volume
- During the sampling, range of E(x) in the live set is narrow

How to pick new points?

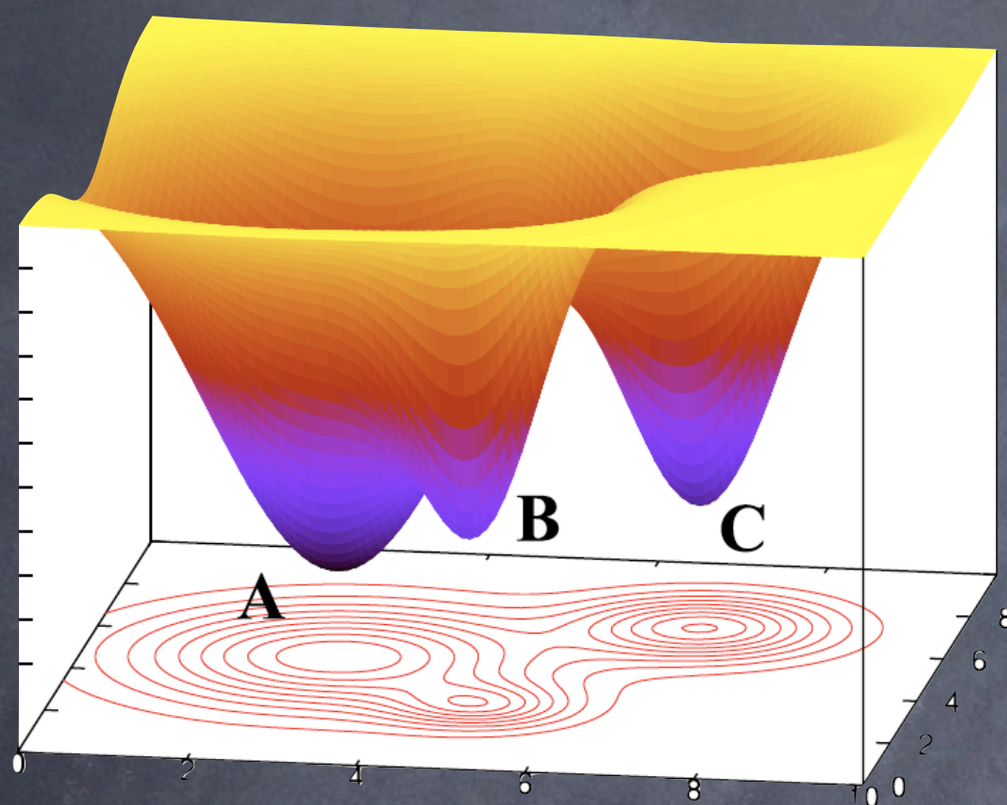
- Need to pick replacement for x_i with uniform probability from $\{x: E(x) < E(x_i)\}$
- MCMC in "flat" space: random walk with ∞ walls starting from x_i



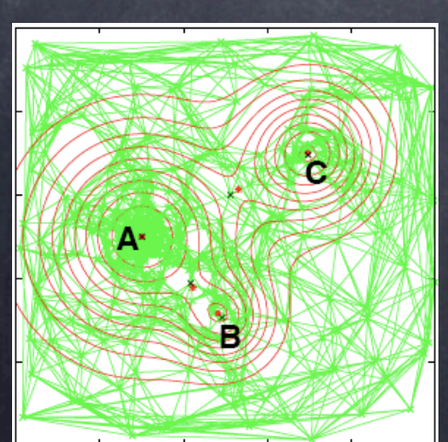
Main points of algorithm

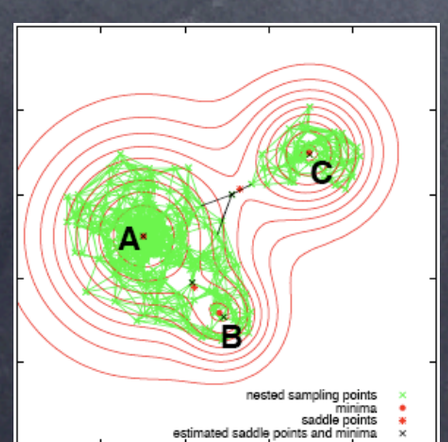
- Converges exponentially
- Independent of temperature β
- Top-down: good ergodicity
- Resolution: $1/N$

Toy model: 3 Gaussians

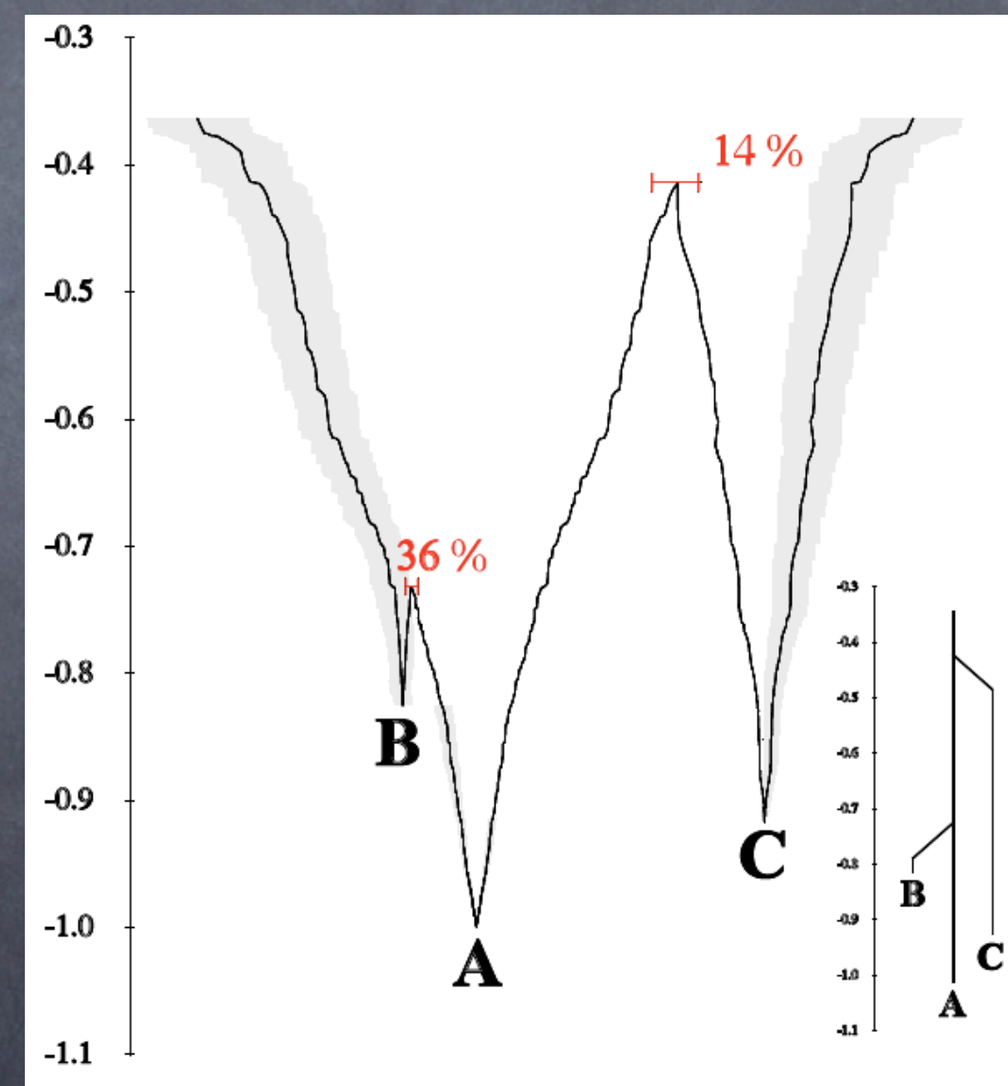


N=100 live points





"Energy Landscape Chart"



phase space volume

Lennard-Jones clusters

$$E_{\mathrm{LJ}} = \sum_{i<j}^{n} 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]$$

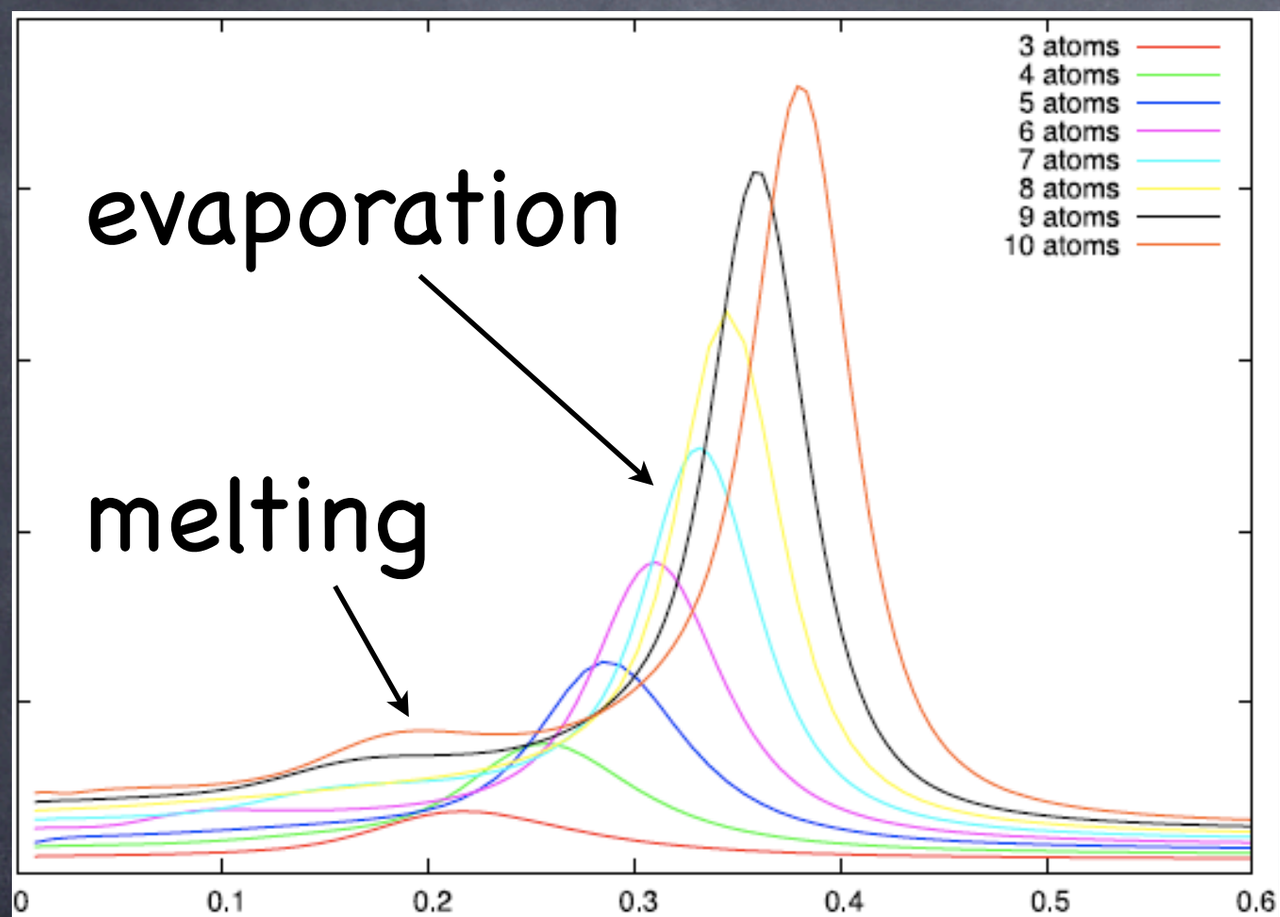
- Partition Function:

$$Z(\beta) = \left(\frac{2\pi m}{\beta} \right)^{3n/2} \frac{V^n}{h^{3n} n!} \sum_i \left[e^{-i/N} - e^{(i+1)/N} \right] e^{-\beta E_i}$$

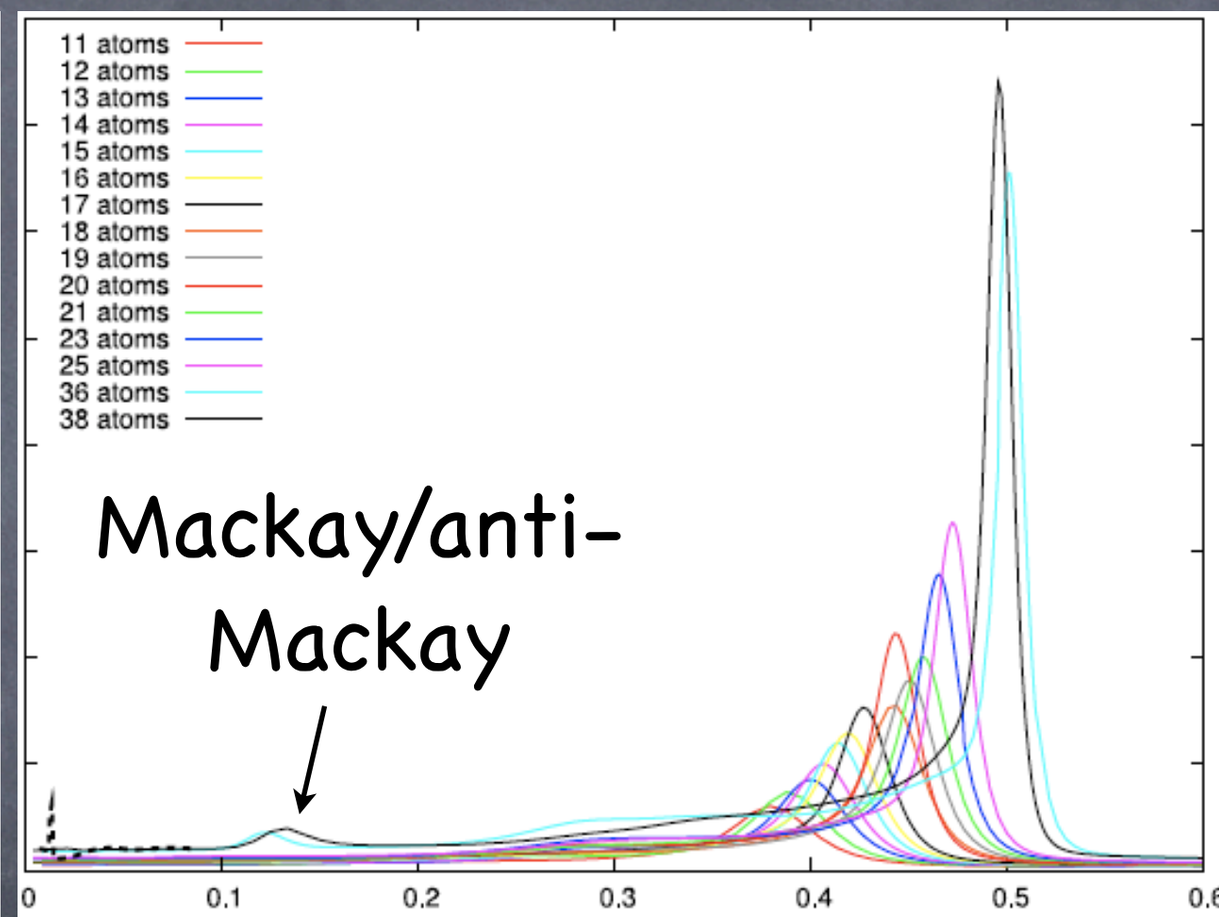
- Internal Energy $U = -\partial \ln Z / \partial \beta$
- Heat capacity $C_V = \partial U / \partial T$

Heat Capacity curves

n = 1–10



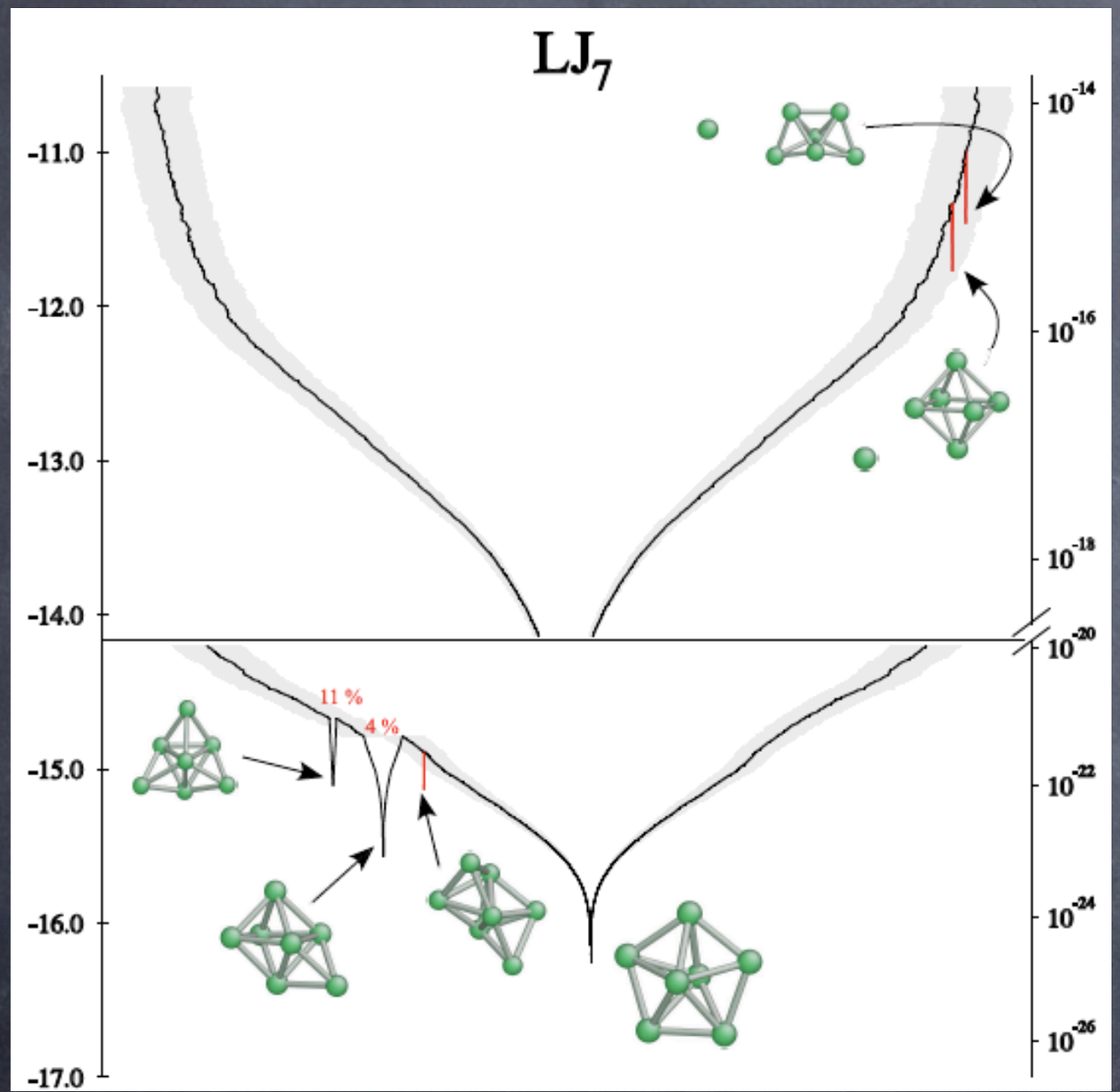
n = 11–38



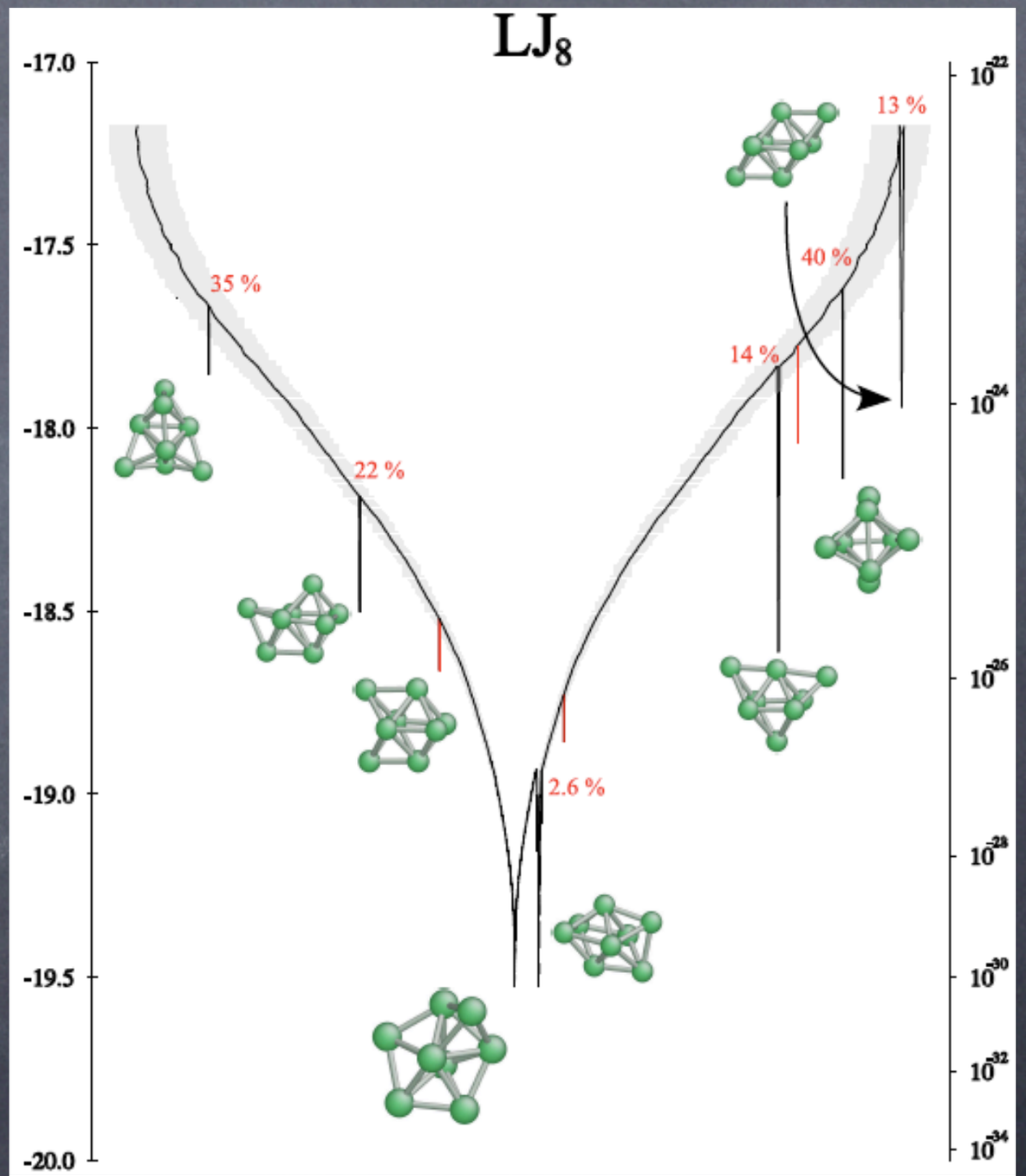
Relative Temperature

$O(10^{10})$ LJ evaluations for largest clusters

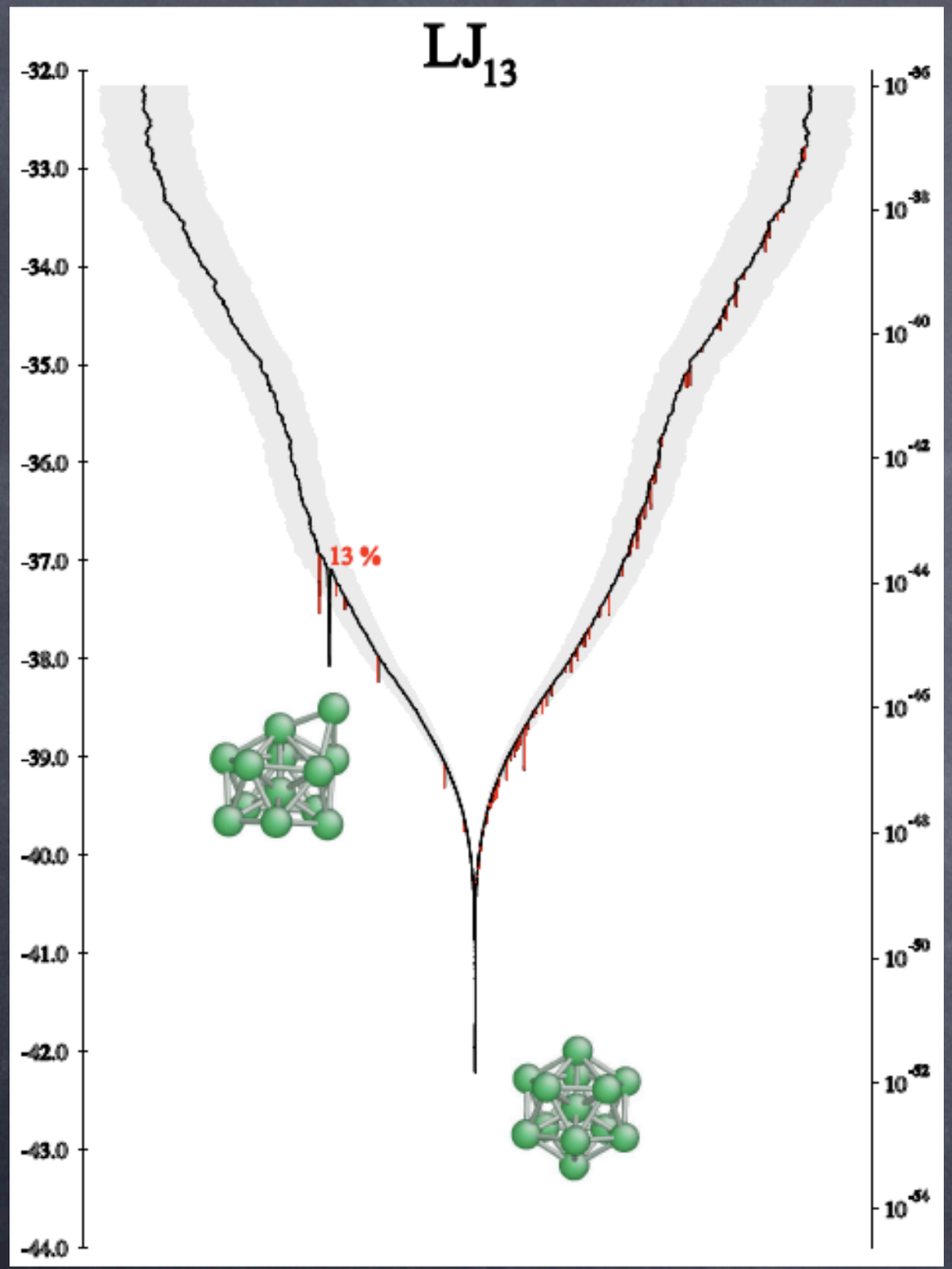
Energy Landscape Charts

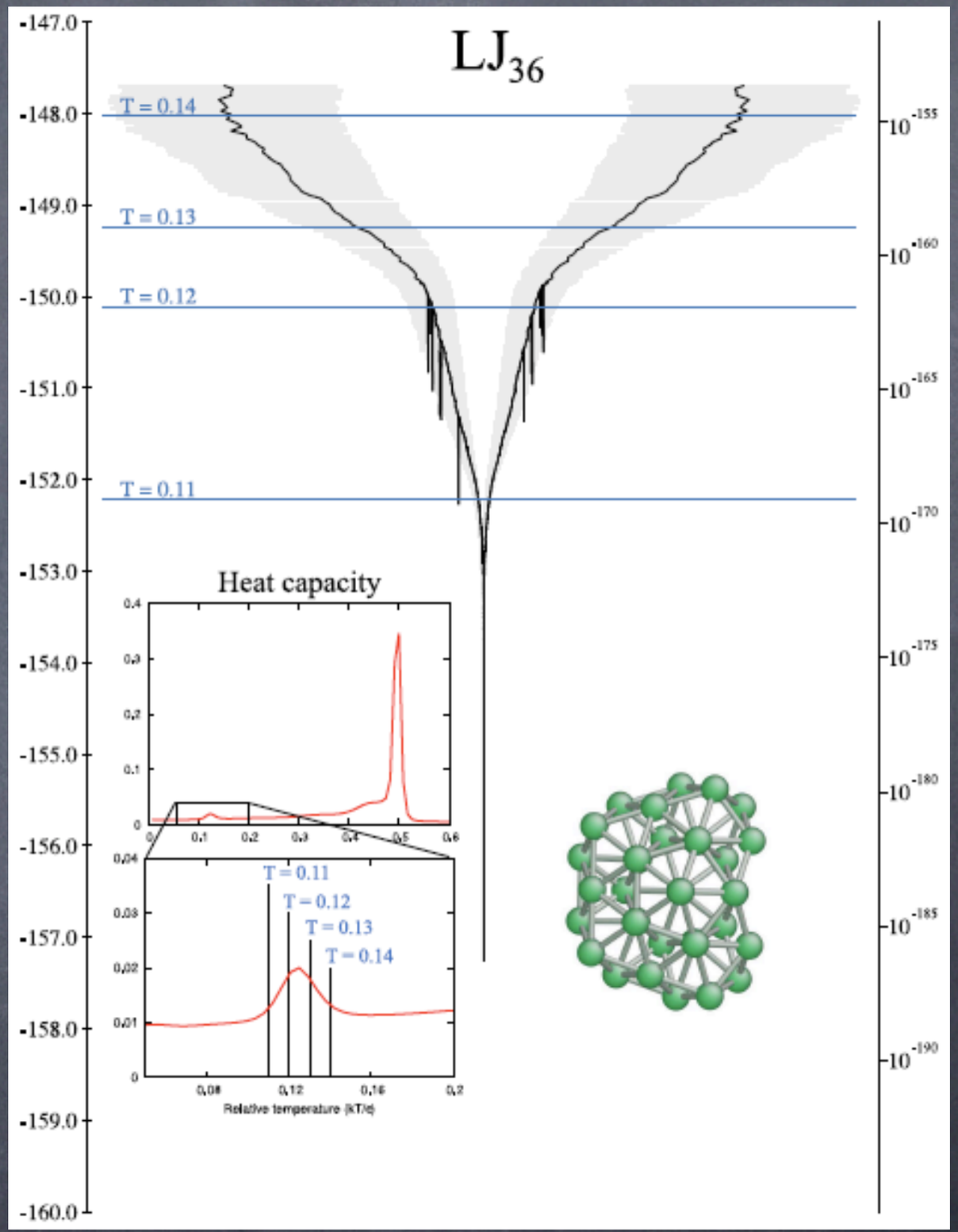


N=5000 live points



N=10000 live points

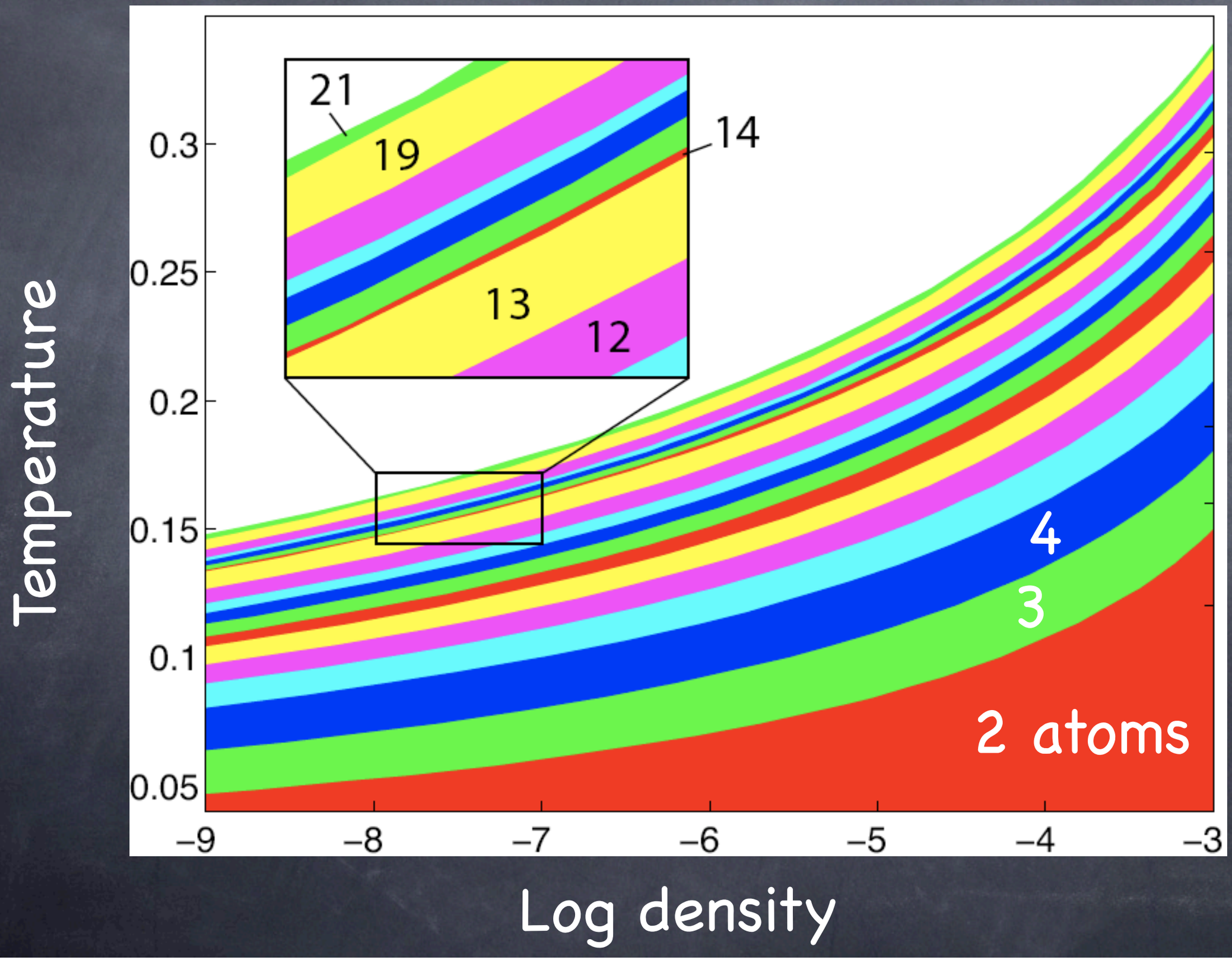




5000 live points

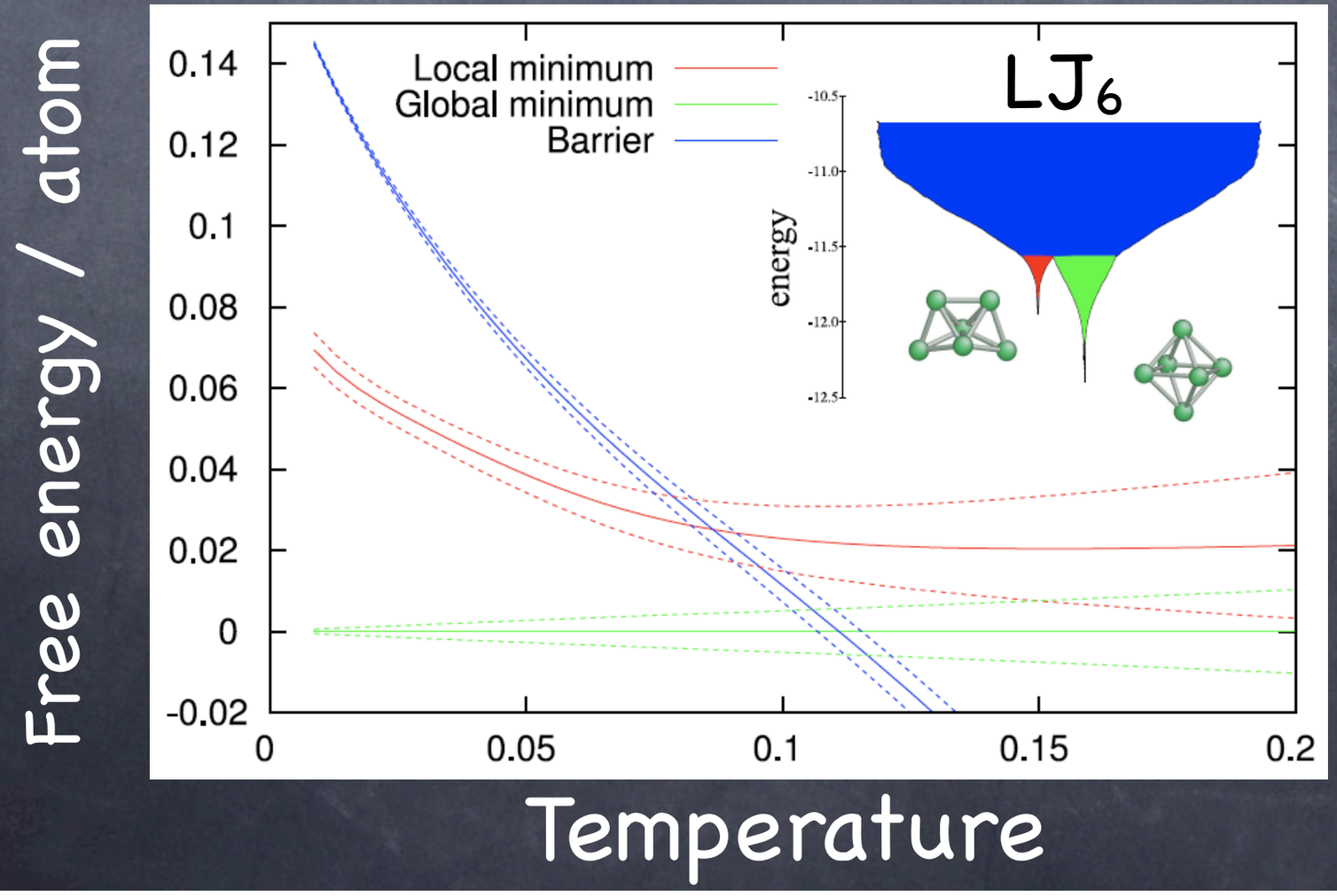
2000 live points

LJ ρ-T phase diagram

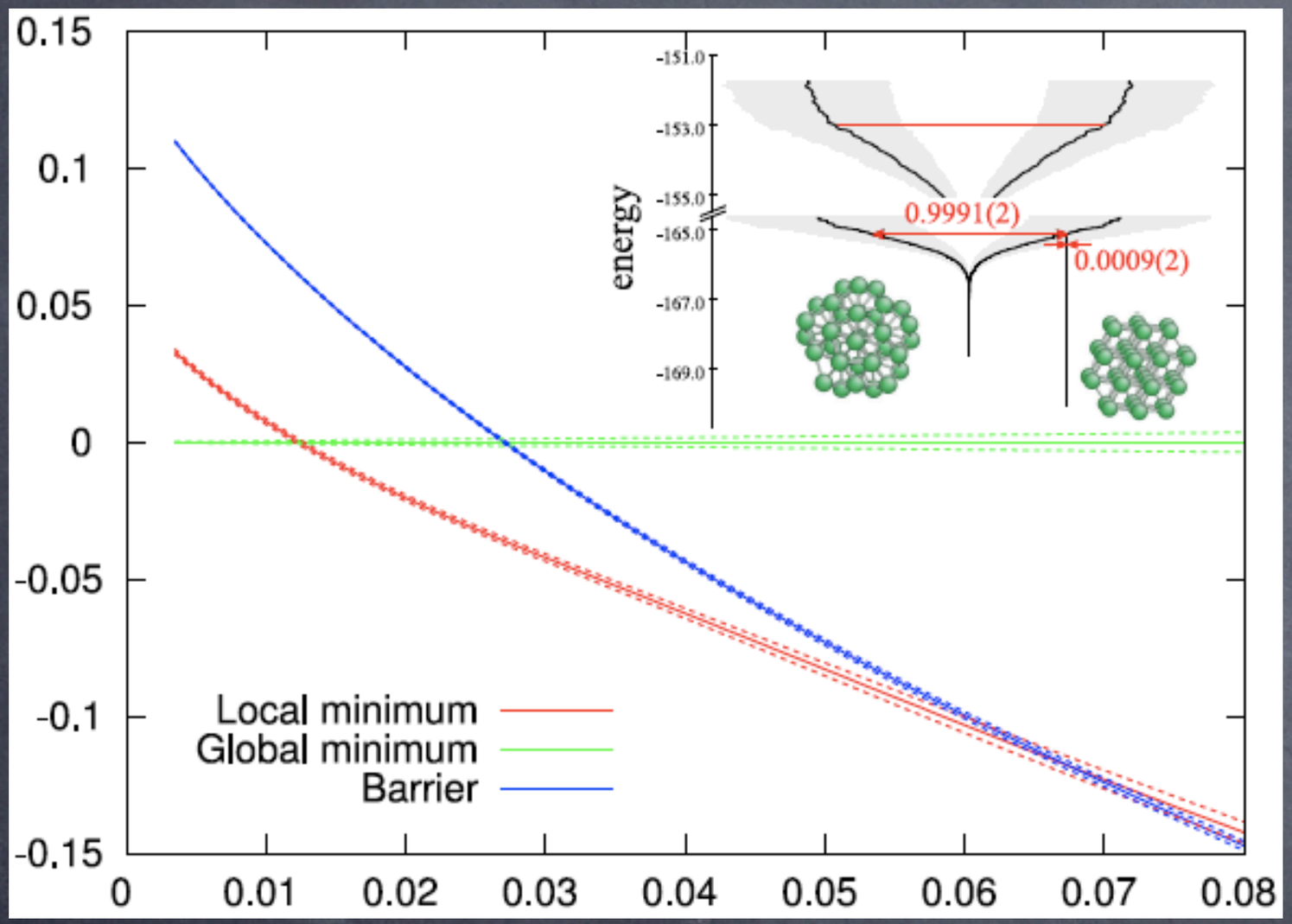


Free energy

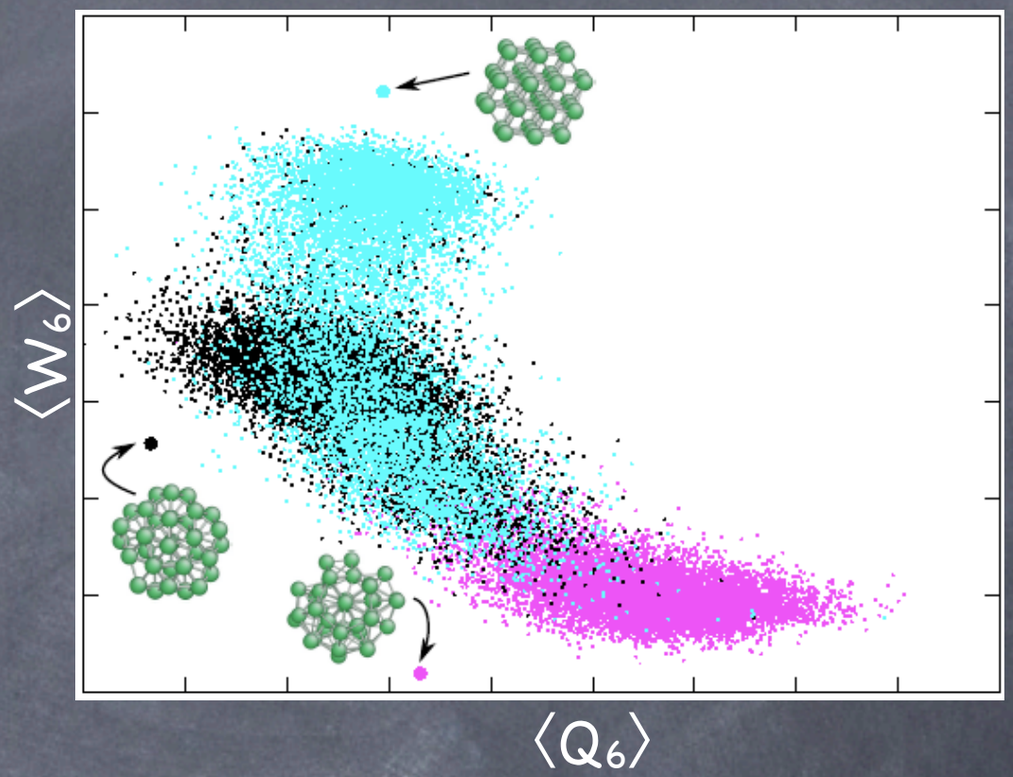
- Macroscopic states : order parameters
- Typically externally defined, ad-hoc
- Microscopically: which basins are occupied?



LJ_{38}

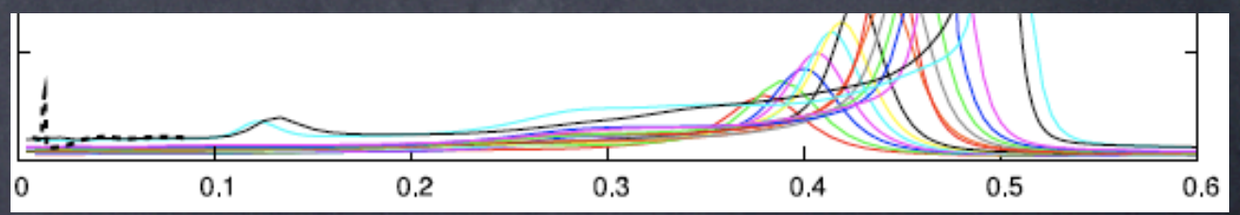


30,000 live points



"Bottom-up" exploration using known minima

Heat capacity peak



Summary

- New ergodic athermal sampling scheme
- Finite resolution Energy Landscape Charts
- Discrete "basin" order parameter: free energy
- Future: smarter ways of picking new points, **build on existing search methods**
- Alternate bottom-up / top-down steps

More acknowledgements

- John Skilling, Farhan Feroz, Mike Hobson
- David Wales, Daan Frenkel