Finding what happens in a space of ever so many dimensions

R. Stephen Berry

The University of Chicago

Brijuni, 28 August–1 September 2006

Some systems surprise us by finding their ways to very unlikely but important 'places'

- Two examples: clusters of alkali halides, such as $(NaCl)_{32}$ or $(KCl)_{32}$, and
- Proteins that fold to active structures, maybe even to unique structures

How do clusters, nanoparticles and proteins find the pathways to where they want to go?

- **This is a question about *real* complexity!**
- **Why? Let’s look at the “world” in which they exist and move**

Think of the effective potential energy surface of the system

- For N particles, 3N dimensions, of which 3N–6 correspond to internal coordinates
- Thus a 6-particle cluster moves on a surface in a space of 12 dimensions
- And the dimension goes up linearly as N,

But...

- The complexity of the surface goes up somewhat (ha, ha) faster.
- Consider a small cluster of argon atoms, say 6 to begin our exploration;
- This has just two minima, an octahedron and a capped pyramid.

This one is simple:

- **The key places on the potential**

J. Chem. Phys., Vol. 110, No. 18, 8 May 1999

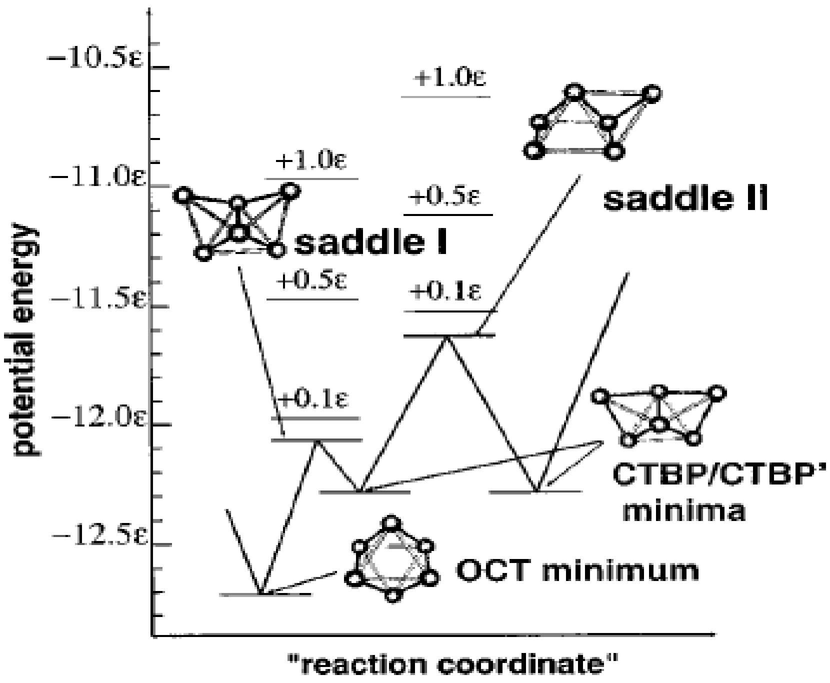


FIG. 1. A schematical picture of the potential energy surface of Ar_6. CTBP′ is a permutational isomer of the CTBP minimum neighboring on the OCT minimum.

How about larger clusters?

- **Here's where the real problem is:**

The picture can't be displayed.

The numbers of minima increase faster than just about anything else

- **The number of geometrically distinct minima increases a bit faster than e^N,**

The numbers of minima increase faster than just about anything else

- **The number of geometrically distinct minima increases a bit faster than e^N,**
- **The number of permutational isomers of each structure goes up approximately as $N!$**

The numbers of minima increase faster than just about anything else

- **The number of geometrically distinct minima increases a bit faster than e^N,**
- **The number of permutational isomers of each structure goes up approximately as $N!$**
- **Do you know anything else that grows as fast as or faster than $N!e^N$?**

Yes! The number of saddles!

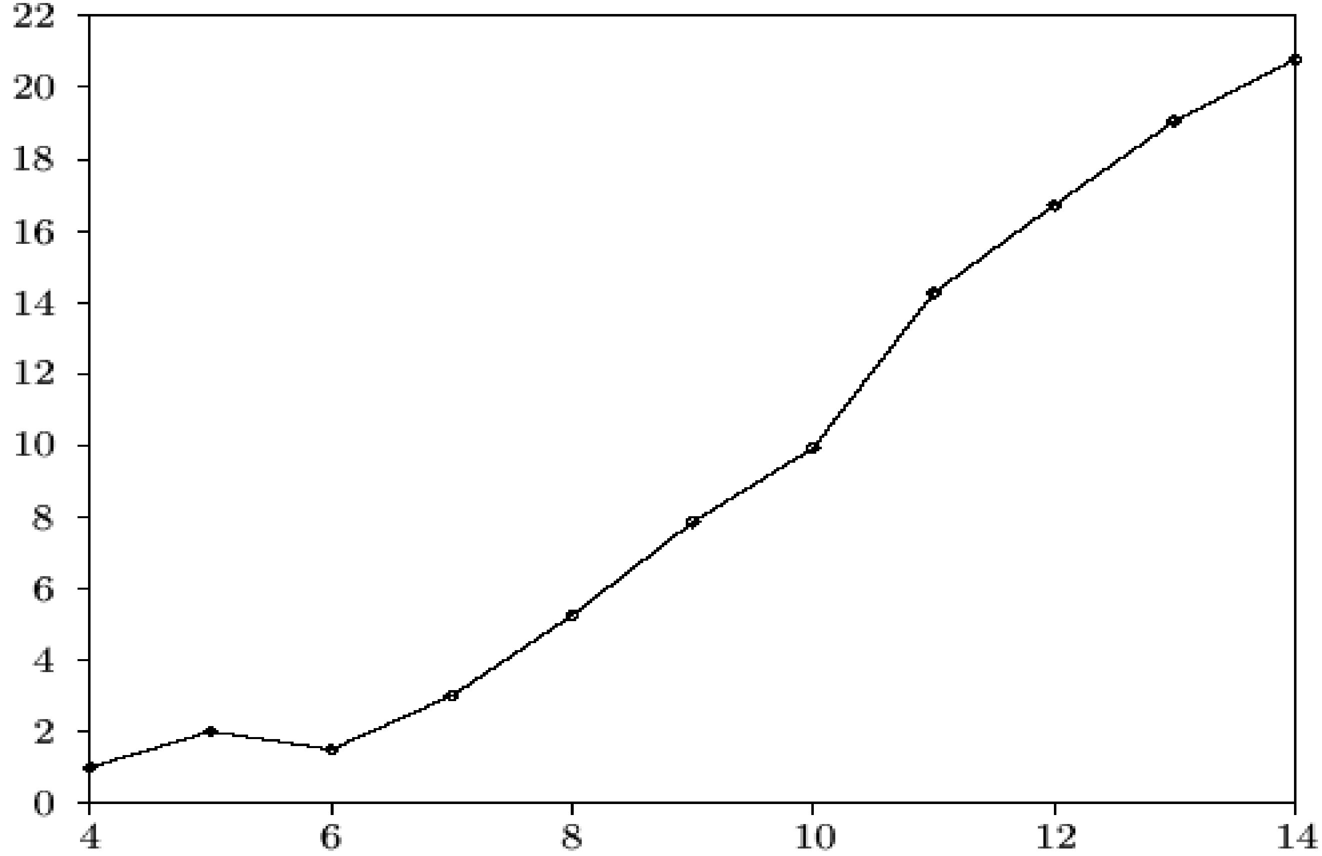


Fig. 5.1. The ratio of the number of transition states to the number of minima for LJ_N clusters as a function of N (8).

David Wales found minima and saddles for small clusters

Table 5.2. *The number of different stationary point structures of index I for LJ_N clusters (8). The numbers in italics are likely to be only weak lower bounds.*

I	N = 4	5	6	7	8	9	10	11	12	13
0	1	1	2	4	8	21	64	170	513	1505
1	1	2	3	12	42	165	635	2423	7868	25653
2	2	4	13	44	179	867	4074	17109	*27957*	*88079*
3	1	6	24	98	494	2820	16407	*47068*		
4	1	6	30	168	1000	6729	46277			
5	0	2	26	190	1458	12093	*97183*			
6	0	1	16	168	1619	16292				
7	0	0	5	101	1334	16578				

A qualitative insight connects topographies with behavior

- We can begin to understand how topography determines whether a system “gets lost” among amorphous structures or finds its way to a “special” structure, one of a very few on the surface
- **Call these two extremes “glass formers” and “structure seekers”**

Find minima and the saddles between them

- Construct the sequences of these stationary points, with the energies of the *minima* in monotonic sequences
- Examine zigzag pathways, the “monotonic sequences” the qualitative patterns of these points

A ‘sawtooth’ potential signifies a ‘glass former’ --Ar_{19}

The picture can't be displayed.

A ‘staircase potential’ gives a ‘structure seeker,’ e.g $(KCl)_{32}$

The picture can't be displayed.

One useful insight already

- **Short-range interparticle forces lead to glass-forming character**
- **Long-range forces *and effective* long-range forces lead to structure-seeking**
- **Polymer chains induce *effective* long-range forces**
- **Alkali halides and foldable proteins are excellent structure-seekers; large rare gas clusters make glasses**

How can we begin to think of what happens when a system relaxes in such a space?

- **Conventional molecular dynamics Is limited to short time spans**
- **Monte Carlo explores the surface but has no information about real motion**
- **One alternative: multiscale molecular dynamics (for another seminar…)**
- **Another: use *kinetics***

We do multiscale dynamics with proteins, but let's discuss kinetics here

- **Kinetic equations: rates of change of populations of local minima, net rate for basin *k* is rate of flow in, less rate of flow out**
- **Assume, with justification, that vibrations thermalize to equilibrium in each local well**
- **Full kinetics is thus described by a set of coupled differential equations, called the "Master Equation"**

Our Master Equation is a set of $\mathcal{N}$ coupled linear, 1st-order equations; $\mathcal{N}$ = # of minima

- **Solve by diagonalizing the (symmetrized) matrix of rate coefficients**
- **For 13 Ar atoms, $\mathcal{N}$ ~1500, quite manageable,**
- **But we certainly wouldn't want to do $\mathcal{N}$ = 25 with all the minima it has!**

What does the Master Equation tell us?

- The eigenvalues, Λ_j, of the matrix of rate coefficients are decay rates for simple exponentials; Λ_0 is the unique eigenvalue for equilibrium, so $\Lambda_0 = 0$. All others are negative. Values near 0 are slow decays.
- The eigenvectors, μ_j, are those particular distributions of concentrations among minima that would decay by simple exponentials.

How can we overcome the problem of immense matrices?

- **We certainly don’t want to diagonalize a 10^8 x 10^8 matrix!**
- **One way: Use matrices that are statistical samples of the full matrix**
- **Another way: extract other information from the master equation**
- **We’re trying both.**

What is important to learn?

- The *slow* processes are usually the most important for us,
- So how can we construct statistical samples of a very rich, large, but rather sparse matrix that will give reliable approximations to the *slow* eigenvalues?
- It’s an unsolved problem in statistics, it turns out.

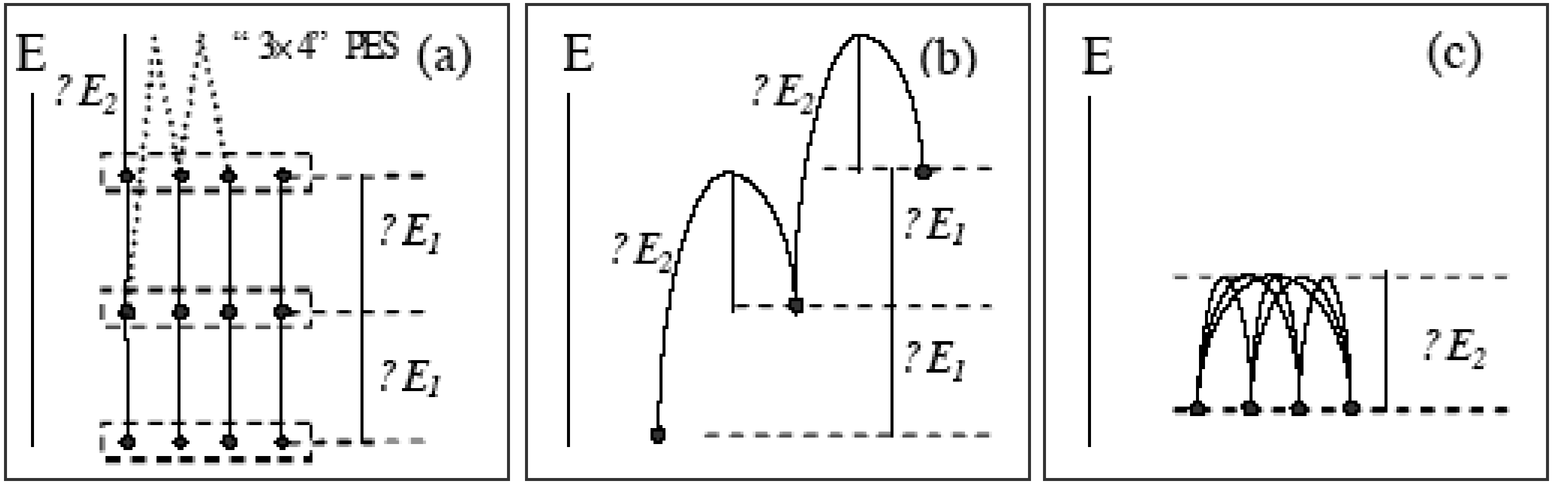
Michael Stein: “Gee, do the experiment!

- **And so we are doing…**
- **Test various kinds of sampling on a model “landscape,” really a net of stationary points representing minima and saddles**
- **Alternatively, study time autocorrelation functions and fluctuations inherent in the chosen master equation**

The Sampled-Network approach

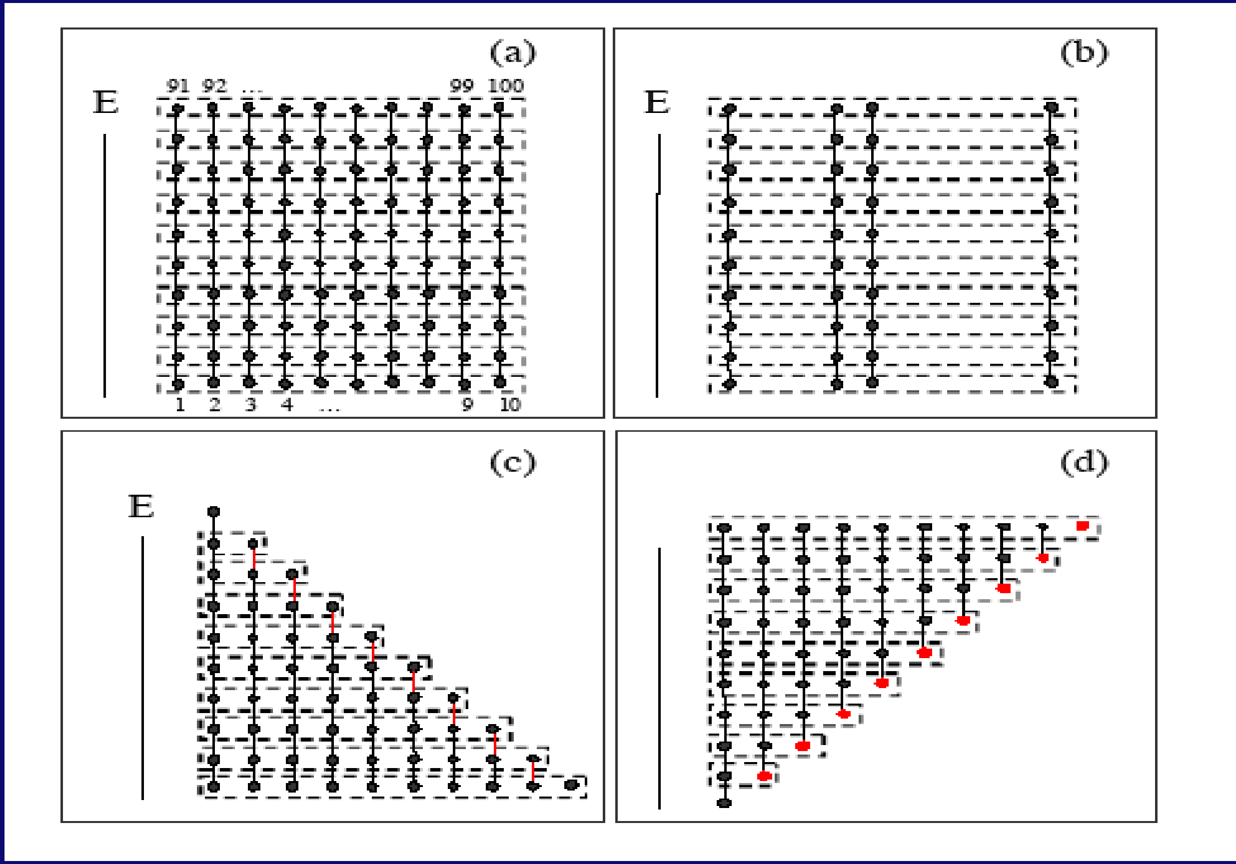
- **A little of the network representing stationary points:**

Fig. 1

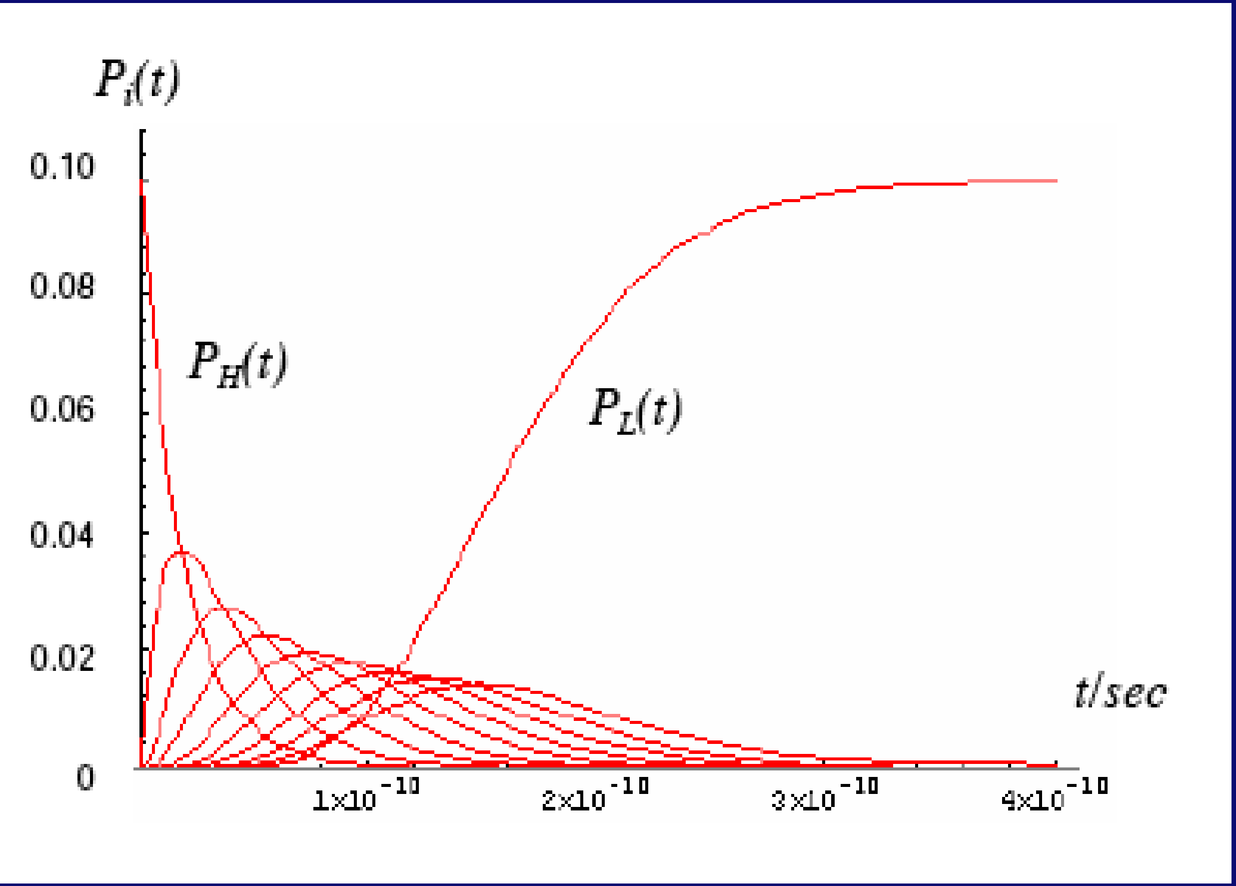


A simple example: 10 x 10

- The whole net and three sampling patterns, 10x10, 10x4, “10x10 tri,” and “10x10 inv”

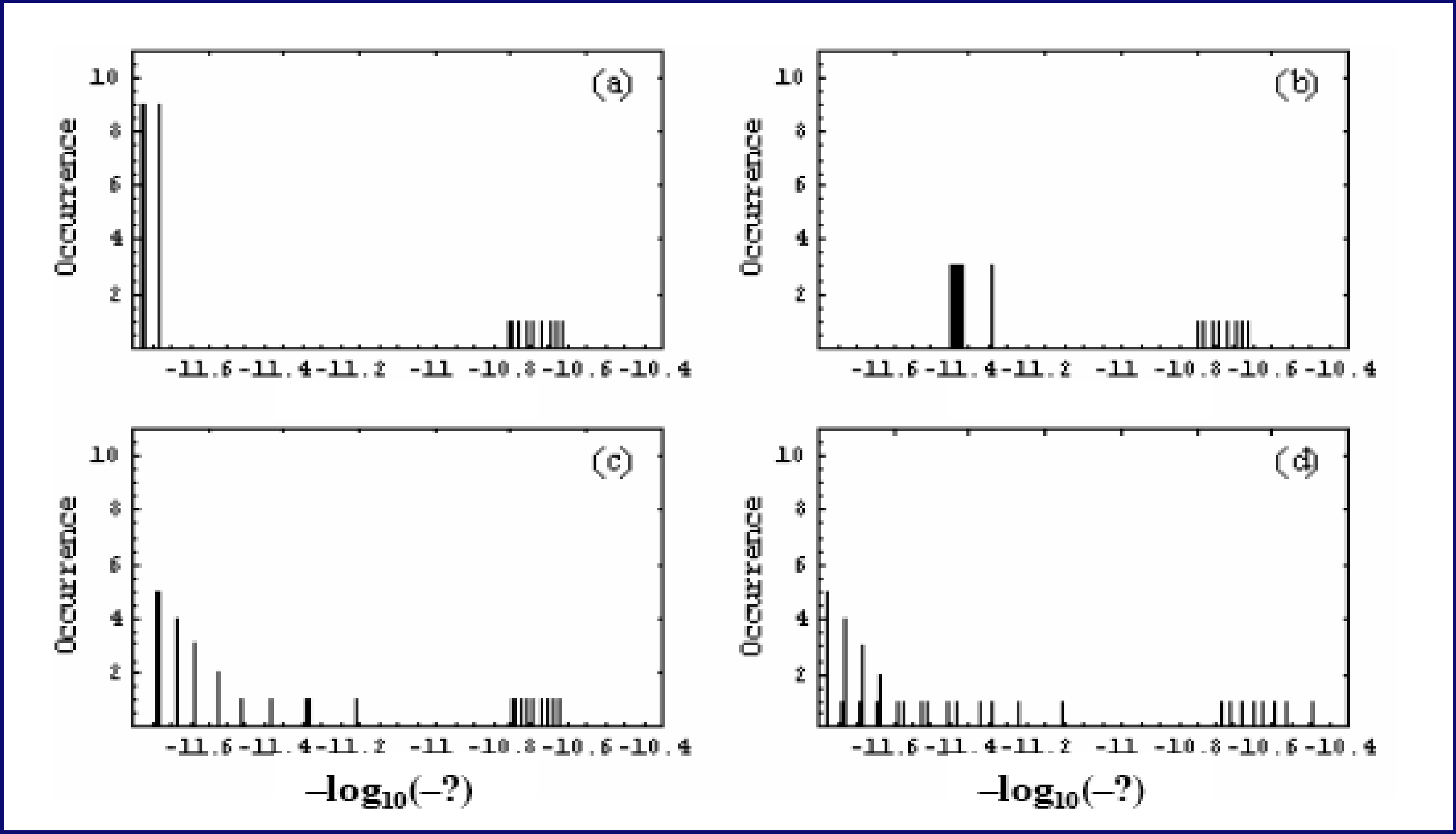


Watch relaxation from an initial distribution uniform in the top layer



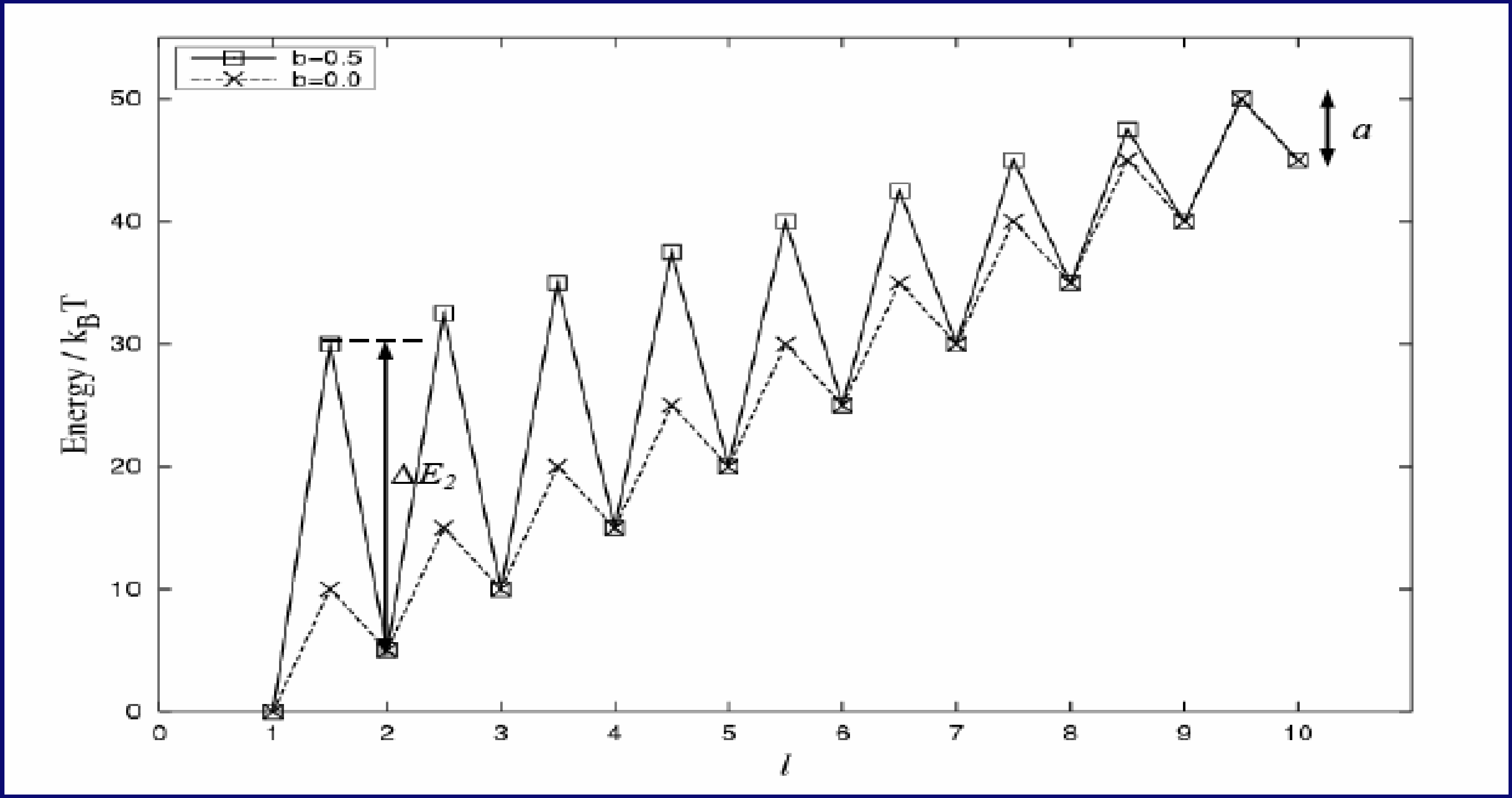
What are the eigenvalue spectra of these four?

- **a-full net; b-4x10; c-tri; d-inv**



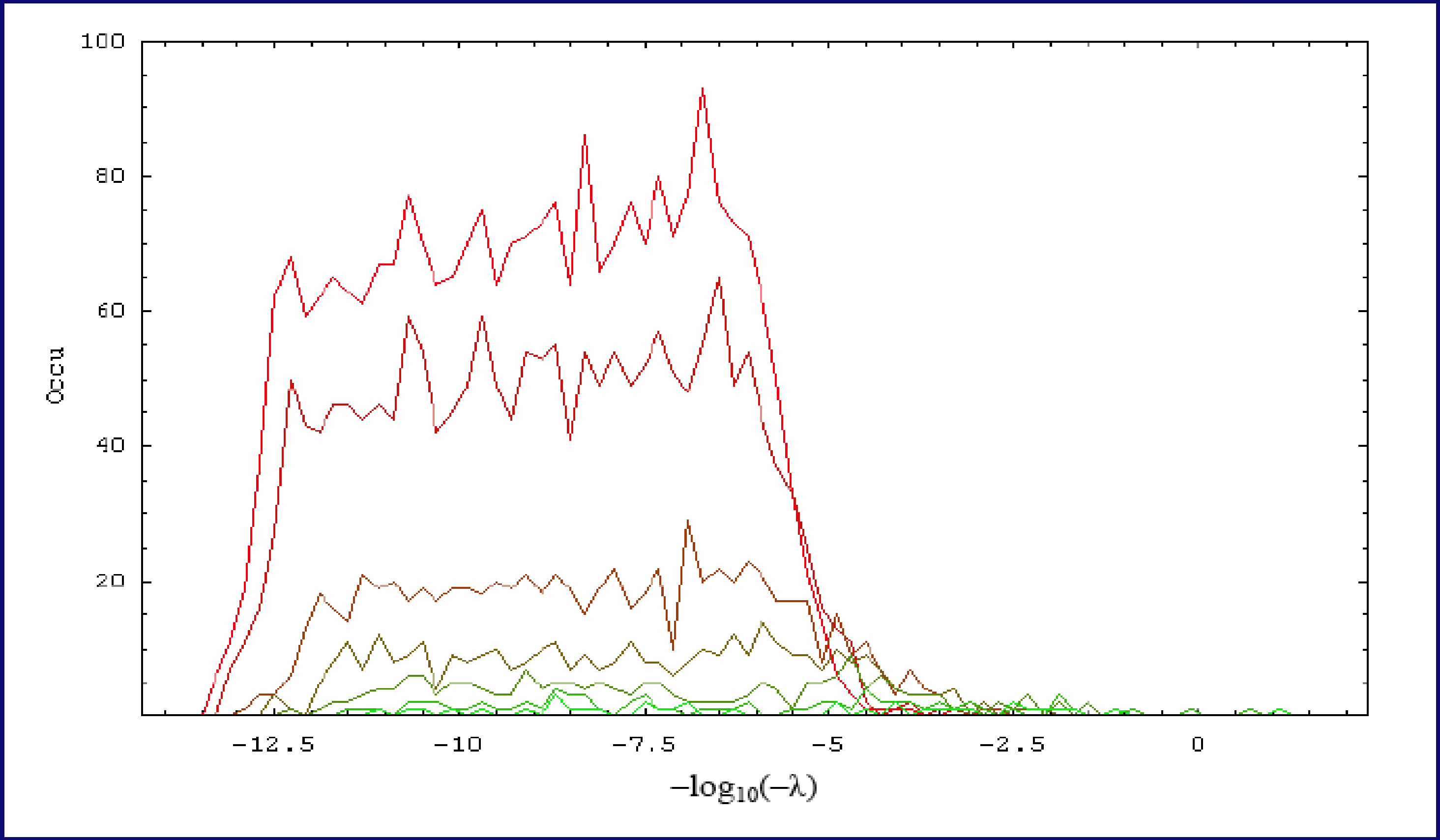
Now make the net more realistic

- **A "perturbed" monotonic sequence of minima**



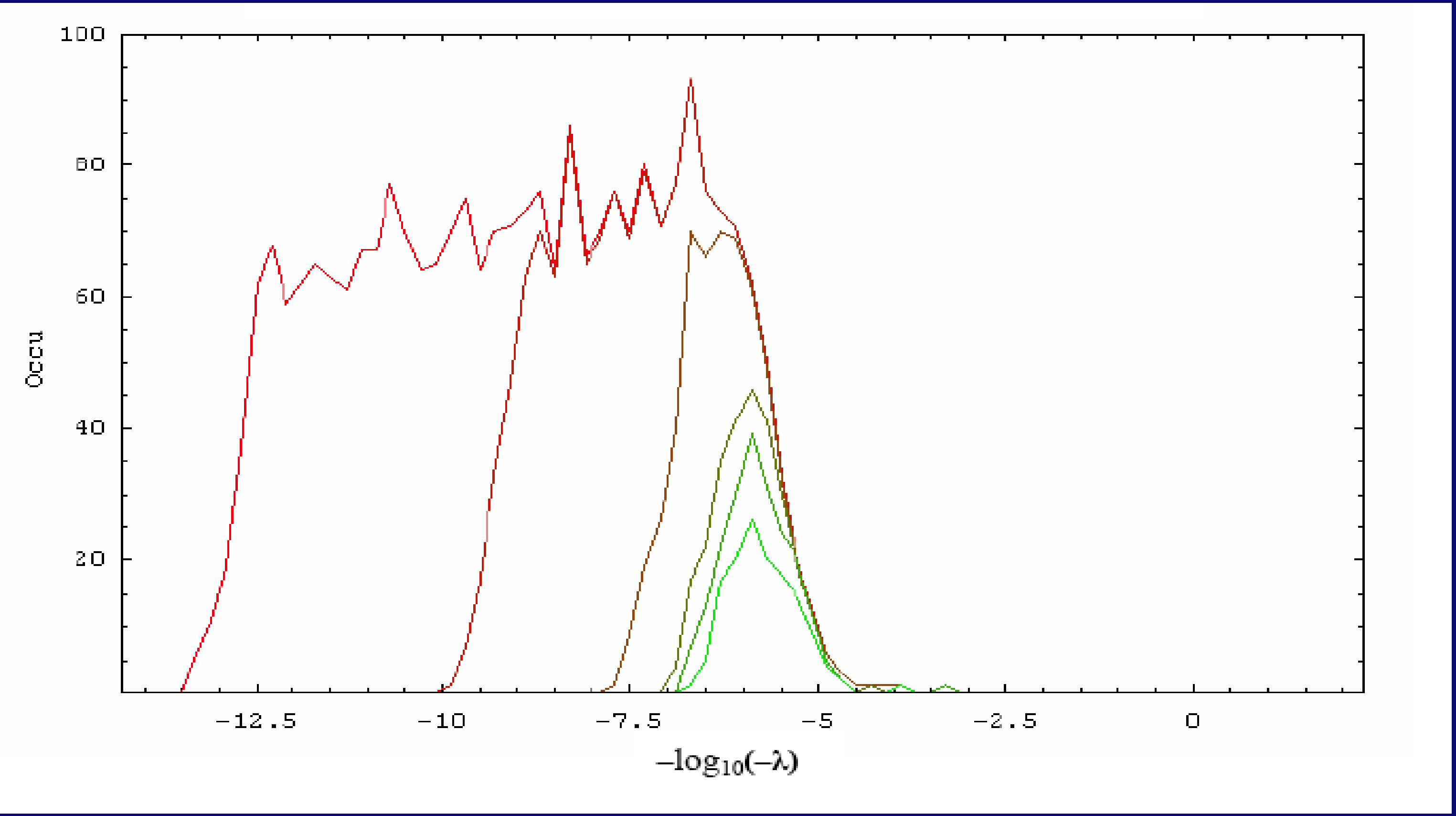
Now compare sampling methods for a bigger system

- **For a "big" (37x70) system, by *sequence*,**



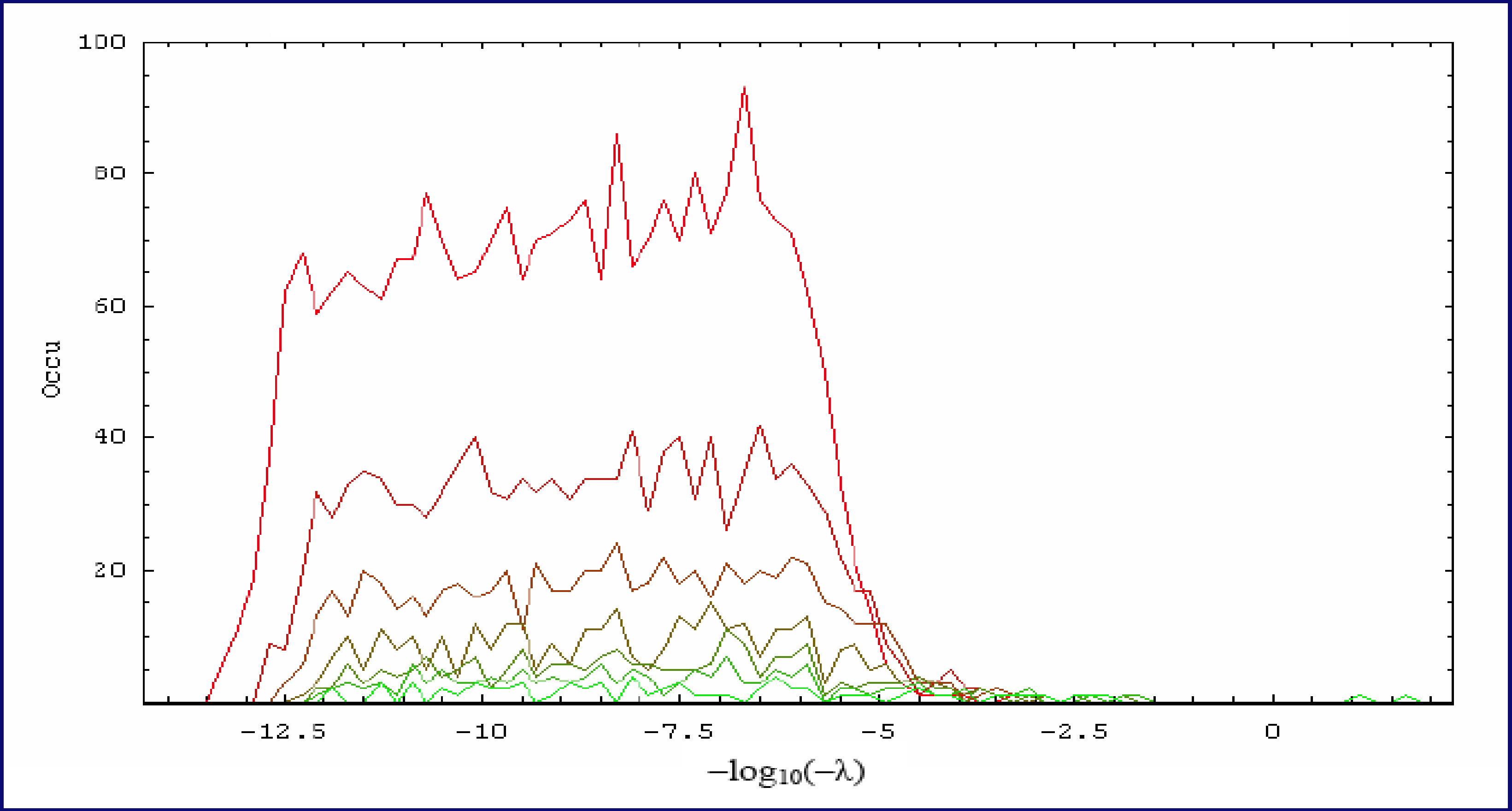
Next method; watch right end for slow eigenvalues

- By "rough topography" method



Third method: "low barrier"

- **Highest, with largest area, is full system**

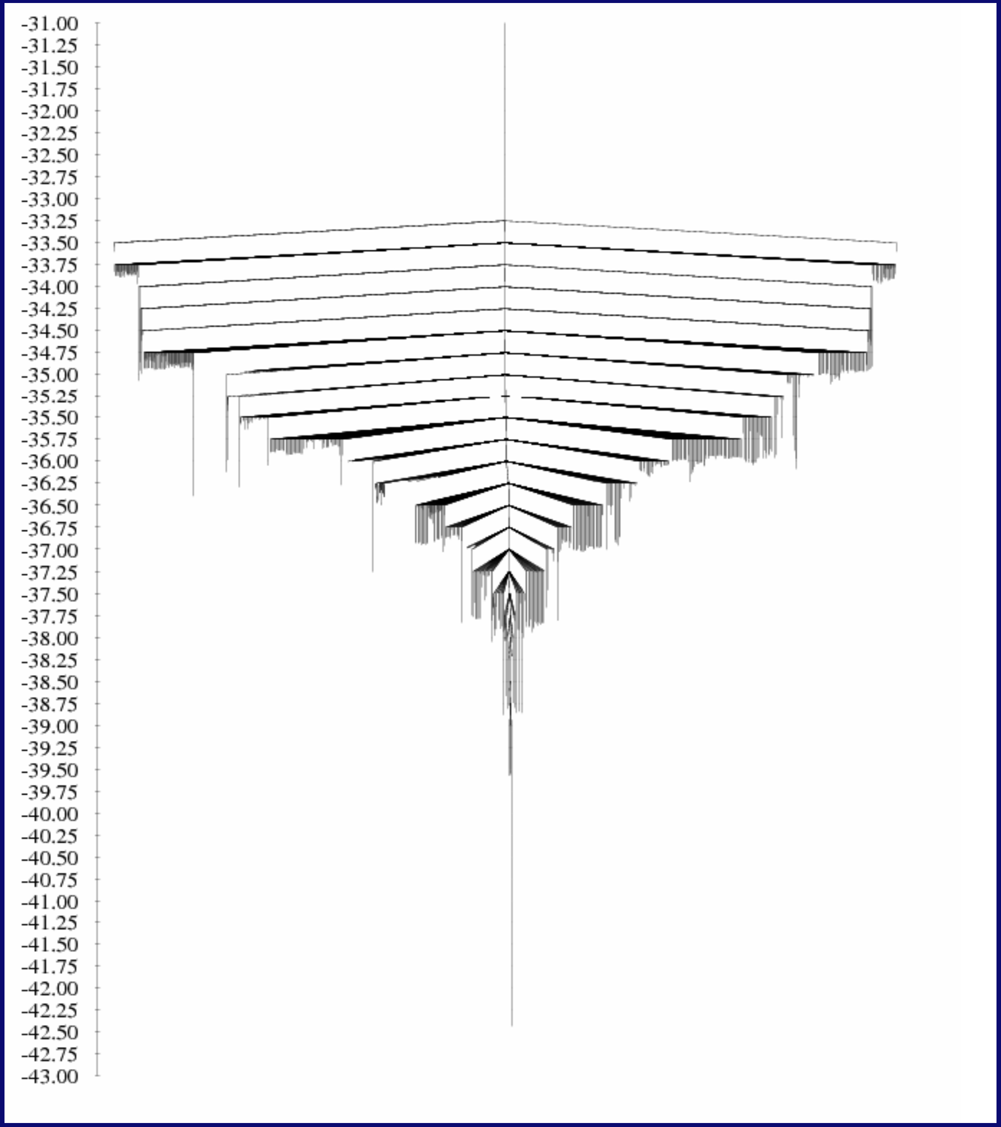


So we apparently have one pretty good sampling method-- 'rough topography'

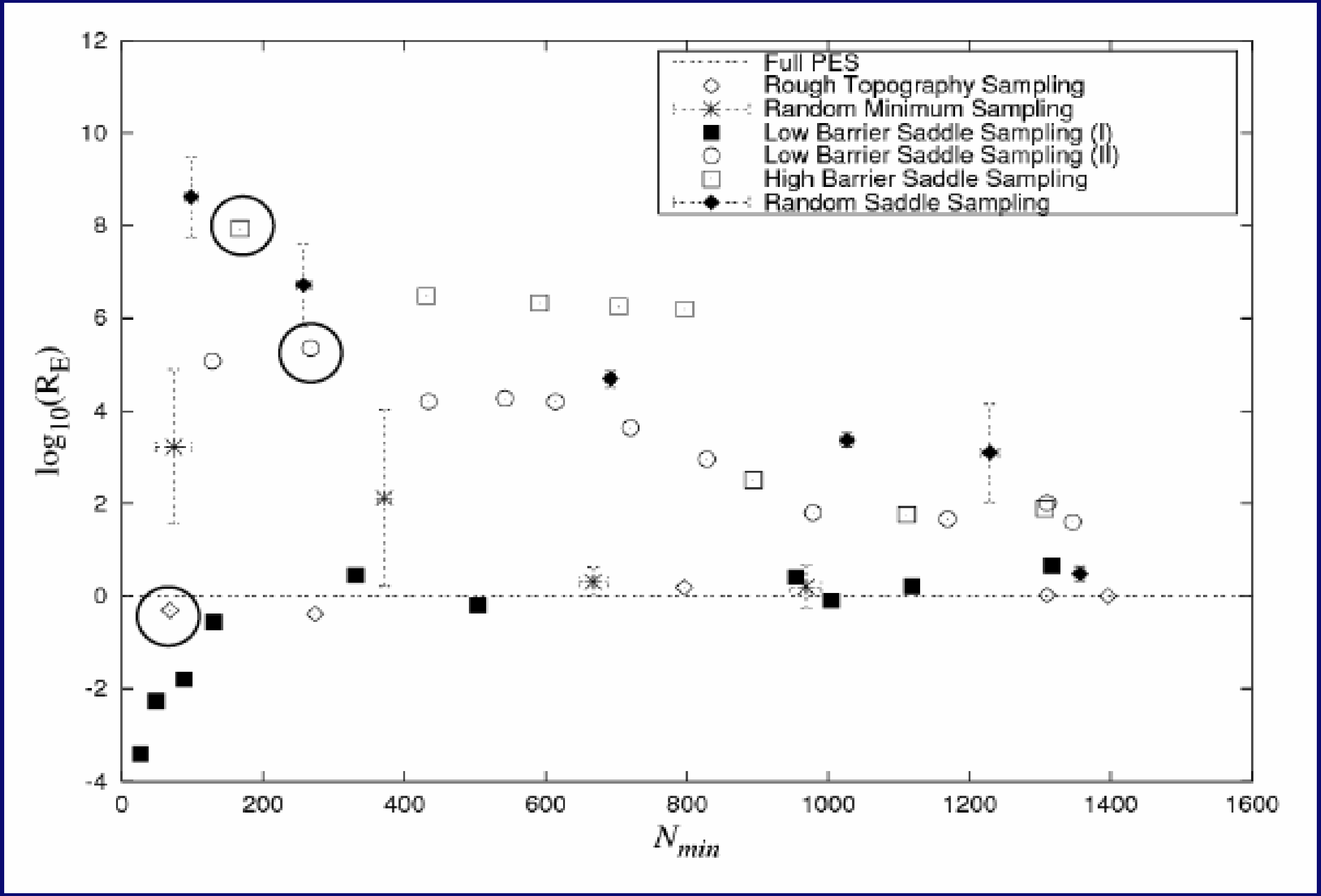
- **The others distort the eigenvalue spectrum at the end toward 0, that of the slow processes**

Try with a realistic system: Ar_{13}

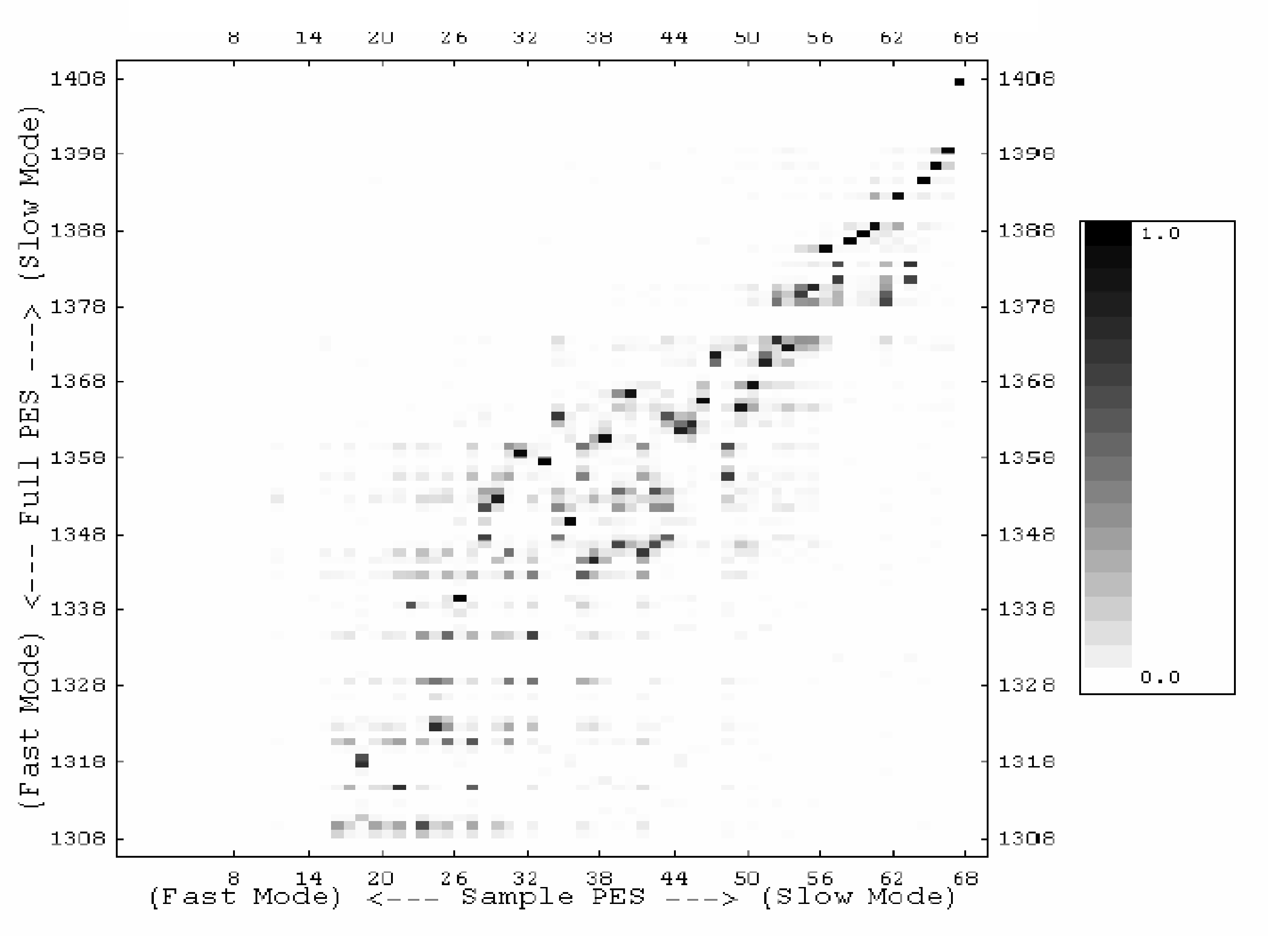
- **Its disconnection diagram:**



Results: relaxation times, as functions of sample size



Rough topography wins again; how about eigenvectors?



What is the minimum sample size to give reliable slow Λ's?

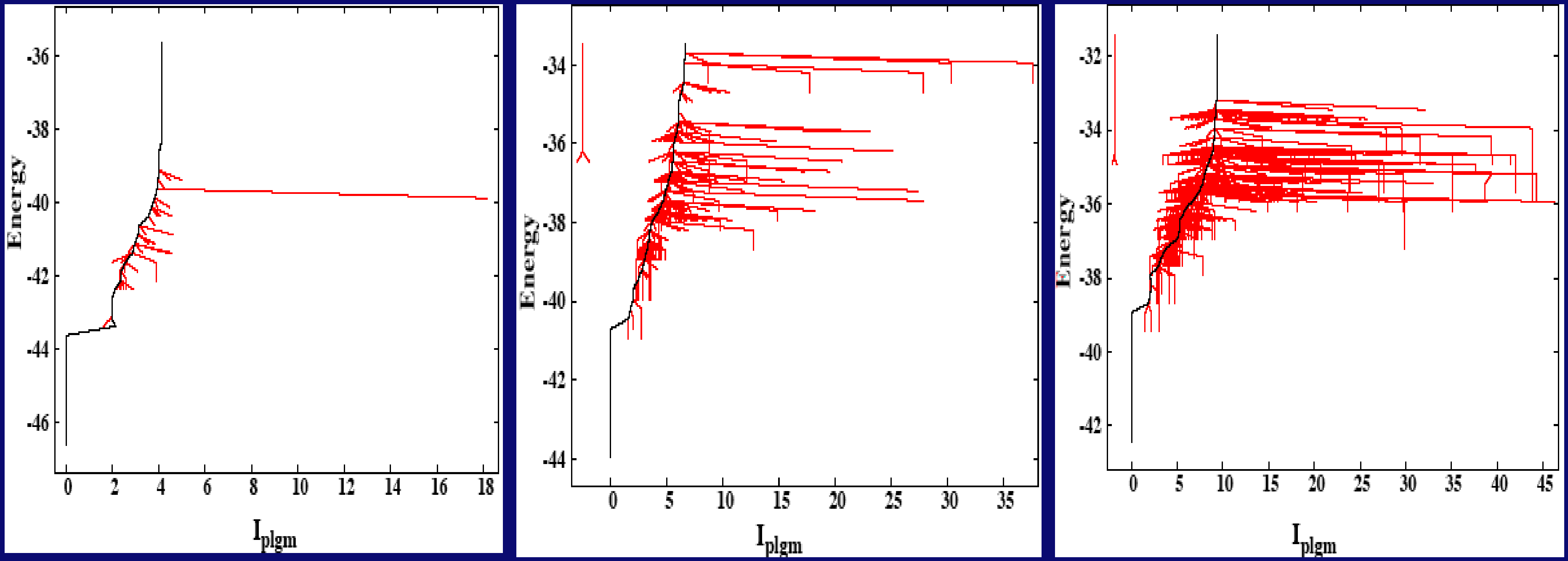
- **This is a big, unanswered question; can we use, say, only 0.1% of the minima on the landscape, even with the best sampling method?**
- **Application to *really* large, complex systems is yet to be done**

How does local topography guide a system to a structure?

- **How is the distribution of energies of local minima related to structure-seeking character?**
- **How is the distribution of barrier energies related to this character?**
- **How is the distribution of barrier *asymmetries* related?**

Shorter range means more deep basins, more complex surface: extended disconnection diagrams

- ρ = 4 ρ = 5 ρ = 6



Shorter range means more complex topography

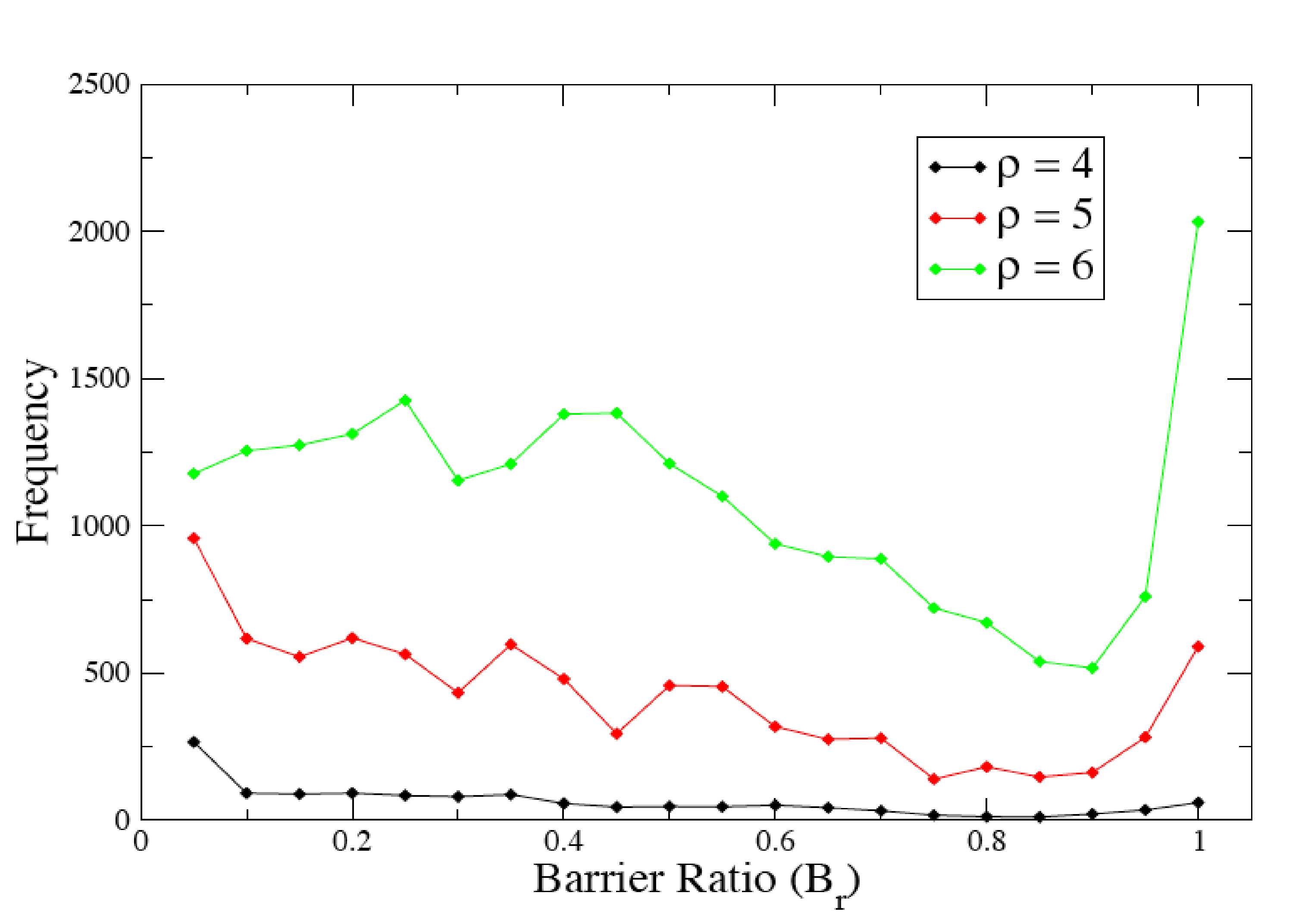
Table 2: Surface statistics, comparing transitions to the GM with all transition in the database (DB).

Statistic	$\rho = 4$	$\rho = 5$	$\rho = 6$
Average GM B_h	1.544	1.670	1.657
Number of GM transitions	636	1480	1824
Number of DB transitions	1267	8403	21844
Ratio GM/DB transitions	0.502	0.176	0.084
Number of GM-directly connected minima	143	459	658
Number of minima in DB	160	712	1410
Ratio GM/DB minima	0.894	0.645	0.467

Look at barrier asymmetry

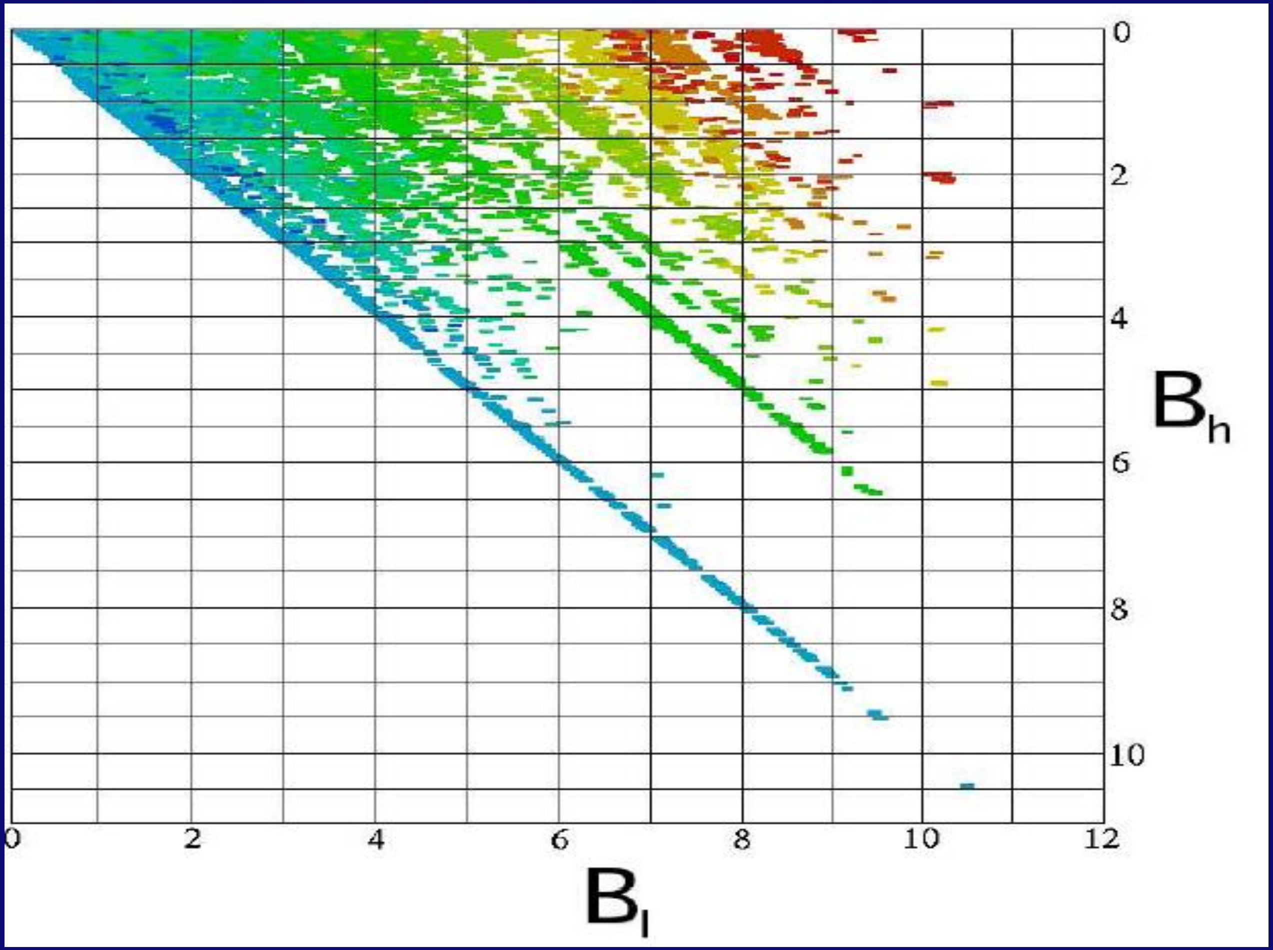
- Model system: 13-particle cluster with Morse interparticle interactions, with three values of the range parameter, $\rho = 4, 5, 6$; smaller ρ means longer range; all are structure-seekers, finding the icosahedral global minimum
- With $\rho = 4$, structure-seeking is strongest
- Examine distribution of asymmetries of saddles, $\Delta E_{highside}/\Delta E_{lowside} = B_r$ for all three parameter values

The barrier asymmetry distribution B_r



Minima distribute in bands, and barrier heights correlate

- Example, for = 6: (red=low E; blue=high E)



The illustration shows that

- Deep-lying saddles are the most asymmetric;
- The high levels have many interlinks; the lower the levels, the fewer links
- The system is small enough that the energy minima fall in distinguishable bands

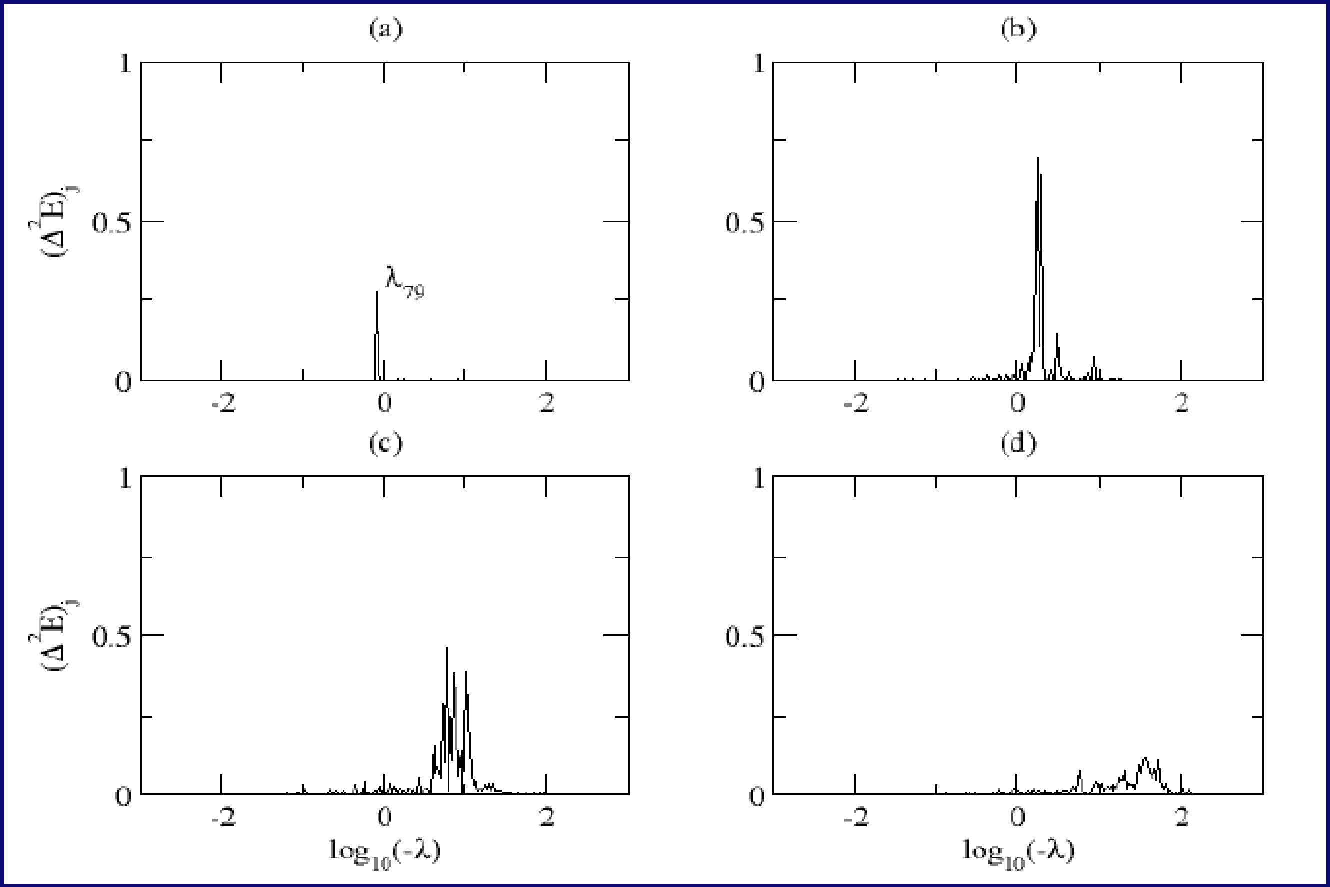
Can we quantify a scale, based on topography, that reveals the degree of structure-seeking or glass-forming character for any chosen system, whether cluster, protein, or…?

Autocorrelation based on the Master Equation

- **Determine the energy fluctuation $\Delta^2 E$, the autocorrelation function over time $\kappa(\tau)$, and its Fourier transform, the fluctuation spectrum $S(\omega)$**
- **Find the contribution of each eigenvector to these functions**
- **Apply the analysis to two atomic clusters, M_{13} and LJ_{38}**

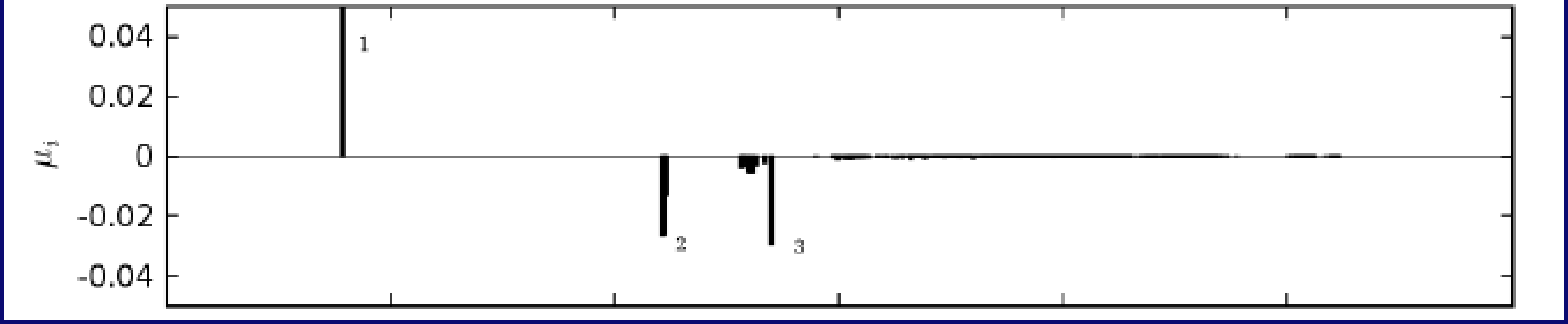
The fluctuations at equilibrium reveal which modes contribute

- At four increasing T's in the coexistence range, the modes that contribute; slow modes to the left



Which minima contribute to motion at lowest T'*s?*

- **At the low end of coexistence, relaxation is very slow, ~2 orders of magnitude below vibrational frequencies.**
- **The important structures are the icosahedral global min and "singly-excited" configurations**



A few more challenges

- **How can we make a topographical map of a landscape in 500 dimensions?**
- **How can we enumerate the *deep basins* on such a surface?**
- **How can we use topographical information to carry out quantum control?**

The people making it happen

- Jun Lu
- Chi Zhang
- and Graham Cox

Thank You!