Extremal Optimization for Low-Energy Excitations of very large Spin-Glasses

EMORY UNIVERSITY

Stefan Boettcher
Dept of Physics

- Collaborators:

Senior:
Allon Percus (Los Alamos)
Paolo Sibani (Odense-DK)
Michelangelo Grigni (Emory/CS)

Students:
Gabriel Charvat (Emory/Phy)
Susannah Cooke (Emory/Phy)
Martin Frank (Emory/Phy)

- Support: NSF-DMR

(Pre-)Prints: *www.physics.emory.edu/faculty/boettcher*

Summary:

- Applications of Extremal Optimization (EO)

Here:

- ➔ EO for Mean-Field Glasses (SK) (n≤1000)
- ➔ Spin Glass Ground States with $\sim 10^6$ Spins
- ➔ EO for Defect Energies of Glasses in $d=3,..,7$
- ➔ Comparison with Mean-Field ($d \to \infty$) limit

Result: **Mean-Field Prediction fails**

Poster:

- ➔ EO vs Metropolis: Energy-Landscape Explorations
- ➔ EO of MAX-3-Coloring at Phase Transition
- ➔ EO for the Thomson Problem
- ➔ Theory of EO

Hard Optimization Problems:

Consider:

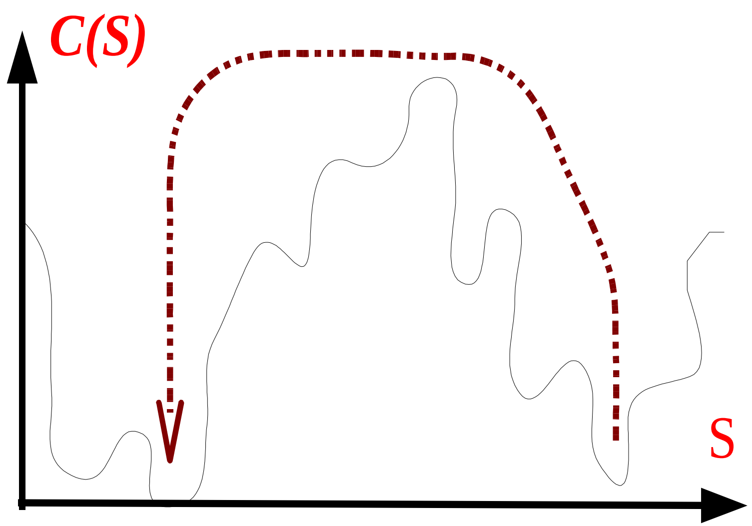
1) System of n variables, $x_i \in I$.
2) Configurations $S=(x_1,\ldots,x_n) \in I^n$.
3) "Cost Function" $C(S)$.
4) Local Search "Neighborhood," $S':=N(S)$.

Problem:

Find <u>*Global Minimum*</u> of $C(S)$ when

1) n is large,
2) $C(S)$ has many local Extrema in N.

=>Search-time $t >> n^k$ for any k .



Need: Seach "Heuristic" to find approximate solution in $t \sim n^k$.

Spin Glass Ground States:

Ferro-magnetic Bonds:

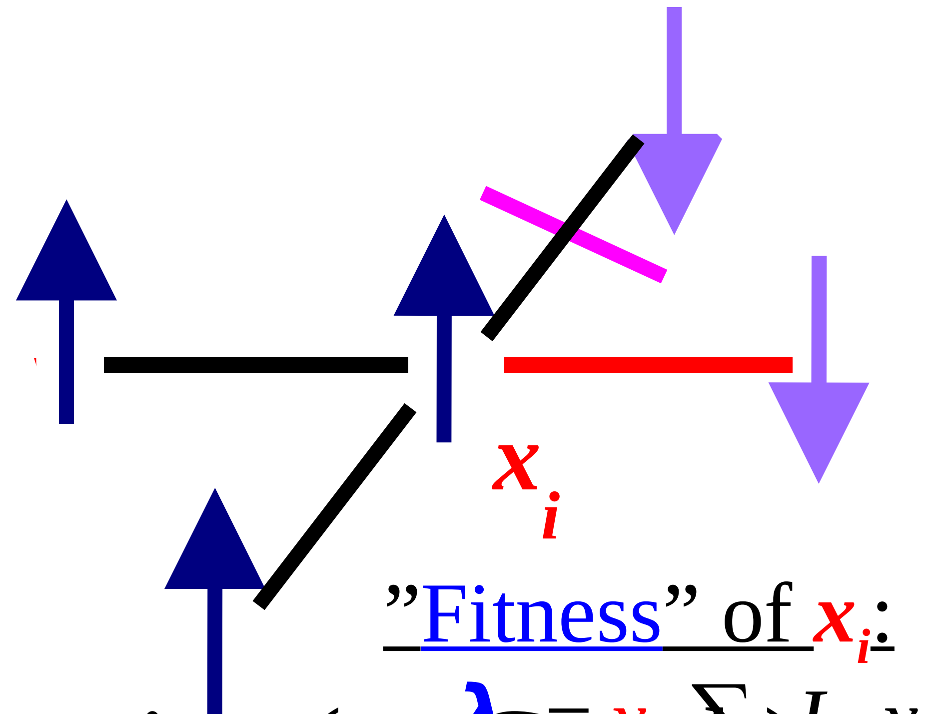
$J = +1$

Anti-ferro-mag Bonds:

$J = -1$

Spin Variables:

$x_i = +1$ $x_i = -1$



"Fitness" of x_i:

$\lambda_i = x_i \sum_j J_{i,j}\, x_j$

Hamiltonian (Cost):

$H = - \sum\sum_{<i,j>} J_{i,j}\, x_i\, x_j$

Lattice (or Graph)

w/ fixed, random Bonds

"FRUSTRATION"

$H = - \sum_i \lambda_i$

Extremal Optimization (EO)

- Motivated by Self-Organized Criticality
  - ➔ Emergent Structure
    - *without* tuning any Control Parameters
    - despite (or because of) Large Fluctuations

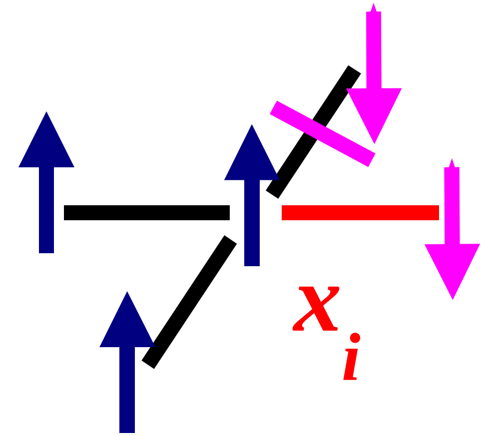
- How can we use it to optimize?
  - ➔ Extremal Driving:
    - Select and eliminate the “bad”,
    - Replace it *at random*,
    - Eventually, only the “good” is left!

Evolutionary Search Heuristic

"Fitness" λ for various Problems:

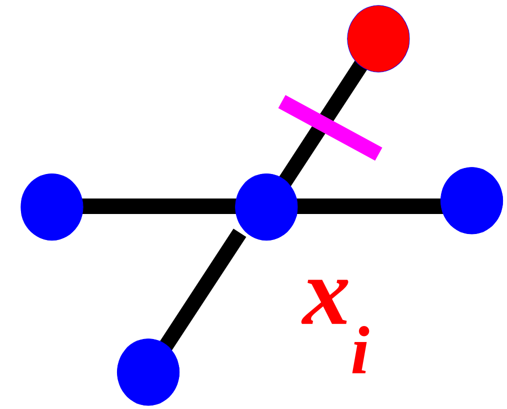
- Spin Glasses (eg. MAX-CUT):

$$\lambda_i = x_i \sum_j J_{i,j}\, x_j$$



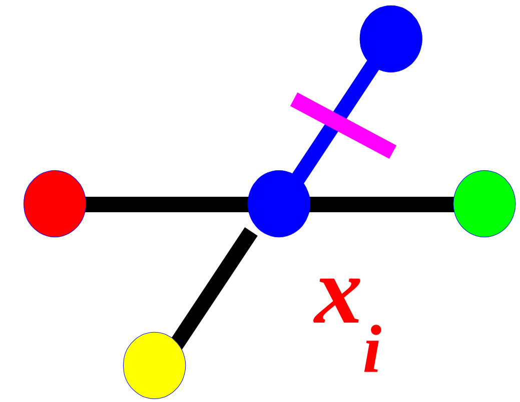
- Partitioning (eg. MIN-CUT):

$$\lambda_i = -\ (\#\text{-cut edges of } x_i)$$



- Coloring (eg. Potts Anti-ferro):

$$\lambda_i = -\ (\#\text{-monochrome edges of } x_i)$$



$$Cost \propto H = -\sum_i \lambda_i$$

>> "Extremal Optimization" (EO): <<

(1) Provide initial Configuration $S=(x_1,\ldots,x_n)$,

(2) Determine "Fitness" λ_i for each Variable x_i ,

(3) Rank all $i=\Pi(k)$ according to

$$\lambda_{\Pi(1)} \le \lambda_{\Pi(2)} \le \ldots \le \lambda_{\Pi(n)}$$

(4) Select x_w w/ $w=\Pi(1)$, i.e. x_w has worst Fitness!

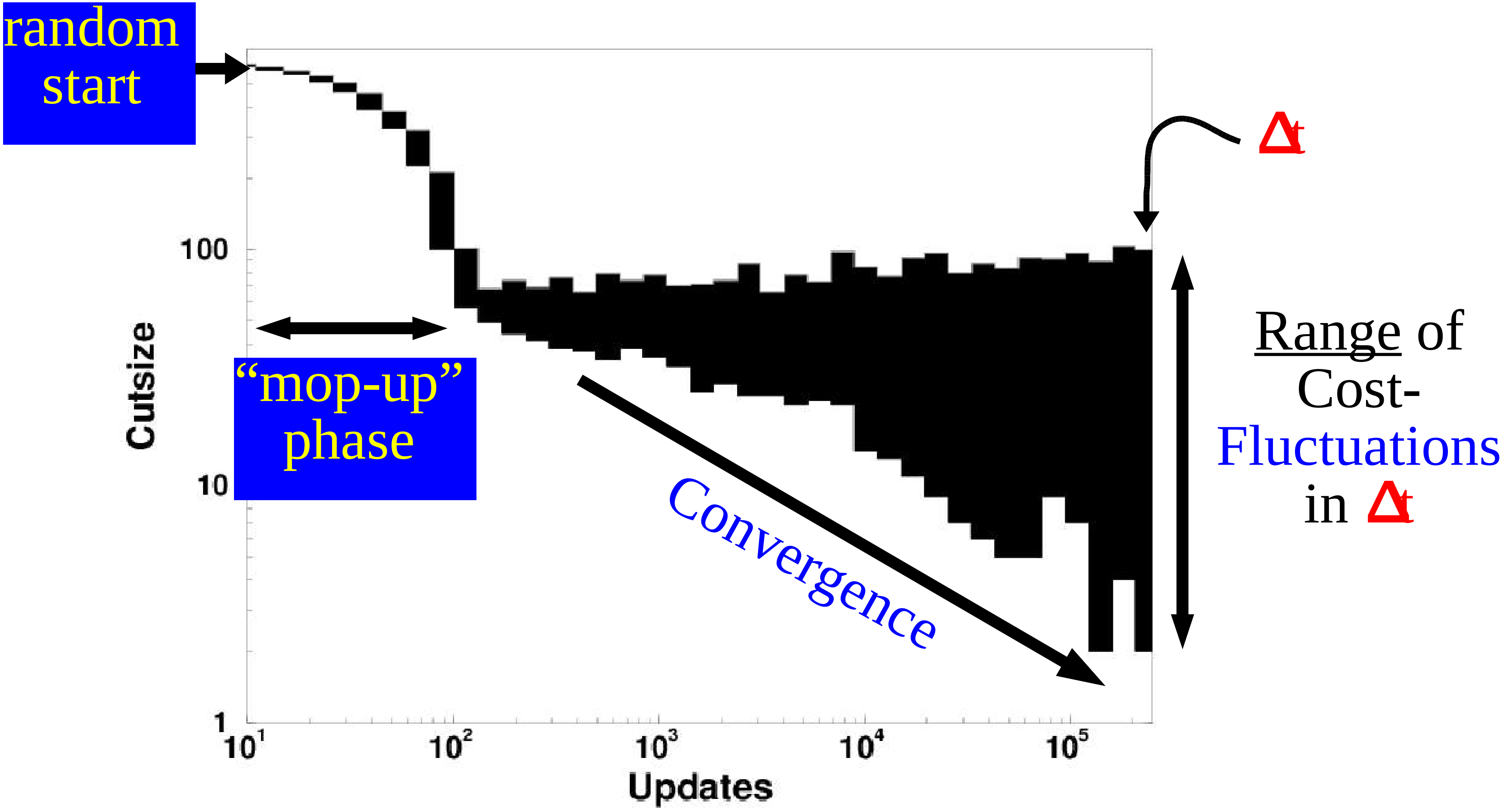
(5) Update x_w unconditionally,

(6) For t_{max} times, Repeat at (2),

(7) Return: Best $Cost(S)$ found along the way!

Typical Extremal Optimization Run:

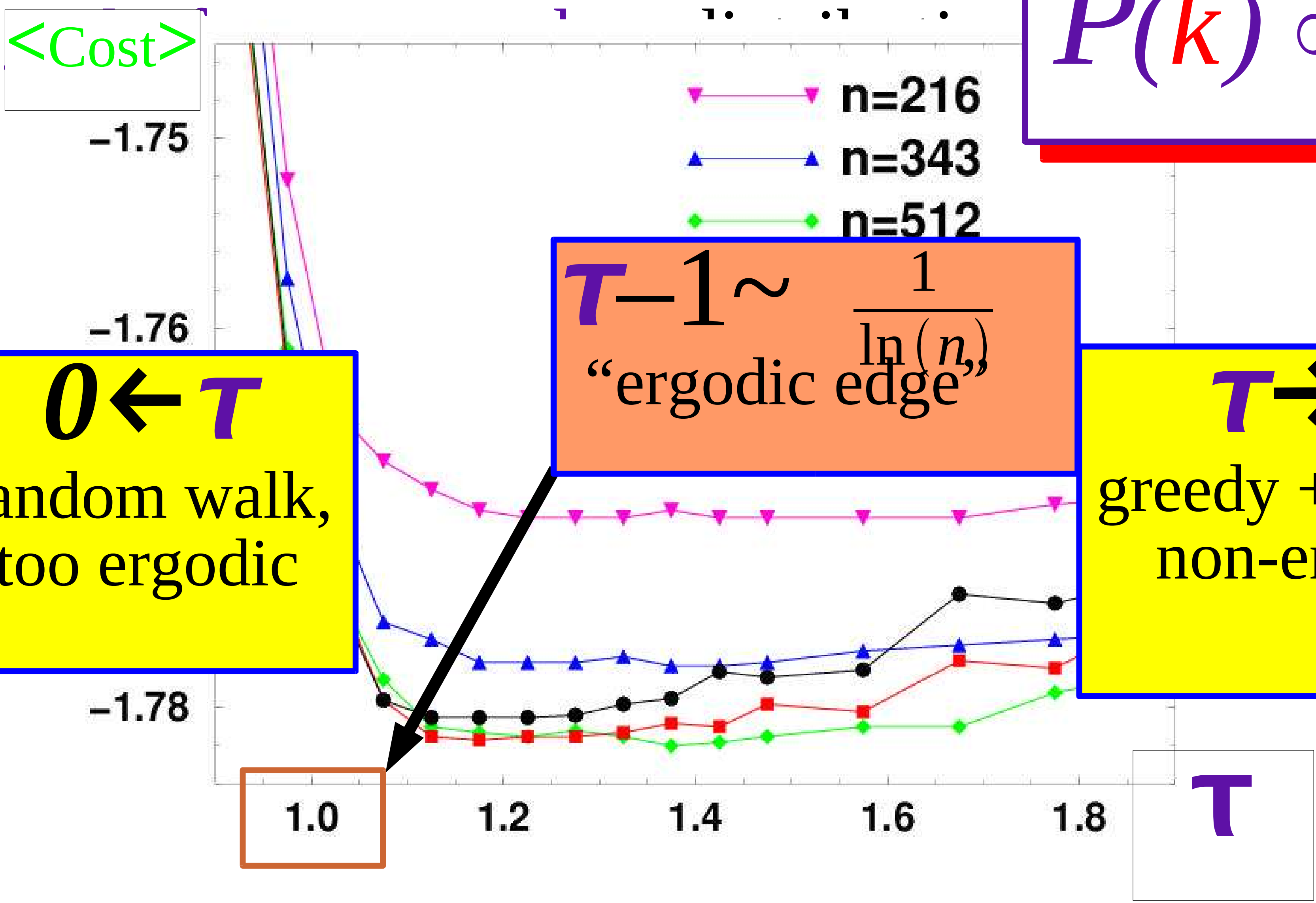
EO-run for Partitioning ($n=500$):



τ-EO - Searching at the "Ergodic Edge":

For Ranks $\lambda_{\Pi(1)} \leq \ldots \leq \lambda_{\Pi(n)}$, update $i=\Pi(k)$ with

$$P(k) \propto k^{-\tau}$$



$$\tau - 1 \sim \frac{1}{\ln(n)}$$

"ergodic edge"

$0 \leftarrow \tau$

random walk, too ergodic

$\tau \rightarrow \infty$

greedy + frozen, non-ergodic

Results for τ-EO:

- Applications by Others:

EO for Image Alignment (Batouche et al, [LNCS2449('02)330])

Aligned Images

Results for τ-EO:

- For Spin Glasses:

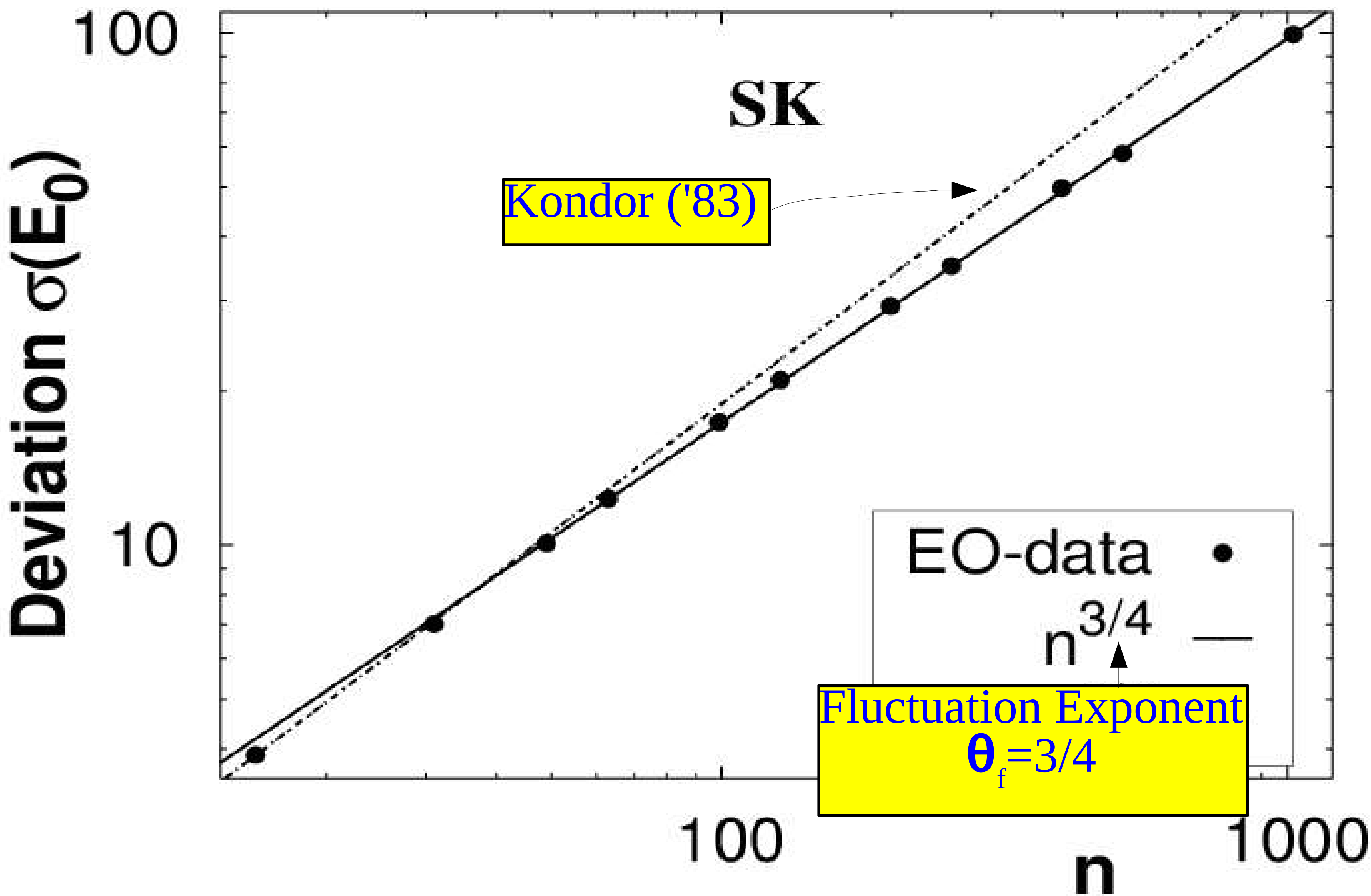
EO for *3d*-Lattice Spin Glasses [PRL**86**('01)5211]

L	t	E/L^3	Pal96	Hartmann97
3	0.0006	-1.6712(6)	-1.67171(9)	-1.6731(19)
4	0.0071	-1.7377(3)	-1.73749(8)	-1.7370(9)
5	0.0653	-1.7609(2)	-1.76090(12)	-1.7603(8)
6	0.524	-1.7712(2)	-1.77130(12)	-1.7723(7)
7	3.87	-1.7764(3)	-1.77706(17)	
8	22.1	-1.7796(5)	-1.77991(22)	-1.7802(5)
9	100	-1.7822(5).		
10	424.	-1.7832(5)	-1.78339(27)	-1.7840(4)
12	9720.	-1.7857(16)	-1.78407(121)	-1.7851(4)
∞	$O(n^4)$	-1.7865(3)	-1.7863(4)	-1.7868(3)

Genetic Algorithms by Pal [PhysicaA**223**('96)283]
and by Hartmann [EPL**40**('97)492]

τ-EO for Sherrington-Kirkpatrick (SK):

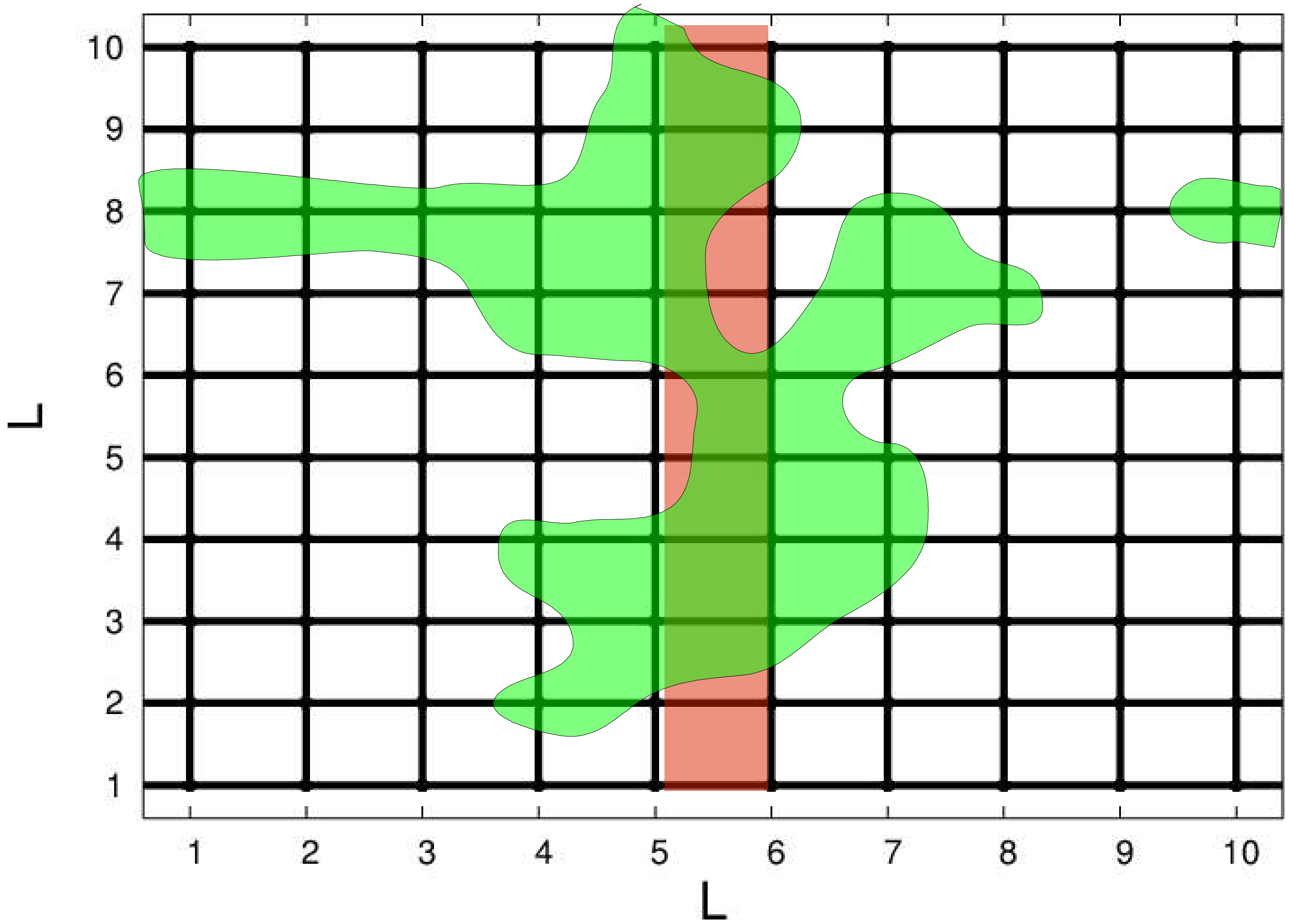
- Mean-Field ($d \rightarrow \infty$) Spin Glasses:



Lattice Spin Glasses (at $T=0$):

Defect-Energy: •Reverse Bonds (Perturb) on scale L

•Measure Energy Fluctuations $\Delta E(L)$

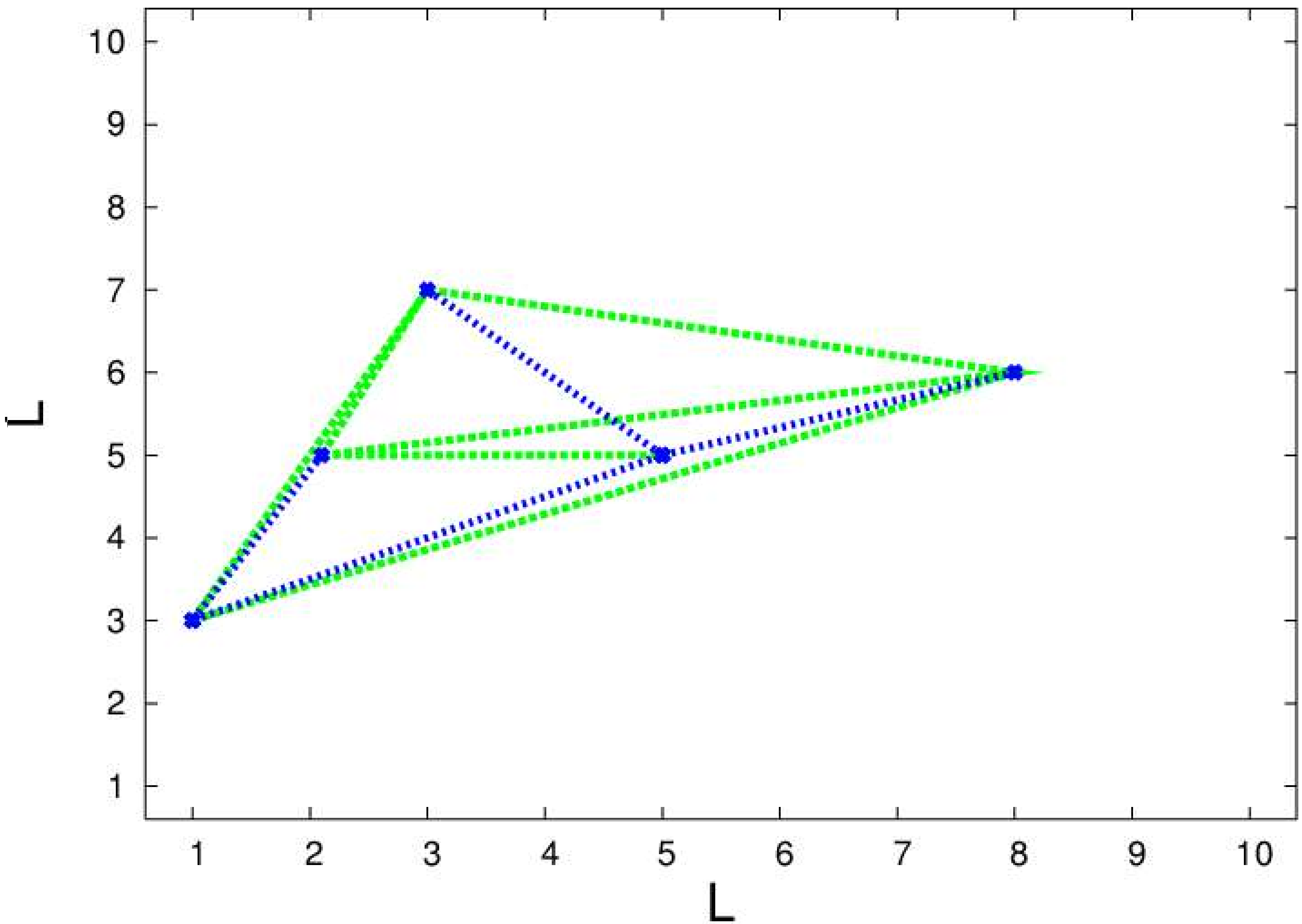


⇨ Low-Energy Excitations (like "small Oscillations")

Lattice Spin Glasses (at $T=0$):

Defect-Energy:

Measure "Stiffness": $\sigma(\Delta E) \sim L^y$



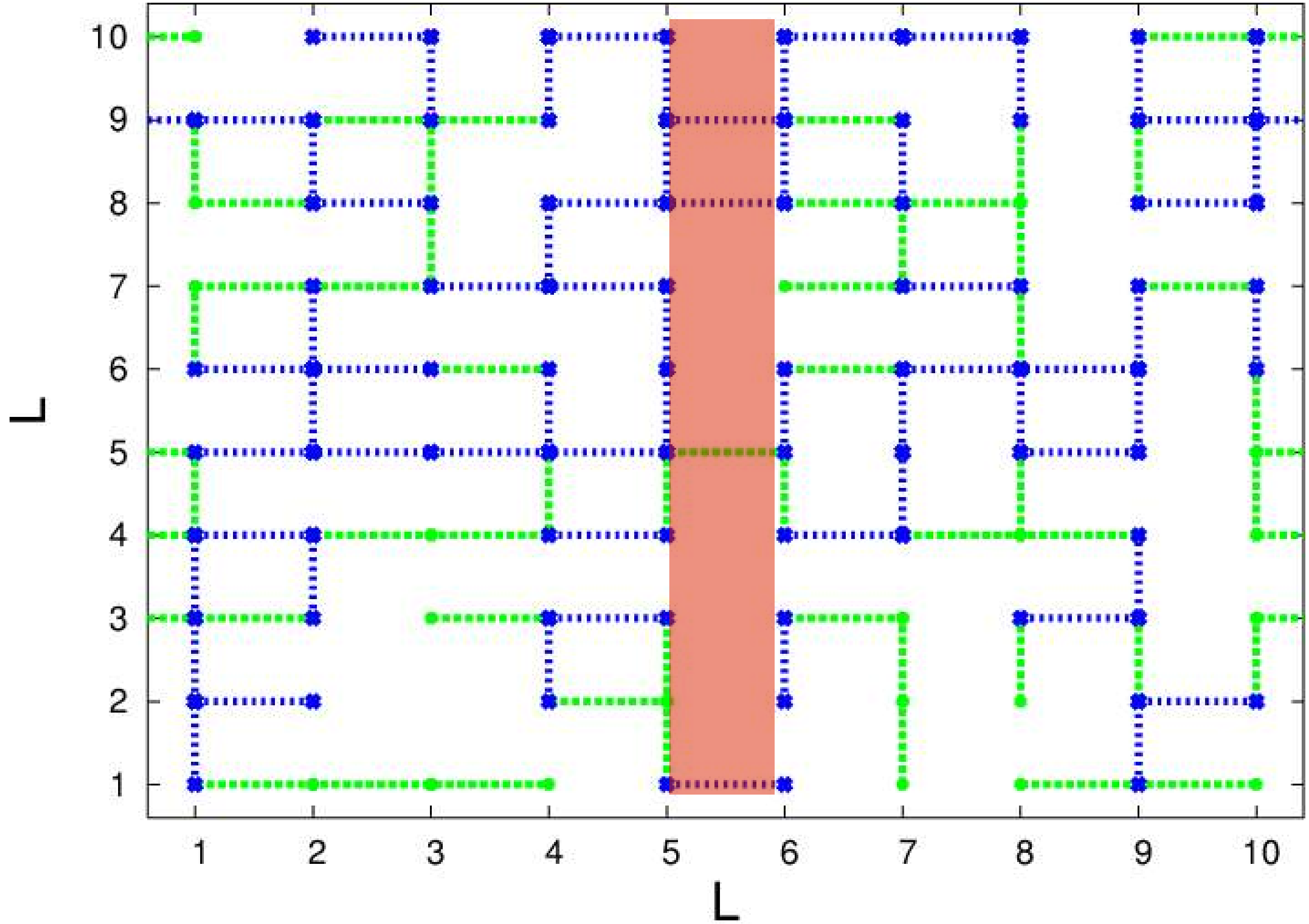
Before:
100 spins

After:
5 spins

⇨ Low-Energy Excitations on ***bond-diluted*** Lattices

Lattice Spin Glasses (at *T=0*):

Defect-Energy: Measure "Stiffness": $\sigma(\Delta E) \sim L^{y}$

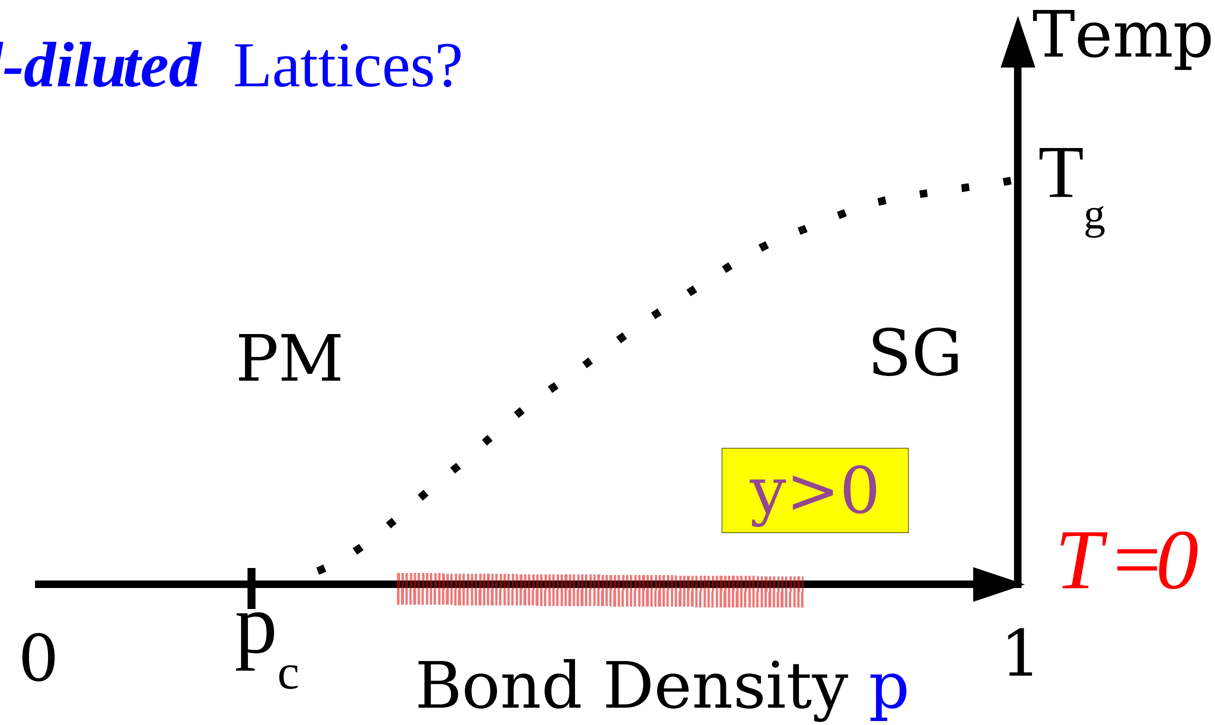


⇨ Low-Energy Excitations on ***bond-diluted*** Lattices

Lattice Spin Glasses (at $T=0$):

Defect-Energy: Measure "Stiffness": $\sigma(\Delta E) \sim L^y$

How ***bond-diluted*** Lattices?



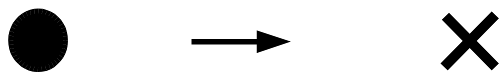
Why ***bond-diluted*** Lattices?

- ➔Simpler Problem
- ➔Larger Sizes L
- ➔Better Scaling

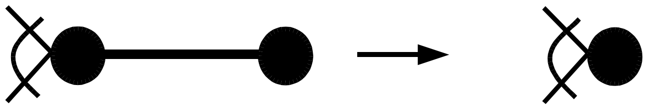
Reduction Method for sparse Graphs:

"Reduce" low-connected Spins, optimize the Remainder ($T=0$ only!):

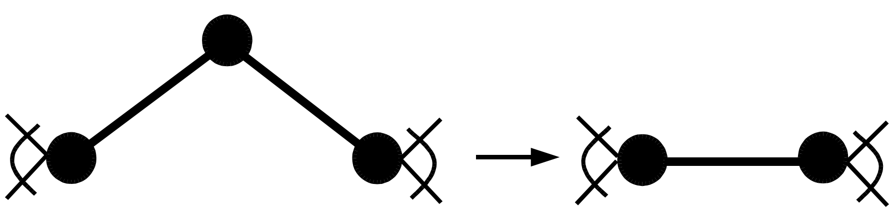
0-con:



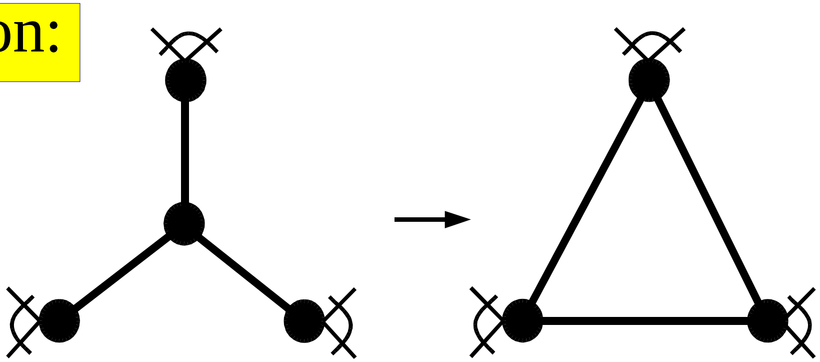
1-con:



2-con:

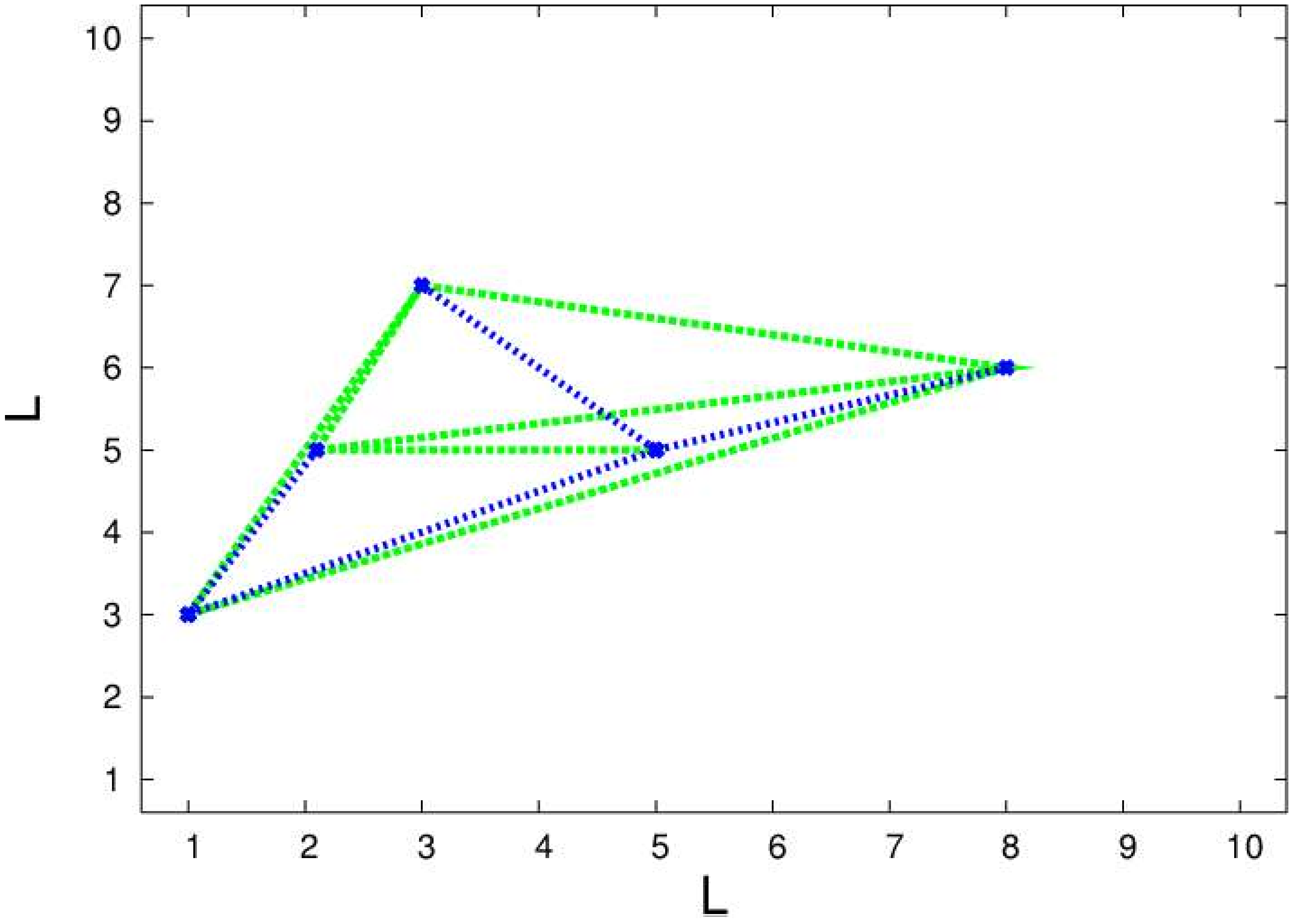


3-con:



Lattice Spin Glasses (at *T=0*):

Applying the Reduction Rules:

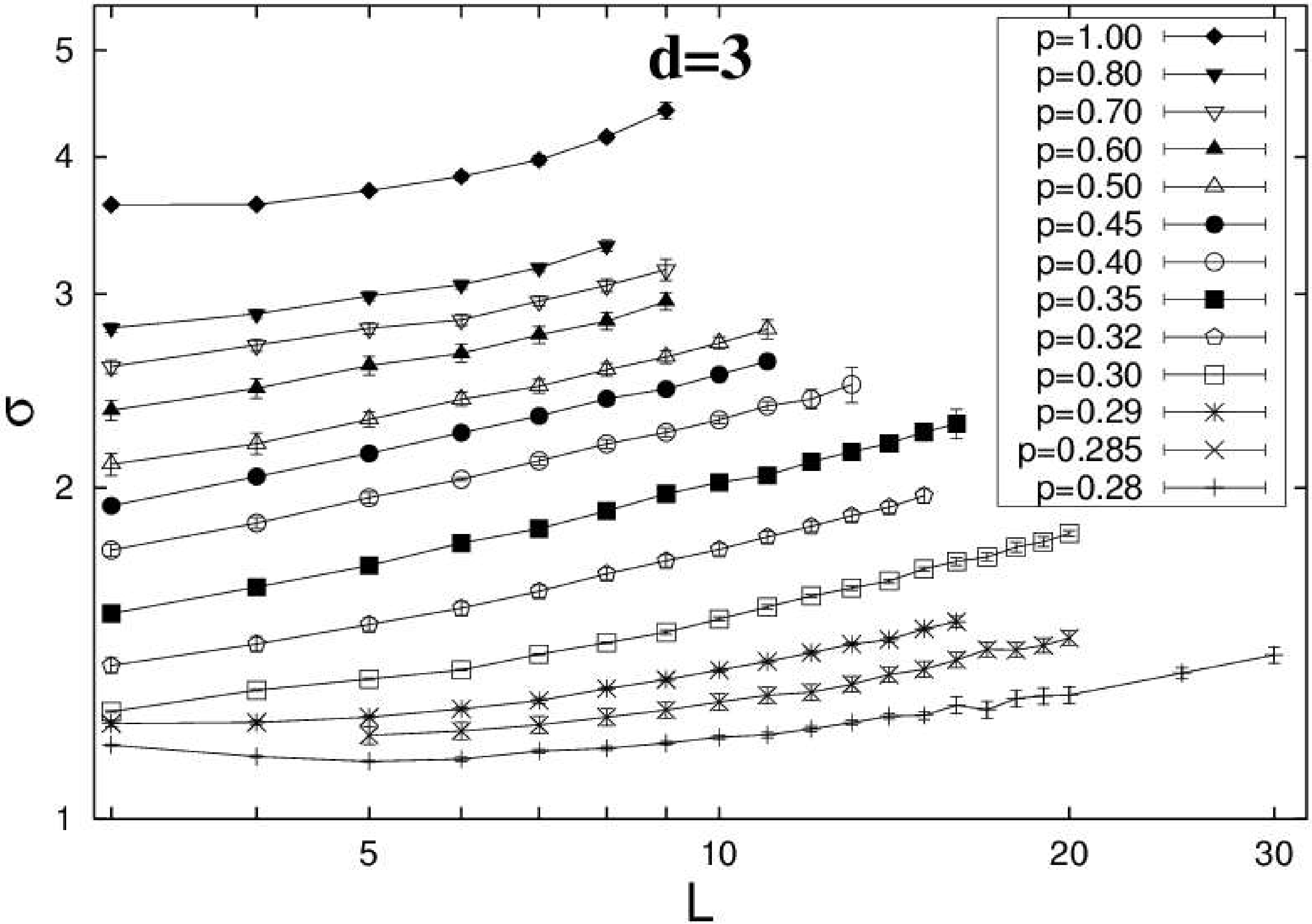


Before:
100 spins

After:
5 spins

Defect-Energy of diluted Lattices:

"Stiffness": $\sigma(\Delta E) \sim L^y$

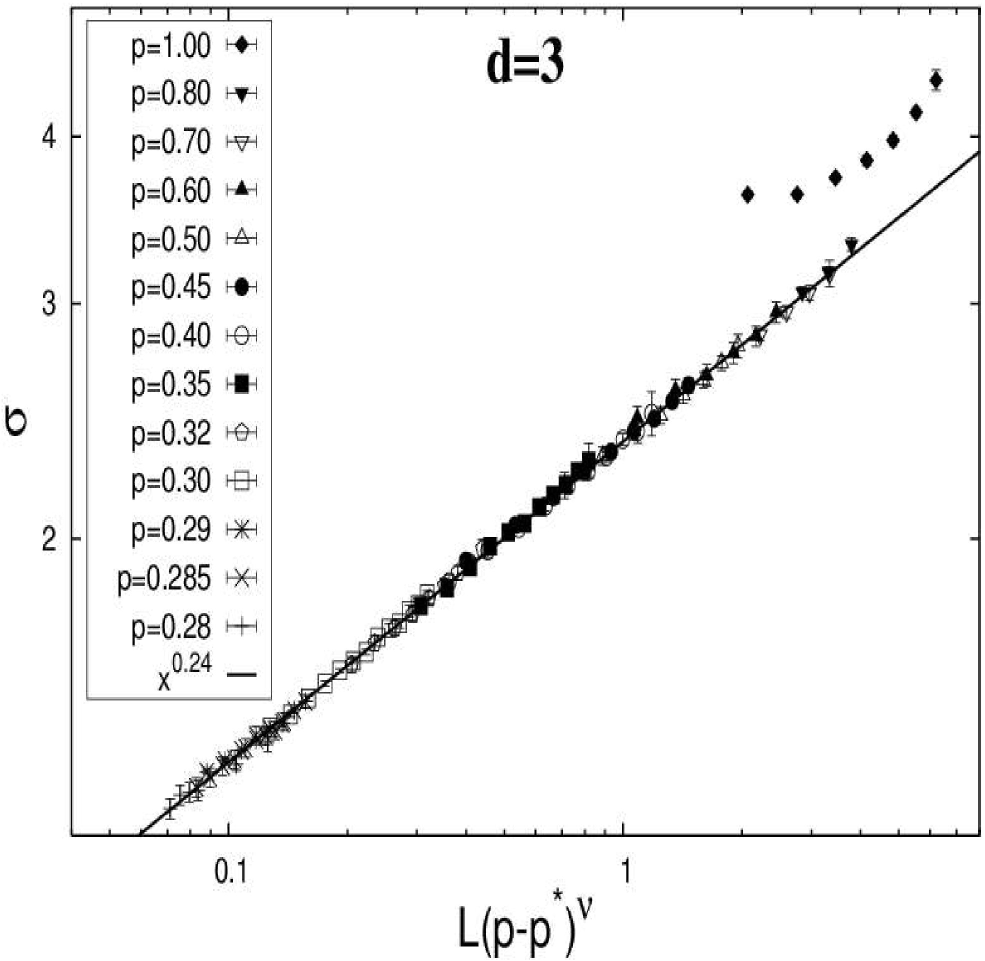


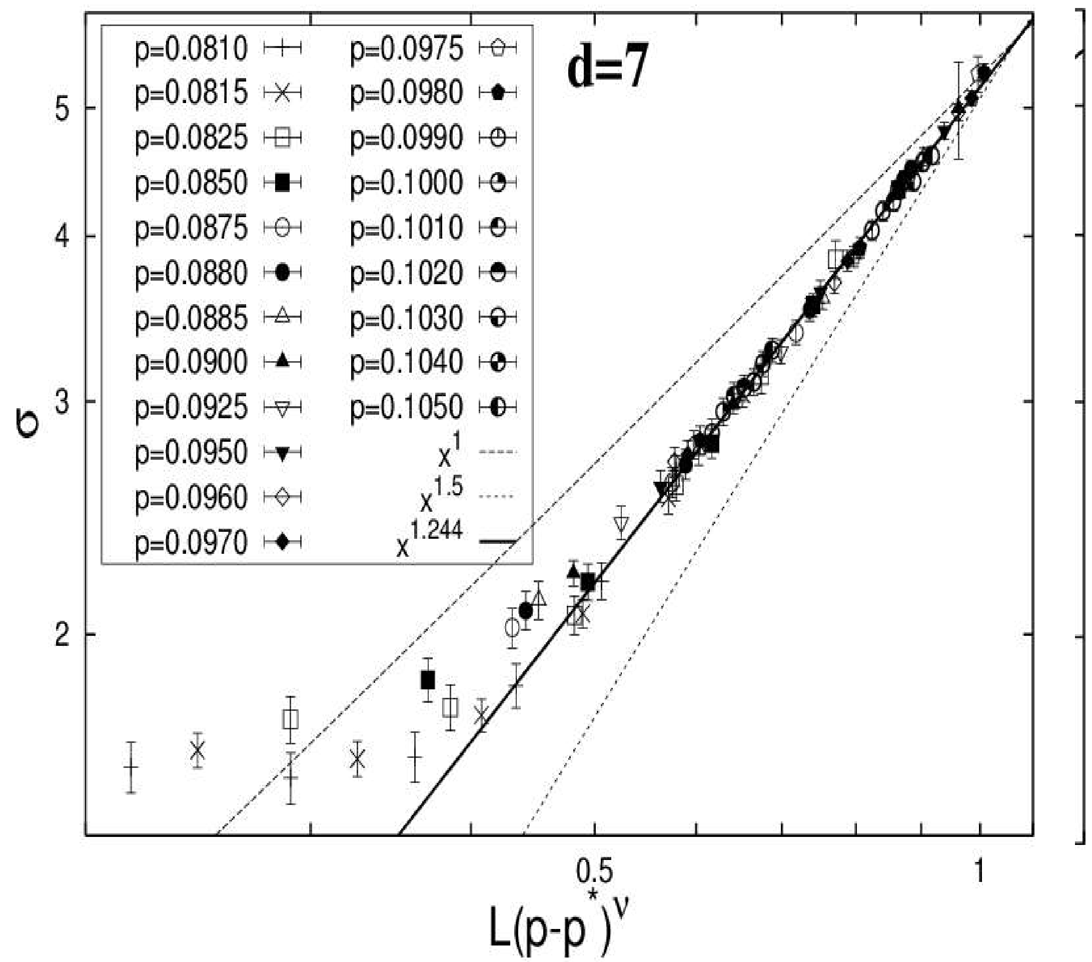
***Stiffness* Exponent *y* for Lattice Glasses:**

- Reduction plus **τ**-EO for dilute-Lattice ±J Glasses:

d=3

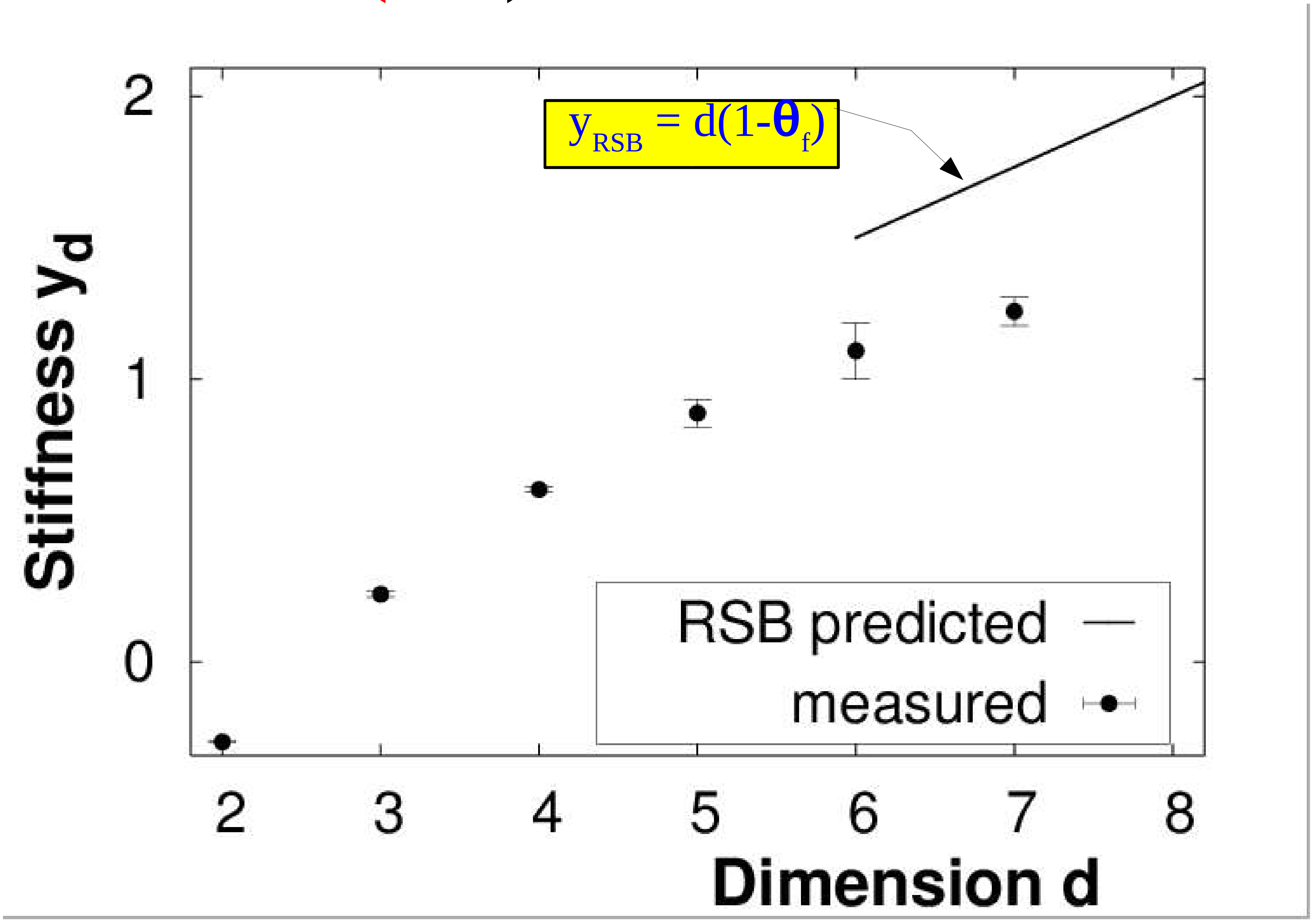
d=7 (!)





Comparing with Mean-Field Theory:

"Stiffness": $\sigma(\Delta E) \sim L^y$



Cinderella Cluster:

- 18 Boxes w/ AthlonXP (2GHz), 512 Mb, no HD, Video, etc
- 1 Server w/ P4, 1Gb, RAID5, Tape, etc
- Diskless NetBoot over fast Switch
- QUANTIAN_4.6.0.9 (Debian) w/
  - TerminalServer
  - OPENMOSIX loadbalancer
  - MPI for parallel processing

Cost: ~$8,000

Conclusions:

- Extremal Optimization:
  - ➔ Selection *against* extremely *Bad* ⇨ Greedy!
  - ➔ *No* Rejection ⇨ Large Fluctuations ⇨ No Trapping!
  - ➔ *Single*, *fixed* Parameter (τ) ⇨ Simple!
  - ➔ τ-EO: Optimizing at the *Ergodic Edge*.
  - ➔ Problems: Definition of Fitness and Sorting Ranks.

- Results:
  - ➔ Works well for Partitioning, Coloring, Spin Glasses, Satisfiability, Pattern Recognition (at least!).
  - ➔ Works poorly for TSP, Polymer Folding, ie. *highly* connected problems!
  - ➔ Theory: "*Jamming*" Model, predicting $\tau_{opt} \searrow 1^{+}$.