Spin glasses and Adiabatic Quantum Computing

A.P. Young

UNIVERSITY OF CALIFORNIA SANTA CRUZ

Talk at the Workshop on Theory and Practice of Adiabatic Quantum Computers and Quantum Simulation, ICTP, Trieste, August 22-26, 2016

Spin Glasses

The problems studied on current quantum annealing hardware are spin glasses. Spin glasses have been studied by physicists for many years and **insights thus obtained are helpful in understanding the difficulties in developing efficient quantum annealers**. I will review these ideas from spin glasses as well as some recent work applying them to quantum annealing, and raise some open questions. (Mainly a review of the work of others.) Will focus on:

- Phase Transitions
- Chaos

The Hamiltonian

We wish to find the **ground state** of the following **classical** Hamiltonian:

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij} S_i S_j - \sum_i h_i S_i$$

where the S_i are Ising spins, ± 1, the J_{ij} are the "frustrated" interactions (random in sign). We may also include random longitudinal fields h_i in which case h will denote their standard deviation. For now we set $h = 0$.

Spin Glass Phase Transition

As the temperature decreases correlations grow. The spin glass correlation length ξ is defined (for $h = 0$) by

$$[\langle S_i S_j \rangle^2] \sim \exp(r_{ij}/\xi)$$

where $\langle \ldots \rangle$ denotes a thermal average and [...] denotes an average over samples (or equivalently an average over different regions of the sample with fixed r_{ij}).
As the transition temperature T_c is approached ξ diverges like

$$\xi \sim (T - T_c)^{-\nu}$$

In dimension = 3, 4, ..., $T_c > 0$. However, in $d = 2$, **including D-Wave's chimera graph**, $T_c = 0$ and we have

$$\xi \sim T^{-\nu}, \quad (d = 2), \quad \text{where } \nu \simeq 3.4$$

Spin Glass Ordering

For $T < T_c$ $\quad \lim_{r_{ij}\to\infty} [\langle S_i S_j \rangle^2] \to (\text{order parameter})^2$

Chaos in spin glasses

As T is lowered the spin glass configuration that that minimizes the free energy can change (quite suddenly, a rounded “transition”) which is called temperature chaos, or T-chaos for short. Spin correlations change at distances greater than ℓ where

$$\ell = c_T \, (\Delta T)^{-\zeta}$$

In addition to T-chaos, there is also sensitivity to small changes in the interactions, called J-chaos, where the length scale is

$$\ell = c_J (\Delta J)^{-\zeta}$$

Numerically $\zeta \simeq 1$ in d = 2, 3, 4 for both J-chaos and T-chaos.

However, the amplitude is much bigger for J-chaos, i.e.

$$c_J \gg c_T$$

Example of T-chaos on the chimera graph

In some spin glass samples temperature chaos will not occur, in others it may occur once, twice etc. Instances where this occurs will be particularly hard to solve. Fraction of instances where this occurs is found to increase with increasing size N.

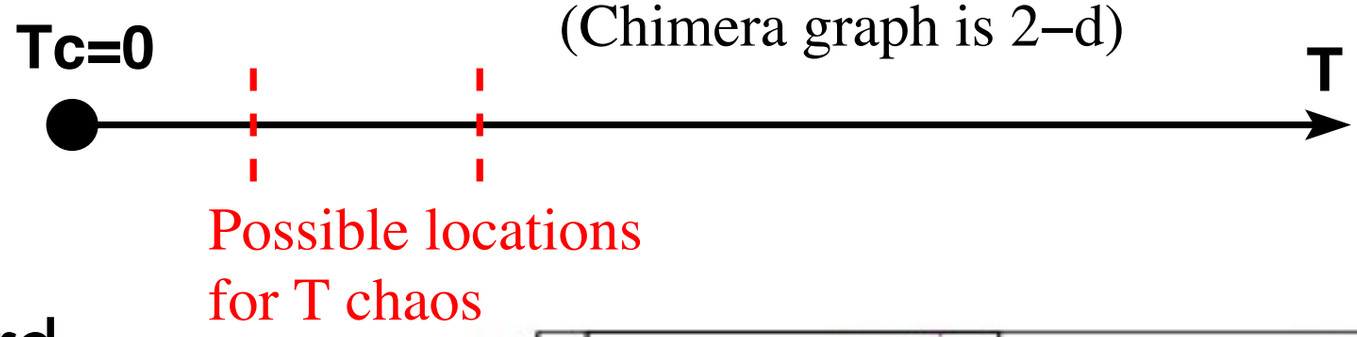
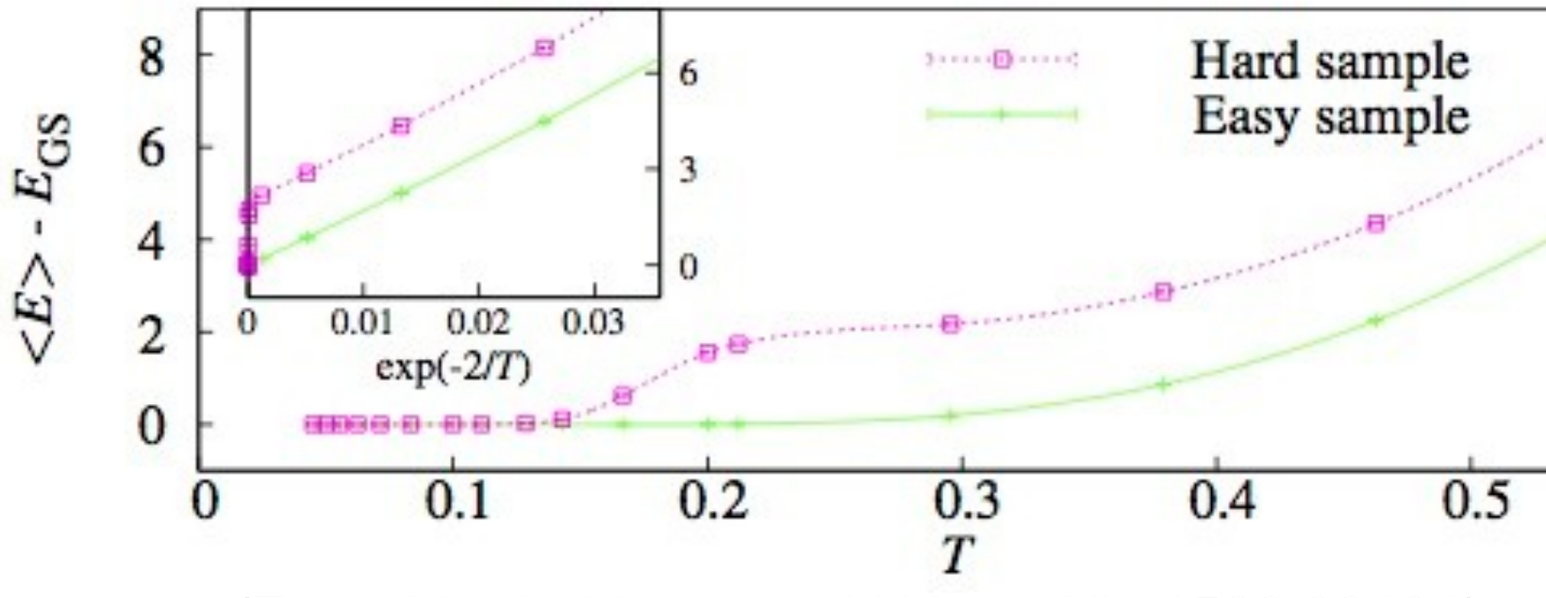


Figure shows a hard sample, in which the energy shows a pronounced change at low-T due to temperature chaos, and an easy sample where this does not occur.



(From Martin-Mayor and Hen, arXiv:1502.02494)

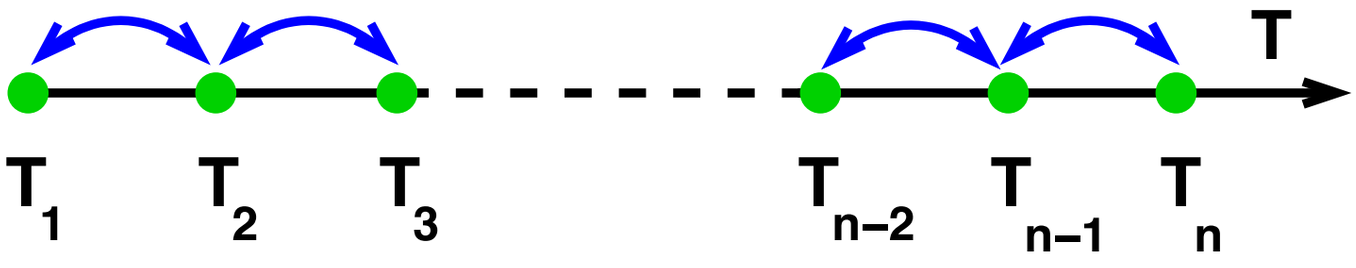
Samples with T-chaos also have low energy excited states which are very different from the ground state (Katzgraber et al, arXiv:1505.01545, also E. Crosson's talk on Tuesday)

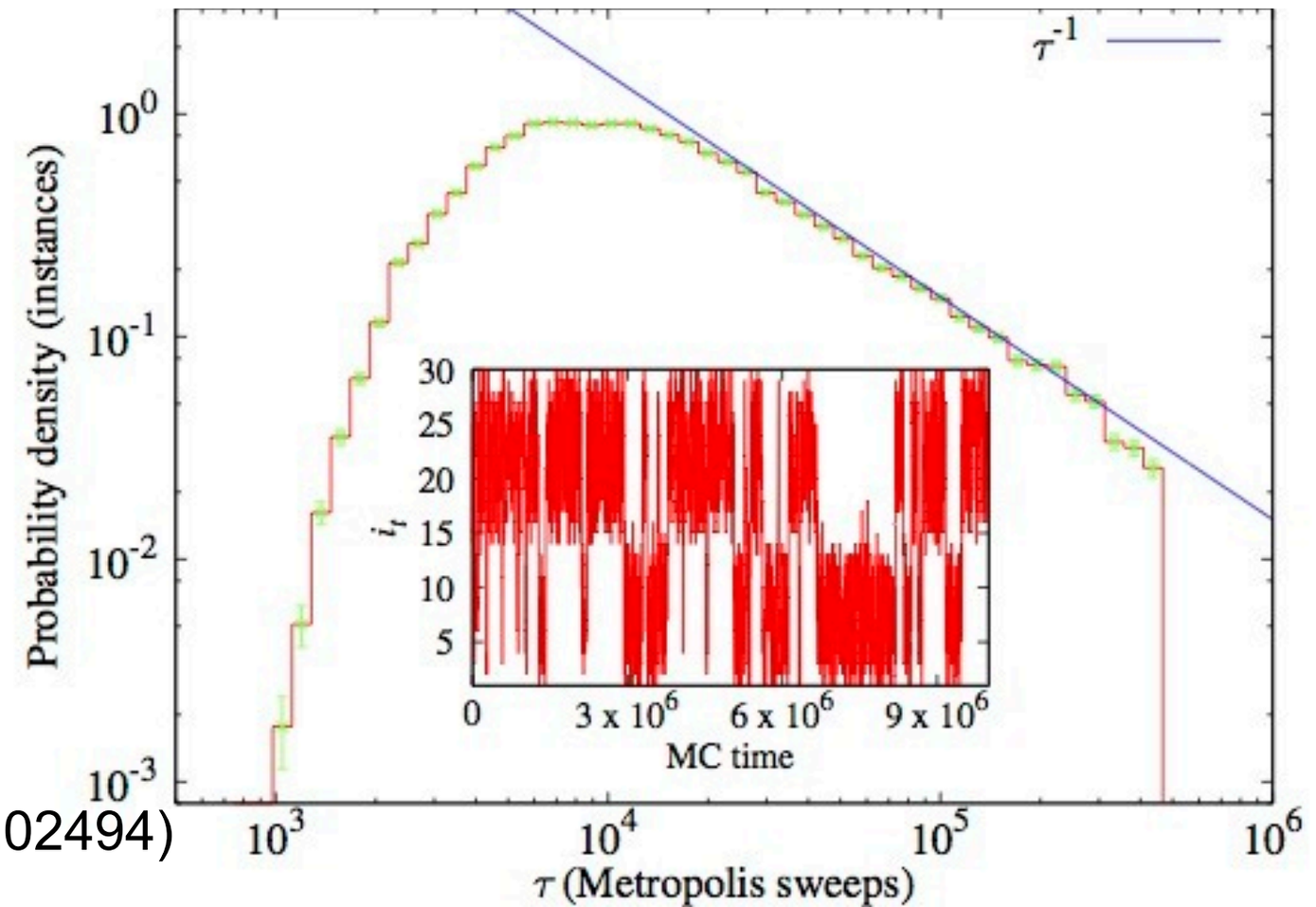
Parallel tempering

The method of choice for simulating spin glasses is called parallel tempering. One simulates copies of the system at n temperatures $T_1 < T_2 < ... < T_n$. Standard MC updates are done at each temperature. In addition there are "swap" moves in which the spin configurations at neighboring temperatures are swapped with an appropriate probability.

Thus, the temperature of each copy does a **random walk** between T_1 and T_n.

The "mixing time" τ is the average time it takes each copy to fully traverse the temperature range. **This is a measure of classical hardness.** There is a broad range, see figure.





(From Martin-Mayor and Hen arXiv:1502.02494)

Sample-to-sample fluctuations

There is a **broad distribution** in the values of the mixing time τ.

Interpretation: samples with small τ presumably have no T-chaos, while those with large τ presumably have one or more temperatures where T-chaos occurs.

One finds that T-chaos is rare for small sizes but happens in most samples for very large sizes.

T-chaos is problematic for classical, annealing-type algorithms.

Is it a problem for other classical algorithms or for quantum annealers?

To see this, for each sample on chimera graph (choice of the J_{ij}) Martin-Mayor and Hen determined τ. This is a measure of the classical hardness.

They then determine the time to solution t_s for each sample on a different classical algorithm and on the D-Wave machine, and ask how the t_s are correlated with the τ (see next slide).

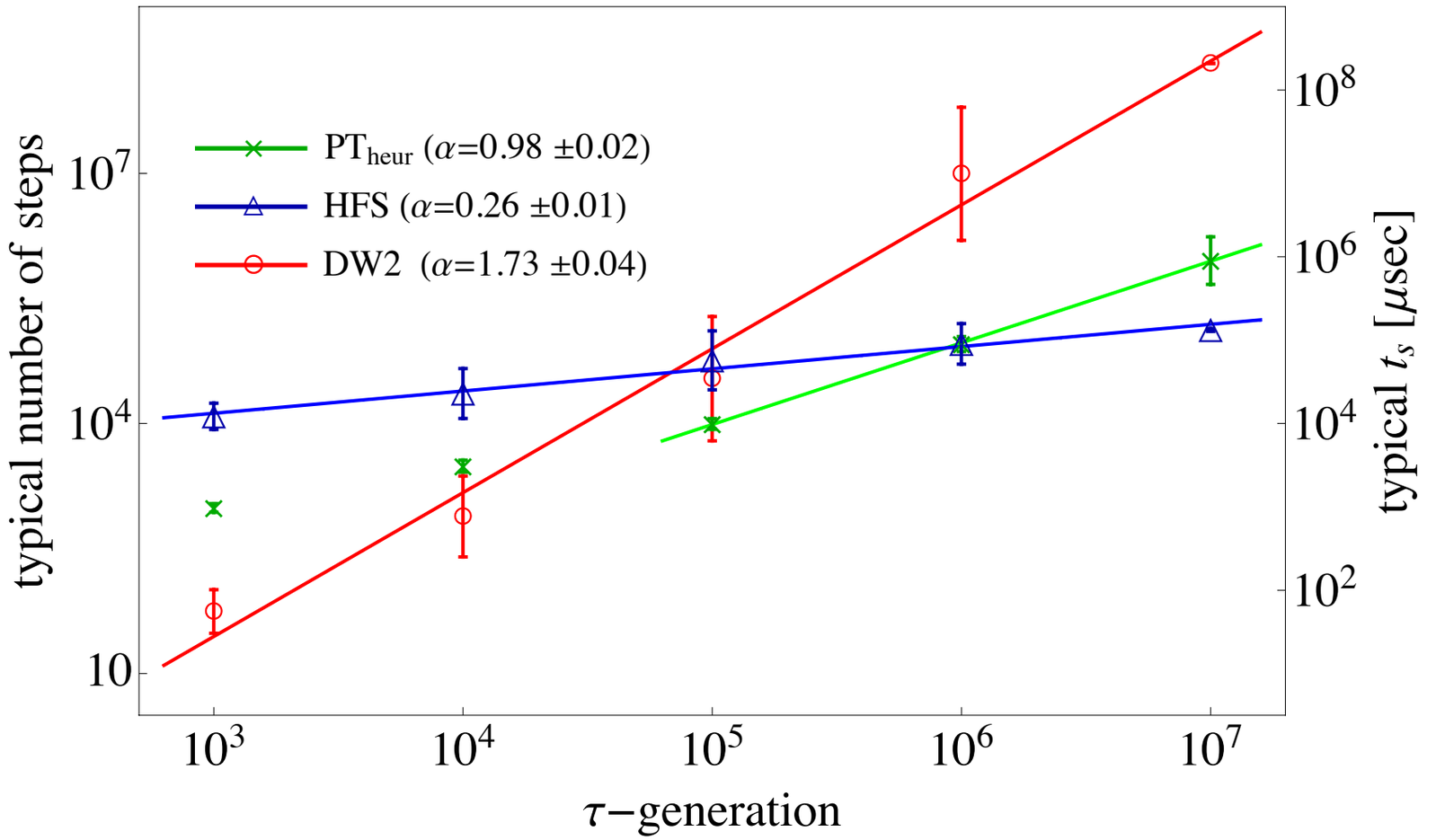
Hamze-de Freitas-Selby algorithm

For each sample, the time to find a solution, t_s, with the HFS algorithm (most efficient known for chimera graph) is determined.

A correlation is found between t_s and the classical mixing time τ.

One finds that $t_s \propto \tau^{\alpha}$ with $\alpha \simeq 0.26$, see figure, blue points, slope 0.26.

By contrast the time to solution for the classical PT algorithm is of order τ, as expected (green points, slope 1). Much stronger correlation with the D-Wave results (red) which will be discussed later.



(From Martin-Mayor and Hen arXiv:1502.02494)

How to generate hard instances?

The classical mixing time τ is a measure of classical hardness.

However, with current quantum hardware, the largest possible size is around N = 1000, for which most problems are not all that hard. How can we generate particularly hard samples?

One possibility: brute force. Compute τ for, e.g. 10^6 samples and study in detail the 10^2 with the largest τ.

Recently, Marshall, Martin-Mayor, Hen, arXiv:1605.03607 proposed a more efficient method to find hard samples by doing "simulated annealing" (SA) in space of couplings.

Standard SA. Want to minimize the energy, E. Add a fictitious temperature so Boltzmann factor is exp(-E/T). Do importance sampling on the spins and slowly decrease T.

M-MM-H. Want to maximize e.g. the mixing time τ. Add a fictitious temperature so "Boltzmann factor" is exp(τ/T). Do importance sampling on the J's and slowly decrease T. The J's evolve to give samples with larger τ.

Now make things quantum

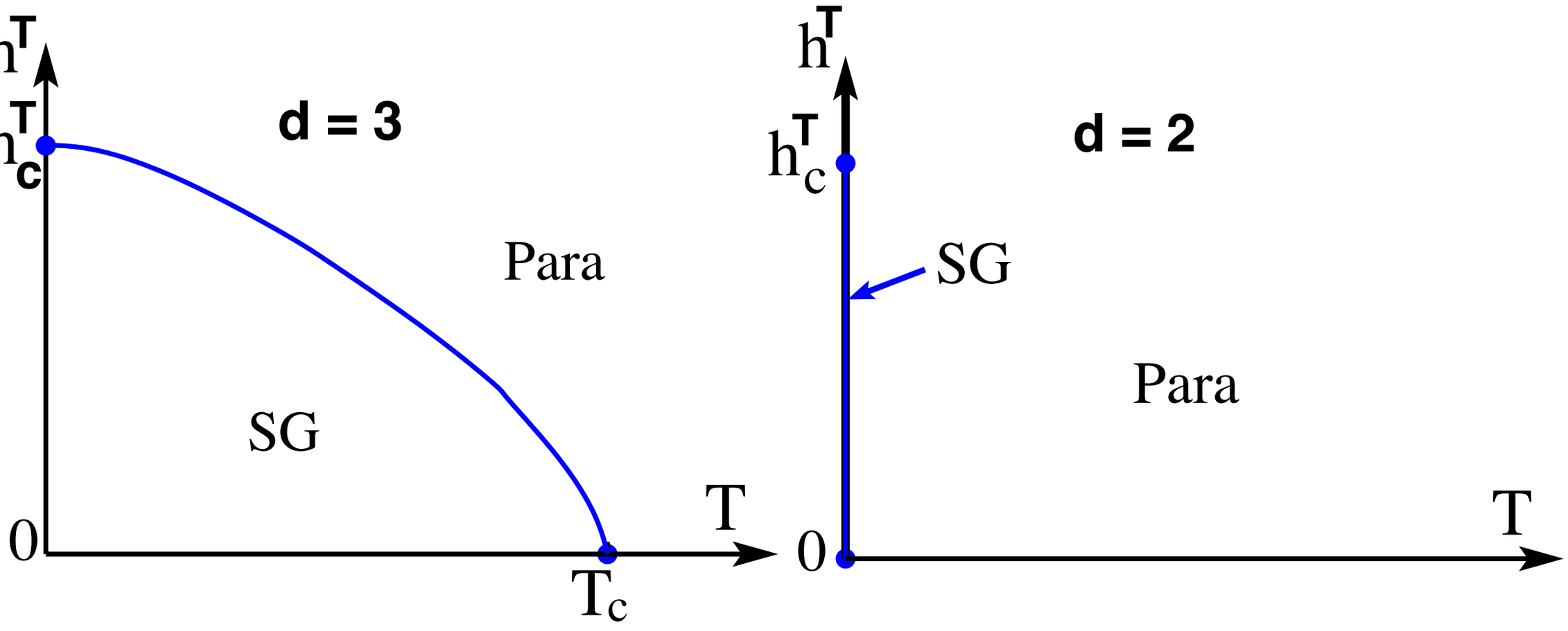
In adiabatic quantum computing the simplest way to add quantum fluctuations to a classical “problem” Hamiltonian is to add a “driver Hamiltonian” consisting of a single transverse field h^T, i.e. for zero longitudinal field, h = 0, we have

$$\mathcal{H} = -\sum_{\langle i,j \rangle} J_{ij}\sigma_i^z\sigma_j^z - h^T\sum_i \sigma_i^x$$

where the Ising spins have been promoted to Pauli spin operators σ^z. We assume the J_{ij} give spin glass behavior and the spins (qubits) are on a lattice in d-dimensions (including chimera for d=2).

Phase boundaries for d = 2 and 3 are shown in the next slide.

Phase boundaries, transverse field spin glass



Note: spin glass phase only at T = 0 for d = 2.

Where's the bottleneck in Quantum Annealing?

Answer: where the energy gap becomes very small (avoided level crossing).

(i) Could be at a quantum critical point(QCP). Characterized by a dynamical exponent z.

$$\Delta E \sim \xi^{-z} \sim (h^T - h_c^T)^{z\nu}, \ (L \to \infty)$$

$$\sim L^{-z}, \ (h^T = h_c^T)$$

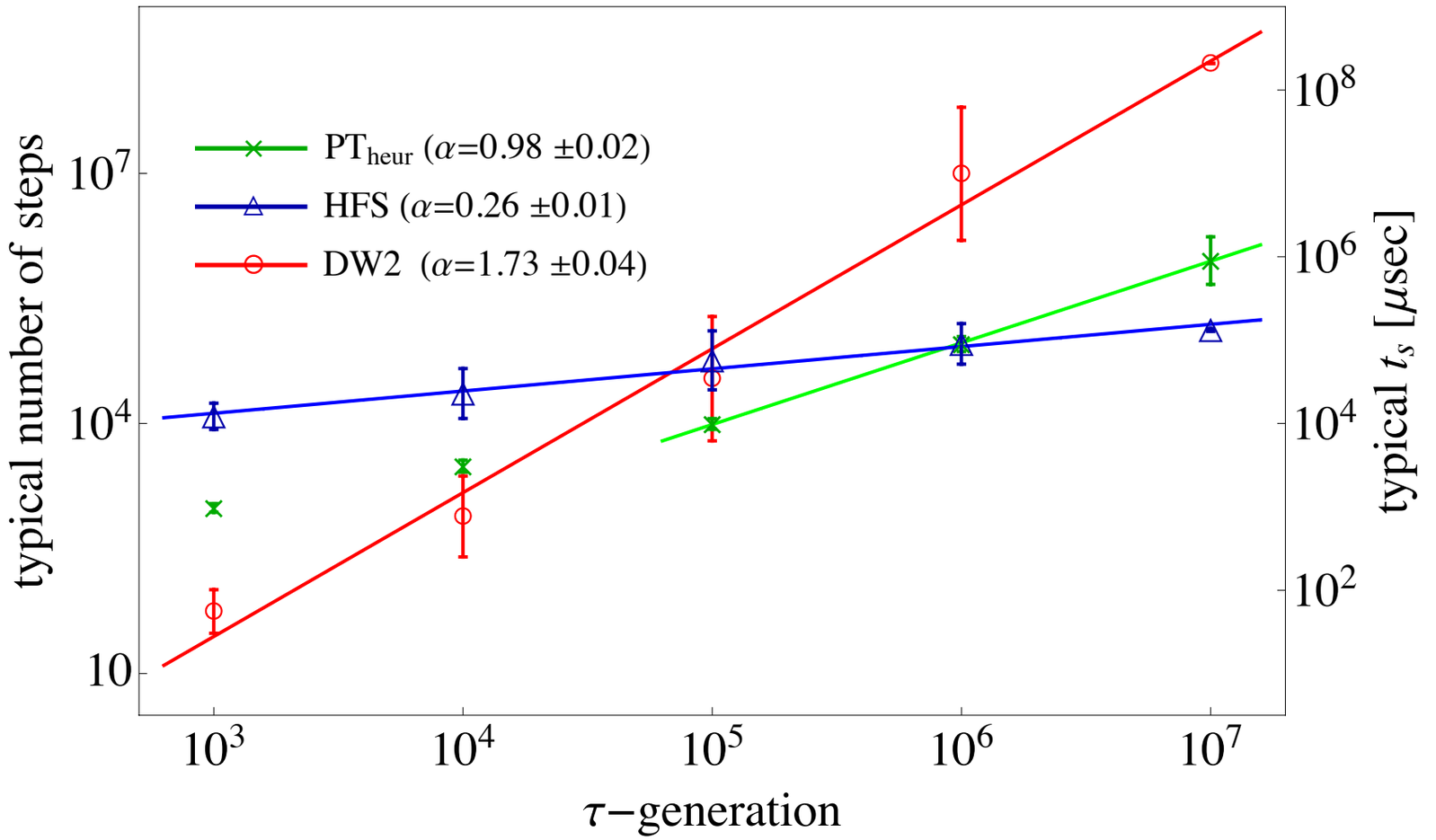
With disorder can have z infinite so gap is exponentially small in size at the QCP (activated dynamical scaling). E.g. d =1 random, transverse field ferromagnet (D.S. Fisher).

(ii) Or could be in the quantum spin glass phase (nature of spin glass state rapidly changes). Analogous to T-chaos. Call this transverse field chaos, or TF-chaos for short. (Avoided level crossing)

D-Wave results

(Martin-Mayor & Hen, arXiv:1502.02494) For each sample, the time to find a solution, t_s, on the D-Wave machine is determined. A strong correlation is found between t_s and the classical mixing time τ. They find that $t_s \propto \tau^{\alpha}$ with $\alpha \simeq 1.73$ see figure (red points, slope 1.73)

By contrast the time to solution for the classical PT algorithm is of order τ, as expected (green points, slope 1). In other words, the hard samples are even harder on the D-Wave machine than classically.



(From Martin-Mayor and Hen)

Interpretation of D-Wave results

Does this mean that quantum annealing is less efficient than classical algorithms? Not necessarily. The observed result that the time to solution on D-Wave varies as the time to solution using the classical parallel tempering (PT) algorithm to a power greater than one could have classical origins:

- The temperature is not low enough. For instances where temperature chaos occurs at a temperature lower than that of the chip then the wrong answer will typically be obtained.
- The strengths of the bonds are not represented exactly in the (analog) D-Wave machine (intrinsic control errors, ICE). Even small changes in the bond strengths can dramatically change the ground state. This is called "J-chaos". Thus D-Wave machine might be getting the right ground state to the wrong problem (some of the time). Do samples with strong T-chaos also have strong J-chaos? Probably, but more work needed to make this precise.

2d vs. 3d

Recent detailed simulations by, e.g. Troyer et al .Science, 345, 420 (2014), Katzgraber et al. arXiv:1505.01545, Martin-Mayor and Hen arXiv: 1502.02494. Mainly for the chimera graph (i.e. 2d) in order to compare with experiment.

How hard is 2d?

Katzgraber et al, arXiv:1401.1546: quite easy because $T_c = 0$. But ξ diverges strongly as $T \rightarrow 0$. On scales less than ξ, is the problem much easier than in 3d? Also one can find hard samples in 2d, (e.g. Martin-Mayor and Hen, arXiv:1502.02494).

Asymptotic scaling of best algorithms seems to be $\exp(c\, N^{1/2})$. Expected since for good algorithms time ~ exp(c' TW) where TW is the ``tree width'' of the graph (~ $N^{1/2}$ in 2d), see e.g. Lidar et al. Science,345, 420 (2014).

How much harder is 3d?

What is the treewidth? Is it $N^{2/3}$? Is the time ~ $\exp(c\, N^{2/3})$?

Would be interesting to have simulations in 3d at at the same level of detail and care that were done in 2d.

Non-zero (random) longitudinal field

Note: for a symmetric distribution of J's, the sign of the field can be gauged away so a uniform field is equivalent to fields with a symmetric distribution. We consider the latter. Remember h is the standard deviation of the field distribution.

First, the effect of a field on classical spin glasses.
The field breaks inversion symmetry and naively should round out the zero-field transition (where this symmetry is broken). But this is not necessarily the case for a spin glass. Back to the definition of correlation length. Recall in zero field, the spin glass correlation function is $C_{ij} \equiv [\langle S_i S_j \rangle^2]$

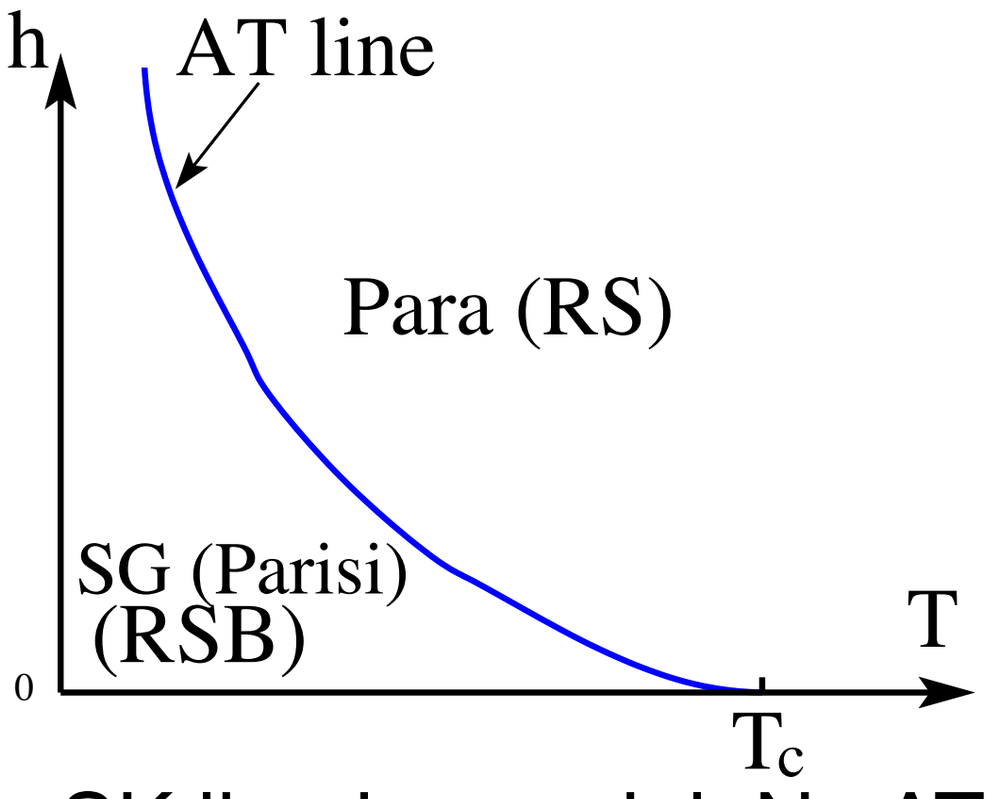
However, in a field we have to replace $\langle S_i S_j \rangle$ by the "connected" correlation function $\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle$ and so

$$C_{ij} \equiv \left[\left(\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \right)^2 \right] \sim e^{-r_{ij}/\xi}$$

provides a definition of the spin glass correlation length in a field (above any possible transition).

Non-zero (random) field

Can ξ diverge in a field? For at least one model the answer is yes. This is the mean-field (infinite-range) Sherrington-Kirkpatrick (SK) model (think of it as infinite-d). There is a line of transitions in a field called the de Almeida-Thouless (AT) line.



This is for the SK model. Is there an AT line in finite-dimensional models? Not clear.

Different hypotheses:

- SK line is special. No AT line in any finite-d (Fisher-Huse)
- An AT line only above the upper critical dimension, $d_u = 6$ (Moore) (some numerical evidence, Katzgraber and APY)
- An AT line everywhere that $T_c > 0$ in *zero* field, i.e. $d \geq 3$ (Parisi et al.)

Note: numerics turns out to be hard.

SG order parameter in non-zero (random) field

For $T < T_c(h)$, i.e. below the AT line (if any)

$$\lim_{r_{ij}\to\infty} \left[\left(\langle S_i S_j \rangle - \langle S_i \rangle \langle S_j \rangle \right)^2 \right] = (\text{order parameter})^2$$

The region where this order parameter ≠ 0 in a longitudinal field, i.e. below the AT line, is called a replica symmetry breaking (RSB) phase. There are many valleys connected by barriers which diverge for $N \to \infty$.

The region where this order parameter = 0, i.e. above the AT line, is called the replica symmetric (RS) phase.

Does a field make the problem harder?

Running parallel tempering simulations in an intermediate-strength field seems harder than in zero field. T-chaos is much larger (similar to J-chaos). Suggests that a field makes the problem harder (at least for annealing-type algorithms).

But, if there is no AT-line, then ξ is always finite. For finite ξ, algorithm of Zintchenko,Hastings and Troyer (2015) finds solution in patches and joins patches together. In 2d, typical instances are solved in polynomial time(!) even down to $h = 0$ where $\xi \to \infty$, at least for the sizes studied. What about 3d? Maybe, then, a field makes things easier?

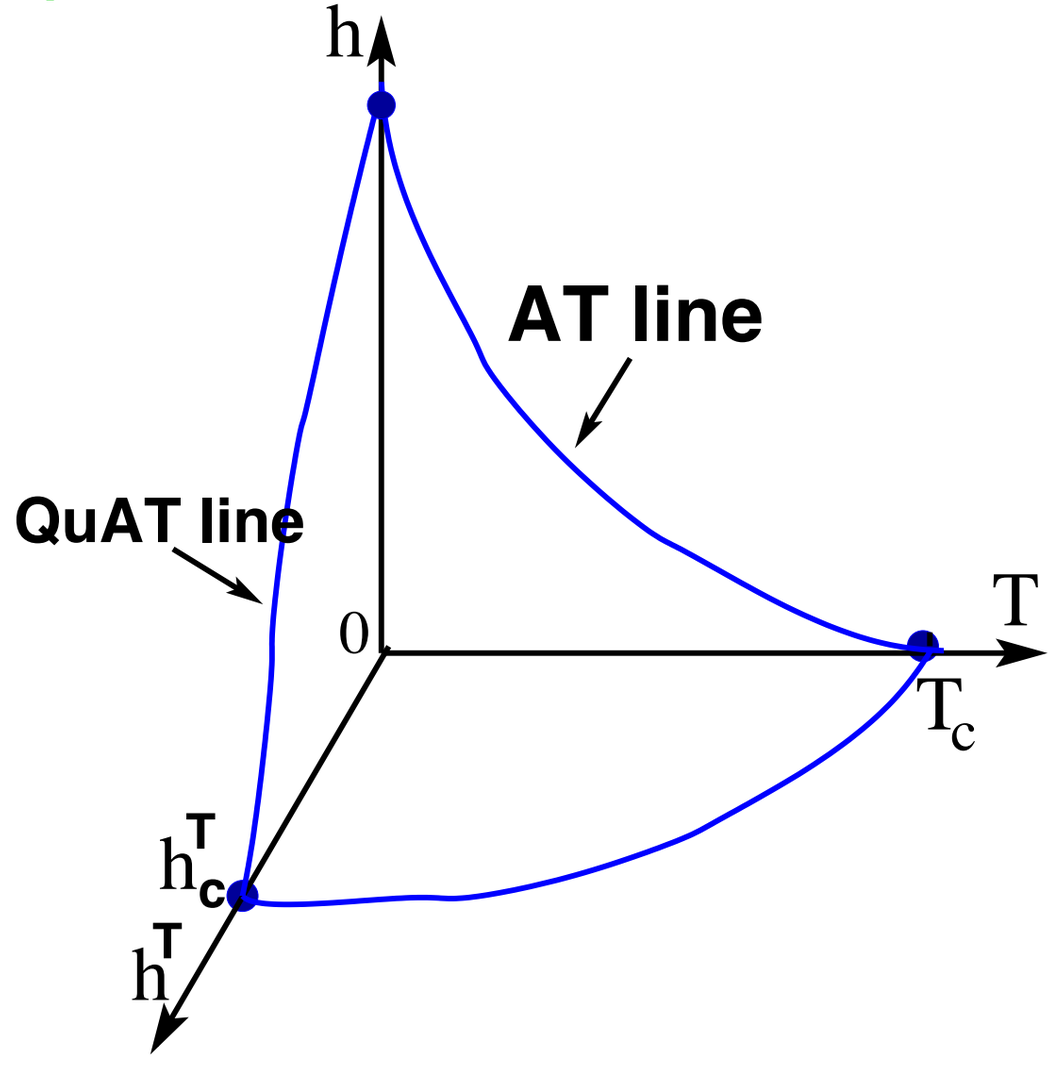
A quantum de Almeida-Thouless (QuAT) line?

For the SK model, there is a (classical) AT line in the h-T plane at $h^T = 0$, below which we have RSB. Is there also RSB for $h^T > 0$?

- **Yes.** Büttner and Usadel (1990) and Goldshmidt and Lai (1990):,
- **No**. Ray and Chakrabarti[2] (1989): See also Chakrabarti's talk on Tuesday).

Does RSB go all the way down to T =0? If so, there is also a quantum AT (QuAT) line in the h-h^T plane at T = 0.

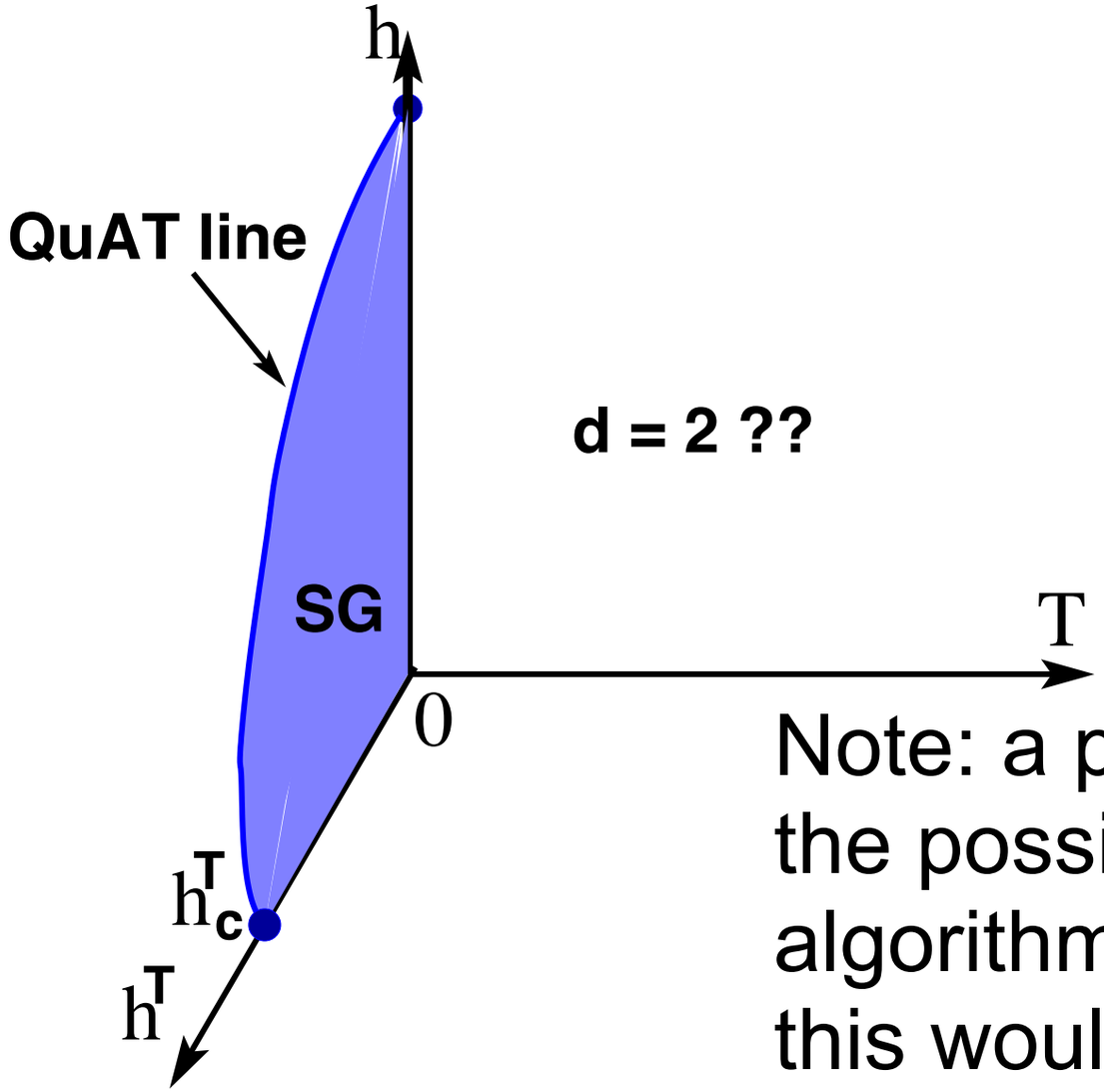
This figure is a surmise for the SK model.



Suppose that there **is** an AT for some range of (finite) dimension, then there may be a QuAT line for some dimensions, not necessarily the same.

Could there even be at QuAT line in d = 2?

In d= 2 there **is** a non-zero critical value of the transverse field, so could there be a QuAT line in this case? Not ruled out but perhaps unlikely.



Recall, below the QuAT line we have RSB, something like the Parisi solution of the SK model.
Above the QuAT line we have an RS phase.

Note: a phase transition provides one of the possible bottlenecks for annealing algorithms. If there is an QuAT line in 2-d, this would provide an additional bottleneck for problems with fields (biases) running on the D-Wave annealer.

Conclusions

- Phase Transitions:
  - They provide one of the bottlenecks in annealing algorithms.
  - If there is a Quantum AT line in a magnetic field at T = 0 this would presumably affect the performance of quantum annealing.
- Chaos:
  - Stressed importance of T-chaos and TF-chaos in making problems hard. To what extent are they correlated? Very broad range of hardness of problems of a given size.
  - Importance of intrinsic control errors (ICE) in getting correct results from analog quantum annealers, e.g. D-Wave, because of J-chaos. Are samples with J-chaos also those with T-chaos?
- Other questions/problems
  - Is h > 0 harder or less hard than h = 0?
  - How much harder is d = 3 than d = 2?

Thank you