The Fair in Unfair Quantum Ground-State Sampling

Itay Hen
Information Sciences Institute, USC

Group Meeting
Jan 13, 2017

preprint:
arXiv:1701.01524

Information Sciences Institute
USC School of Engineering

Collaborators

- Brian Zhang – a high school student, now at Stanford.

- Gene Wagenbreth – formerly at ISI, now at Cray.

- Victor Martin-Mayor – will be visiting during the last week of January and will deliver the QI/Condensed Matter seminar on Jan 27.

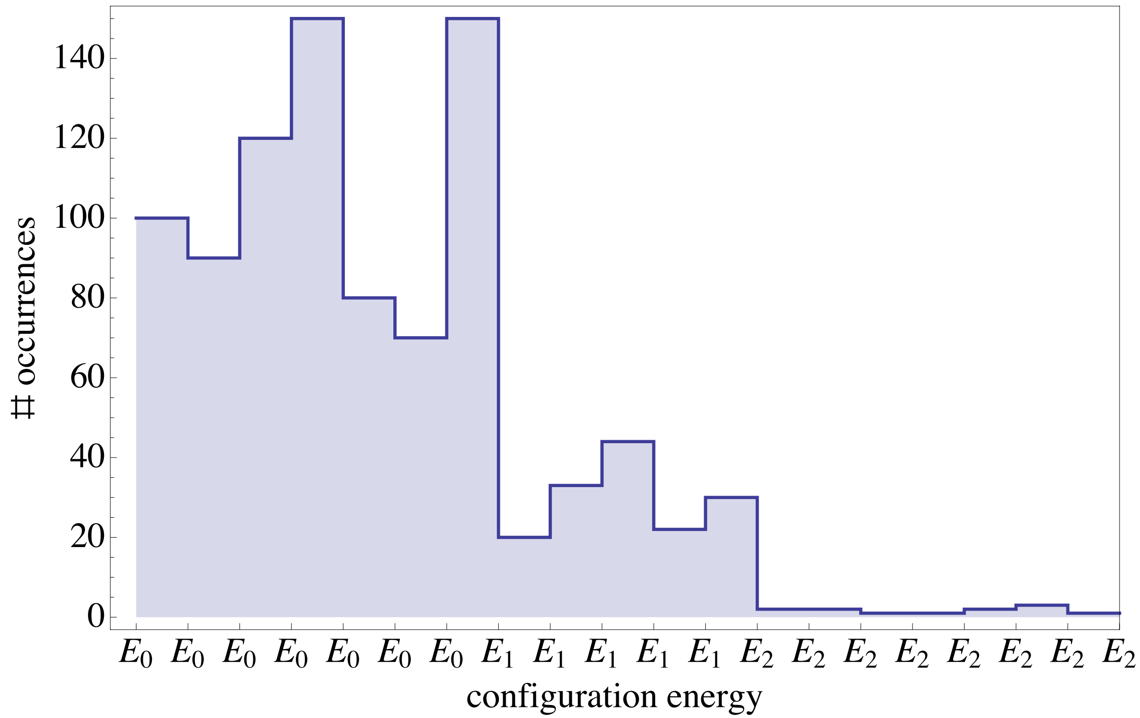
Motivation

Motivation

- many practical problems can be cast as a task of finding all (or as many as possible) minimizing assignments of combinatorial cost functions. some examples are: verification and validation (a longtime goal of our Lockheed-Martin collaborators), circuit fault detection, k-SAT filtering.
- devising algorithms that sample the ground state manifolds of such cost functions in unique manners is therefore of great practical interest in many areas.

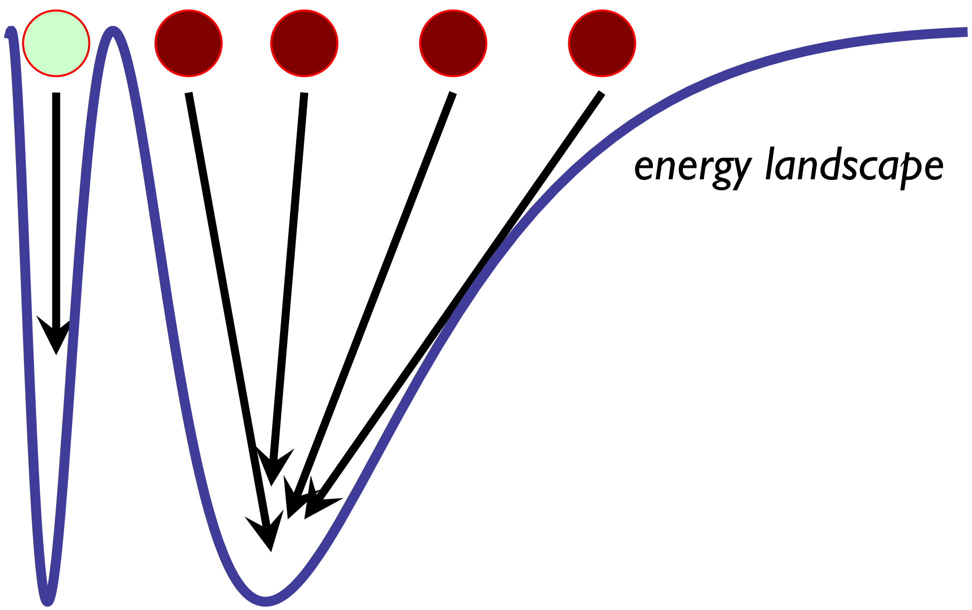
Quantum Annealers as Samplers

- quantum computers are apparently good at sampling.
  - most notably: Shor's QFT (used in almost all non-blackbox quantum algorithms that are superior to the best classical ones).
  - also, boson sampling (Aaronson & Arkhipov).
- quantum annealers are samplers. no reason to believe they always generate "classically-easy" distributions.
- this property may be used in a variety of ways.



Quantum Annealers as Samplers

- consider a (classical) optimization problem with a degenerate ground subspace.
- a thermal annealer functioning as a solver/optimizer, will sample the solution space, favoring some solutions over others (unless it has already equilibrated).
- some configurations may be "classically suppressed".
- when looking at time-to-minimal energy, this often does not matter.
- in other cases, it does. for example, when the problem at hand requires finding all solutions.



Quantum Annealers as Samplers

- when solving optimization problems, quantum annealers too sample the classical solution space.
- considering the generic interpolation:

  $$H(s) = (1 - s)H_d + sH_p$$

  under appropriate conditions, the annealer would choose one specific quantum ground state (as ensured by the adiabatic theorem).
- the amplitudes of the various classical states correspond to a "quantumly generated probability distribution" over solution states.

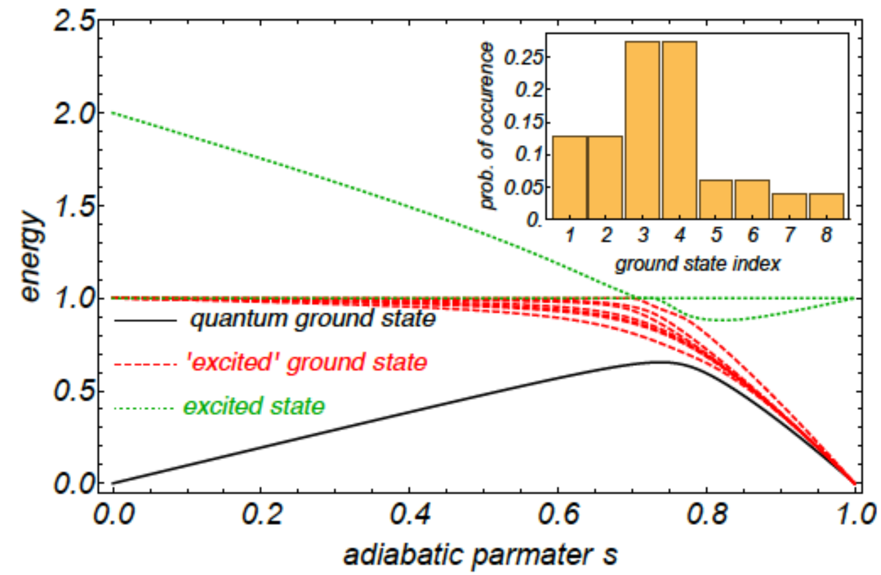
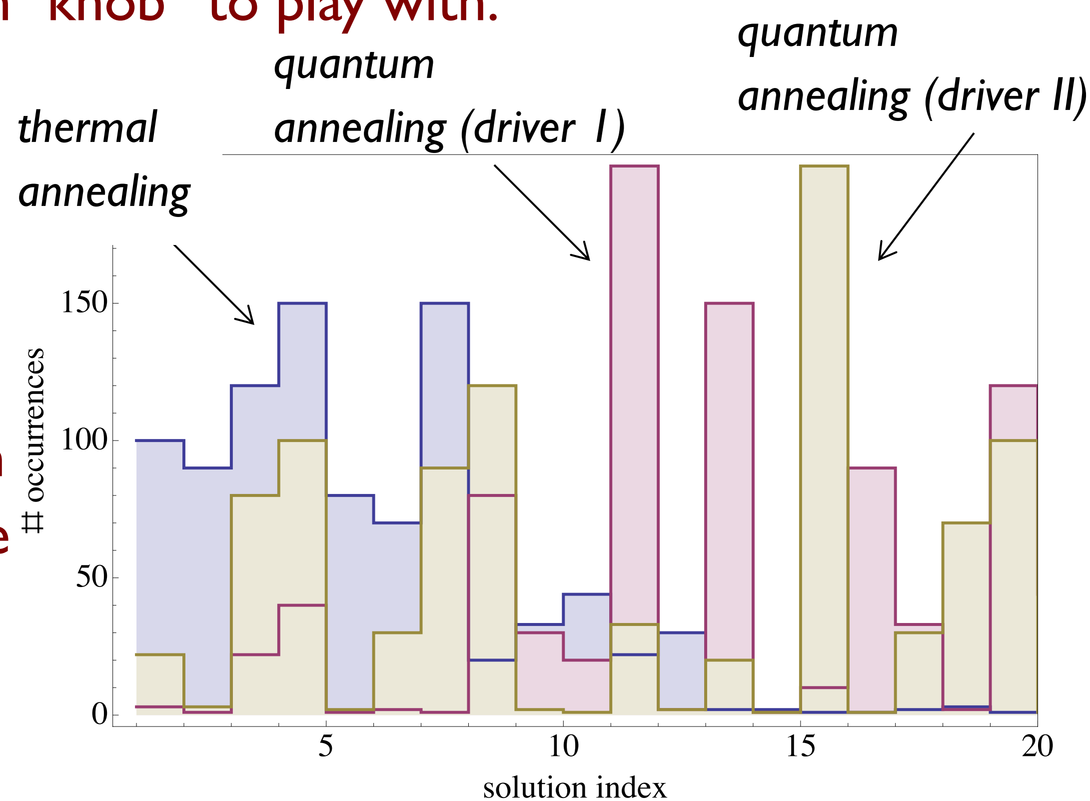


FIG. 1. Ten lowest energy levels of an 8-qubit Hamiltonian interpolating between a transverse field driver Hamiltonian and a randomly generated Ising Hamiltonian. The solid black line indicates the energy of the instantaneous ground-state. The dashed red and dotted green lines indicate excited states that lead to final ground-states and final excited states, respectively. Inset: The probabilities for obtaining the various classical ground-states upon measuring the quantum ground-state at the end of the evolution in the computational basis.

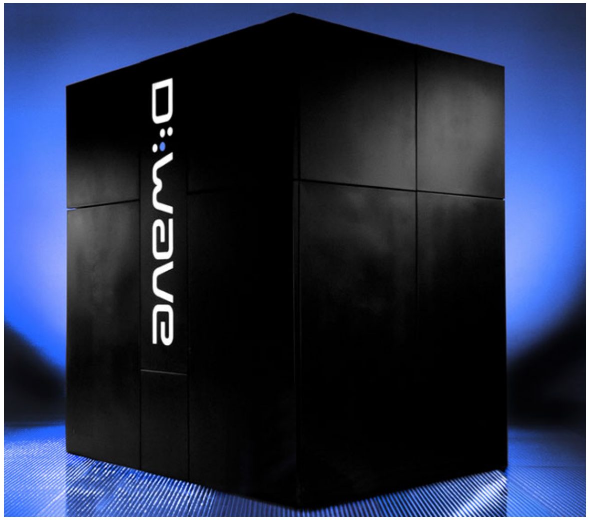
The "extra knob"

- the final distribution over the ground states will not only be generally different than the classical one, but will also depend on the choice of the initial (driver) Hamiltonian (or intermediate ones).
- this gives us an extra quantum "knob" to play with.
- e.g., what happens if the driver Hamiltonian is non-stoquastic?
- if classically suppressed states are found by a quantum annealer, this would constitute as a qualitatively quantum advantage.



Main questions of this study

- do quantum annealers sample the ground-state manifold of optimization problems differently than classical algorithms such as simulated annealing? how differently?
- how differently different drivers sample the ground-state manifold?
- how does the D-Wave processor fit in?
- can the dissimilarity in the sampling mechanisms be leveraged towards practical advantages?



Method

Setup

- to answer the above questions, we study the sampling capabilities of various algorithms as they pertain to the ground-state distributions (or GSDs) of Ising spin glasses:

$$H_p = \sum_{\langle ij \rangle} J_{ij} s_i s_j$$

- we compare four algorithms:
  - i) Simulated annealing (SA) considered as an optimizer.
  - ii) ideal zero-temperature transverse-field quantum annealer (TFQA):

$$H_d^{\mathrm{TF}} = -\sum_i \sigma_i^x$$

  - iii) ideal zero-temperature non-stoquastic quantum annealer (NSQA):

$$H_d^{\mathrm{NS}} = -\sum_i \sigma_i^x + \sum_{\langle ij \rangle} \tilde{J}_{ij} \sigma_i^x \sigma_j^x$$

  - iv) D-Wave Two (DW2) gauge-averaged, total of 10^7 anneals ($20-40\mu s$).

Tested instances

- we generate spin-glass instances with a Chimera connectivity graph.
- the instances we choose are of the planted-solution type. why?
- they enable the a-priori enumeration of all ground-state configurations (which required heavy numerical efforts).
- we end up with nearly 2000 instances all of whose individual ground states we have obtained.

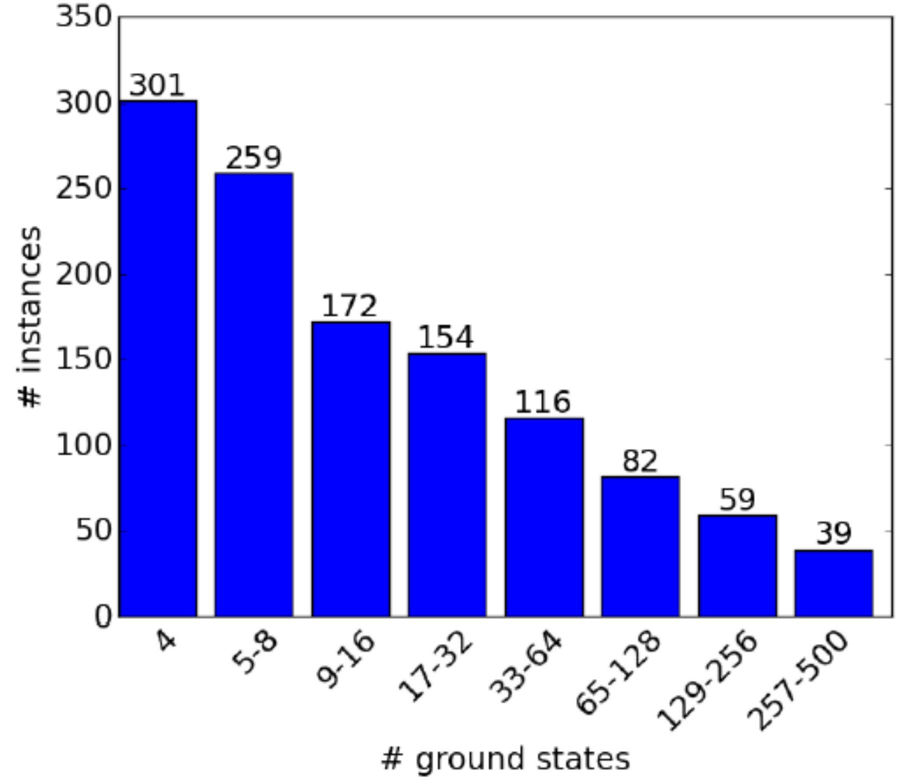
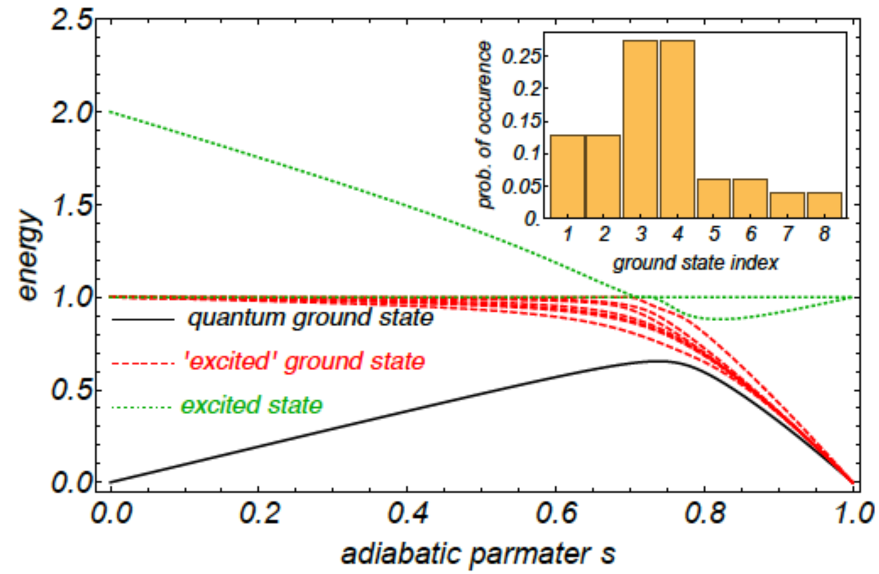


FIG. 9. **Histogram of number of problem instances vs. number of ground-states.** The distribution of number of solutions is a function of clause density after discarding instances with more than 500 solutions.

Calculating the quantum ground state

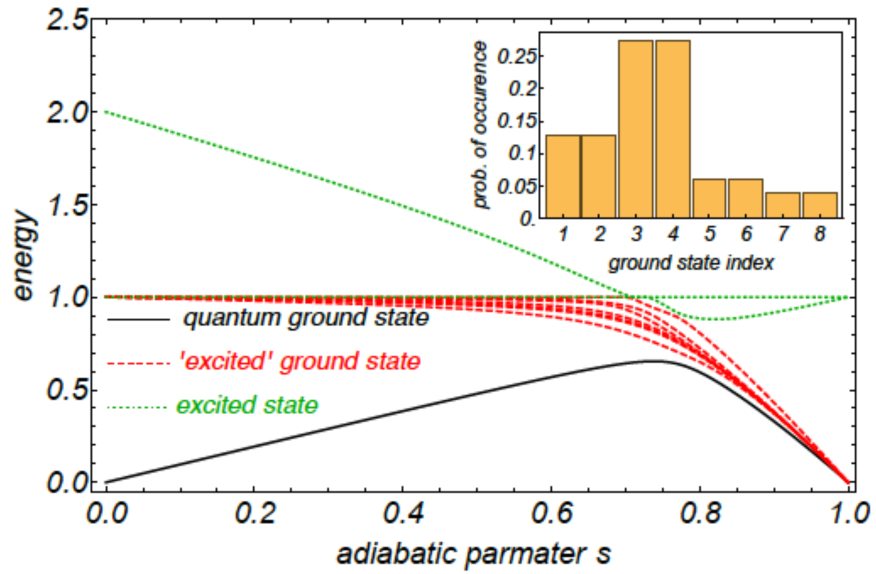
- SA and DW2 GSDs are easily obtained.
- however, calculating the quantum ground state of an $n = 500$ qubit problem Hamiltonian as produced by an ideal quantum annealer is in general not an easy task.



- it requires the diagonalization of H_p together with the application of degenerate quantum perturbation theory (up to n-th order).

Calculating the quantum ground state

- we devise a special algorithm that allows us, based on our knowledge of the individual ground states of H_p, to calculate the quantum ground state, without resorting to perturbation theory.



- the algorithm is based on the observation that the goal is essentially isolating the state of minimal energy.
- we are after the $s \to 1$ limit of $|\psi_{GS}(s)\rangle = \arg \min_{\psi \in V_k^S} \frac{\langle\psi|H(s)|\psi\rangle}{\langle\psi|\psi\rangle}$
- the algorithm is based on a "numerical annealing" of $H(s)$ on restricted sub-spaces.

Hilbert subspaces increasing in size

Results

Simulated annealing

- running simulated annealing (SA) as an optimizer, we indeed find that some minimizing configurations are reached much sooner than others.
- runtimes to individual solutions often differ by orders of magnitude.
- SA is run as an optimizer (to minimum of average runtime). if run as “thermalizer”, i.e., for infinitely long time, equipartition theorem prescribes same probability for all same-energy configurations (as also evident in figure).

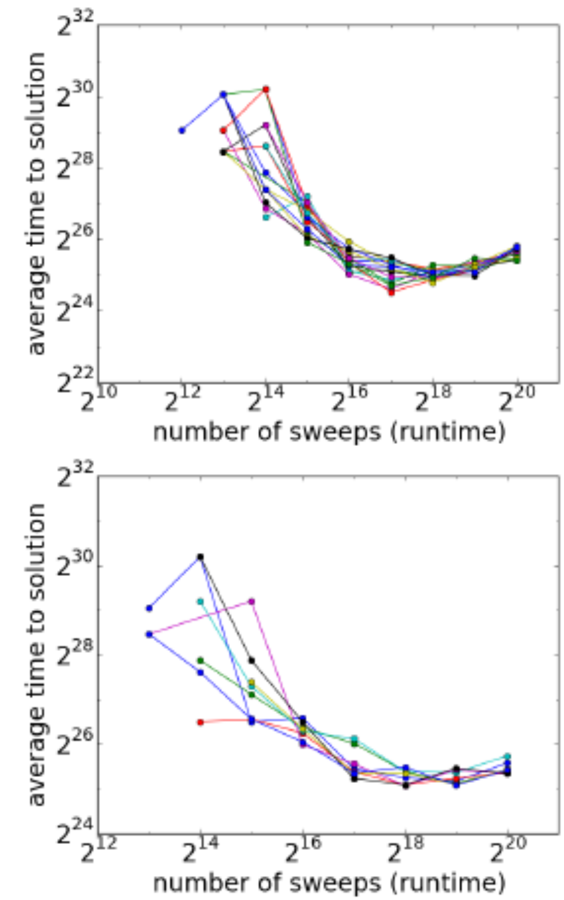


FIG. 2. **Simulated annealing average time to solution—the ratio of success probability to number of sweeps, vs. number of sweeps for different solutions (log-log scale).** Each line represents a different solution. **Top:** A 504-bit instance with 16 solutions. **Bottom:** A 504-bit instance with 4 solutions. For most instances, certain solutions are reached considerably sooner than others.

Are the probability distributions similar?

- we feed our generated instances to all four algorithms.
- we obtain GSDs for every instance (and for every algorithm).
- we ask: what is the probability that two GSDs of a given instance obtained by two different algorithms come from the same underlying distribution?
- to answer the question, we employ a Kolmogorov-Smirnov test.
- we find in general, that the distributions are significantly different from one another.

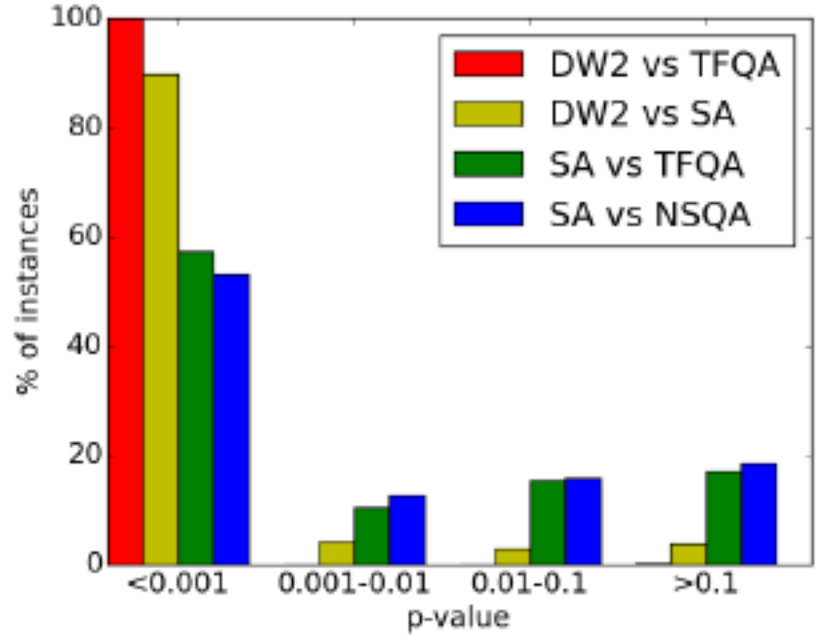


FIG. 3. The p-values generated by the Kolmogorov-Smirnov tests to quantify the differences between pairs of algorithms.

SA vs TFQA vs DW2

- some examples: looking at the GSDs as obtained by SA, TFQA and DW2 for three representative instances,
- we find qualitatively varying results.

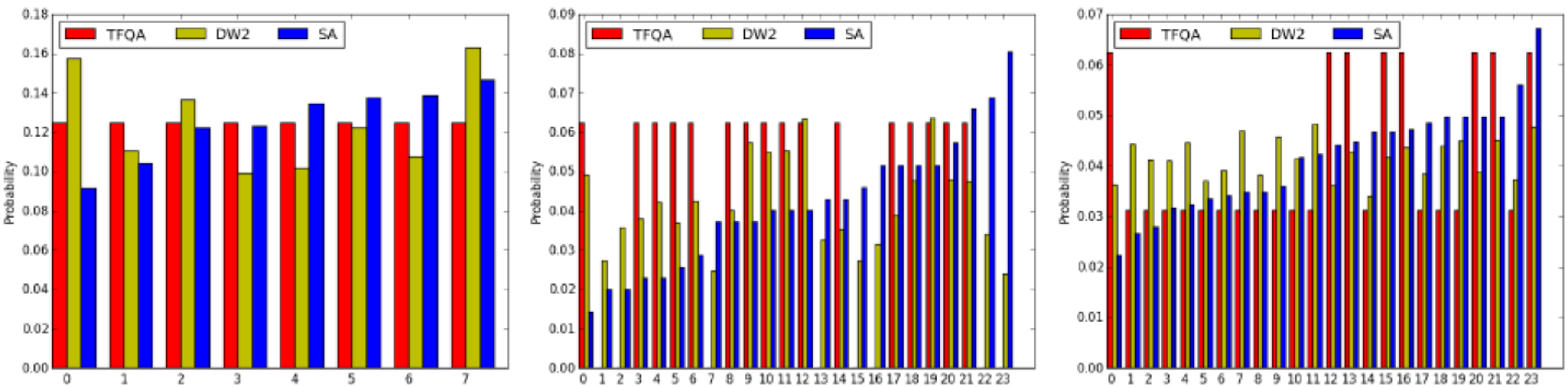


FIG. 4. **Three representative GSDs for simulated annealing (SA - blue), ideal transverse-field quantum annealer (TFQA - red) and the D-Wave Two experimental processor (DW2 - yellow).** In the left instance, probabilities for obtaining the various ground-states predict that all solutions are about equally likely. In the middle instance, we observe that those ground-states that our analytic TFQA predicts should appear more often, and do indeed appear more often in the experimental DW2 data. In the right instance, there is no clear relationship between the various algorithms.

Transverse-field QA vs non-stoquastic QA

- do different drivers generate different GSDs? yes!
- specifically, small probability configurations in one could have high-probabilities of being detected in the other.
- also, zero-probability configurations in one may have a non-zero probability in the other.

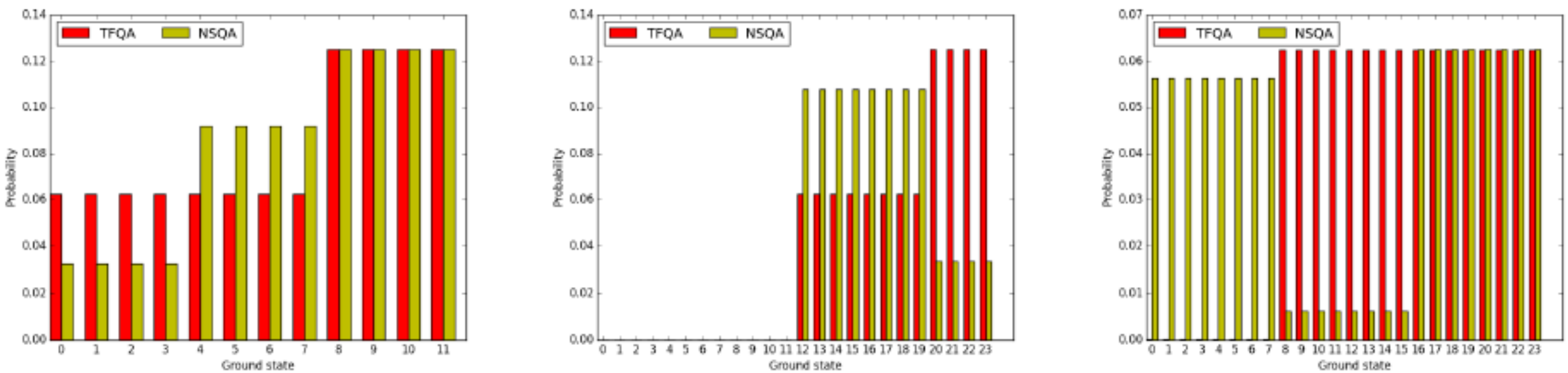


FIG. 7. **Three sample GSD comparisons between the two drivers: the stoquastic TFQA (red) and non-stoquastic NSQA (yellow).** For the instance on the left both drivers sample the ground-state manifold similarly; for the middle instance less probable configurations for one method become more probable in the other and vice versa; on the right, we find an instance for which states with zero probability of occurring with one type of quantum fluctuations have distinct positive probabilities of occurring in the other.

The “fair” sides of unfair ground-state sampling

The bright sides of biased sampling

- we define a measure for the bias of a GSD (denoted $\boldsymbol{p}$) obtained for a given instance:

$$b(\mathbf{p}) = \sqrt{\frac{D}{D-1}\sum_{i=1}^{D}(p_i - 1/D)^2} = \sqrt{\frac{D\sum p_i^2 - 1}{D-1}}$$

- here, D is the degeneracy of the H_p ground-state manifold and p_i denotes the probability/ frequency of the i-th ground state.

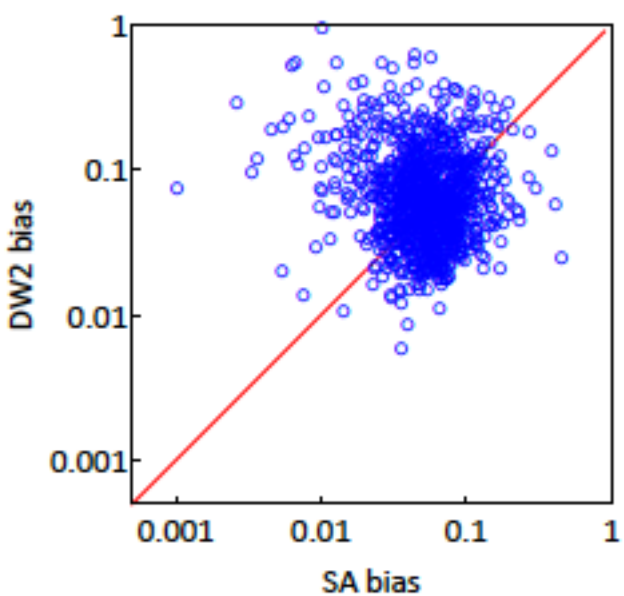
- the bias is zero for a flat distribution and is one for an extremely biased distribution.

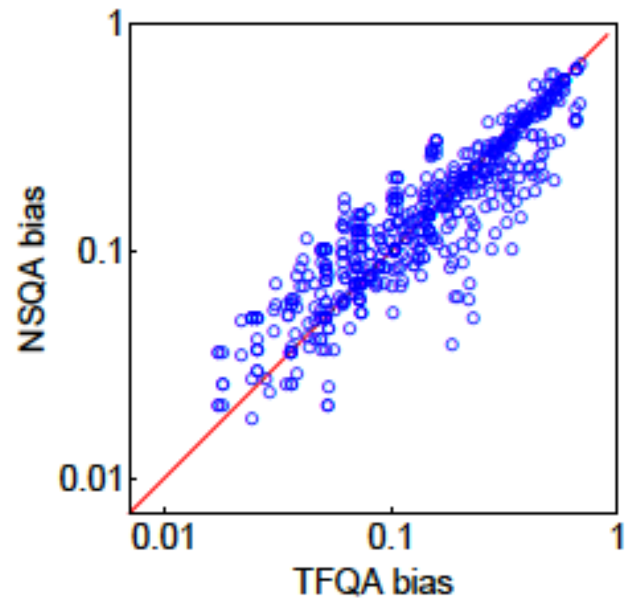
The bright sides of biased sampling

- we define a measure for the bias of a GSD (denoted $\boldsymbol{p}$) obtained for a given instance:

$$b(\mathbf{p}) = \sqrt{\frac{D}{D-1}\sum_{i=1}^{D}(p_i - 1/D)^2} = \sqrt{\frac{D\sum p_i^2 - 1}{D-1}}$$

- we can compare the biases of pairs of algorithms, e.g., (each point is a different instance):





The bright sides of biased sampling

- of more interest would be to ask: what is the bias associated with running two samplers for half the duration (generating $n/2$ samples for each) as compared to running only one algorithm from which we obtain n samples?
- assuming one algorithm generates the GSD $\boldsymbol{p}_1$ and another generates $\boldsymbol{p}_2$,we compare $b(\boldsymbol{p}_1)$ against $b(\frac{\boldsymbol{p}_1+\boldsymbol{p}_2}{2})$.
- the smaller the bias is the better the algorithm.
- this answers the question: is there any merit in using a combination of two algorithms versus using only one. can the bias be leveraged towards generating flatter samples?

The fair in unfair

- some results: the figure on the right shows scatter plots of the bias of SA (horizontal axis) over the various instances vs those generated by SA+TFQA/SA+NSQA.

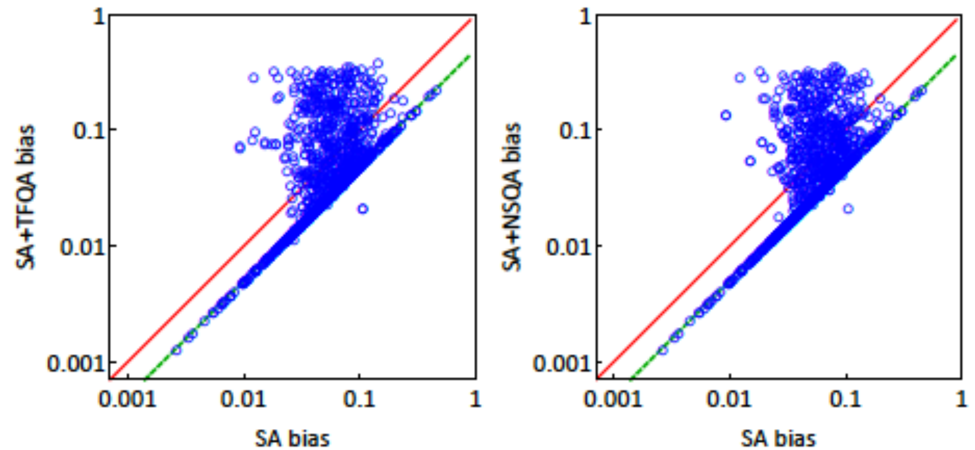


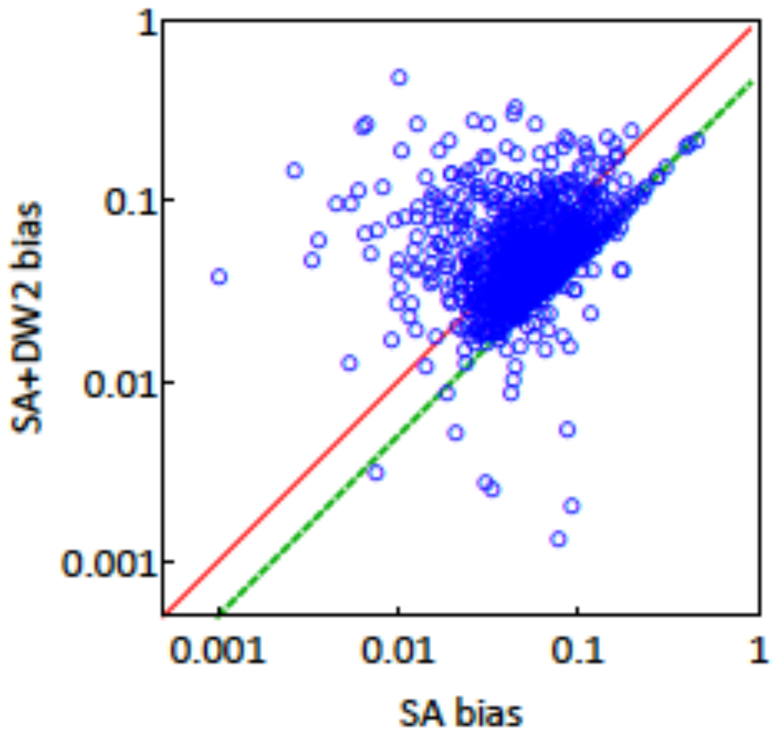
FIG. 5. **Scatter plots of biases of the tested instances' GSDs as obtained by SA vs. a combination of SA with ideal quantum annealers.** Left: SA vs. SA+TFQA. Right: SA vs. SA+NSQA. In the majority of cases, a combination of the methods leads a smaller overall bias, i.e., a lesser degree of unfairness. The solid red line is the 'equal bias' $y = x$ line, whereas the dashed green $y = x/2$ line is the bias obtained when SA is combined with a flat, unbiased GSD, in which case the bias is halved.

- points below the $y = x$ (red) line indicate reduced bias. this happens to be true for most instances.

- also noticeable is the $y = x/2$ (green) line, points along which indicate a flat QA GSD.

What about D-Wave?

- comparing SA to SA+DW2, results are generally mixed.
- for about half of the instances the bias is found to be larger and for the rest smaller.
- since the DW2 GSDs are uncorrelated in general with QA GSDs (nor with SA GSDs), the mechanism generating the distributions is unclear.



- is it quantum in nature? thermal? ICE? answer is important because it tells us whether a classical injection of errors for example is enough to reproduce the D-Wave GSDs.

Conclusions and outlook

Conclusions and outlook

- we find that quantum annealers sample the ground state manifolds of spin glasses differently than classical thermal algorithms. in addition, different drivers produce qualitatively difference distributions.

- unclear where D-Wave distributions come from. how can one find out?

- when complementing classical algorithms, quantum annealers reduce the bias associated with classical GSD sampling. seemingly produce quantum advantages.

- we have not discussed the *speed* with which GSDs are generated. should be discussed as well.

- is this a path to quantum supremacy? could be.

Thank You!

The Fair in
Unfair Quantum Ground-State Sampling

Itay Hen
Information Sciences Institute, USC

Group Meeting
Jan 13, 2017

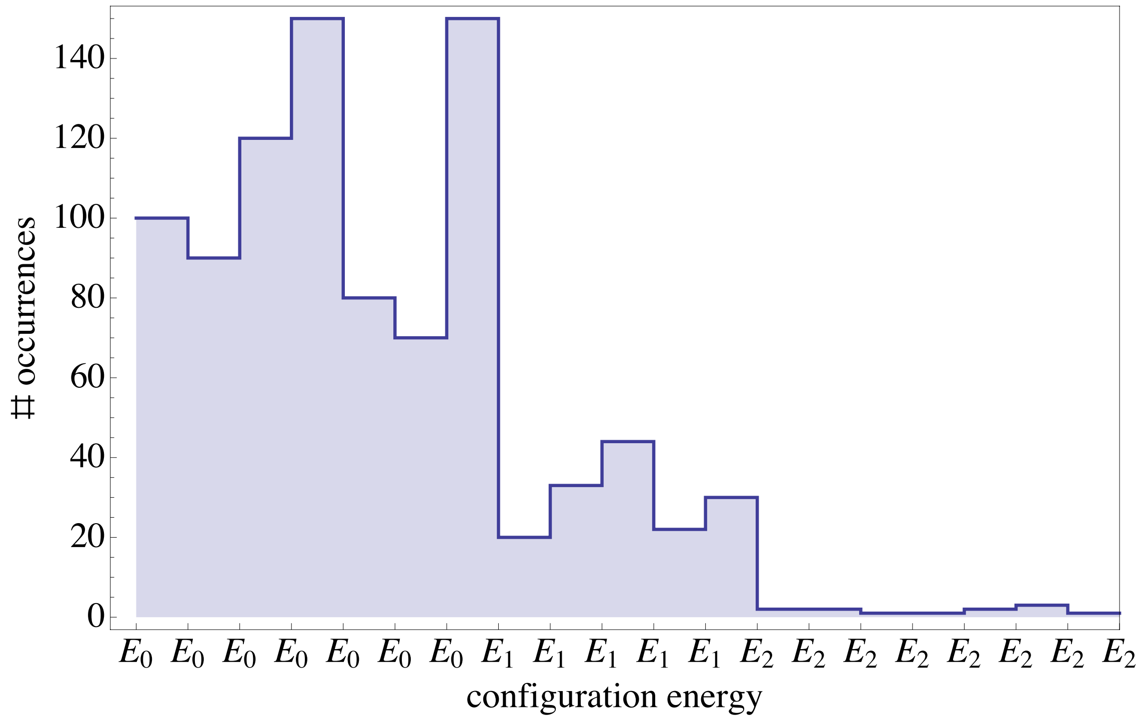
preprint:
arXiv:1701.01524

Information Sciences Institute
USC School of Engineering

Additional slides

Quantum Annealers as Samplers

- quantum annealing optimizers produce more than just hit/miss answers (leading eventually to "success probabilities").
- quantum annealers produce distributions of configurations (some of which are valid optimal solutions, others are not).
- outcomes contain additional valuable information that may not be easily accessible classically.
- this "quantitative difference" may be exploited.



The extra quantum "knob"

-

- A unique device that sample
manifolds of discrete cost func
way is a quantum annealer. Inte
Hamiltonian H_d, which suppli
problem Hamiltonian H_p wh

- the quantum ground state re
quantum anneal is a linear com
below). The probability distribu
assignments is inherently quan
of the quantum fluctuations.

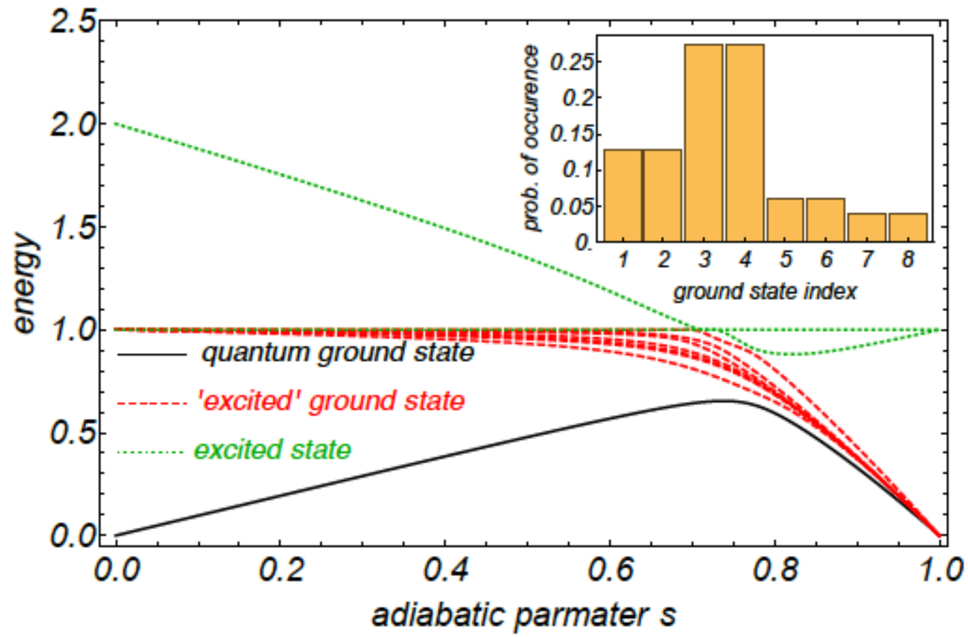


FIG. 1. Ten lowest energy levels of an 8-qubit Hamiltonian interpolating between a transverse field driver Hamiltonian and a randomly generated Ising Hamiltonian. The solid black line indicates the energy of the instantaneous ground-state. The dashed red and dotted green lines indicate excited states that lead to final ground-states and final excited states, respectively. **Inset:** The probabilities for obtaining the various classical ground-states upon measuring the quantum ground-state at the end of the evolution in the computational basis.