Title: Different Strategies for Quantum Adiabatic Optimization, and the Computational Power of Simulated Quantum Annealing

Date: Dec 03, 2014 04:00 PM

URL: http://pirsa.org/14120029

Abstract: Quantum Adiabatic Optimization proposes to solve discrete optimization problems by mapping them onto quantum spin systems in such a way that the optimal solution corresponds to the ground state of the quantum system. The standard method of preparing these ground states is using the adiabatic theorem, which tells us that quantum systems tend to remain in the ground state of a time-dependent Hamiltonian which transforms sufficiently slowly. In this talk I'll show that alternative strategies using non-adiabatic effects can in some cases be dramatically faster for instances which are hard for the traditional adiabatic method.

I will also discuss Simulated Quantum Annealing (SQA), which is a classical simulation of adiabatic optimization at non-zero temperature based on Path-Integral Quantum Monte Carlo. SQA is widely used in practice to study adiabatic optimization, but relatively little is known about the rate of convergence of the markov chain that underlies the algorithm. By focusing on a family of instances which adiabatic optimization solves in polynomial time, but require exponential time to solve using classical (thermal) simulated annealing, I will present numerical evidence as well as a work-in-progress proof that SQA can be exponentially faster than classical simulated annealing.

Different Strategies for Adiabatic Optimization and The Computational Power of Simulated Quantum Annealing

Elizabeth Crosson

December 3rd, 2014

Different Strategies for Optimization Using the Quantum Adiabatic Algorithm
E.C., Edward Farhi, Cedric Yen-Yu Lin, Han-Hsuan Lin, Peter Shor. **arXiv:1401.7320**

Tunneling Through High Energy Barriers in Simulated Quantum Annealing
E.C., Mingkai Deng. **arXiv:1410.8484**

Quantum Adiabatic Optimization (QAO) solves difficult combinatorial optimization problems by encoding them into quantum spin systems!

- Combinatorial optimization problems involve minimizing (or maximizing) a function f on a discrete domain like the set of n-bit strings $\{0,1\}^n$.
- Many important problems in computer science are of this form e.g. boolean satisfaction, the traveling salesman problem, graph partitioning, etc.
- Optimization algorithms are considered efficient if they minimize the function in $\mathcal{O}(\text{poly}(n))$ time, and inefficient if they require exponential time.
- Many optimization problems are NP-Complete, so we strongly suspect that (in the worst case) they are computationally hard for both classical and quantum computers.

Boolean satisfaction (SAT) is the problem of finding bit strings that satisfy logical formulas built out of AND, OR, and NOT.

- 2-SAT example: (b_1 OR b_2) AND (NOT b_2 OR b_3)
- A clause like b_1 OR b_2 is satisfied as long as b_1 and b_2 aren't both false.
- If we let the $+1$ eigenstate of σ^z stand for true, and -1 eigenstate stand for false, then the formula is satisfied only if $(1-\sigma_1^z)(1-\sigma_2^z)+(1+\sigma_2^z)(1-\sigma_3^z) = 0$.
- By turning the logical formula into an equation, we've re-cast the problem of finding a satisfying assignment into the problem of finding the ground state of a 2-local Hamiltonian!

Overview of QAO

- Minimize a function $f : \{0,1\}^n \to \mathbb{R}$ by mapping it onto a quantum system:

$$H_P = \sum_{z \in \{0,1\}^n} f(z)|z\rangle\langle z|$$

- If w is the unique string that minimizes f, then $|w\rangle$ is the ground state of H_P.

- Produce the ground state of H_P using a time-dependent transverse field:

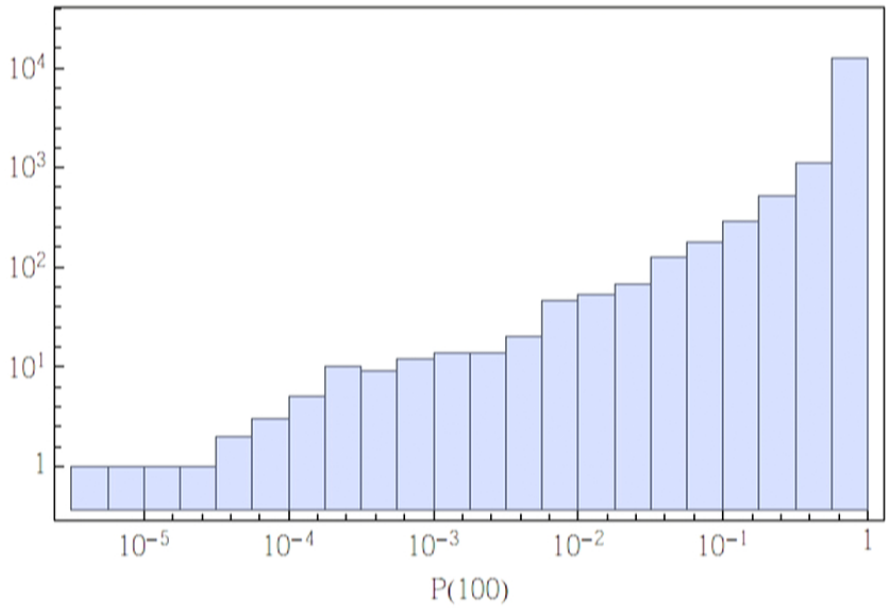
$$H(t) = \left(1 - \frac{t}{T}\right) \sum_{i=1}^{n} \left(\frac{1 - \sigma_i^x}{2}\right) + \frac{t}{T} H_P \quad , \quad 0 \leq t \leq T$$

How slowly do we need to change the QAO Hamiltonian to find the optimal state $|w\rangle$ with high probability?

- **Success Probability:** $P(T) = |\langle w|\psi(t=T)\rangle|^2$.
- **Run-time** primarily depends on $g_{\min}$, the minimum energy gap between the ground state and first excited state.
- **Adiabatic theorem:** $T > \mathcal{O}\left(\frac{\text{poly}(n)}{g_{\min}^2}\right) \Rightarrow P(T) \approx 1$.
- **Hard instances** for QAO are those with exponentially small $g_{\min}$.

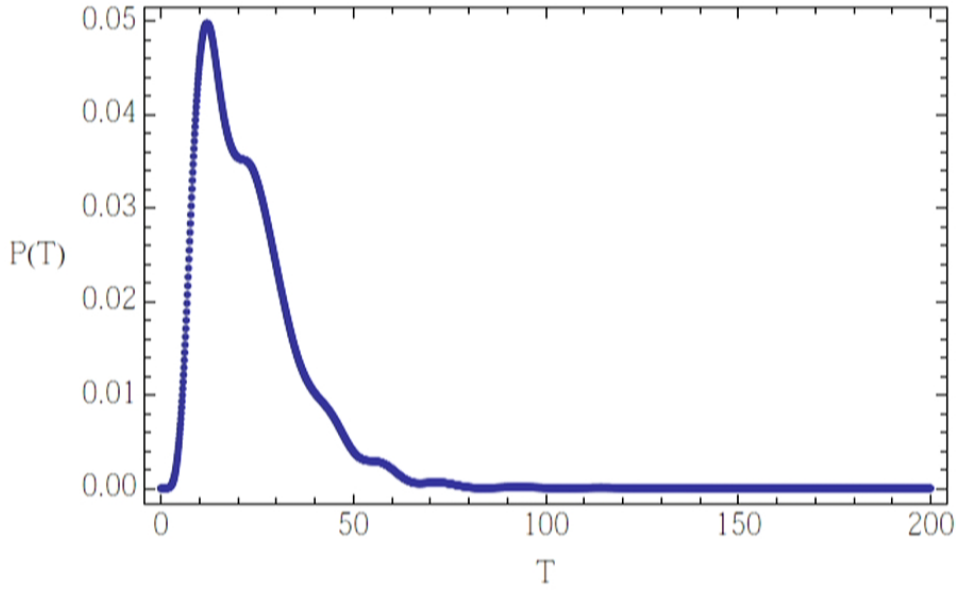
Can we increase $P(T)$ for hard instances, when $T < \mathcal{O}(g_{\min}^{-2})$?

Searching for hard random instances: we generated over 200,000 instances of MAX 2-SAT at $n = 20$ bits to find an ensemble of 137 hard instances.

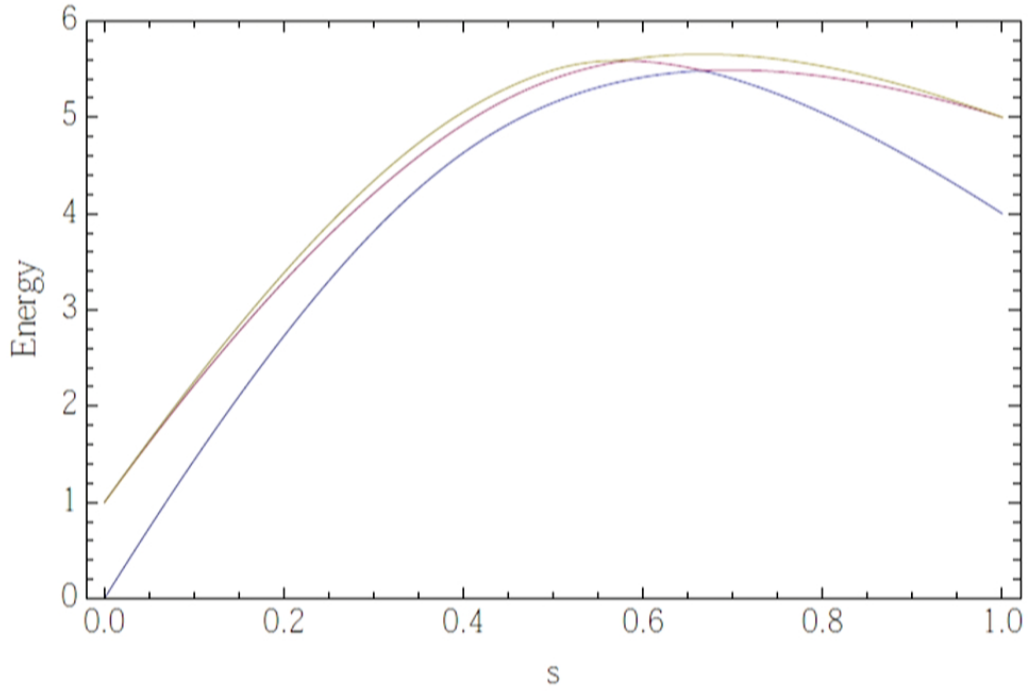


Distribution of QAO success probabilities with $T = 100$.

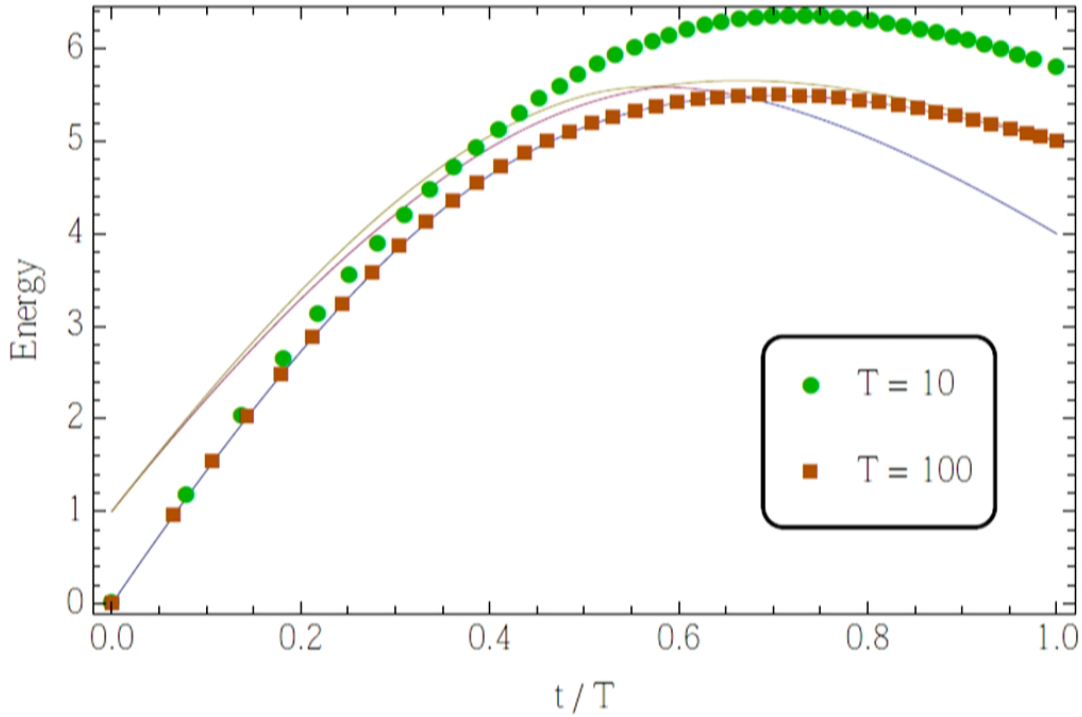
Discovery: hard instances have peaks in $P(T)$ at smaller values of T!



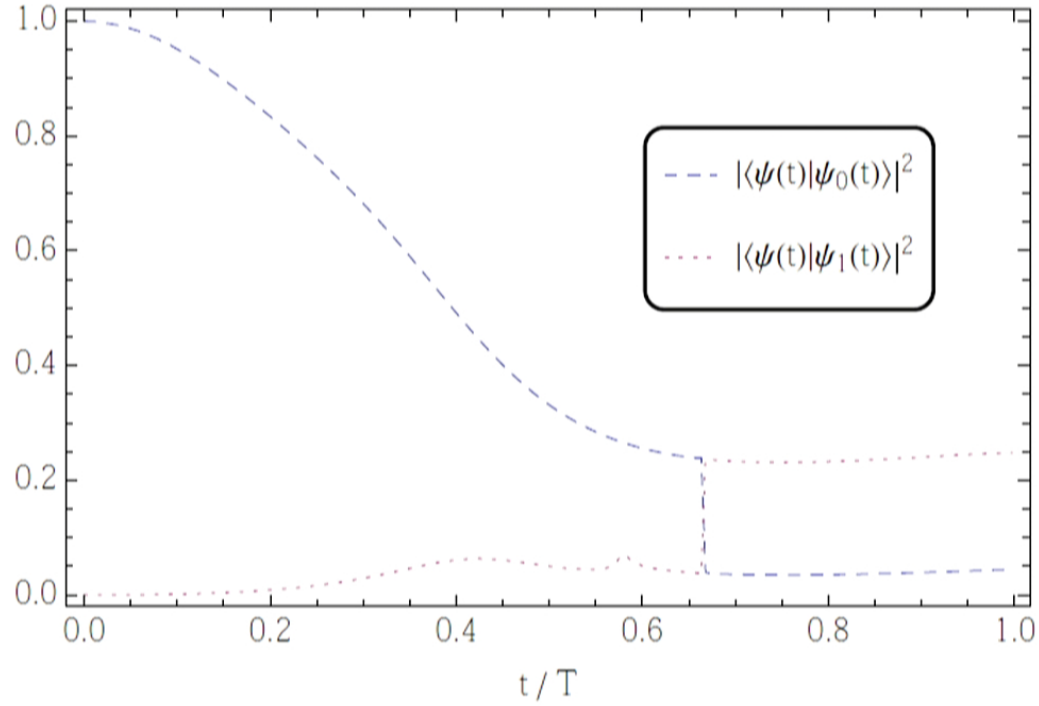
Success probability as a function of total evolution time T for an instance with $P(100) = 5 \times 10^{-5}$.



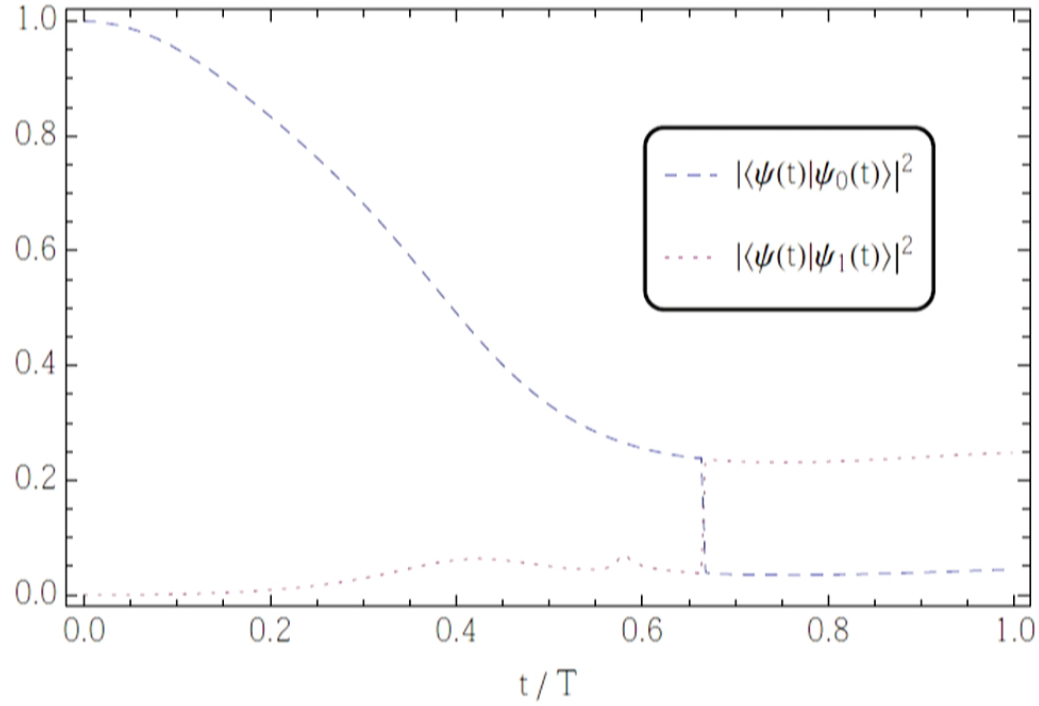
The lowest 3 energy levels for this instance as a function of $s = t/T$ shows an avoided crossing near $s = 0.66$ which gives $g_{min} = 10^{-5}$.



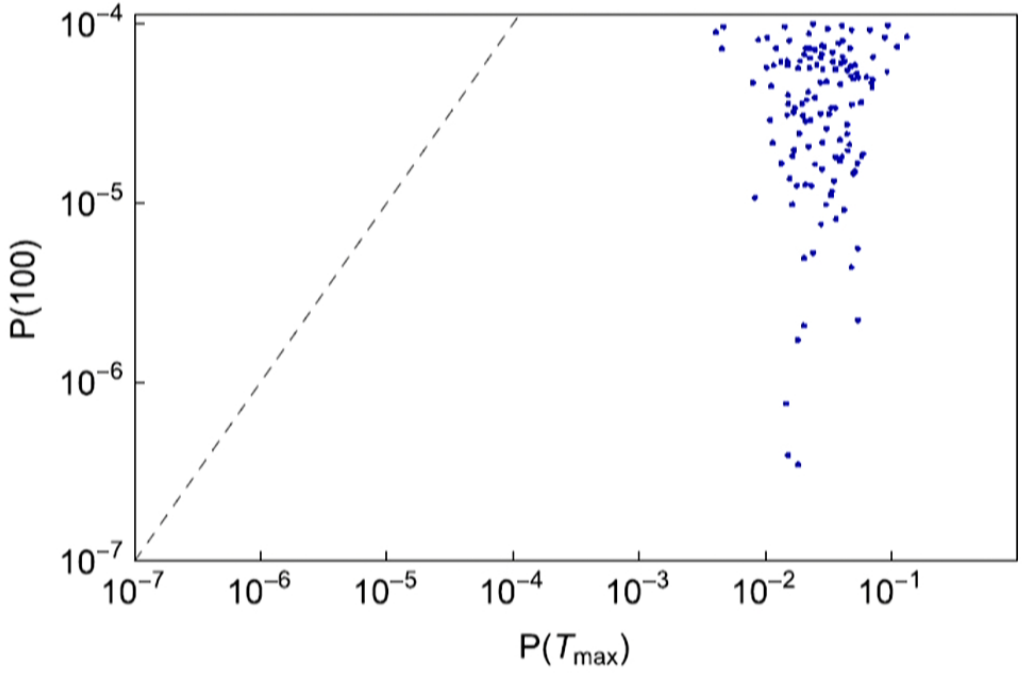
The three lowest energy levels for this instance shown together with the instantaneous expectation of the energy for $T = 10$ and $T = 100$.



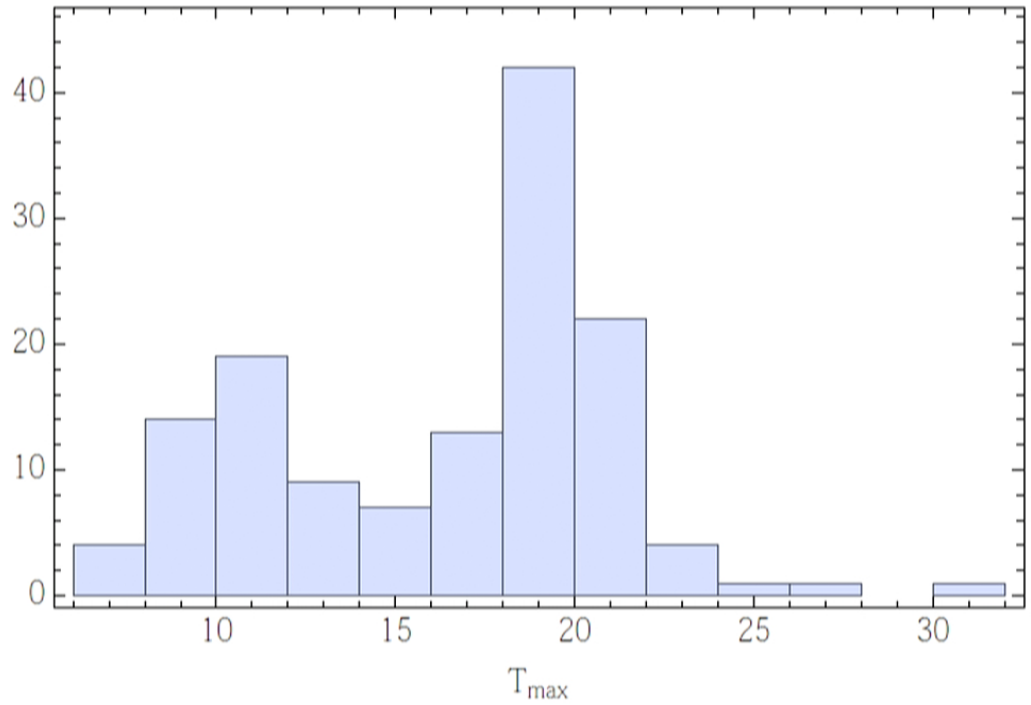
The overlap of the rapidly evolved system ($T = 10$) with the first two instantaneous energy eigenstates of $H(t)$.



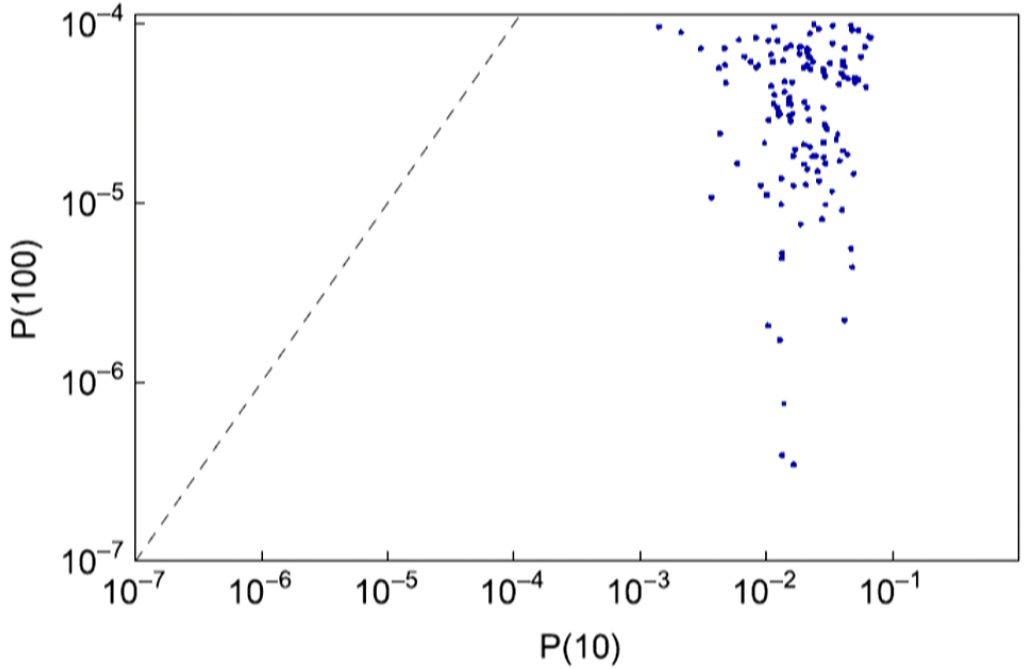
The overlap of the rapidly evolved system ($T = 10$) with the first two instantaneous energy eigenstates of $H(t)$.



Comparing $P(T_{\max})$ with $P(100)$. Note that the value of $T_{\max}$ depends on the instance.



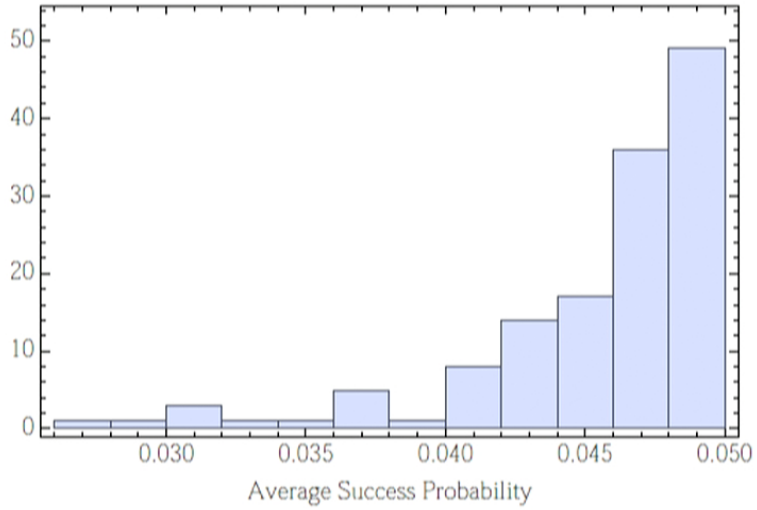
The distribution of the times $T_{\max}$ at which the success probabilities of our hard instances are maximized in the interval $[0, 40]$.



Comparing $P(10)$ with $P(100)$. The success probability is increased for **all** of our hard instances!

Going lower by aiming higher!

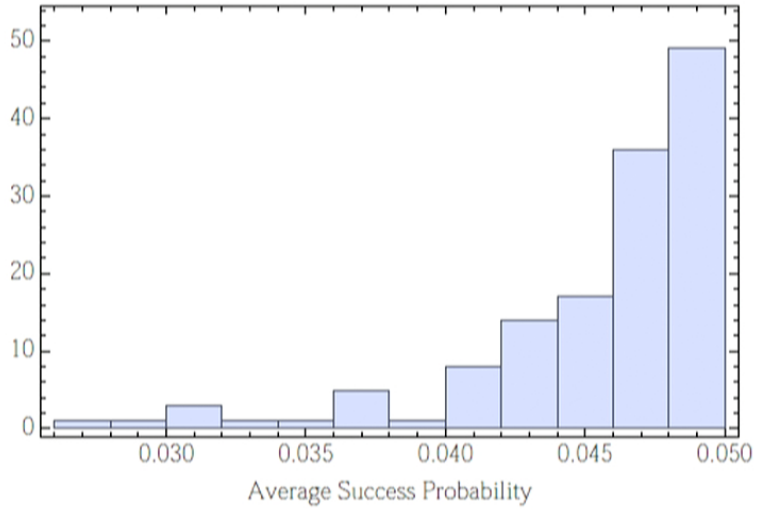
- Can we exploit this effect directly by preparing the system in a random first excited state of the uniform transverse field?



- Average success probability near 1/20, which saturates the upper bound given by conservation of probability!

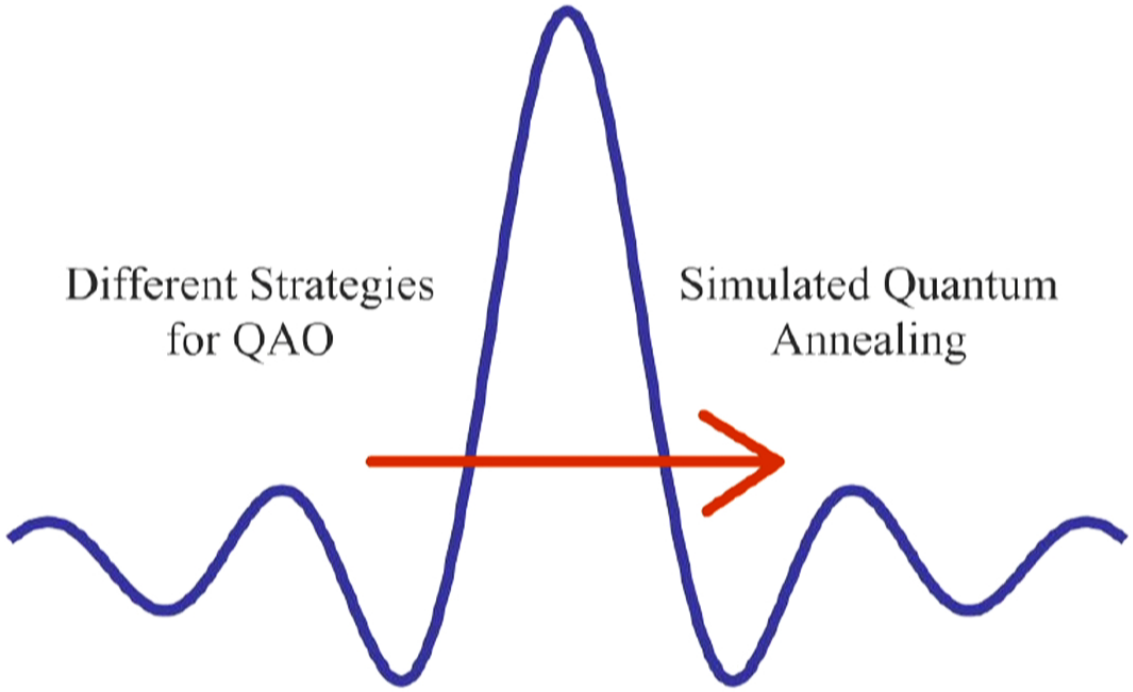
Going lower by aiming higher!

- Can we exploit this effect directly by preparing the system in a random first excited state of the uniform transverse field?



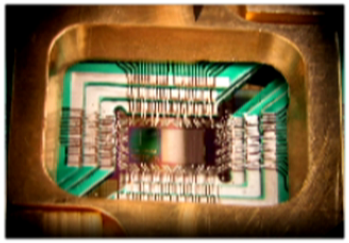
- Average success probability near 1/20, which saturates the upper bound given by conservation of probability!

Tunneling to a different topic!

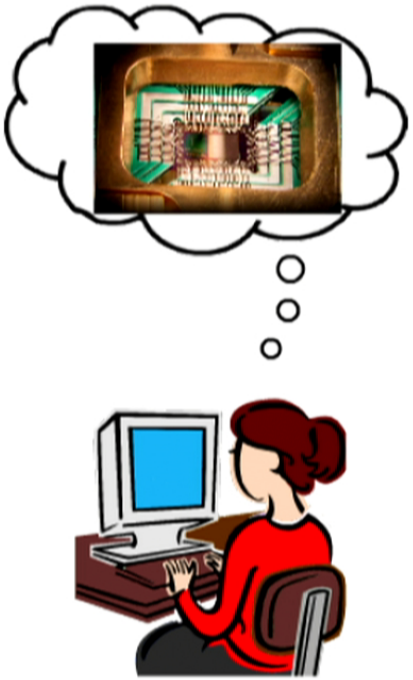


Simulated Annealing

Quantum Annealing



Simulated Quantum Annealing



Path-Integral Quantum Monte Carlo method underlies SQA

- Trotterize the QAO Hamiltonian: $\exp(A+B) \approx \left(\exp\left(\frac{A}{L}\right)\exp\left(\frac{B}{L}\right)\right)^L$
- Quantum-to-Classical mapping with "imaginary-time direction"

$$\rho_{QA} = \frac{e^{-\beta H}}{\mathrm{Tr}\; e^{-\beta H}} \quad \Longleftrightarrow \quad \pi(\mathbf{z}) = \frac{e^{-\beta E_C(\mathbf{z})}}{\sum_{\mathbf{z}\in\Omega} e^{-\beta E_C(\mathbf{z})}}$$

$$\exp(H_z + H_x) \approx \prod_{i=1}^{L} \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)$$

$$= \sum_{z_1,\ldots,z_L} \prod_{i=1}^{L} \langle z_i| \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)|z_{i+1}\rangle$$

Path-Integral Quantum Monte Carlo method underlies SQA

- Trotterize the QAO Hamiltonian: $\exp\left(A+B\right) \approx \left(\exp\left(\frac{A}{L}\right)\exp\left(\frac{B}{L}\right)\right)^{L}$
- Quantum-to-Classical mapping with "imaginary-time direction"

$$\rho_{QA} = \frac{e^{-\beta H}}{\mathrm{Tr}\; e^{-\beta H}} \quad \Longleftrightarrow \quad \pi(\mathbf{z}) = \frac{e^{-\beta E_C(\mathbf{z})}}{\sum_{\mathbf{z}\in\Omega} e^{-\beta E_C(\mathbf{z})}}$$

$$\exp\left(H_z + H_x\right) \approx \prod_{i=1}^{L} \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)$$

$$= \sum_{z_1,\ldots,z_L} \prod_{i=1}^{L} \langle z_i| \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)|z_{i+1}\rangle$$

Path-Integral Quantum Monte Carlo method underlies SQA

- Trotterize the QAO Hamiltonian: $\exp(A+B) \approx \left(\exp\left(\frac{A}{L}\right)\exp\left(\frac{B}{L}\right)\right)^L$
- Quantum-to-Classical mapping with "imaginary-time direction"

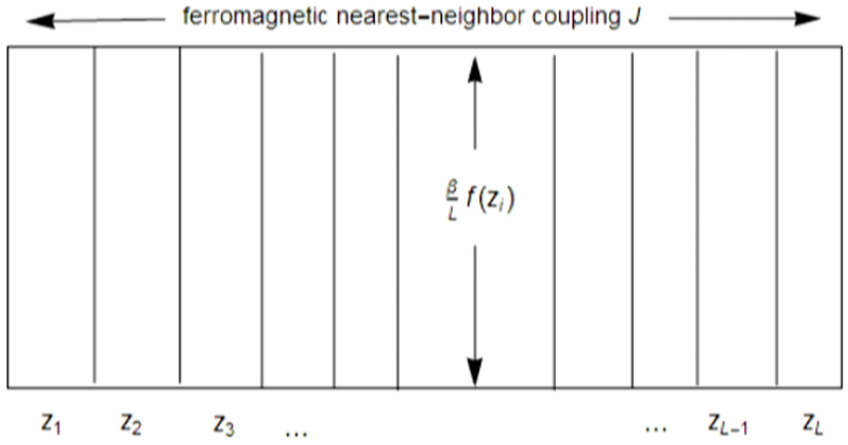
$$\rho_{QA} = \frac{e^{-\beta H}}{\mathrm{Tr}\; e^{-\beta H}} \quad \Longleftrightarrow \quad \pi(\mathbf{z}) = \frac{e^{-\beta E_C(\mathbf{z})}}{\sum_{\mathbf{z}\in\Omega} e^{-\beta E_C(\mathbf{z})}}$$

$$\exp(H_z + H_x) \approx \prod_{i=1}^{L} \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)$$

$$= \sum_{z_1,\ldots,z_L} \prod_{i=1}^{L} \langle z_i| \exp\left(\frac{H_z}{L}\right)\exp\left(\frac{H_x}{L}\right)|z_{i+1}\rangle$$

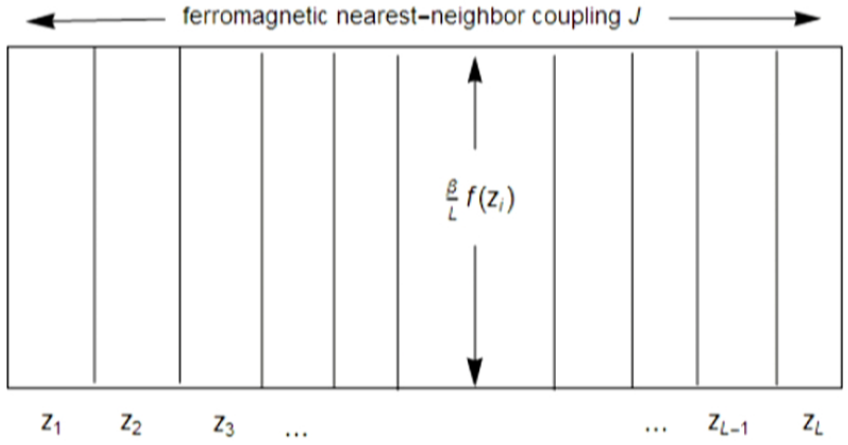
$$\beta E_C(\mathbf{z}) = \beta E_C(z_1, ..., z_L) = \sum_{i=1}^{L} \left(\frac{\beta}{L} f(z_i) + J \sum_{i=1}^{n} z_{i,j} z_{i+1,j} \right)$$

$$J \approx \mathcal{O}\left(\frac{1}{1-s} \right) \text{ as } s \to 1$$



$$\beta E_C(\mathbf{z}) = \beta E_C(z_1, ..., z_L) = \sum_{i=1}^{L} \left(\frac{\beta}{L} f(z_i) + J \sum_{i=1}^{n} z_{i,j} z_{i+1,j} \right)$$

$$J \approx \mathcal{O}\left(\frac{1}{1-s} \right) \text{ as } s \to 1$$



Sample from π using a **Markov Chain!**

- Local moves (bit flips) that satisfy detailed balance:

$$P(\mathbf{z}, \mathbf{z}') = \min\left\{1, \frac{\pi(\mathbf{z}')}{\pi(\mathbf{z})}\right\} = \min\left\{1, e^{E_C(\mathbf{z}') - E_C(\mathbf{z})}\right\}$$

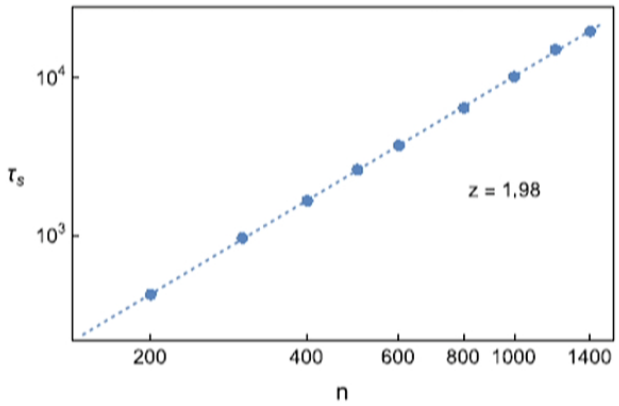
- Run-time of the algorithm $\Longleftrightarrow$ Markov Chain mixing time
- Bounding the mixing time can be a challenging problem!
- Spectral analysis: $\tau = \mathcal{O}\left(\frac{1}{\lambda(P)} \log \frac{1}{\pi_{\min}}\right)$

Hamming Weight with a Spike

$$f(z) = \begin{cases} |z| & : |z| \neq n/4 \\ n & : |z| = n/4 \end{cases}$$

- SA takes exponential time to find the minimum at $z = 00...00$
- QA has $g_{\min} = \mathcal{O}(n^{-1/2})$, and finds the minimum in time $\mathcal{O}(n)$
- How does SQA fare?

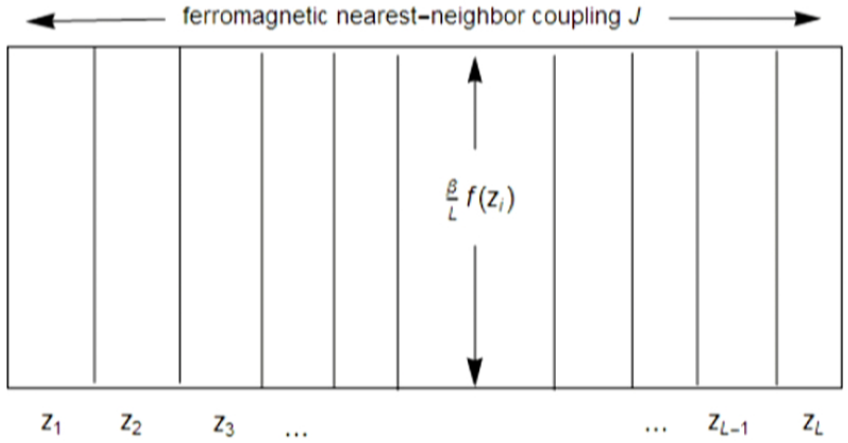
Numerical results: SQA finds the minimum in polynomial time!



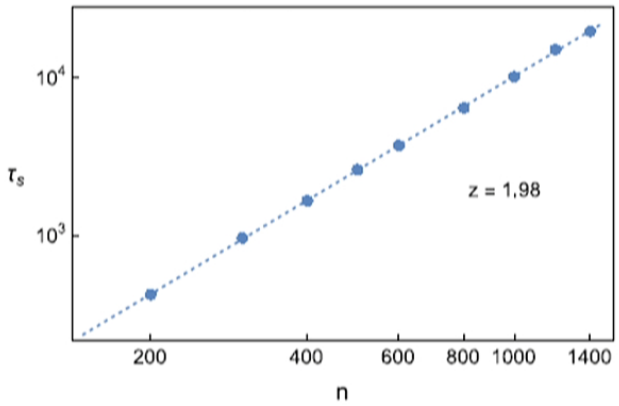
- Number of sweeps needed to mix scales like $\mathcal{O}(n^z)$ with $z \approx 2$.
- Probability of stepping on the spike is $\mathcal{O}(\exp(\frac{-\beta n}{L}))$, therefore $L = O(n)$ trotter slices are required for tunneling, otherwise the mixing time is exponential.
- Critical slowdown as $s \to 1$ can be understood in terms of the zero-temperature phase transition for the classical 1D Ising model.

$$\beta E_C(\mathbf{z}) = \beta E_C(z_1, ..., z_L) = \sum_{i=1}^{L} \left(\frac{\beta}{L} f(z_i) + J \sum_{i=1}^{n} z_{i,j} z_{i+1,j} \right)$$

$$J \approx \mathcal{O}\left(\frac{1}{1-s} \right) \text{ as } s \to 1$$



Numerical results: SQA finds the minimum in polynomial time!



- Number of sweeps needed to mix scales like $\mathcal{O}(n^z)$ with $z \approx 2$.
- Probability of stepping on the spike is $\mathcal{O}(\exp(\frac{-\beta n}{L}))$, therefore $L = O(n)$ trotter slices are required for tunneling, otherwise the mixing time is exponential.
- Critical slowdown as $s \to 1$ can be understood in terms of the zero-temperature phase transition for the classical 1D Ising model.

Proof idea: "time spent on the spike decomposition"

$$b(\mathbf{z}) = \#\text{ of } \ z_i \text{ in } \mathbf{z} \text{ with } |z_i| = n/4$$

$$B_k = \{\mathbf{z} \in \Omega : \frac{kL}{m} \leq b(\mathbf{z}) \leq \frac{(k+1)L}{m}\} \ , \ L \gg k = 1, ..., m \gg n$$

- Decompose P into P_{B_k}, which moves within each B_k, and a block chain $\bar{P}$ that moves between the B_k.
- Show rapid mixing within each block by a Rayleigh quotient comparison to the Hamming weight without the spike.
- The block chain is a birth-and-death chain, need to show that there are no bottlenecks.

Proof idea: "time spent on the spike decomposition"

$$b(\mathbf{z}) = \#\text{ of } \ z_i \text{ in } \mathbf{z} \text{ with } |z_i| = n/4$$

$$B_k = \{\mathbf{z} \in \Omega : \frac{kL}{m} \leq b(\mathbf{z}) \leq \frac{(k+1)L}{m}\} \ , \ L \gg k = 1, ..., m \gg n$$

- Decompose P into P_{B_k}, which moves within each B_k, and a block chain $\bar{P}$ that moves between the B_k.
- Show rapid mixing within each block by a Rayleigh quotient comparison to the Hamming weight without the spike.
- The block chain is a birth-and-death chain, need to show that there are no bottlenecks.

Conclusion

- Non-adiabatic effects can significantly increase the success probabilities for QAO on hard instances of MAX 2-SAT at $n = 20$ bits.
- **Open question:** How well do these strategies work at larger n, where typical random instances are expected to be hard for QAO?
- Simulated Quantum Annealing can inherit some of the advantages of Quantum Annealing, and can be exponentially faster than Simulated Annealing.
- **Open question:** Can Quantum Annealing generally be simulated on a classical computer with polynomial overhead?

Conclusion

- Non-adiabatic effects can significantly increase the success probabilities for QAO on hard instances of MAX 2-SAT at $n = 20$ bits.
- **Open question:** How well do these strategies work at larger n, where typical random instances are expected to be hard for QAO?
- Simulated Quantum Annealing can inherit some of the advantages of Quantum Annealing, and can be exponentially faster than Simulated Annealing.
- **Open question:** Can Quantum Annealing generally be simulated on a classical computer with polynomial overhead?