Counterexamples to Efficient Simulated Quantum Annealing

In principle Quantum Monte Carlo allows us to simulate the Hamiltonians used in Quantum Annealing, so it's natural to seek conditions for which this simulation is efficient.

Today we will discuss some recent examples, due to Hastings, which show an exponential separation in run times between a particular form of quantum annealing and a particular form of Quantum Monte Carlo.

Generalized Quantum Annealing

Last time we considered the standard version of quantum annealing which is based on qubits with arbitrary Z interactions and uniform transverse magnetic field. A more general version of QA could be defined:

$$H(p_1, \ldots, p_m) = H_{\text{kinetic}}(p_1, \ldots, p_m) + H_{\text{potential}}(p_1, \ldots, p_m)$$

where $p_i = p_i(t)$ are parameters which depend on the time t, so that the ground state of H can be efficiently prepared adiabatically and encodes some information of interest as $t \to \infty$. For our purposes H_{kinetic} will consist of hopping terms on graphs (e.g. $|u, v\rangle\langle v, u| + h.c.$ for adjacent vertices u and v), and $H_{\text{potential}}$ will be diagonal in the basis of vertex labels.

In standard quantum annealing the graph is the n-dimensional hypercube, the single-site magnetic fields create a hoping term on each vertex (since $X = |0\rangle\langle 1 + h.c.$) , and $H_{\text{potential}}$ is diagonal in the computational basis (i.e. the basis of vertex labels).

The Laplacian of a Graph G

Form the Laplacian from the adjacency matrix by adding the vertex degrees on the diagonal and negating the off-diagonal terms:

$$L(G) = \deg(G) - \text{adj}(G) = \sum_{v \in V(G)} \deg(v) \; |v\rangle\langle v| \; - \sum_{(u,w) \in E(G)} (\; |u\rangle\langle v| + h.c.\;)$$

The Laplacian contains the same hopping information as adj(G), but the rescaling makes it more suitable as a kinetic term. This should be familiar to physicists because the Laplacian of a discretized line is exactly the discretized second derivative:

$$\partial^2 = \begin{pmatrix} 0 \ldots & -1 & 2 & -1 & 0 \ldots & 0 \ldots \\ 0 \ldots & 0 \ldots & -1 & 2 & -1 & 0 \ldots \end{pmatrix}$$

From a quantum computer scientist's point of view the important property of L is that it's stoquastic: which means the off-diagonal terms are non-positive and the diagonal terms are real.

Stoquastic Hamiltonians

Stoquastic Hamiltonians have simple ground states: there is choice of phase for which every component of the ground state is non-negative. This means the ground state can be completely characterized by a classical probability distribution.

In terms of computational complexity, the version of Kitaev's ground state problem restricted to stoquastic Hamiltonians has a classical interactive proof, in contrast with the problem being QMA-complete for general Hamiltonians.

The importance of stoquasticity for quantum monte carlo is that all the off-diagonal components having the same sign lets us relate them to markov chain transition probabilities.

Path-Integral Quantum Monte Carlo

Recall that we rewrite the quantum partition function $Z = \text{tr}\, e^{-\beta H}$ in terms of L slices of imaginary time:

$$\text{Tr}\ e^{-\beta H} = \text{Tr}\left(e^{-\beta \frac{H}{L}}\right)^L = \sum_{\{\sigma_1, \ldots, \sigma_L\}} \prod_{i=1}^{L} \left\langle \sigma_i \left| \exp\left(-\frac{\beta H}{L}\right) \right| \sigma_{i+1} \right\rangle$$

which suggests a markov chain with states $(\sigma_1, \ \ldots, \ \sigma_L)$ and weights given by

$$\pi(\sigma_1, \ \ldots, \ \sigma_L) = \frac{1}{Z} \prod_{i=1}^{L} \left\langle \sigma_i \left| \exp\left(-\frac{\beta H}{L}\right) \right| \sigma_{i+1} \right\rangle$$

which is why the property that H is stoquastic is necessary for π to be a probability distribution.

PIMC for a Particle in $\mathbb{R}^k$

In the position basis $|\sigma\rangle = |x_1, \ \ldots, \ x_k\rangle$ the Hamiltonian is:

$$H = \frac{1}{m} \nabla^2 + V(x_1, \ \ldots, \ x_k)$$

The states of the QMC markov chain are length L sequences of position vectors $(\sigma_1, \ \ldots, \ \sigma_L)$, and the off-diagonal terms of the Hamiltonian couple σ_i and σ_{i+1} iff $\left|\vec{x}_{i+1} - \vec{x}_i\right| \leq 1$, so the states of the markov chain look like *worldlines through imaginary time*.

The local transition rules have to take us from one connected worldline to another. A 1-dimensional

example of such a transition is illustrated below:

Since the sequence $(\sigma_1, \ldots, \sigma_L)$ is a sample from Tr $e^{-\beta H}$ it is natural to require periodic boundary conditions $\sigma_1 = \sigma_L$ (although there are tricks for getting around this). In this case the worldline is a closed path in imaginary time, and the time evolution of the markov chain looks like a continuous deformation of this path by wiggling i.e. a homotopy.

First Counterexample: The Circle

We begin with a single particle in a bounded 2-dimensional grid. Initially the particle is free, but we gradually turn on the potential $V(r, \theta) = \mu r^2 + g r^4$ (polar coordinates).

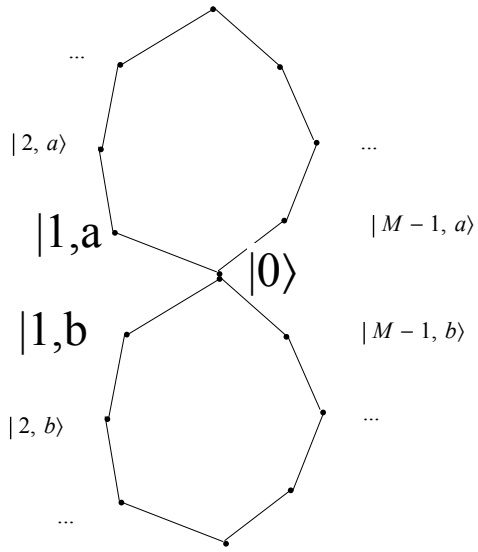
This has the effect of confining the particle to a circle of radius $r_{\text{min}} = \mu/2$, confined by a harmonic potential in the radial direction. Since the particle is free, it will wind around the circle some number of times proportional to $\sqrt{\beta}$ (β is imaginary time, so this follows from the $\sqrt{t}$ spreading in the kernel of the heat equation i.e. the propagator of a free particle).

Already we see that the QMC has trouble equilibrating, as once the potential is sufficiently strong to confine the imaginary time trajectory $(\sigma_1, \ldots, \sigma_L)$ to points on the circle (up to exponentially small corrections), the state of the QMC cannot move from one winding number to another.

If instead we worked with a free particle on S^1 , then the markov chain state space would simply be disconnected. Instead working in $\mathbb{R}^2$ allows for some exponentially small probability of transitioning between winding numbers, so we can make the claim that the markov chain is slowly mixing.

Second Counterexample: Boquet of Circles

This example is very similar to the previous one, but it introduces a geometric space that will be used in subsequence examples:

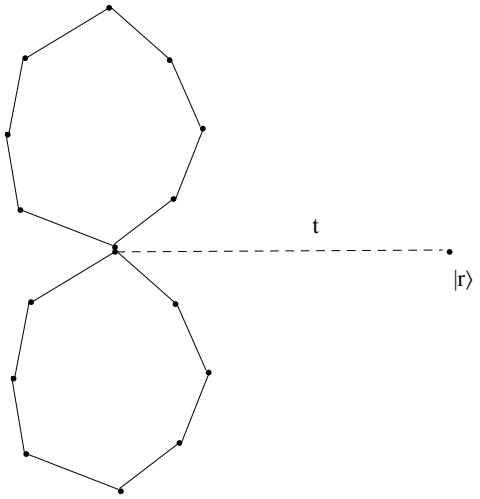


Starting again from a discretized region in $\mathbb{R}^2$, a potential is introduced which confines the particle to the space illustrated above. The homotopy group for this space is the free group on 2 generators $G = \langle a, b : a^M = b^M = 1 \rangle$, so instead of having trouble equilibrating between winding numbers, the closed imaginary time path QMC has trouble equilibrating between topological sectors corresponding to words which are not equivalent.

The punchline is that the typical wordlength is proportional to β (perhaps because the point $|0\rangle$ has four directions to diffuse in, and the heat kernel spreads as $t^{d/2}$ in dimension d), so non-trivial topological sectors should contribute to the partition function.

Third Counterexample: “Too Short a Word”

The purpose of this example is for the QMC to not only mix slowly, but to give the wrong answer as well. As before we have the boquet of circles, and we append a state $|r\rangle$ with the additional Hamiltonian terms $H_r = -t(|0\rangle\langle r| + h.c.) + E|r\rangle\langle r|$.



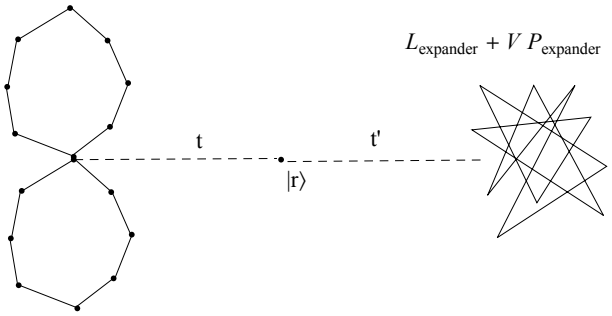
This time we do not start from a grid in $\mathbb{R}^2$, instead the entire space is as illustrated above. We can see outright that the markov chain state space is disconnected (i.e. a closed path can never change its topological sector), but the other gives an interesting but convoluted argument that we will go through anyway.

Initially we make E negative and large in absolute value and set the tunneling $t = 0$, so that the ground state is $|r\rangle$. The ground state energy of the boquet is 0, and the first excited state has energy $c > 0$. We then tune $0 < t < c$ and $0 < E < c$ so that the system experiences an avoided crossing: the energy E of the state $|r\rangle$ crosses zero (the ground state energy of the bouquet), but since $t > 0$ the crossing is avoided and the ground state will be a superposition of $|r\rangle$ and the ground state of the bouquet, with most of the amplitude on the bouquet. Finally, we turn off t and turn on a term $-h\,|0\rangle\langle 0|$ with $h \to \infty$, so that the ground state at the end of the protocol is $|0\rangle$.

To see that the QMC returns the wrong answer, we note that it cannot access not trivial topological sectors, so instead of the bouquet it only sees the universal cover of the bouquet, which is a tree T with 4 branches of length $M - 1$ eminating from the root vertex $|0\rangle$. The point is that the ground state energy on T is $c' > 0$, so if instead of $E < c$ in the previous step we take $E < \min(c, c')$, then the $|r\rangle$ remains the ground state throughout the annealing procedure, as far as the QMC is concerned. Therefore in the end it produces $|r\rangle$ instead of $|0\rangle$.

Fourth Counterexample: “Opening the Path”

Since all the examples so far have been based on non-trivial homotopy, the point of this example is to allow the imaginary time path to have open boundary conditions $\sigma_1 \neq \sigma_L$.



The idea is to attach an expander graph with N vertices labeled $|1, G\rangle, \ldots, |N, G\rangle$ and couple it to the previous example in such a way that the open boundary points have a high probabilty of never leaving the expander, so that the PIMC path on the rest of the space acts like a closed path just as before.

The new Hamiltonian terms are a coupling $-t(|r\rangle\langle 1, G| + h.c.)$, the laplacian L_{expander}, and the potential P_{expander} which is 1 on the expander and 0 elsewhere. We start with $t = t' = 0$ and E large and negative so that the ground state is $|r\rangle$. The ground state of $L_{\text{expander}} + V\, P_{\text{expander}}$ has energy V, and when we make $t' > 0$ the ground state is a superposition of $|r\rangle$ and the ground state of the expander. Taking $t' << 1$ and using perturbation theory puts probability proportional to $|t'|^2 / |E - V|$ on the expander state, and the rest on $|r\rangle$.

Even though most of the probability is on $|r\rangle$, by taking E close to V we can concentrate an arbitrary amount of amplitude on the expander. Once this is done we keep t' and $|E - V|$ fixed and run the protocol from the previous example. The point is that the boundaries of the PIMC state σ_1 and σ_L will approximately follow the distribution:

$$P(\sigma_1, \sigma_L) \approx \psi_0(\sigma_1)\, \psi_0(\sigma_L)$$

where ψ_0 is the ground state, so by taking the size of the expander to be large the probability that the boundary points will ever leave can be made exponentially small. It's not clear whether we need an expander graph for this application, or whether something simpler like a line would suffice.