When is Quantum Monte Carlo comparable to Quantum Adiabatic Optimization?

Adiabatic Optimization

Seeks the solution to a classical optimization problem: minimizing a function $f(z) \in \mathbb{R}$ over all n-bit strings $z \in \{0, 1\}^n$.

Samples the solution from a prepared quantum state $\sum_{z \in \{0,1\}^n} f(z) \quad |z\rangle \langle z|$ as the ground state of some H_f

Prepares the ground state by starting in the ground state of $H_0 = -\sum_{i=1}^{n} X_i$ and adiabatically dragging to H_f :

$$H(s) = (1 - s) H_0 + s H_f \quad , \quad s \text{ goes from } 0 \text{ to } 1$$

Works if $|ds/dt| < 1/\Delta(s)^2$, where $\Delta(s)$ is the energy gap of $H(s)$.

We sometimes think physically in terms of a magnetic field Γ which starts very large and anneals to zero $H = H_f + \Gamma H_0$.

Markov Chain Monte Carlo

A markov chain is a random walk on a state space graph Ω, with stationary probabilities $\pi(v)$ for each vertex $v \in \Omega$, and transition probabilities $P(i, j)$ between adjacent vertices.

The random walk converges to the stationary distribution if $\pi(i) P(i, j) = \pi(j) P(j, i)$. (reversibility, detailed balance)

The rate of convergence is determined by the conductance $C = \min_{S: |S| < \Omega/2} \left| P\left(S, \overline{S}\right) \right| / \pi(S)$

Cheeger's inequality relates the conductance the eigenvalue gap of P. The markov chain converges in polynomial time iff the eigenvalue gap is lower bounded by some $1/\text{poly}(n)$.

Path-Integral Quantum Monte Carlo

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

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

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

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

The transition probabilities are given by the metropolis rule $P(v, w) = \min[1, \ \pi(w)/\pi(v)]$.

PIMC, Continued

In the case of adiabatic optimization the Hamiltonian is a sum of diagonal and off-diagonal terms $H = H_O + H_D$.

Since L is large we have $\left\langle \sigma_i \left| \exp\left(-\frac{\beta H}{L}\right) \right| \sigma_{i+1} \right\rangle = e^{-\beta \frac{E(\sigma_i)}{L}} \left\langle \sigma_i \left| \exp\left(-\frac{\beta H_O}{L}\right) \right| \sigma_{i+1} \right\rangle$.

When $H_O = -\sum X_i$, the off-diagonal term can be split as a product over individual spins $\sigma = (S_1, \ \ldots, \ S_n)$

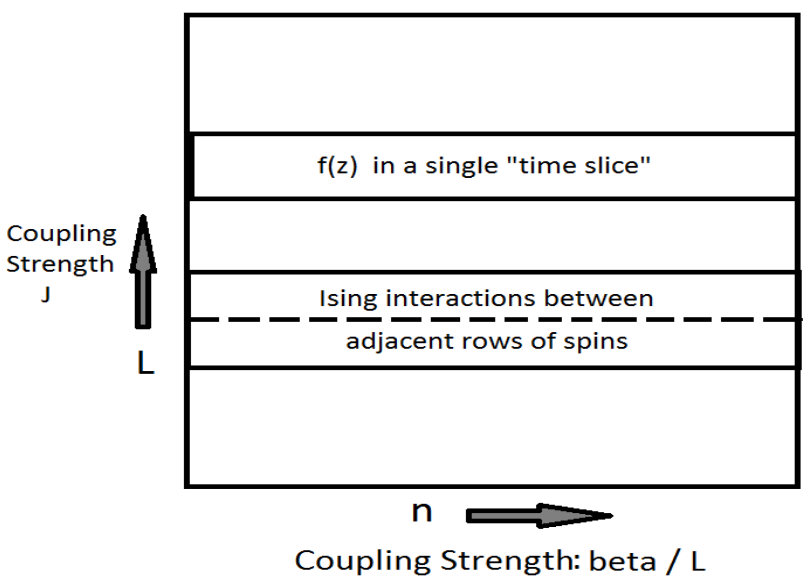
$$\langle \sigma \,|\, \exp(a\, H_O) \,|\, \sigma' \rangle = \prod_{i=1}^{n} \left\langle S_i \,\middle|\, e^{a X_i} \,\middle|\, S_i' \right\rangle = \text{const} * \prod_{i=1}^{n} \exp\left(-J\, S_i\, S_i'\right) \ , \ J = \frac{1}{2} \log(\coth(a))$$

All together the partition function can be rewritten in terms of an effective classical Hamitlonian:

$$H_{\text{eff}} = \sum_{i=1}^{L} \sum_{j=1}^{n} \exp\left(-\frac{\beta}{L} f(S_{i1}, \ \ldots, \ S_{in}) + J\, S_{ij}\, S_{i+1\,j}\right) \ , \ J = \frac{1}{2} \log\left(\coth\left(\frac{\beta\, \Gamma}{L}\right)\right)$$

Effective Ising Model

$$H_{\text{eff}} = \sum_{i=1}^{L} \sum_{j=1}^{n} \exp\left(-\frac{\beta}{L} f(S_{i1}, \ \ldots, \ S_{in}) + J\, S_{ij}\, S_{i+1\,j}\right) \ , \ J = \frac{1}{2} \log\left(\coth\left(\frac{\beta\, \Gamma}{L}\right)\right)$$



Specific Problem Instances

Find problems with a large adiabatic gap for which QMC has provably fast or slow mixing.

The “Hamming weight with a spike” and the “bush of implications” are two examples which have a large adiabatic gap and provably slow convergence for simulating annealing. The behavior of QMC on these is still unknown.

Hastings has recently given multiple topology-based examples with a large adiabatic gap and slow mixing for PIMC, but there are still many open questions.

Hamming weight with a spike

The objective function $f(z)$ is equal to the Hamming weight $h(z)$, except at z^* such that $h(z^*) = n/4$ where we set $f(z^*) = n^\alpha$.

Simulated annealing fails to find the true minimum $f(z) = 0$ because at high temperatures the state is likely to have Hamming weight $n/2$, and as the temperature decreases the system is unable to cross the spike because of the exponentially small transition probability $e^{-\beta n^\alpha}$.

Adiabatic optimization can be shown to have a large gap, essentially because the spread in the ground state wave function allows it to tunnel through the barrier.

A single time slice of PIMC can step onto the spike with probability $\exp\left(-\frac{\beta}{L} n^\alpha\right)$, which is not exponentially small. However, the smallness of β/L means that entropy dominates the configuration of the rows when Γ is large (and hence the coupling along the columns is low). It remains to be seen whether some combination of schedules for β and Γ can allow PIMC to step through the spike and find the true ground state.

Bush of Implications

The objective function is $f(z) = (1 - z_0) - \sum_{i=1}^{n} z_0(1 - z_i)$. This can be thought of as z_0 being a single bit, which prefers to be “true”, which is connected to the other bits by imply clauses.

The true ground state has all the bits set to 1 (“true”), but simulated annealing gets stuck because at high temperatures half of the conclusion bits are set to “false”, and so the premise bit is strongly biased to be false. As the temperature is lowered the conclusion bits feel no influence to become “true”, and so a substantial proportion of them are “false” at all temperatures and the premise bit has an exponentially low probability to be set to “true.”

As in the previous example, a single time slice of PIMC can set the premise bit to true with probability $\exp\left(-\frac{\beta}{L} O(n)\right)$, which is not exponentially small, so there is some hope that PIMC can tunnel through the obstruction and converge efficiently.

Hasting's Circle

In this example the states are position vectors marking a single particle in $\mathbb{R}^2$. This means we should think of PIMC states as connected paths in a (bounded region of) 2 + 1 (discretized) euclidean dimensions.

Initially the particle is free, and we gradually turn on a potential which confines it to a circle of radius R centered around the origin. The particle is free within the circle, so it will wind around a number of times proportional to $\sqrt{\beta}$ (from random walk diffusion).

After a certain amount of time the potential in the circle is raised so that the particle is confined by a harmonic potential to the point where the circle intersects the x-axis. Hastings argues that the PIMC can no longer transition between states of different winding numbers, which prevents the markov chain from equilibrating.

Transverse Ising Gadgets

To represent positions in a discretized euclidean dimension using an Ising-like model on the hypercube we apply an effective Hamiltonian:

$$H_{\text{gadget}} = J \sum_{i,j} Z_i Z_j + h \sum_i Z_i + J' \sum_{|i-j|<R-1} Z_i Z_j - B \sum_{i=1}^{n} X_i$$

By tuning J and h appropriately, the first two terms determine a degenerate ground space where R spins are equal to 1 and the other $N - R$ spins are 0. Then by tuning J' we restrict the ground states to strings of the form 00...011...100...0 where there is exactly one length R sequence of spins equal to 1, and all other spins are equal to 0. Finally, by taking $|B|$ to be a small perturbation we create a hopping between adjacent states of the ground space. To introduce a potential we need to use localized magnetic field $h_i Z_i$.

Further Work

Determine the mixing properties of QMC on the Hamming weight with a spike, and the bush of implications. These are the only two known examples of an exponential seperation between adiabatic optimization and simulated annealing.

Determine whether PIMC fails to converge on the hypercube version of Hasting's circle example, and whether this can be resolved by non-local update rules which are frequently used in practical PIMC implementations.

Seek more examples of problem instances for which PIMC is slow to converge which are (1) naturally on the hypercube without gadgets and (2) resemble realistic optimization problems which may or may not be classically difficult.

Find more complete characterizations and understanding of QMC mixing behavior.