Simulated Quantum Annealing

- Optimization
- Markov Chain Monte Carlo
- Simulated Thermal Annealing
- Quantum Annealing

Optimization

Our goal is to minimize a function $f : \{0, 1\}^n \to \mathbb{R}$ that maps n-bit strings to real numbers.

A string which attains the minimum value of f is called an exact solution. A string σ for which $| f(\sigma) - f_{\min} | = \epsilon$ is an approximate solution with error ϵ.

This class of optimization problems includes NP-hard problems such as SAT: given a formula φ on n variables, let f_φ map strings (considered as truth assignments) to the number of unsatisfied clauses in φ, then the minimum of f_φ is zero iff φ has a satisfying assignment.

Monte Carlo Sampling

Monte Carlo methods solve optimization problems by sampling from a probability distribution π which assigns more weight to strings that have low values of f.

Since the domain of f is exponentially large, strings with lower values of f need to have exponentially more weight:

$$f(\sigma_1) < f(\sigma_2) \;\Rightarrow\; \pi(\sigma_1)/\pi(\sigma_2) > c^{f(\sigma_2)-f(\sigma_1)} \quad , \quad \text{for some constant } c \in \mathbb{R}$$

This condition will be satisfied if we restrict f to be non-negative and sample from a distribution of the following form:

$$\pi(\sigma) = \frac{c^{-f(\sigma)}}{Z} \quad , \quad Z = \sum_{\sigma \in \{0,1\}^n} c^{-f(\sigma)}$$

How can we sample this distribution? Rejection sampling will take too long, since the acceptance probability is exponentially small. One solution is to sample from a sequence of distributions $\pi_0, \pi_1, \ldots$ that converges to π, and this is called Markov Chain Monte Carlo (MCMC).

Markov, Metropolis, Glauber

Starting from an initial string σ_0, a markov chain produces a sequence of strings $\sigma_0, \sigma_1, \ldots$ by a random walk on the state space $\{\sigma\}$. The random walk is defined by a stochastic matrix M, where the matrix element $M(\sigma', \sigma)$ is the probability of sending σ to σ'. If we choose M correctly, these σ_i will be samples from a sequence of distributions π_i that is approaching π.

The main condition for $\{\pi_i\} \to \pi$ is for M to satisfy detailed balance: $M(\sigma', \sigma)\, \pi(\sigma') = M(\sigma, \sigma')\, \pi(\sigma)$.

The metropolis rule is a standard way of constructing a transition matrix that will automatically satisfy detailed balance:

$$M(\sigma', \sigma) = \text{Min}[1, \pi(\sigma')/\pi(\sigma)]$$

In general σ_i and σ_{i+1} could be quite far apart, in practice most markov chains use local moves (such as flipping a single bit). This is called "Glauber dynamics."

Mixing Time

To get performance guarantees for algorithms based on MCMC it is necessary to bound the mixing time. There are many closely related definitions:

$$t_{\text{mix}} = \max_{\pi_0} \min_{t} \left| M^t \pi_0 - \pi \right| < 1/4$$

The mixing time is related to the second eigenvalue λ of M as:

$$t_{\text{mix}} = \frac{1}{\lambda} \log\left(\frac{1}{\pi_{\text{min}}}\right) \quad , \quad \pi_{\text{min}} = \min_{\sigma} \pi(\sigma)$$

Minimizing Energy

Some physical systems solve optimization problems naturally by finding configurations that minimize the energy:

Ex. A hanging chain takes the shape of a hyperbolic cosine, because of all possible curves it is the one which minimizes the gravitational potential energy.

In physics language, the strings $\sigma \in \{0, 1\}^n$ are the *states* of the system, and the goal is to find states which minimize the energy $E : \{0, 1\}^n \to \mathbb{R}$.

Thermal Equilibration

In classical physics, any system in equilibrium with an environment at temperature T is described by the Boltzmann distribution:

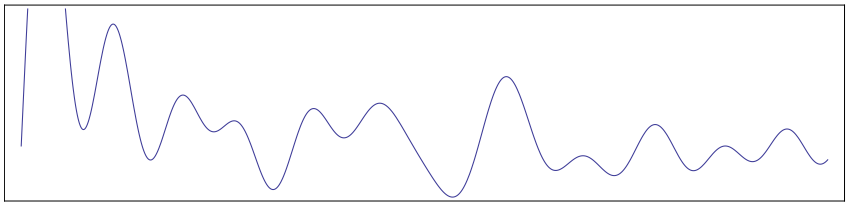
$$\pi(\sigma) = \frac{e^{-E(\sigma)/T}}{Z} \quad , \quad Z = \sum_{\{\sigma\}} e^{-E(\sigma)/T}$$

This has the same form as the MCMC distribution we considered earlier!

The glauber dynamics of single bit flips can also be thought of as a model for the action of the environment on the system. We propose a bit flip, and if it lowers the energy we accept it, and if raises the energy by Δ then we accept it with probability $e^{-\Delta}$.

Thermal Annealing and Metastable States

In metallurgy, annealing refers to heating a metal and then slowly cooling it, which makes it stronger and improves its properties. It works because metals only equilibrate very slowly at room temperature, they get stuck in metastable states which contain defects and internal stresses, whereas they equilibrate much faster at high temperatures when they become red-hot. The idea is that the metal will come to equilibrium at high temperature, and then if it is cooled down slowly enough it will stay in equilibrium throughout the process until it reaches room temperature.



For MCMC this corresponds to rapid mixing at high temperatures, and torpid mixing at low temperatures, which suggests that we run a MCMC which alternately raises and lowers the temperature to escape from local minima.

Quantum Optimization

To translate this problem to the quantum domain, our strings σ become computational basis states $|\sigma\rangle$ on n-qubits, and our energy function becomes the diagonal Hamiltonian:

$$H = \sum_{\sigma \in \{0,1\}^n} E(\sigma) \; |\sigma\rangle\langle\sigma|$$

Since the Hamiltonian is completely classical so far, the non-zero temperature density matrix reproduces the Boltzmann distribution:

$$\langle \rho(T) \,|\, \sigma \rangle = \frac{e^{-H/T}}{\mathrm{Tr}\,(e^{-H/T})} \; |\sigma\rangle = \frac{e^{-E(\sigma)/T}}{Z}$$

Quantum annealing proposes that instead of thermal fluctuations, we turn on off-diagonal quantum terms:

$$H_{\mathrm{QA}} \equiv H + \Gamma V \quad , \quad V = -\sum X_i$$

Just as with thermal annealing, we start with large fluctuations where equilibration happens quickly, then decrease Γ slowly enough to remain in the ground state of H_{QA} until it tends towards the ground state of H.

Adiabatic Theorem

The adiabatic theorem guarantees that the system will remain in the ground state of H_{QA} as long as Γ decreases sufficiently slowly: we want $|d\Gamma/dt| < O(1/\Delta)$ where Δ is the spectral gap of H_{QA}.

As long as the spectral gap is larger than $O(1/\mathrm{poly}(n))$ then the optimization can be performed efficiently on a quantum computer.

Just as we simulate thermal annealing on a classical computer, we can also try to simulate quantum annealing.

Quantum Monte Carlo

The MCMC we applied to sampling the Boltzmann distribution can also be adapted for sampling quantum ground states. The idea is to simulate the Shrodinger time evolution using Trotterization:

$$Z = \mathrm{Tr}\ e^{-\beta H_{\mathrm{QA}}} = \mathrm{Tr}\ e^{-\beta (H_0 + \Gamma V)}$$

$$= \lim_{n \to \infty} \mathrm{Tr} \left[e^{-\frac{\beta}{N} H_0}\, e^{\frac{-\beta}{N} \Gamma V} \right]^N$$

Suzuki recognized the form $Z = \mathrm{Tr}\, T^n$ where T is a transfer matrix for a classical partition function in 1 extra dimension:

$$H\mathrm{eff}[\{\sigma\}] = \sum_{i=1}^{n} \sum_{j=1}^{N} \left\{ f(\sigma_i) - \tilde{\Gamma}\ \sigma_{i,j}\, \sigma_{i+1,j} \right\}$$

$$\tilde{\Gamma} = \frac{1}{2\beta} \log\left(\coth\left(\frac{\beta \Gamma}{M} \right) \right)$$