Quantum Fourier transform beyond Shor's algorithm

András Gilyén

Alfréd Rényi Institute of Mathematics
Budapest, Hungary

Quantum Computing Graduate Summer School, Park City Mathematics Institute, 2023 July 21

Day 5 – Operator Fourier Transform & Metropolis / Gibbs Sampling

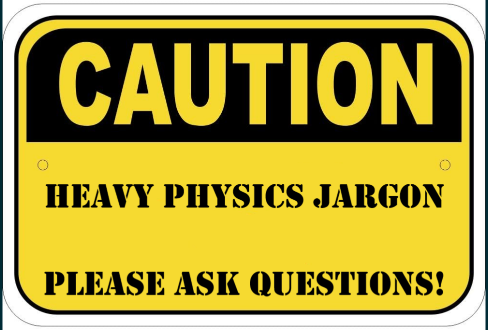
The Gibbs State

The Gibbs Distribution (classical)

- Important in (statistical) physics, describes distribution of states at temperature $T = 1/\beta$.
- Given an "energy" function $E\colon [d] \to \mathbb{R}$, the Gibbs distribution is $\propto \sum_{i=1}^{d} e^{-\beta E(i)}$.

Quantum Gibbs State – corresponding to system Hamiltonian H

$$\propto \sum_{i=1}^{d} e^{-\beta H}$$



Classical (discrete) Metropolis-Hastings algorithm

The Objective

- We want to sample from a target distribution $\propto \tau \in \mathbb{R}_+^N$
  Think about Gibbs sampling of an n-spin Ising model $z \in \{-1, +1\}^n$:

$$H(z) = -\sum_{i,j} \alpha_{ij} z_i z_j - \sum_j \mu_j z_j, \qquad \tau_z = \exp(-\beta H(z)), \qquad N = 2^n$$

The Algorithm

- Suppose we have some symmetric "exploratory" Markov chain $P \in \mathbb{R}_+^{N\times N}$
  For example: pick a random spin and flip it
- Metropolis-Hastings algorithm: from z make a transition to z' according to P
  - If $\tau_{z'} \geq \tau_z$ accept the move
  - If $\tau_{z'} < \tau_z$ reject the move with probability $1 - \frac{\tau_{z'}}{\tau_z}$

Why Does it Work?

- This modified Markov chain $P^{(\tau)}$ has nice properties:
  - The stationary distribution is $\propto \tau$ (+we don't need to know the normalization!)
  - $P^{(\tau)}$ is detailed balanced with respect to τ (a.k.a. reversible)
  - In some sense $P^{(\tau)}$ is the closest such Markov chain to P (Billera and Diaconis'01)
  - Often converges rapidly in physically motivated examples

Continuous-time variant of Metropolis-Hastings

Continuous-time Markov Chains

- We have a continuous-time Markov chain $\exp(t\,L)$ with symmetric generator L
  - The off-diagonal entries of L are the (non-negative) jump rates
  - The diagonal entry is minus the sum of the off-diagonal elements in the column
  - I.e., L is the Laplacian of a weighted directed graph

Continuous-time Metropolis-Hastings

- We modify the jump rates similarly
  - If $\tau_j \geq \tau_i$ then $L^{(\tau)}_{ji} := L_{ji}$, i.e., accept the move
  - If $\tau_j < \tau_i$ then $L^{(\tau)}_{ji} := \frac{\tau_j}{\tau_i} L_{ji}$, i.e., reject the move with probability $1 - \frac{\tau_j}{\tau_i}$

Properties of the Metropolis-Hastings Generator

- This modified generator $L^{(\tau)}$ has nice properties:
  - The stationary distribution is $\propto \tau$ (+we don't need to know the normalization!)
  - $L^{(\tau)}$ is detailed balanced with respect to τ (a.k.a. reversible)
  - In some sense $L^{(\tau)}$ is the closest such generator to L (Diaconis and Miclo'09)
  - Often converges rapidly in physically motivated examples

Quantum Metropolis sampling?

The Objective

What if the objective function is a (non-commuting) quantum Hamiltonian?
For example transverse-field Ising model:

$$H = -\sum_{i,j} \alpha_{ij} Z_i \cdot Z_j - \sum_j \mu_j X_j, \qquad \tau = \exp(-\beta H), \qquad N = 2^n$$

The Discrete-time Algorithm

Suppose we have some symmetric “exploratory” quantum process (channel) Q
For example: pick a random spin and flip it (apply X_j for random $j \in [n]$)
Quantum Metropolis! (Temme, Osborne, Vollbrecht, Poulin, Verstraete Nature’11)

- If $E_{\psi'} \leq E_\psi$ accept the move (where $H = \sum_\psi E_\psi |\psi\rangle\langle\psi|$)
- If $E_{\psi'} > E_\psi$ reject the move with probability $1 - \frac{\exp(-\beta E_{\psi'})}{\exp(-\beta E_\psi)}$

This is just a walk on the eigenstates!

- The stationary distribution is $\propto \tau$ (+we don’t need to know the normalization!)
- Hopefully converges rapidly in physically motivated examples
- Need to compute energy, but phase estimation has finite precision!
- Need to revert state if step is rejected (complicated Marriott-Watrous rewinding)!

How to handle ambiguity in phase estimation?

- Temme, Osborne, Vollbrecht, Poulin, Verstraete Nature'11:
  - Use shift-invariant boosted phase estimation → provably impossible
- Yung and Aspuru-Guzik'12
  - Just assume phase estimation is perfect → unphysical
- Wocjan and Temme'21 (continuous-time quantum Metropolis ↔ Davies generator)
  - Assume spectrum has periodic gaps ("rounding promise") → unphysical
- Rall, Wang, Wocjan'22 (builds on WT'21 – continuous-time)
  - Apply random shifts to remove ambiguity with high probability → large overheads
- Chen, Kastoryano, Brandão, G'23 (builds on WT'21 – continuous-time)
  - Solution: Apply Gaussian damped phase estimation & operator Fourier transform → ☺

Continuous-time quantum Metropolis

- Infinitesimal generator, a.k.a., Lindbladian superoperator $\mathcal{L}^\dagger[\cdot]$:

$$\mathcal{L}^\dagger[\rho] = \sum_{j=0}^{m} \underbrace{K_j \rho K_j^\dagger}_{\text{transition}} - \frac{1}{2}\underbrace{\left(K_j^\dagger K_j \rho + \rho K_j^\dagger K_j\right)}_{\text{decay}}$$

- After time t the induced quantum channel is the superoperator

$$\exp(t\mathcal{L}^\dagger[\cdot])$$

- Metropolis modification of the jumps, a.k.a., Davis generator

$$\sum_{j,\Delta} \min\left\{1, \exp(-\beta \underbrace{\Delta}_{E_{\psi'} - E_\psi})\right\} K_j^{(\Delta)}[\cdot]\left(K_j^{(\Delta)}\right)^\dagger - \frac{1}{2} \ldots \text{(decay part)},$$

where

$$K^{(\Delta)} := \sum_{\psi,\psi' :\ E_{\psi'} - E_\psi = \Delta} |\psi'\rangle\langle\psi'|K|\psi\rangle\langle\psi|.$$

Reduce “jump rates” according to Metropolis weights

- The energy differences $\Delta = E_{\psi'} - E_{\psi}$ are called Bohr frequencies. We can decompose K according to the set of Bohr frequeinces B:

$$K = \sum_{\Delta \in B} K^{(\Delta)} \quad \text{where} \quad K^{(\Delta)} = \sum_{\psi,\psi' :\ E_{\psi'} - E_{\psi} = \Delta} |\psi'\rangle\langle\psi'|K|\psi\rangle\langle\psi|.$$

- We want to decompose K to many jump operators labeled by energy change

$$\left|\bar{0}\right\rangle \otimes K \to \sum_{\Delta \in B} |\Delta\rangle \otimes K^{(\Delta)},$$

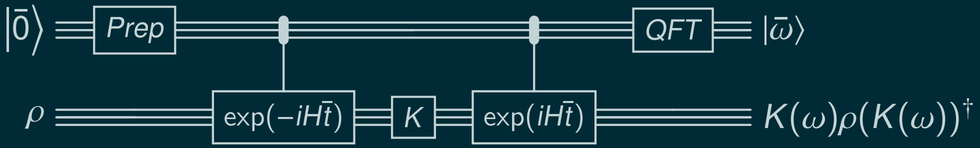
then reduce jump intensity according to the energy difference

$$\sum_{\Delta \in B} |\Delta\rangle \otimes K^{(\Delta)} \to \sum_{\Delta} \min\left\{1, \exp(-\beta\Delta)\right\} |\Delta\rangle \otimes K^{(\Delta)}.$$

- Leading to the Metropolis modification of the jumps:

$$\sum_{j,\Delta} \min\left\{1, \exp(-\beta \underbrace{\Delta}_{E_{\psi'} - E_{\psi}})\right\} K_j^{(\Delta)}[\cdot]\left(K_j^{(\Delta)}\right)^{\dagger} - \frac{1}{2} \ldots \text{(decay part)}.$$

Operator Fourier transform



- Understanding operator Fourier transform:

$$\underbrace{\sum_t f(t)|t\rangle \otimes K}_{\text{peak at } 0} \to \sum_t f(t)|t\rangle \exp(iHt) \otimes K \exp(-iHt) = \sum_t f(t)|t\rangle \otimes \sum_{\Delta \in B} \exp(i\Delta t) K^{(\Delta)}$$

because

$$\exp(iHt)|\psi'\rangle\langle\psi| \exp(-iHt) = \exp(-i(E_{\psi'} - E_\psi)t)|\psi'\rangle\langle\psi|.$$

Finally we apply Fourier transform:

$$\sum_{\Delta \in B} \sum_t f(t) \exp(i\Delta t)|t\rangle \otimes K^{(\Delta)} \overset{QFT}{\to} \sum_\Delta \underbrace{\sum_\omega \hat{f}(\omega - \Delta)|\omega\rangle}_{\text{peak at } \Delta} \otimes K^{(\Delta)} =: \sum_\omega |\omega\rangle \otimes \underbrace{K(\omega)}_{\approx K^{(\Delta)} \text{ for } \Delta = \omega}$$

Weak measurement scheme for Lindbladians

Block-encoding of Lindblad generators

We say that the unitary U is a block encoding of the generator $\mathcal{L}$ consisting of Lindblad operators K_j if $(\langle 0^b| \otimes I)U(|0^a\rangle \otimes I) = \sum_{j=0}^{m} |j\rangle \otimes K_j$.

Using operator Fourier transform we get a block-encoding of $\sum_{j=0}^{m} \sum_{\omega} |j, \omega\rangle \otimes K_j(\omega)$.

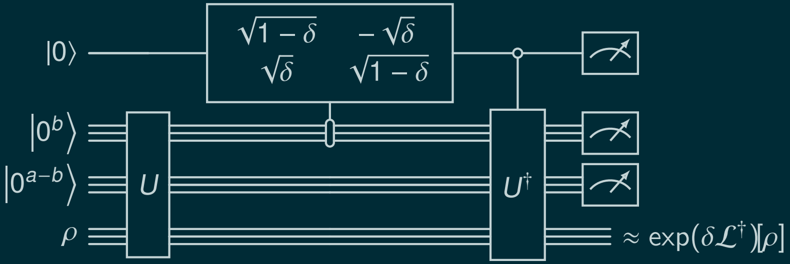


Figure: Quantum circuit implementation of an approximate δ-time step via a weak measurement scheme. Y denotes the Pauli-Y matrix and the gate $e^{-i\theta Y}$ is a rotation by angle θ.

Derivation

Assuming the system register is in the pure state $|\psi\rangle$, this circuit C acts as follows:

$$
\begin{aligned}
|0\rangle\cdot\left|0^a\right\rangle|\psi\rangle &\overset{(1)}{\to} |0\rangle\cdot U\left|0^a\right\rangle|\psi\rangle \\
&\overset{(2)}{\to} \left(\sqrt{1-\delta}|0\rangle + \sqrt{\delta}|1\rangle\right)\cdot\left(|0^b\rangle\langle 0^b|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \;+\; |0\rangle\cdot\left(I-|0^b\rangle\langle 0^b|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \\
&= |0\rangle\cdot U\left|0^a\right\rangle|\psi\rangle \;+\; \sqrt{\delta}|1\rangle\cdot\left|0^b\right\rangle\underbrace{\left(\left\langle 0^b\right|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle}_{|\psi_0'\rangle:=} \;-\; (1-\sqrt{1-\delta})|0\rangle\cdot\left(|0^b\rangle\langle 0^b|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \\
&\overset{(3)}{\to} |0\rangle\cdot\left|0^a\right\rangle|\psi\rangle \;+\; \sqrt{\delta}|1\rangle\cdot\left|0^b\right\rangle\left|\psi_0'\right\rangle \;-\; (1-\sqrt{1-\delta})|0\rangle\cdot U^\dagger\left(|0^b\rangle\langle 0^b|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \\
&= |0\rangle\cdot\left|0^a\right\rangle|\psi\rangle \;+\; \sqrt{\delta}|1\rangle\cdot\left|0^b\right\rangle\left|\psi_0'\right\rangle \;-\; (1-\sqrt{1-\delta})|0\rangle\cdot\left|0^a\right\rangle\left(\left\langle 0^a\right|\otimes I\right)U^\dagger\left(\left|0^b\right\rangle\otimes I\right)\cdot\left(\left\langle 0^b\right|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \\
&\qquad\qquad -\; (1-\sqrt{1-\delta})|0\rangle\cdot\left(I-|0^a\rangle\langle 0^a|\otimes I\right)U^\dagger\left(|0^b\rangle\langle 0^b|\otimes I\right)U\left|0^a\right\rangle|\psi\rangle \\
&= |0\rangle\cdot\left|0^a\right\rangle\left(I-\underbrace{(1-\sqrt{1-\delta})}_{\frac{\delta}{2}+O(\delta^2)}\sum_{j\in J}K_j^\dagger K_j\right)|\psi\rangle + \sqrt{\delta}|1\rangle\cdot\left|0^b\right\rangle\sum_{j\in J}|j\rangle K_j|\psi\rangle - \underbrace{(1-\sqrt{1-\delta})}_{\frac{\delta}{2}+O(\delta^2)}|0\rangle\cdot\left|0^a\perp\right\rangle,
\end{aligned}
$$

where $|0^a\perp\rangle$ is some quantum state such that $\||0^a\perp\rangle\|\leq 1$ and $(\langle 0^a|\otimes I)\cdot|0^a\perp\rangle = 0$. Tracing out the first $a+1$ qubits we get that the resulting state is $O\left(\delta^2\right)$-close to the desired state.

Open questions

- In which (physical) systems can we expect rapid convergence?
- How to bound the gap of the generator or the mixing time?
- How noise resilient is this algorithm?
- Finally a quadratic improvement for carbon capture? ☺