PREPARING THERMAL STATES ON A DIGITAL QUANTUM COMPUTER

by

Matthew Hagan

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Physics
University of Toronto

Preparing Thermal States on a Digital Quantum Computer

Matthew Hagan
Doctor of Philosophy

Department of Physics
University of Toronto
2025

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et.

This thesis is dedicated to...

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The goal of this thesis is to serve as a blueprint for creating quantum channels that can prepare thermal states of arbitrary systems. This framework was primarily created as an algorithmic process to prepare input states of the form $\rho(\beta) = \frac{e^{-\beta H}}{\text{tr}}\left(e^{-\beta H}\right)$, known as Gibbs, Boltzmann, or thermal states, for simulations on a digital, fault-tolerant quantum computer. As $\rho(\beta)$ approaches the ground state as the inverse temperature β diverges, meaning Gibbs states serve as useful proxies for problems in which ground states dictate features of dynamics, such as the electron wavefunction in banded condensed matter systems or in chemical scenarios. One of the key features of our algorithm is the addition of only one single extra qubit outside of those needed to store the state of quantum system being simulated. This may seem like a minor technical achievement, given the existence of Gibbs samplers that also have only one extra qubit or at worst a constant number of qubits overhead, but the way in which this single qubit is utilized in our algorithm highlights the connections between our channel as an algorithmic tool to prepare states of interest and the model of thermalization that we believe the physical world may actually follow. The rest of this introduction is to provide context for how the technical results contained in later chapters of this thesis contribute to the growing interplay between Physics and Theoretical Computer Science within the realm of Quantum Computing.

The easiest way to grasp the context of this thesis is to first understand classical thermal states. In classical mechanics we typically have access to a Hamiltonian H that is a function on phase space (x, p) to the reals $\mathbb{R}$. We can turn this function into a probability distribution via the canonical ensemble $p_\beta(x, p) = \frac{e^{-\beta H(x,p)}}{\int dx dp e^{-\beta H(x,p)}}$, which can be thought of as the "density" of particles

near a particular (x, p) in phase space for a thermodynamically large collection of non-interacting particles under the same Hamiltonian H. The inverse temperature β, typically taken to be $\frac{1}{k_\beta T}$ where k_β is Boltzmann's constant and T the temperature, plays an important role in shaping the distribution $p_{\beta(x,p)}$. For example, in the $\beta \to \infty$ the distribution becomes concentrated at the minimum energy points of $H(x,p)$, which correspond to points with zero momentum $p = 0$ and are minimums of the potential energy $V(x)$. This shows that the problem of preparing classical thermal states somehow contains'' the problem of function optimization, and in fact being able to sample from classical thermal states for arbitrarily large β allows us to approximate the partition function $\mathcal{Z} = \int e^{-\beta H(x,p)} dx dp$, which is known to be a #P-Hard computational problem.

In order to understand when classical systems can reach thermal equilibrium we need to understand where does the concept of temperature come from? We can use the following thought experiment of two systems A and B that are completely isolated from their surroundings, each initially has some internal energy U_A and U_B. We could define entropies σ_A and σ_B from the distribution of the energy of each particle, averaged over some time window, and declare that an equilibrium has been reached between the two systems when the total entropy is constant

$$d\sigma = d\sigma_A + d\sigma_B = 0.$$

Then we know that the total energy of the two systems must remain constant, due to thermal isolation, so $dU = dU_A + dU_B = 0$. However, in this scenario the energy of the system is only a function of the entropy, so we have

$$dU = \frac{\partial U_A}{\partial \sigma_A} d\sigma_A + \frac{\partial U_B}{\partial \sigma_B} d\sigma_B = \left(\frac{\partial U_A}{\partial \sigma_A} - \frac{\partial U_B}{\partial \sigma_B} \right) d\sigma_A = 0.$$

This implies that when the two systems are in equilibrium $\frac{\partial U_A}{\partial \sigma_A} - \frac{\partial U_B}{\partial \sigma_B} = 0$ and the property $\frac{\partial U_A}{\partial \sigma_A}$ is equal to that of B. This is a good notion of temperature, so we define $T = 1/\frac{\partial U}{\partial \sigma}$.

The rationale for introducing this experiment is two-fold. On the first, it gives us a very rigorous definition of temperature by isolating exactly the condition meant behind "thermal equilibrium". We find that we can still define temperature even if we are not sure of how the two systems exchange energy. All we really need to know is that the systems are completely isolated and that energy is exchanged *somehow*, and that this energy exchange is captured by the entropy. The second main point of this experiment is that it gives a very concrete way to thermalize a system to some temperature T: if you have another system at that temperature just put the two

into contact! Eventually they will equilibrate to some intermediate temperature and the process can be repeated.

Although the above result tells us that two systems in thermal contact with each other will eventually reach thermal equilibrium it doesn't exactly tell us *how* it does so. Empirically, we have found a few different ways in which two systems can exchange heat:

- They can trade photons at various wavelengths, described by the Plank law, which then get "absorbed" by the various parties. This is how the earth reaches a thermal equilibrium with the sun and the vacuum of space.

- The nuclei of molecules can vibrate, which can then cause nearby nuclei and electrons to oscillate as well through Coulomb's law. This is the process of conduction, and can be extended to rigid crystals via phonon theory.

- Convection allows for hot parts of a liquid or gas to mix with colder parts. This brings the two extremal parts of a system in closer contact and allows for their temperatures to average out quicker.

These three mechanisms are typically sufficient for classical systems to equilibrate. Typically rates of heat exchange are measured empirically and the random microscopic effects are essentially "averaged out" at the macroscopic into a single coefficient of heat transfer.

If one has complete control over some system, such as a refrigerator, and can prepare it in a thermal state of temperature T_{cold}, we can then utilize this to cool down other systems via equilibration. Given our understanding that preparing thermal states is an incredibly challenging problem in the worst case, this raises the question: can we mimick this cooling process algorithmically on digital computers to sample from thermal distributions? If nature is able to cool systems down efficiently, we should be able to do so as well on computer simulations.

When trying to simulate a classical system we can typically compute the energy of the system $H(x,p)$ given a configuration of position and momenta. For example if we are trying to prepare the thermal state for a gas of non-interacting molecules with no background potential then the energy is just $\sum_i p_i^2/2m$. If we encode the positions into standardized 64-bit floating point numbers, computing the energy can be done in order $O(N)$ time, where N is the number of particles simulated. This is referred to as having oracle access to H, once we have a collection (x,p) we can then compute H as a function call. The challenging problem is to then take this oracle H and

output a list of samples $\{s_1, s_2, ..., s_S\}$, where s_i is a pair of position and momenta $s_i = (x_i, p_i)$, that replicate the thermal statistics. We want to use these samples to compute difficult thermal quantities, given some observable $O(x, p)$ (e.g. the velocity of a particle $v = |\vec{p}|$) we would like the following to be as close to each other as possible

$$\frac{1}{S} \sum_{i=1}^{S} O(x_i, p_i) \frac{e^{-\beta H(x_i, p_i)}}{\sum_{j=1}^{S} e^{-\beta H(x_j, p_j)}} \approx \int dx dp\ O(x, p) \frac{e^{-\beta H(x, p)}}{\int e^{-\beta H(x', p')} dx' dp'}.$$

This is known as Monte Carlo integration, as the integral on the right is replaced by the random process on the left.

The earliest technique for attacking such problems computationally was the Metropolis-Hastings algorithm. This generated samples $\{s_i\}$ using a two-step process that was repeated over and over. Starting with some random initial sample s_0 one then generates a proposed sample $\tilde{s}_1$ using a random transition $T(x' |\ x)$ over the state space. For example, if the system of interest is a collection of spins $\vec{\mu}_j$ then we can pick one spin uniformly and then randomly rotate it. Once we have the proposed new sample $\tilde{s}_1$ we decide to accept it, in which case $s_1 = \tilde{s}_1$ or reject it, in which case we go back to the previous point and set $s_1 = s_0$, with the randomized filter

$$\Pr[s_i = \tilde{s}_i] = \min\left(1, \frac{e^{-\beta H(\tilde{s}_i)}/\mathcal{Z}}{e^{-\beta H(s_{i-1})}/\mathcal{Z}}\right) = \min\left(1, e^{-\beta(H(\tilde{s}_i) - H(s_{i-1}))}\right).$$

This is known as the Metropolis filter and one key insight is that it only depends on the difference in energy between the previous state and the proposed state, the dependence on the partition function vanishes when taking the ratio. Since the initial sample can be random and the transition step is also randomized, this leads to an efficiently computable algorithm for generating new samples.

The last remaining question is if these samples are actually representative of the thermal distribution $p_\beta(x, p)$. So long as the transition matrix is ergodic, meaning any starting state (x', p') has a finite expected hitting time to reach any other arbitrary state (x'', p''), then the thermal distribution $p_\beta(x, p)$ satisfies a condition known as Detailed Balance which then guarantees convergence to $p_\beta(x, p)$. As the above process is a Markov chain, since the distribution over the next sample only depends on the state of the current sample, we can define the Markov transition probability (which is different from the state transition probabilities) as $\Pr[(x, p) \rightarrow (x', p')]$ and

we say that a state probability distribution π satisfies Detailed Balance if the following equation holds

$$\Pr[(x_1, p_1) \to (x_2, p_2)]\pi(x_1, p_1) = \Pr[(x_2, p_2) \to (x_1, p_1)]\pi(x_2, p_2).$$

This condition, along with ergodicity, is sufficient to prove that the probability s_i, for $i \gg 1$, will be distributed according to π. The amount of samples needed to converge to π is an incredibly difficult question to answer and the subject of much study on Markov chains.

One of the most recent improvements to this algorithm came about in the late 1980's and was formalized in Radford Neal's thesis in the Computer Science Department here at the University of Toronto [1]. This new algorithm, now called Hamiltonian Monte Carlo (HMC) but previously known as Hybrid Monte Carlo, modified the existing Metropolis-Hastings algorithm by changing the state transition function. Instead of choosing the next proposed sample randomly, in HMC the momentum variable is chosen from a Gaussian distribution with variance proportional to $1/\beta$ and then the time dynamics generated by $H(x, p)$ are used to propose a new sample. Since time evolution does not change the energy, the Metropolis filter then accepts every sample in the limit of perfect numerical integration of Hamilton's equations of motion. This technique allows for much higher dimensional state spaces to be explored and empirically leads to less correlated samples.

We introduce HMC for the conceptual changes it makes to the original Metropolis-Hastings method. In Metropolis-Hastings the filter step is really what fixes the distribution to mimic the Gibbs distribution and the transition step is simply to guarantee ergodicity. The filtration step in this sense is rather artificial and more computational in spirit. In HMC the filtration step is virtually eliminated by instead utilizing the dynamics provided by nature. We are only able to take advantage of this because in classical mechanics the position and momentum variables "commute", meaning that the thermal distribution is really a product of distributions over position and momenta as

$$e^{-\beta H(x,p)} = e^{-\beta \frac{p^2}{2m} - \beta V(x)} = e^{-\beta \frac{p^2}{2m}} \cdot e^{-\beta V(x)},$$

and the partition function also splits as a product $\mathcal{Z} = \mathcal{Z}_x \cdot \mathcal{Z}_p$. In a sense, HMC is using time dynamics to provide equilibration not between multiple particles but instead between momenta and position, and due to the simple nature of momentum thermal states as Gaussian random variables we can prepare these states at arbitrary temperatures relatively easy.

This crucial step of commutativity of x and p proves to be a difficult obstacle to overcome when extending HMC to a quantum mechanical setting. If one takes a naïve adoption to a continuous, single variable quantum Hamiltonian $H(\hat{x}, \hat{p})$ and utilizes Gaussian distributed momentum *kicks* $e^{ip_{\text{kick}}\hat{x}}|\psi\rangle$, where p_{kick} is Gaussian and $\hat{x}$ is the position operator that generates momentum translations, one can then show that the maximally mixed state is the unique fixed point of the dynamics. This says that our model of a thermalizing environment as providing random momentum shifts is fundamentally **wrong**. Instead, one needs to come up with a simpler system that can be prepared in the thermal state at the desired temperature and use that to cool (or heat) the system to the target temperature.

The smaller environment that we ended up landing on is a single additional two level system, or qubit. The next difficulty is then to choose an interaction between the qubit and the system of interest that leads to thermalization. When looking to physics for inspiration, the situation for thermalization is much less clear in the quantum setting compared to the clear picture we have classically. Even determining a model for equilibrium is not as clear cut. If we look at a closed setting with two Hilbert spaces $\mathcal{H}_A \otimes \mathcal{H}_B$, similar to our classical scenario, one needs to determine the interaction model between the two halves. Further, the energy of each half is no longer a single function but rather a random variable dependent on the density matrix ρ. Even computing derivatives of the mean energy is difficult as one needs an expression for ρ to compute the von Neumann entropy $-\operatorname{Tr}[\rho \log \rho]$.

There are two main approaches physicists have taken to avoid these complications: one involves a conjecture on large, closed quantum systems and the other involves modelling open quantum systems. The first involves a conjecture known as the Eigenstate Thermalization Hypothesis (ETH) [2] and outlines conditions in which a very large closed quantum system may replicate thermal expectations for *local* observables on single particles. This framework has many connections to quantum chaos and has been proven to hold in such chaotic systems, but remains unproven for generic quantum systems, hence the "hypothesis" in the name. The second connection involves studying the effects of a quantum environment on a quantum system and essentially ignoring the state of the environment. By making several assumptions, such as an infinitely large environment, weak coupling, and an assumption known as Markovianity (where the system and environment remain in product states after each interaction), one can show that the thermal state is the unique fixed point for generic open systems. This is done using the Davies' Generators [3] that give

rise to a Linbladdian evolution on the system with the desired fixed point. However, there exist many scenarios in nature where these assumptions are not met, such as finite sized environments that retain some correlations with the system. Finding Linblad evolutions that allow for generic thermalization for arbitrary systems remains an open question.

This leads to our model which completely obliterates its opponents.

Chapter 2

Composite Simulations

The simulation of time-independent Hamiltonian dynamics is a fundamental primitive in quantum computing. To start, the computational problem of approximating the time dynamics of even k-local Hamiltonians (where k is a small constant) is BQP-Complete. This means that any computational problem that can be solved efficiently on a quantum computer can be efficiently reduced to a simulation problem.

The simulation of quantum systems remains one of the most compelling applications for future digital quantum computers [4], [5], [6], [7], [8], [9]. As such, there are a plethora of algorithm options for compiling a unitary evolution operator $U(t) = e^{\{-iHt\}}$ to circuit gates [10], [11], [12], [13], [14], [15], [16], [17]. Some of the simplest such algorithms are product formulas in which each term in a Hamiltonian $H = \sum_i h_i H_i$ is implemented as $e^{iH_i t}$. A product formula is then a particular sequence of these gates that approximates the overall operator $U(t)$. Two of the most well known product formula include Trotter-Suzuki Formulas [11], [13], [18], [19] and the QDrift protocol in which terms are sampled randomly [17], [20]. These two approaches are perhaps the most popular ancilla-free simulation methods yet discovered.

One of the main drawbacks of Trotter-Suzuki formulas is that each term in the Hamiltonian has to be included in the product formula regardless of the magnitude of the term. This leads to a circuit with a depth that scales at least linearly with the number of terms in H, typically denoted L. QDrift avoids this by randomly choosing which term to implement next in the product formula according to an importance sampling scheme in which higher weight terms have larger probabilities. The downside to QDrift is that it has the same asymptotic scaling with $\frac{t}{\varepsilon}$ as a

first-order Trotter formula, meaning it is outperformed at large $\frac{t}{\varepsilon}$ by even a second-order Trotter formula.

In this paper we present a framework for combining simulation channels in a way that allows one to flexibly interpolate the gate cost tradeoffs between the individual channels. The primary example we study is the composition of Trotter-Suzuki and QDrift channels. This is motivated in some part as an effort to extend randomized compilers to include conditional probabilities and in some part to encapsulate progress in chemistry simulations of dropping small weight terms or shuffling terms around different time steps [21]. This latter concept was first developed with the idea of "coalescing" terms into "buckets" by Wecker et al. [21] and further explored by Poulin et al. [22]. They showed that grouping terms of similar sizes together to be skipped during certain Trotter steps led to negligible increases in error and reduced gate counts by about a factor of 10. Similar improvements are also seen in the randomized setting of [23]. In this work we extend on these ideas by placing a specific set of terms into a Trotter partition and the rest in a QDrift partition. This simple division can then be studied analytically and we are able to provide sufficient conditions on asymptotic improvements over completely Trotter or completely QDrift channels. Although we are not able to develop the idea of conditional samples in QDrift protocols, our procedure can be viewed as a specific subset of what a generic Markovian QDrift would look like. We briefly mention these generalizations in Section .

Chapter 3

Preparing Quantum Thermal States

Testing.

3.1 intro

Chapter 4

Appendix

Chapter

Bibliography

[1] R. M. Neal, "Probabilistic inference using Markov chain Monte Carlo methods," 1993.

[2] J. M. Deutsch, "Eigenstate Thermalization Hypothesis," *Reports on Progress in Physics*, vol. 81, no. 8, p. 82001, Aug. 2018, doi: 10.1088/1361-6633/aac9f1.

[3] E. B. Davies, "Markovian master equations," *Communications in mathematical Physics*, vol. 39, pp. 91–110, 1974.

[4] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, "Simulation of electronic structure Hamiltonians using quantum computers," *Molecular Physics*, vol. 109, no. 5, pp. 735–750, 2011, [Online]. Available: https://doi.org/10.1080/00268976.2011.552441

[5] S. P. Jordan, K. S. Lee, and J. Preskill, "Quantum algorithms for quantum field theories," *Science*, vol. 336, no. 6085, pp. 1130–1133, 2012, [Online]. Available: https://doi.org/10.1126/science.1217069

[6] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, "Elucidating reaction mechanisms on quantum computers," *Proceedings of the national academy of sciences*, vol. 114, no. 29, pp. 7555–7560, 2017, [Online]. Available: https://doi.org/10.1073/pnas.1619152114

[7] R. Babbush, D. W. Berry, and H. Neven, "Quantum simulation of the Sachdev-Ye-Kitaev model by asymmetric qubitization," *Phys. Rev. A*, vol. 99, no. 4, p. 40301, Apr. 2019, doi: 10.1103/PhysRevA.99.040301.

[8] Y. Su, D. W. Berry, N. Wiebe, N. Rubin, and R. Babbush, "Fault-Tolerant Quantum Simulations of Chemistry in First Quantization," *PRX Quantum*, vol. 2, no. 4, p. 40332, Nov. 2021, doi: 10.1103/PRXQuantum.2.040332.

[9] T. E. O'Brien *et al.*, "Efficient quantum computation of molecular forces and other energy gradients," *Phys. Rev. Res.*, vol. 4, no. 4, p. 43210, Dec. 2022, doi: 10.1103/PhysRevResearch.4.043210.

[10] D. Aharonov and A. Ta-Shma, "Adiabatic quantum state generation and statistical zero knowledge," in *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing,* 2003, pp. 20–29. [Online]. Available: https://doi.org/10.1145/780542.780546

[11] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, "Efficient quantum algorithms for simulating sparse Hamiltonians," *Communications in Mathematical Physics*, vol. 270, no. 2, pp. 359–371, 2007, [Online]. Available: https://doi.org/10.1007/s00220-006-0150-x

[12] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma, "Simulating Hamiltonian Dynamics with a Truncated Taylor Series," *Phys. Rev. Lett.*, vol. 114, no. 9, p. 90502, Mar. 2015, doi: 10.1103/PhysRevLett.114.090502.

[13] A. M. Childs, A. Ostrander, and Y. Su, "Faster quantum simulation by randomization," *Quantum*, vol. 3, p. 182, Sep. 2019, doi: 10.22331/q-2019-09-02-182.

[14] G. H. Low and I. L. Chuang, "Hamiltonian Simulation by Qubitization," *Quantum*, vol. 3, p. 163, Jul. 2019, doi: 10.22331/q-2019-07-12-163.

[15] G. H. Low, V. Kliuchnikov, and N. Wiebe, "Well-conditioned multiproduct Hamiltonian simulation," *arXiv preprint arXiv:1907.11679*, 2019, [Online]. Available: https://doi.org/10.48550/arXiv.1907.11679

[16] G. H. Low and N. Wiebe, "Hamiltonian Simulation in the Interaction Picture," *arXiv preprint arXiv:1897.10070*, 2019, doi: 10.48550/arXiv.1805.00675.

[17] E. Campbell, "Random Compiler for Fast Hamiltonian Simulation," *Phys. Rev. Lett.*, vol. 123, no. 7, p. 70503, Aug. 2019, doi: 10.1103/PhysRevLett.123.070503.

[18] N. Wiebe, D. Berry, P. Høyer, and B. C. Sanders, "Higher order decompositions of ordered operator exponentials," *Journal of Physics A: Mathematical and Theoretical*, vol. 43, no. 6, p. 65203, Jan. 2010, doi: 10.1088/1751-8113/43/6/065203.

[19] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu, "Theory of Trotter Error with Commutator Scaling," *Phys. Rev. X*, vol. 11, no. 1, p. 11020, Feb. 2021, doi: 10.1103/ PhysRevX.11.011020.

[20] D. W. Berry, A. M. Childs, Y. Su, X. Wang, and N. Wiebe, "Time-dependent Hamiltonian simulation with L^1-norm scaling", *Quantum*, vol. 4, p. 254, Apr. 2020, doi: 10.22331/ q-2020-04-20-254.

[21] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer, "Gate-count estimates for performing quantum chemistry on small quantum computers," *Physical Review A*, vol. 90, no. 2, Aug. 2014, doi: 10.1103/physreva.90.022305.

[22] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer, "The Trotter step size required for accurate quantum simulation of quantum chemistry," *arXiv preprint arXiv:1406.4920*, 2014, [Online]. Available: https://doi.org/10.48550/arXiv.1406.4920

[23] I. D. Kivlichan, C. E. Granade, and N. Wiebe, "Phase estimation with randomized Hamiltonians," *arXiv preprint arXiv:1907.10070*, 2019, doi: 10.48550/arXiv.1907.10070.