

Adiabatic times for Markov chains and applications

Kyle Bradford and Yevgeniy Kovchegov

Department of Mathematics
Oregon State University
Corvallis, OR 97331-4605, USA
bradfork, kovchegy @math.oregonstate.edu

Abstract

We state and prove a generalized adiabatic theorem for Markov chains and provide examples and applications related to Glauber dynamics of the Ising model over $\mathbb{Z}^d/n\mathbb{Z}^d$. The theorems derived in this paper describe a type of adiabatic dynamics for $\ell^1(\mathbb{R}_+^n)$ norm preserving, time inhomogeneous Markov transformations, while quantum adiabatic theorems deal with $\ell^2(\mathbb{C}^n)$ norm preserving ones, i.e. gradually changing unitary dynamics in \mathbb{C}^n .

Keywords: time inhomogeneous Markov processes; ergodicity; mixing times; adiabatic; Glauber dynamics; Ising model

AMS Subject Classification: 60J10, 60J27, 60J28

1 Introduction

The long-term stability of time inhomogeneous Markov processes is an active area of research in the field of stochastic processes and their applications. See [9] and [10], and references therein. The adiabatic time, as introduced in [5], is a way to quantify the stability for a certain class of time inhomogeneous Markov processes. In order for us to introduce the reader to the type of adiabatic results that we will be working with in this paper, let us first mention earlier results that were published in [5], thus postponing a more elaborate discussion of the matter until subsection 1.2.

1.1 Preliminaries

The mixing time quantifies the time it takes for a Markov chain to reach a state that is close enough to its stationary distribution. For the discrete-time finite state case we will look at the evolution of the Markov chain through its probability transition matrix. See [6] for a systematized account of mixing time theory and examples. Let $\|\cdot\|_{TV}$ denote the total variation distance.

Definition 1. Suppose P is a discrete-time finite Markov chain with a unique stationary distribution π , i.e. $\pi P = \pi$. Given an $\epsilon > 0$, the mixing time $t_{mix}(\epsilon)$ is defined as

$$t_{mix}(\epsilon) = \inf \{t : \|\nu P^t - \pi\|_{TV} \leq \epsilon, \text{ for all probability distributions } \nu\}.$$

To define the adiabatic time in its first and simplest form (that we will expand and generalize a few pages down) we have to consider a time inhomogeneous Markov chain whose probability transition matrix evolves linearly from an initial probability transition matrix $P_{initial}$ to a final probability transition matrix P_{final} . Namely, we consider two transition probability operators, $P_{initial}$ and P_{final} , on a finite state space Ω , and we suppose there is a unique stationary distribution π_f of P_{final} . We let

$$P_s = (1 - s)P_{initial} + sP_{final} \quad (1)$$

We use (1) to define a time inhomogeneous Markov chain $P_{\frac{t}{T}}$ over $[0, T]$ time interval. The adiabatic time quantifies how gradual the transition from $P_{initial}$ to P_{final} should be so that at time T , the distribution is ϵ close to the stationary distribution π_f of P_{final} .

Definition 2. Given $\epsilon > 0$, a time T_ϵ is called the adiabatic time if it is the least T such that

$$\max_\nu \|\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 - \pi_f\|_{TV} \leq \epsilon$$

where the maximum is taken over all probability distributions ν over Ω .

With these definitions one would naturally ask how adiabatic and mixing times compare. This will be especially relevant given the emergence of quantum adiabatic computation and some instances of using adiabatic algorithms to solve certain classical computation problems. See [1] and [8]. It can be speculated that there may be scenarios in which the adiabatic time is more convenient to compute than mixing times. If we find the relationship between the two, it will give us an understanding of the adiabatic transition (which is more prevalent in a context of physics) in terms of mixing times and vice versa. The following adiabatic theorem was proved in [5].

Theorem (Kovchegov 2009). Let t_{mix} denote the mixing time for P_{final} . Then the adiabatic time

$$T_\epsilon = \mathcal{O}\left(\frac{t_{mix}(\epsilon/2)^2}{\epsilon}\right)$$

In subsection 1.4 we will give an example that shows the order of t_{mix}^2/ϵ is the best bound for the adiabatic time in this setting. There $\Omega = \{0, 1, 2, \dots, n\}$ and

$$P_{initial} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad P_{final} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Similar adiabatic results hold in the case of continuous-time Markov chains. There, the concept of an adiabatic time is defined within the same setting and a relationship with the mixing time is shown. Let us state a continuous adiabatic result from [5], and then prove a more general statement of the theorem in the next section.

Once again we define the mixing time as a measurement of the time it takes for a Markov chain to reach a state that is close enough to its stationary distribution. For the continuous-time, finite-state case we look at the evolution of the Markov chain through its probability transition matrix as a function over time.

Definition 3. *Suppose $P(t)$ is a finite continuous-time Markov chain with a unique stationary distribution π . Given an $\epsilon > 0$, the mixing time $t_{mix}(\epsilon)$ is defined as*

$$t_{mix}(\epsilon) = \inf \{t : \|\nu P(t) - \pi\|_{TV} \leq \epsilon, \text{ for all probability distributions } \nu\}.$$

To define an adiabatic time we have to look at the linear evolution of a generator for the initial probability transition matrix to a generator for the final probability transition matrix. Suppose $Q_{initial}$ and Q_{final} are two bounded generators for continuous-time Markov processes on a finite state space Ω , and π_f is the unique stationary distribution for Q_{final} . Let us define a time inhomogeneous generator

$$Q[s] = (1 - s)Q_{initial} + sQ_{final} \quad (2)$$

Given $T > 0$ and $0 \leq t_1 \leq t_2 \leq T$, let $P_T(t_1, t_2)$ denote a matrix of transition probabilities of a Markov process generated by $Q[\frac{t}{T}]$ over the time interval $[t_1, t_2]$. With this new generator we define the adiabatic time to be the smallest transition time T such that regardless of our starting distribution, the continuous-time Markov chain generated by $Q[\frac{t}{T}]$ arrives at a state close enough to our stationary distribution π_f .

Definition 4. *Given $\epsilon > 0$, a time T_ϵ is called the adiabatic time if it is the least T such that*

$$\max_{\nu} \|\nu P_T(0, T) - \pi_f\|_{TV} \leq \epsilon$$

where the maximum is taken over all probability distributions ν over Ω .

The above definition for continuous-time Markov chains is similar to the one in the discrete time setting. The corresponding adiabatic theorem for the continuous times case was proved in [5].

Theorem (Kovchegov 2009). *Let t_{mix} denote the mixing time for Q_{final} . Take λ such that $\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{i,j}^{initial}$ and $\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{i,j}^{final}$, where $q_{i,j}^{initial}$ and $q_{i,j}^{final}$ are the rates in $Q_{initial}$ and Q_{final} respectively. Then the adiabatic time*

$$T_\epsilon \leq \frac{\lambda t_{mix}(\epsilon/2)^2}{\epsilon} + \theta$$

where $\theta = t_{mix}(\epsilon/2) + \epsilon/(4\lambda)$.

This is once again the best bound as can be shown through the corresponding example.

In the next section we will state the adiabatic results for Markov chains that generalize the above mentioned theorems in [5] and provide examples of applications in statistical mechanics. Section 2 is dedicated to proofs.

1.2 Results and discussion

Here we extend the results from [5], and thus expand the range of problems that can be analyzed with these types of adiabatic theorems. One such problem that we will discuss in subsection 1.3 deals with adiabatic Glauber dynamics for the Ising model. Now, in order to solve a larger class of problems, we redefine the adiabatic transition for both the discrete and continuous cases.

We consider an adiabatic dynamics where transition probabilities change gradually from $P_{initial} = \{p_{i,j}^{initial}\}$ to $P_{final} = \{p_{i,j}^{final}\}$ so that, for each pair of states i and j , the corresponding mutation of $p_{i,j}$ from $p_{i,j}^{initial}$ to $p_{i,j}^{final}$ is implemented differently and not always linearly. In the case of discrete time steps, this means defining

$$p_{i,j}[s] = (1 - \phi_{i,j}(s))p_{i,j}^{initial} + \phi_{i,j}(s)p_{i,j}^{final}, \quad (3)$$

where $\phi_{i,j} : [0, 1] \rightarrow [0, 1]$ are continuous functions such that $\phi_{i,j}(0) = 0$ and $\phi_{i,j}(1) = 1$ for all locations (i,j) . We require functions $\phi_{i,j}$ to be such that the new operators $P_s = \{p_{i,j}[s]\}$ are Markov chains for all $s \in [0, 1]$, e.g. $\phi_{i,j} \equiv \phi_{i,k} \forall j, k \in \Omega$.

The above definition generalizes (1). If we suppose there is a unique stationary distribution π_f for P_{final} , then the Definition 2 of **adiabatic time** T_ϵ given in the previous section will hold for the adiabatic dynamics defined in (3). The new T_ϵ is related to mixing time via the following adiabatic theorem, that we will prove in section 2.

Theorem 1 (Discrete Adiabatic Theorem). *Let $P_{\frac{t}{T}} = \{p_{i,j}[\frac{t}{T}]\}$ be an inhomogeneous discrete-time Markov chain over $[0, T]$. Let $\phi(s) = \min_{i,j} \phi_{i,j}(s)$ be the pointwise minimum function of all of the $\phi_{i,j}$ functions. If $m \geq 1$ is an integer such that ϕ is $m + 1$ times continuously differentiable in a neighborhood of 1,*

$$\phi^{(k)}(1) = 0 \text{ for all integers } k \text{ such that } 1 \leq k < m$$

and $\phi^{(m)}(1) \neq 0$, then

$$T_\epsilon = \mathcal{O} \left(\frac{t_{mix}^{\frac{m+1}{m}}(\epsilon/2)}{\epsilon^{\frac{1}{m}}} \right)$$

The above is, in fact, the best bound in the new setting as shown through the example given later. See subsection 1.4.

Now we extend the notion of adiabatic dynamics for the continuous-time Markov generators as follows. We let

$$q_{i,j}[s] = (1 - \phi_{i,j}(s))q_{i,j}^{initial} + \phi_{i,j}(s)q_{i,j}^{final} \quad \text{for all pairs } i \neq j, \quad (4)$$

where once again $\phi_{i,j} : [0, 1] \rightarrow [0, 1]$ are continuous functions such that $\phi_{i,j}(0) = 0$ and $\phi_{i,j}(1) = 1$ for all locations (i,j) . Also, we let $Q[s]$ denote the corresponding Markov operator.

If there is a unique stationary distribution π_f for Q_{final} , then the Definition 4 of **adiabatic time** will apply for the extended adiabatic dynamics in (4), and the new T_ϵ can be again related to mixing time.

Theorem 2 (Continuous Adiabatic Theorem). *Let $Q \left[\frac{t}{T} \right]$ ($t \in [0, T]$) generate the inhomogeneous discrete-time Markov chain. Let $\phi(s) = \min_{i,j} \phi_{i,j}(s)$ be the pointwise minimum function of all of the $\phi_{i,j}$ functions. Suppose $m \geq 1$ is an integer such that ϕ is $m + 1$ times continuously differentiable in a neighborhood of 1,*

$$\phi^{(k)}(1) = 0 \text{ for all integers } k \text{ such that } 1 \leq k < m$$

and $\phi^{(m)}(1) \neq 0$. If we take λ such that

$$\lambda \geq \max_{i \in \Omega} \sum_{j:j \neq i} q_{i,j}^{initial} \quad \text{and} \quad \lambda \geq \max_{i \in \Omega} \sum_{j:j \neq i} q_{i,j}^{final},$$

where $q_{i,j}^{initial}$ and $q_{i,j}^{final}$ are the rates in $Q_{initial}$ and Q_{final} respectively. Then

$$T_\epsilon = \mathcal{O} \left(\left[\frac{\lambda}{\epsilon} \right]^{\frac{1}{m}} t_{mix}^{\frac{m+1}{m}} (\epsilon/2) \right)$$

The reader can reference the proof of this theorem in section 2. Again this is the best bound in the new setting as can be shown through the same example. See subsection 1.4. Observe that we could take $\phi : [0, 1] \rightarrow [0, 1]$ such that $\phi(s) \leq \min_{i,j} \phi_{i,j}(s)$ in both adiabatic theorems thus guaranteeing the function is nice enough.

Now we check that the above continuous adiabatic theorem is scale invariant. For a positive M , we scale the initial and final generators to be $\frac{1}{M}Q_{initial}$ and $\frac{1}{M}Q_{final}$ respectively. Then the adiabatic evolution is slowed down M times, and the new adiabatic time should be of order $M \left[\frac{\lambda}{\epsilon} \right]^{\frac{1}{m}} t_{mix}^{\frac{m+1}{m}} (\epsilon/2)$ with the old t_{mix} and λ taken before scaling. On the other hand the new mixing time will be Mt_{mix} , and the new λ is $\frac{\lambda}{M}$ as the rates are M times lower. Plugging the new parameters into the expression in the theorem, we obtain

$$\left[\frac{\lambda}{M\epsilon} \right]^{\frac{1}{m}} (Mt_{mix})^{\frac{m+1}{m}} = M \left[\frac{\lambda}{\epsilon} \right]^{\frac{1}{m}} t_{mix}^{\frac{m+1}{m}}$$

confirming the theorem is invariant under time scaling.

Let us revisit adiabatic theorems in physics and quantum mechanics. The reader can find a version of the quantum adiabatic theorem in [7] and multiple other sources.

The adiabatic results in physics consider a system that transitions from one state to another, while the energy function changes from an initial $\mathcal{H}_{initial}$ to \mathcal{H}_{final} . If the change in the energy function happens slowly enough, for the system that is initially at one of the equilibrium states (i.e. an eigenstate of the initial energy function $\mathcal{H}_{initial}$), the resulting state will end up ϵ -close to the corresponding eigenvector of the final energy function \mathcal{H}_{final} . That is, provided the change in the external conditions is

gradual enough, the j th eigenstate of $\mathcal{H}_{initial}$ is carried to an ϵ -proximity of the j th eigenstate of \mathcal{H}_{final} .

Often the adiabatic results concern with one eigenstate, the ground state. Thinking of Schrödinger equation as an $\ell^2(\mathbb{C}^n)$ norm preserving linear dynamics, and a finite Markov process as a natural description of an $\ell^1(\mathbb{R}_+^n)$ norm preserving linear dynamics, the ground state of one would correspond to the stationary state of the other. It is important to mention that in addition to all of the above properties, the quantum adiabatic theorems often require the transition to be gradual enough for the state to be within an ϵ -proximity of the corresponding ground state at each time during the transition. Taking this into account, the complete analogue of the quantum adiabatic theorem for $\ell^1(\mathbb{R}_+^n)$ would be the one in which the initial distribution is $\mu_0 = \pi_{initial}$ and

$$\|\mu_t - \pi_t\| < \epsilon \quad \forall t \in [0, T],$$

where $\mu_t = \mu_0 P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{t}{T}}$ is the distribution of the time inhomogeneous Markov chain at time $t \in [0, T]$, $\pi_{initial}$ is the stationary distribution of $P_{initial}$, and π_t is the stationary distribution $P_{\frac{t}{T}}$. See [8] for a related result. While we are currently working on proving the above mentioned complete analogue in both discrete and continuous cases, the adiabatic results of this section are sufficiently strong for answering our questions concerning adiabatic Glauber dynamics as stated in the following subsection. Observe that the results of the next subsection could not be obtained using the adiabatic theorems of [5].

Finally, we would like to point out that the models of the adiabatic Markov evolution considered in this paper are similar to simulated annealing. See [3], [4], and [2]. We expect some of the adiabatic Markov chain results to be used for a class of optimization problems by introducing Monte Carlo Markov Chains with the self-adjusting rates.

1.3 Applications to Ising models with adiabatic Glauber dynamics

Let us first state a version of the quantum adiabatic theorem. Given two Hamiltonians, $\mathcal{H}_{initial}$ and \mathcal{H}_{final} , acting on a quantum system. Let

$$\mathcal{H}(s) = (1 - s)\mathcal{H}_{initial} + s\mathcal{H}_{final} \tag{5}$$

Suppose the system evolves according to $\mathcal{H}(t/T)$ from time $t = 0$ to time T . Then if T is large enough, the final state of the system will be close to the ground state of \mathcal{H}_{final} . They are ϵ close in the ℓ^2 norm whenever $T \geq \frac{C}{\epsilon \Delta^3}$, where Δ is the least spectral gap of $\mathcal{H}(s)$ over all $s \in [0, 1]$, and C depends linearly on a square of the distance between $\mathcal{H}_{initial}$ and \mathcal{H}_{final} .

Now, switching to canonical ensembles of statistical mechanics will land us in a Gibbs measure space with familiar probabilistic properties, i.e. the Markov property of statistical independence. We consider a nearest-neighbor Ising model. There the spins can be of two types, -1 and +1. The spins interact only with nearest neighbors. A

Hamiltonian determines the energy-value of the interactions of the configuration of spins.

Here, for a microstate, we multiply its energy by the thermodynamic beta and call it the Hamiltonian of the microstate. In other words, letting x be a configuration of spins, the Hamiltonian we use in this paper will be defined as

$$\mathcal{H}(x) = -\frac{\beta}{2} \sum_{i \neq j} M_{i,j} x(i)x(j)$$

where β is the thermodynamic beta, i.e. its inverse is the temperature times Boltzmann's constant, $M = \{M_{i,j}\}$ is a symmetric matrix and for locations i and j , $M_{i,j} = 0$ if i is not a nearest neighbor to j and $M_{i,j} = 1$ if i is a nearest neighbor to j .

The Markov property of statistical independence is reflected through the *local Hamiltonian* defined at every location j as follows

$$\mathcal{H}^{loc}(x(j)) = -\beta \sum_{i: i \sim j} x(i)x(j),$$

where $i \sim j$ means i and j are nearest neighbors on the graph.

In the original, non-adiabatic case, the Glauber dynamics is used to generate the following Gibbs distribution

$$\pi(x) = \frac{1}{Z(\beta)} e^{-\mathcal{H}(x)}$$

over all spin configurations $x \in \{-1, +1\}^S$, where S denotes all the sites of a graph, and $Z(\beta)$ is the normalization constant. Let us describe how the Glauber dynamics works in the case when each vertex of the connected graph is of the same degree. There, for each location j , we have an independent exponential clock with parameter one associated with it. When the clock rings, the spin $x(j)$ of configuration x at the site j on the graph is reselected using the following probability

$$P(x(j) = +1) = \frac{e^{-\mathcal{H}^{loc}(x_+(j))}}{e^{-\mathcal{H}^{loc}(x_-(j))} + e^{-\mathcal{H}^{loc}(x_+(j))}} = 2 - 2 \tanh \left\{ \mathcal{H}^{loc}(x_+(j)) \right\}$$

where $x_+(i) = x_-(i) = x(i)$ for $i \neq j$, $x_+(j) = +1$ and $x_-(j) = -1$. Here

$$P(x(j) = -1) = 1 - P(x(j) = +1)$$

Also $\mathcal{H}^{loc}(x_-(j)) = -\mathcal{H}^{loc}(x_+(j))$. Now we have a continuous-time Markov process, where the state space is the collection of the configurations of spins.

Now, consider an adiabatic evolution of Hamiltonians as in (5). There at each time t ,

$$\mathcal{H}(s) = (1 - s)\mathcal{H}_{initial} + s\mathcal{H}_{final},$$

where $s = \frac{t}{T}$. The local Hamiltonians must therefore evolve accordingly,

$$\mathcal{H}_s^{loc} = (1 - s)\mathcal{H}_{initial}^{loc} + s\mathcal{H}_{final}^{loc}$$

and the *adiabatic Glauber dynamics* is the one where when the clock rings, the spin $x(j)$ is reselected with probabilities

$$P_s(x(j) = +1) = \frac{e^{-\mathcal{H}_s^{loc}(x_+(j))}}{e^{-\mathcal{H}_s^{loc}(x_-(j))} + e^{-\mathcal{H}_s^{loc}(x_+(j))}}$$

Here too, $\mathcal{H}_s^{loc}(x_-(j)) = -\mathcal{H}_s^{loc}(x_+(j))$.

The stationary distribution of the Q_{final} -generated Markov process, i.e. Glauber dynamics with $\mathcal{H}_{initial}$ energy, is, for a configuration x ,

$$\pi(x) = \frac{e^{-\mathcal{H}_{final}(x)}}{\sum_{\text{all config. } x'} e^{-\mathcal{H}_{final}(x')}}$$

Let β_0 and β_1 denote the values of thermodynamic beta for $\mathcal{H}_{initial}$ and \mathcal{H}_{final} respectively.

1.3.1 Adiabatic Glauber dynamics on $\mathbb{Z}^2/n\mathbb{Z}^2$

Consider nonlinear adiabatic Glauber dynamics of an Ising model on a two-dimensional torus $\mathbb{Z}^2/n\mathbb{Z}^2$. There any two neighboring spin configurations x and y in $\{-1, +1\}^{n^2}$ differ at only one site on the graph, say $v \in \mathbb{Z}^2/n\mathbb{Z}^2$. That is $y(u) = \begin{cases} x(u) & \text{if } u \neq v \\ -x(u) & \text{if } u = v \end{cases}$.

The transition rates evolve according to the adiabatic Glauber dynamics rules, and the transition rates can be represented as

$$q_{x,y}[s] = (1 - \phi_{x,y}(s))q_{x,y}^{initial} + \phi_{x,y}(s)q_{x,y}^{final}$$

as in (4). Here the functions $\phi_{x,y}(s)$ for two neighbors x and y depend entirely on the spins around the discrepancy site v . Namely if all four neighbors of v are of the same spin (+1 or -1), then

$$\phi_{x,y}(s) = \frac{\cosh(-4\beta_1) \cdot \sinh(s(4\beta_0 - 4\beta_1))}{\sinh(4\beta_0 - 4\beta_1) \cdot \cosh(-4\beta_0 + s(4\beta_0 - 4\beta_1))}$$

If it is three of one kind, and one of the other (i.e three +1 and one -1, or three -1 and one +1) as illustrated below

$$\begin{array}{ccccc} & & -1 & & \\ & & | & & \\ +1 & - & v & - & +1 \\ & & | & & \\ & & +1 & & \end{array}$$

then

$$\phi_{x,y}(s) = \frac{\cosh(-2\beta_1) \cdot \sinh(s(2\beta_0 - 2\beta_1))}{\sinh(2\beta_0 - 2\beta_1) \cdot \cosh(-2\beta_0 + s(2\beta_0 - 2\beta_1))}$$

If there are two of each kind, any function works, as both, the initial and the final, local Hamiltonians produce the same transition rates $q_{x,y}^{initial} = q_{x,y}^{final} = 1/2$.

Suppose $\tanh(2\beta_1) < \frac{1}{2}$. Observe that

$$\frac{\cosh(-4\beta_1) \cdot \sinh(s(4\beta_0 - 4\beta_1))}{\sinh(4\beta_0 - 4\beta_1) \cdot \cosh(-4\beta_0 + s(4\beta_0 - 4\beta_1))} \geq \frac{\cosh(-2\beta_1) \cdot \sinh(s(2\beta_0 - 2\beta_1))}{\sinh(2\beta_0 - 2\beta_1) \cdot \cosh(-2\beta_0 + s(2\beta_0 - 2\beta_1))}$$

for $s \in [0, 1]$. Therefore by Theorem 15.1 in [6] and Theorem 2 of this paper, the adiabatic time

$$T_\epsilon = \mathcal{O} \left(C \frac{n^2}{\epsilon} \left[\log(n) + \log \left(\frac{2}{\epsilon} \right) \right]^2 \right),$$

where $C = \frac{(2\beta_0 - 2\beta_1)[\coth(2\beta_0 - 2\beta_1) - \tanh(-2\beta_1)]}{[1 - \tanh(2\beta_1)]^2}$. Here, at every vertex on the torus we attached a Poisson clock with rate one, and therefore we can take $\lambda = n^2$. Also $m = 1$ in the theorem, and one can find the expression for t_{mix} in [6].

1.3.2 Adiabatic Glauber dynamics on $\mathbb{Z}^d/n\mathbb{Z}^d$

The adiabatic Glauber dynamics of an Ising model on a d -dimensional torus $\mathbb{Z}^d/n\mathbb{Z}^d$ solves similarly. There the minimum function $\phi(s)$ of the Theorem 2 is same as in the case of $d = 2$

$$\phi(s) = \frac{\cosh(-2\beta_1) \cdot \sinh(s(2\beta_0 - 2\beta_1))}{\sinh(2\beta_0 - 2\beta_1) \cdot \cosh(-2\beta_0 + s(2\beta_0 - 2\beta_1))}$$

and the adiabatic time

$$T_\epsilon = \mathcal{O} \left(C \frac{n^d}{\epsilon} \left[\log(n) + \log \left(\frac{2}{\epsilon} \right) \right]^2 \right) \quad \text{if} \quad \tanh(2\beta_1) < \frac{1}{d},$$

where again $C = \frac{(2\beta_0 - 2\beta_1)[\coth(2\beta_0 - 2\beta_1) - \tanh(-2\beta_1)]}{[1 - \tanh(2\beta_1)]^2}$.

Notice that the time scaling argument that followed the statement of Theorem 2 works here as well. That is, if we use one Poisson clock of rate one for all vertices, or equivalently place Poisson clocks of rate n^{-d} at every individual vertex, the new adiabatic time will be

$$T'_\epsilon = \mathcal{O} \left(C \frac{n^{2d}}{\epsilon} \left[\log(n) + \log \left(\frac{2}{\epsilon} \right) \right]^2 \right)$$

as $\lambda' = 1$ here.

1.4 Optimality of the bound

In this subsection we give examples that show that the t_{mix}^2 order of adiabatic time given in Kovchegov [5], and $t_{mix}^{\frac{m+1}{m}}$ order for more general settings of this current paper are in fact optimal. We consider discrete probability transition matrices

$$P_{initial} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad P_{final} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

over $n + 1$ states, $\{0, 1, 2, \dots, n\}$, and let the discrete-time adiabatic probability transition matrix to be

$$P_s = (1 - s)P_{initial} + sP_{final}$$

as in [5]. Let π_f again denote the stationary distribution for P_{final} . Here $\pi_f = (0, \dots, 0, 1)$ and the mixing time $t_{mix}(\epsilon) = n$ for any $\epsilon \in (0, 1)$.

Now, since $\mu P_{initial} = \rho$ for any probability distribution μ , we have the following inequality

$$\|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} \geq \left\| \rho \left(\sum_{j=0}^T \left(1 - \frac{j}{T}\right) \frac{T!}{j! \cdot T^{T-j}} \right) (P_{final}^{T-j} - \pi_f) \right\|_{TV}$$

Observe that $\rho P_{final}^{T-j} = \pi_f$ for any $0 \leq j \leq T - n$. Therefore

$$\begin{aligned} \|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} &\geq \sum_{j=T-n+1}^T \left(1 - \frac{j}{T}\right) \frac{T!}{j! T^{T-j}} \\ &= \sum_{j=T-n+1}^T \left(\frac{T!}{j! \cdot T^{T-j}} - \frac{T!}{(j-1)! \cdot T^{T-(j-1)}} \right) \\ &= 1 - \frac{T!}{(T-n)! \cdot T^n} \\ &= 1 - \frac{T-n+1}{T} \cdots \frac{T-1}{T}. \end{aligned}$$

Now, because $\frac{T-n+1}{T} \cdots \frac{T-1}{T} \leq \left(\frac{T-n}{T}\right)^{\frac{n}{2}}$ for $n \geq 2$, we see that

$$\|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} \geq 1 - \left(\frac{T-n}{T}\right)^{\frac{n}{2}} \geq 1 - e^{-\left(\frac{n^2}{4T}\right)}$$

Thus $\epsilon \geq \|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} \geq 1 - e^{-\left(\frac{n^2}{4T}\right)}$ implies $T \geq \frac{-n^2}{4 \log(1-\epsilon)} \approx \frac{n^2}{4\epsilon} = \frac{t_{mix}^2}{4\epsilon}$, proving that the order of adiabatic time $T_\epsilon = \mathcal{O}\left(\frac{t_{mix}(\epsilon/2)^2}{\epsilon}\right)$ in [5] is optimal.

1.4.1 Optimal bound for general settings

The same example works in the more general setting considered in this paper. For the same $P_{initial}$ and P_{final} , let

$$p_{i,j}[s] = (1 - \phi_{i,j}(s))p_{i,j}^{initial} + \phi_{i,j}(s)p_{i,j}^{final}$$

as in (4). Suppose $\phi_{i,j}(s) = \phi(s)$ for all pairs of states (i, j) , and suppose $m \geq 1$ is an integer such that ϕ is $m + 1$ times continuously differentiable in a neighborhood of 1,

$$\phi^{(k)}(1) = 0 \text{ for all integers } k \text{ such that } 1 \leq k < m$$

and $\phi^{(m)}(1) \neq 0$. Then

$$\|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} \geq \left\| \rho \left(\sum_{l=0}^{T-1} (1 - \phi(l/T)) \prod_{j=l+1}^T \phi(j/T) \right) (P_{final}^{T-l} - \pi_f) \right\|_{TV}$$

and therefore

$$\begin{aligned} \|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} &\geq \sum_{l=T-n+1}^{T-1} (1 - \phi(l/T)) \prod_{j=l+1}^T \phi(j/T) \\ &= \sum_{l=T-n+1}^{T-1} \left(\prod_{j=l+1}^T \phi(j/T) - \prod_{j=l}^T \phi(j/T) \right) \\ &= 1 - \prod_{j=T-n+1}^T \phi(j/T). \end{aligned}$$

The minimum function $\phi(x) = 1 + \frac{\phi^{(m)}(1)(x-1)^m}{m!} + O(|x-1|^{m+1})$ and

$$\begin{aligned} \|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV} &\geq 1 - \prod_{j=T-n+1}^T \left(1 + \frac{(-1)^m \phi^{(m)}(1) \cdot (T-j)^m}{T^m \cdot m!} + O((1-j/T)^{m+1}) \right) \\ &= 1 - e^{\sum_{j=T-n+1}^T \log \left(1 + \frac{(-1)^m \phi^{(m)}(1) \cdot (T-j)^m}{T^m \cdot m!} + O((1-j/T)^{m+1}) \right)} \\ &\geq 1 - e^{\frac{(-1)^m \phi^{(m)}(1)}{T^m} \cdot \sum_{j=1}^{n-1} j^m + O((n/T)^{m+1})} \end{aligned}$$

as $\log(1+x) \leq x$.

It is a well known fact that

$$\sum_{j=1}^{n-1} j^k = \sum_{j=0}^k \frac{B_j}{(k+1)-j} \binom{k}{j} n^{(k+1)-j}, \quad (6)$$

where B_j is the j th Bernoulli number.

Suppose $\epsilon \geq \|\nu P_{\frac{1}{T}} \cdots P_{\frac{T}{T}} - \pi_f\|_{TV}$, then

$$\epsilon \approx -\log(1-\epsilon) \geq \frac{(-1)^{m+1} \phi^{(m)}(1)}{T^m} \cdot \sum_{j=0}^m \frac{B_j}{(m+1)-j} \binom{m}{j} n^{(m+1)-j} + O((n/T)^{m+1})$$

Thus confirming that the order of adiabatic time $T_\epsilon = \mathcal{O}\left(\frac{m+1}{\frac{t}{\epsilon} \frac{m}{x}}\right)$ in Theorem 1 is optimal.

Naturally, there is a similar example in the continuous case. There

$$Q_{initial} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ 1 & 0 & -1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{pmatrix} \quad \text{and} \quad Q_{final} = \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & -1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}$$

2 Proofs

In this section we give formal proofs to both adiabatic theorems, Theorem 1 and Theorem 2.

2.1 Proof of Theorem 1

Proof. We write

$$p_{i,j}[s] = (1 - \phi_{i,j}(s))p_{i,j}^{(initial)} + (\phi_{i,j}(s) - \phi(s))p_{i,j}^{(final)} + \phi(s)p_{i,j}^{(final)}$$

and define transition probability matrix $\hat{P} = \{\hat{p}_{i,j}\}$ to be such that

$$(1 - \phi(s))\hat{p}_{i,j} = (1 - \phi_{i,j}(s))p_{i,j}^{(initial)} + (\phi_{i,j}(s) - \phi(s))p_{i,j}^{(final)}$$

We will thus have that

$$P_s = (1 - \phi(s))\hat{P} + \phi(s)P_{final}$$

Observe that

$$\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 = \left[\prod_{j=N+1}^T \phi(j/T) \right] \nu_N P_{final}^{T-N} + \mathcal{E}$$

where $\nu_N = \nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T}{T}}$, and \mathcal{E} is the rest of the terms, and both T and N are natural numbers with $N < T$.

By the triangle inequality, we have

$$\max_{\nu} \|\nu P_{\frac{1}{T}} P_{\frac{2}{T}} \cdots P_{\frac{T-1}{T}} P_1 - \pi_f\|_{TV} \leq \max_{\nu} \|\nu P_{final}^{T-N} - \pi_f\|_{TV} \cdot \left[\prod_{j=N+1}^T \phi(j/T) \right] + S_N$$

where $0 \leq S_N \leq 1 - \left[\prod_{j=N+1}^T \phi(j/T) \right]$.

Let set $T - N = t_{mix}(\epsilon/2)$, where $\epsilon > 0$ is small. Then we have that

$$\max_{\nu} \|\nu P_{final}^{T-N} - \pi_f\|_{TV} \cdot \left[\prod_{j=N+1}^T \phi(j/T) \right] \leq \epsilon/2$$

Setting $1 - \left[\prod_{j=N+1}^T \phi\left(\frac{j}{T}\right) \right] \leq \epsilon/2$ we obtain

$$\log(1 - \epsilon/2) \leq \sum_{j=N+1}^T \log \phi(j/T)$$

We plug in the approximation of the minimum function ϕ around $x = 1$

$$\phi(x) = 1 + \frac{\phi^{(m)}(1)(x-1)^m}{m!} + O(|x-1|^{m+1})$$

obtaining

$$-\log(1 - \epsilon/2) \geq - \sum_{j=N+1}^T \log \left(1 + \frac{(-1)^m \phi^{(m)}(1)(T-j)^m}{T^m \cdot m!} + O((1-j/T)^{m+1}) \right)$$

Therefore

$$-\log(1 - \epsilon/2) \geq \frac{(-1)^{m+1} \phi^{(m)}(1)}{T^m \cdot m!} \sum_{j=1}^{T-N-1} j^m + O\left(\frac{(T-N)^{m+2}}{T^{m+1}}\right)$$

Observe that $(-1)^{m+1} \phi^{(m)}(1) \geq 0$ as $\phi : [0, 1] \rightarrow [0, 1]$ and $\phi(1) = 1$.

By (6), $\sum_{j=1}^{t_{mix}(\epsilon/2)-1} j^m = \sum_{k=0}^m \frac{B_k}{(m+1-k)} \binom{m}{k} t_{mix}(\epsilon/2)^{(m+1)-k}$, where B_k is the k th Bernoulli number, and therefore

$$\epsilon > -\log(1 - \epsilon/2) \geq \frac{(-1)^{m+1} \phi^{(m)}(1)}{T^m \cdot m!} \sum_{k=0}^m \frac{B_k}{(m+1-k)} \binom{m}{k} t_{mix}(\epsilon/2)^{(m+1)-k} + O\left(\frac{(T-N)^{m+2}}{T^{m+1}}\right)$$

In order for the right hand side of the above equation to be $-\log(1 - \epsilon/2)$ close to zero, it is sufficient for T to be of order of $\mathcal{O}\left(\frac{t_{mix}^{\frac{m+1}{m}}(\epsilon/2)}{\epsilon^{\frac{1}{m}}}\right)$. \square

2.2 Proof of Theorem 2

Proof. Define \hat{Q} to be a Markov generate with off-diagonal entries

$$\hat{q}_{i,j} = \frac{1 - \phi_{i,j}(s)}{1 - \phi(s)} q_{i,j}^{(initial)} + \frac{\phi_{i,j}(s) - \phi(s)}{1 - \phi(s)} q_{i,j}^{(final)}$$

Then writing

$$q_{i,j}[s] = (1 - \phi_{i,j}(s)) q_{i,j}^{(initial)} + (\phi_{i,j}(s) - \phi(s)) q_{i,j}^{(final)} + \phi(s) q_{i,j}^{(final)}$$

would imply

$$Q[s] = (1 - \phi(s)) \hat{Q} + \phi(s) Q_{final}$$

Observe that

$$\lambda \geq \max_{i \in \Omega} \sum_{j:j \neq i} \hat{q}_{i,j} \quad \text{and} \quad \lambda \geq \max_{i \in \Omega} \sum_{j:j \neq i} q_{i,j} \left[\frac{t}{T} \right]$$

as

$$\lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{i,j}^{(initial)} \quad \text{and} \quad \lambda \geq \max_{i \in \Omega} \sum_{j: j \neq i} q_{i,j}^{(final)}$$

Let $P_{final}(t) = e^{tQ_{final}}$ denote the transition probability matrix associated with the generator Q_{final} , and let $P_0 = I + \frac{1}{\lambda} \hat{Q}$ and $P_1 = I + \frac{1}{\lambda} Q_{final}$.

The P_0 and P_1 are discrete Markov chains. Conditioning on the number of arrivals within the $[N, T]$ time interval

$$\nu P_T(0, T) = \nu_N P_T(N, T) = \nu_N \left(\sum_{n=0}^{\infty} \frac{(\lambda(T-N))^n}{n!} e^{-\lambda(T-N)} I_n \right)$$

where $\nu_N = \nu P_T(0, N)$ and

$$I_n = \frac{n!}{(T-N)^n} \int \cdots \int_{N < s_1 < \cdots < s_n < T} \left[\left(1 - \phi\left(\frac{s_1}{T}\right)\right) P_0 + \phi\left(\frac{s_1}{T}\right) P_1 \right] \cdots \left[\left(1 - \phi\left(\frac{s_n}{T}\right)\right) P_0 + \phi\left(\frac{s_n}{T}\right) P_1 \right] ds_1 \cdots ds_n$$

i.e. the order statistics of the n .

Therefore, combining the terms with P_{final} , we obtain

$$\begin{aligned} \nu P_T(0, T) &= \nu_N \left(\sum_{n=0}^{\infty} \frac{\lambda^n P_{final}^n}{n!} e^{-\lambda(T-N)} \int_N^T \cdots \int_N^T \phi\left(\frac{s_1}{T}\right) \cdots \phi\left(\frac{s_n}{T}\right) ds_1 \cdots ds_n \right) + \mathcal{E} \\ &= e^{-\lambda(T-N)} \nu_N \left(\sum_{n=0}^{\infty} \frac{\lambda^n T^n}{n!} P_{final}^n \left(\int_{\frac{N}{T}}^1 \phi(x) dx \right)^n \right) + \mathcal{E}, \end{aligned}$$

where \mathcal{E} denotes the rest of the terms.

Take $K > 0$ and define

$$T = \left(\int_{\frac{K-1}{K}}^1 \phi(x) dx \right)^{-1} t_{mix}(\epsilon/2)$$

and

$$N = \frac{(K-1)}{K} \left(\int_{\frac{K-1}{K}}^1 \phi(x) dx \right)^{-1} t_{mix}(\epsilon/2)$$

Recall the approximation of the minimum function ϕ around $x = 1$

$$\phi(x) = 1 + \frac{\phi^{(m)}(1)(x-1)^m}{m!} + O(|x-1|^{m+1})$$

and therefore

$$\int_{\frac{K-1}{K}}^1 \phi(x) dx = \frac{1}{K} \left(1 + \frac{\gamma(K)}{K^m} \right),$$

where $\gamma(K) = (-1)^m \frac{\phi^{(m)}(1)}{(m+1)!} + O(K^{-1})$. Thus we can write

$$\nu P_T(0, T) = e^{-\lambda(T-N)} \nu_N \left(\sum_{n=0}^{\infty} \frac{\lambda^n (T-N)^n}{n!} P_{final}^n \left[1 + \gamma(K) \left(\frac{T-N}{T} \right)^m \right]^n \right) + \mathcal{E}$$

We see, using standard uniformization argument, that

$$\begin{aligned} \nu P_T(0, T) &= e^{-\lambda \left(1 + \frac{\gamma(K)}{K^m}\right)^{-1} t_{mix}(\epsilon/2)} \nu_N \left(\sum_{n=0}^{\infty} \frac{\lambda^n t_{mix}(\epsilon/2)^n}{n!} P_{final}^n \right) + \mathcal{E} \\ &= e^{\lambda \left(\frac{\gamma(K)}{K^m + \gamma(K)}\right) t_{mix}(\epsilon/2)} \nu_N \exp \{Q_{final} \cdot t_{mix}(\epsilon/2)\} + \mathcal{E} \end{aligned}$$

Now, since $(-1)^m \phi^{(m)}(1) \leq 0$, we have that, for any probability distribution ν ,

$$\|\nu P_T(0, T) - \pi_f\|_{TV} = \|\nu \exp \{Q_{final} \cdot t_{mix}(\epsilon/2)\} - \pi_f\|_{TV} \cdot e^{\lambda \left(\frac{\gamma(K)}{K^m + \gamma(K)}\right) t_{mix}(\epsilon/2)} + S_N,$$

where, by the triangle inequality,

$$0 \leq S_N \leq 1 - e^{\lambda \left(1 + \frac{\gamma(K)}{K^m}\right)^{-1} t_{mix}(\epsilon/2)} \left(\sum_{n=0}^{\infty} \frac{\lambda^n (t_{mix}(\epsilon/2))^n}{n!} \right)$$

and, by definition of t_{mix} ,

$$\|\nu \exp \{Q_{final} \cdot t_{mix}(\epsilon/2)\} - \pi_f\|_{TV} \cdot e^{\lambda \left(\frac{\gamma(K)}{K^m + \gamma(K)}\right) t_{mix}(\epsilon/2)} < \epsilon/2$$

Taking $K = c(\lambda/\epsilon)^{\frac{1}{m}} t_{mix}(\epsilon/2)^{\frac{1}{m}}$ with constant $c \gg (-1)^{m+1} \frac{\phi^{(m)}(1)}{(m+1)!}$, we obtain

$$\epsilon > -\log(1 - \epsilon/2) \geq \lambda \left(\frac{-\gamma(K)}{K^m + \gamma(K)} \right) t_{mix}(\epsilon/2)$$

and therefore

$$0 \leq S_N \leq 1 - e^{\lambda \left(\frac{\gamma(K)}{K^m + \gamma(K)}\right) t_{mix}(\epsilon/2)} < \epsilon/2$$

Thus

$$T = \frac{K t_{mix}(\epsilon/2)}{1 + \frac{\gamma(K)}{K^m}} = \mathcal{O} \left(\left[\frac{\lambda}{\epsilon} \right]^{\frac{1}{m}} t_{mix}^{\frac{m+1}{m}}(\epsilon/2) \right)$$

□

References

- [1] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd and O. Regev, *Adiabatic Quantum Computation Is Equivalent to Standard Quantum Computation* SIAM Review, Vol. **50**, No. 4., (2008), 755-787
- [2] V. Cerny, *A thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm* Journal of Optimization Theory and Applications, **45**, (1985), 41-51

- [3] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *Optimization by Simulated Annealing* Science, Vol. **220**, Number 4598, (1983), 671680
- [4] S. Kirkpatrick, *Optimization by simulated annealing: Quantitative studies* Journal of Statistical Physics, Vol. **34**, Numbers 5-6, (1984), 975-986
- [5] Y. Kovchegov, *A note on adiabatic theorem for Markov chains* Statistics & Probability Letters, **80**, (2010), 186-190
- [6] D. A. Levin, Y. Peres and E. L. Wilmer, *Markov Chains and Mixing Times* Amer. Math. Soc., Providence, RI, (2008)
- [7] A. Messiah, *Quantum mechanics* John Wiley and Sons, NY, (1958)
- [8] S. Rajagopalan, D. Shah and J. Shin, *Network Adiabatic Theorem: An efficient randomized protocol for contention resolution* Proc. of the eleventh intl. joint conference on measurement and modeling of computer systems, (2009), 133-144
- [9] L. Saloff-Coste and J. Zúñiga, *Merging and stability for time inhomogeneous finite Markov chains* arXiv: 1004.2296v1 [math.PR] (2010)
- [10] L. Saloff-Coste and J. Zúñiga, *Time inhomogeneous Markov chains with wave like behavior* Annals of Applied Probability, Vol. **20**, Number 5, (2010), 1831-1853