

Stable adiabatic times for Markov chains

Kyle Bradford^{a,*}, Yevgeniy Kovchegov^a, Thinh Nguyen^b

^a*Department of Mathematics
Oregon State University
368 Kidder Hall*

Corvallis, OR 97331 541-737-4686, USA

^b*School of Electrical Engineering and Computer Science*

*Oregon State University
1148 Kelley Engineering Center
Corvallis, OR 97331-5501, USA*

Abstract

In this paper we continue our work on **adiabatic time** of time-inhomogeneous Markov chains first introduced in [9] and [3]. Our study is an analog to the well-known Quantum Adiabatic (QA) theorem which characterizes the quantum adiabatic time for the evolution of a quantum system as a result of applying of a series of Hamilton operators, each is a linear combination of two given initial and final Hamilton operators, i.e. $\mathbf{H}(s) = (1 - s)\mathbf{H}_0 + s\mathbf{H}_1$. Informally, the quantum adiabatic time of a quantum system specifies the speed at which the Hamiltonian operators changes so that the ground state of the system at any time s will always remain ϵ -close to that induced by the Hamilton operator $\mathbf{H}(s)$ at time s . Analogously, we derive a sufficient condition for the stable adiabatic time of a time-inhomogeneous Markov evolution specified by applying a series of transition probability matrices, each is a linear combination of two given irreducible and aperiodic transition probability matrices, i.e., $\mathbf{P}_t = (1 - t)\mathbf{P}_0 + t\mathbf{P}_1$. In particular we show that the stable adiabatic time $t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = O(t_{mix}^4(\epsilon/2)/\epsilon^3)$, where t_{mix} denotes the maximum mixing time over all \mathbf{P}_t for $0 \leq t \leq 1$.

Keywords: *time-inhomogeneous Markov chain, mixing time, stability, adiabatic time*

1. INTRODUCTION

In this paper we study the stability of time-inhomogeneous Markov chains via the notion of stable adiabatic time, an extension of the adiabatic time first

*Corresponding author.

Email address: `bradfork@science.oregonstate.edu` (Kyle Bradford),
`kovchegy@math.oregonstate.edu` (Yevgeniy Kovchegov), `thinhq@eecs.oregonstate.edu`
(Thinh Nguyen)

introduced in [9] and [3]. Our study is motivated in part by the well-known Quantum Adiabatic (QA) theorem which characterizes the quantum adiabatic time for the evolution of a quantum system as a result of applying of a series of Hamiltonian operators, each is a linear combination of two pre-specified initial and final Hamilton operators, i.e., $\mathbf{H}(s) = (1-s)\mathbf{H}_0 + s\mathbf{H}_1$. Quantum adiabatic time of a quantum system, to be discussed in detail shortly, specifies the rate at which Hamiltonian operators change so that the ground state of the system at any time s will always remain ϵ -close to that induced by the Hamilton operator $\mathbf{H}(s)$ at time s . The first Quantum Adiabatic theorem was stated in the 1920s by M. Born and V.A. Fock [5], and have been subsequently studied in [8] among others. Recently, the quantum adiabatic time plays an important role in the development of quantum adiabatic computing. Specifically, quantum adiabatic algorithms are constructed as a sequence of Hamilton operators applied to a quantum system in such a way that drives the system to the desirable state or output, see for example [10]. Thus, the quantum adiabatic time is a natural choice for characterizing the running times of adiabatic quantum algorithms.

We analogously derive a sufficient condition for the stable adiabatic time of a time-inhomogeneous Markov evolution specified by applying a series of transition probability matrices, each is a linear combination of two given irreducible and aperiodic transition probability matrices, i.e., $\mathbf{P}_t = (1-t)\mathbf{P}_0 + t\mathbf{P}_1$. In particular we show that the stable adiabatic time $t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = O(t_{mix}^4(\epsilon/2)/\epsilon^3)$, where t_{mix} denotes the maximum mixing time induced over all the transition probability matrices during the evolution. We note the stable adiabatic time for time-inhomogeneous Markov has recently found practical applications in network design. We refer to the recent work of Rajagoplan et al. [12] where adiabatic time were used to design optimal medium access protocols in wireless networks. Recently, the time-inhomogeneous evolution $\mathbf{P}_t = (1-t)\mathbf{P}_0 + t\mathbf{P}_1$ has also been used to describe the performance of queuing models [17] for networks. Specifically, in this setting, the arrival rate of packet at the queue is assumed to be unknown and is estimated progressively. Appropriate sending rate is then determined based on this estimation. As a result, \mathbf{P}_t describes the a queuing policy (or sending rate) which varies with time based on the new statistics. The adiabatic time is then used to characterize the performance of the queuing model under uncertainty due error in estimation. To motivate our work, we now provide a short overview of the Quantum Adiabatic theorem as discussed in [2].

1.1. QUANTUM ADIABATIC THEOREM

Let \mathbf{H}_0 and \mathbf{H}_1 be two given Hamiltonian operators. Let $T > 0$ be a positive integer. For $t \in [0, T]$, denote $\mathbf{H}(s) = (1-s)\mathbf{H}_0 + s\mathbf{H}_1$, then $\mathbf{H}(s)$ is also a Hamiltonian operator dependent on a time parameter $s = t/T$. In general, a Hamiltonian operator described above, is not required to have a finite number of physical pure states. In this paper we only consider transition probability matrices of finite dimension, which are analogous the Hamiltonian operators with a finite number of n physical pure states.

The ground state in quantum mechanics refers to the lowest-energy state. The quantum adiabatic theorem concerns one eigenstate of the energy function, the ground state. Here we denote $\Phi(s)$ as the ground state of $\mathbf{H}(s)$ and we let $\gamma(s)$ be the eigenvalue associated with it. For a given $T > 0$, when we say that we apply the adiabatic evolution given by \mathbf{H} and Φ for time T we mean to initialize the system in the state $\Phi(0)$ and then apply the continuously varying Hamiltonian $\mathbf{H}(t/T)$ for time $t \in [0, T]$.

Given $\epsilon > 0$ the quantum adiabatic theorem informally says that if, by selecting a large enough value of T , we assume that the change in the Hamiltonian happens slowly enough, then when we apply the adiabatic evolution given by $\mathbf{H}(s)$ and Φ for time T we will be in an ϵ -ball around $\Phi(1)$ with respect to the $l^2(\mathbb{C}^n)$ -norm. This leads us to the following definition.

Definition 1. *Given $\epsilon > 0$ the quantum adiabatic time, denoted as $t_{qad}(\mathbf{H}, \Phi, \epsilon)$, is equal to the smallest positive time, T , required to make the application of the adiabatic evolution given by \mathbf{H} and Φ for time T arrive in an ϵ -ball around $\Phi(1)$ with respect to the $l^2(\mathbb{C}^n)$ -norm.*

The quantum adiabatic theorem gives a sufficient condition for the quantum adiabatic time. The version of quantum adiabatic theorem as proved in [2] motivated our work in [9] and [3]. Supposing that all eigenvalues of $\mathbf{H}(s)$ are either smaller than $\gamma(s) - \Delta$ or larger than $\gamma(s) + \Delta$ (i.e. there is a spectral gap of Δ around $W_1(s)$), then [2] tells us that

$$t_{qad}(\mathbf{H}, \Phi, \epsilon) = O\left(\frac{1}{\epsilon^2 \Delta^4}\right).$$

In [5] the authors suggest that the quantum adiabatic theorem can be described as follows: for an infinitely slow change of the system, i.e., at an infinitely large value of T , the probability of a quantum state changing energy levels remains infinitely small, even for finite values of $s = t/T$ so that $s \in [0, 1]$.

While the result in [2] was used to motivate the work in [9] and [3], the latter notion in [5] is what motivated the work in this paper. We describe an analogue to this quantum system in the context of the classical time-inhomogeneous Markov chains.

1.2. MAIN RESULT

Markov chains over a finite number of n states are $l^1(\mathbb{R}_+^n)$ -norm preserving processes. We denote $\|\cdot\|_1$ as the $l^1(\mathbb{R}^n)$ norm throughout the paper. We assume that the reader has a prior understanding of basic types of Markov chains, such as irreducible, aperiodic, time-homogeneous and time-inhomogeneous Markov chains (see [6], [7] and [11]). Throughout the paper we focus on discrete-time Markov chains and use $\|\cdot\|_{TV}$ to denote the total variation norm. We begin with the definition of mixing time.

Definition 2. For $\epsilon > 0$ the mixing time of a time-homogeneous, irreducible and aperiodic Markov chain governed by a probability transition matrix \mathbf{P} , which has unique stationary distribution π , is defined as:

$$t_{mix}(\mathbf{P}, \epsilon) = \inf\{T \in \mathbb{N} : \|\nu\mathbf{P}^T - \pi\|_{TV} \leq \epsilon\} \quad (1)$$

over all distributions ν .

Mixing times of time-homogeneous irreducible and aperiodic Markov chains have been well studied, see for example [1] and [11]. Mixing time characterizes how fast a chain converges to its stationary distribution. It is particularly important for bounding the running times of many randomized algorithms, for example the simulated annealing algorithm and metropolis-hasting algorithm as explained in [13]. There have also been recent studies on mixing times for time-inhomogeneous Markov chain [14], [15] and [16] which is more closely related to our work. For example, Saloff-Coste and Zúñiga [14] consider spectral bounds of the mixing time for time-inhomogeneous Markov chains on a finite state space when each step in transition corresponds to an ergodic Markov kernel with the same stationary measure. In their subsequent work [15], Saloff-Coste and Zúñiga employed spectral techniques to obtain asymptotic behavior of time-inhomogeneous Markov chains. In this work, the concept of c-stability is introduced, which is an abstraction of requiring all Markov kernels to have the same stationary measure, and bounds for this kind of stability are obtained.

Our work, on the other hand, studies a specific class of discrete-time, time-inhomogeneous Markov chains that was first constructed in [9]. Specifically, we consider the probability transition matrices for two discrete-time, time-homogeneous, irreducible and aperiodic Markov chains over n states. We denote these matrices as \mathbf{P}_0 and \mathbf{P}_1 throughout the paper and call these the initial and the final transition matrices respectively. We define for $t \in [0, 1]$ a class of probability transition matrices $\{\mathbf{P}_t\}_{t \in [0,1]}$ such that

$$\mathbf{P}_t = (1 - t)\mathbf{P}_0 + t\mathbf{P}_1.$$

For $t \in [0, 1]$ we also define π_t to be the stationary distribution of \mathbf{P}_t . Given $T \in \mathbb{N}$, the specific time-inhomogeneous Markov chain being considered in our paper is the one such that the probability transition matrix at time k is $\mathbf{P}_{\frac{k}{T}}$ for $0 \leq k \leq T$. We consider the class of all time-inhomogeneous Markov chains of this type over all $T \in \mathbb{N}$. We will say that any Markov chain in this class is governed by an adiabatic evolution between \mathbf{P}_0 and \mathbf{P}_1 .

Stability of these kinds of Markov chains were described in [9] and [3] using the notion of adiabatic time defined as follows:

Definition 3. For $\epsilon > 0$ the adiabatic time of a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between \mathbf{P}_0 and \mathbf{P}_1 , is defined as:

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = \inf\{T^* \in \mathbb{N} : \max_{\nu} \|\nu\mathbf{P}_0\mathbf{P}_{\frac{1}{T}}\mathbf{P}_{\frac{2}{T}} \cdots \mathbf{P}_1 - \pi_1\|_{TV} \leq \epsilon \text{ for } T \in \mathbb{N}, T \geq T^*\}, \quad (2)$$

where ν is a probability distribution.

Next, we recall a result from [9] that compares the adiabatic time of a time-inhomogeneous Markov chain and the mixing time of the time-homogeneous Markov chain governed by the final transition matrix.

Theorem 1.

Given a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between the two irreducible and aperiodic \mathbf{P}_0 and \mathbf{P}_1 , for $\epsilon > 0$

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = O\left(\frac{t_{mix}^2(\mathbf{P}_1, \epsilon/2)}{\epsilon}\right). \quad (3)$$

We showed that this bound is tight in [3] by finding a pair of matrices, \mathbf{P}_0 and \mathbf{P}_1 , with the following property: as $\epsilon \rightarrow 0$, there exists a positive constant C such that

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = \frac{Ct_{mix}^2(\mathbf{P}_1, \epsilon/2)}{\epsilon}.$$

In the following Proposition, we provide an upper bound on the adiabatic time using the square of the mixing time. Although this is a minor improvement, it is necessary for our main result. The proof of this Proposition is given in Section 4.

Proposition 1. Given a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between the two irreducible and aperiodic \mathbf{P}_0 and \mathbf{P}_1 , for $\epsilon > 0$

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) \leq \frac{2t_{mix}^2(\mathbf{P}_1, \epsilon/2)}{\epsilon}. \quad (4)$$

Definition 3 suggests that for $\epsilon > 0$ and any $T \geq t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon)$ any probability distribution will evolve under consecutive applications of $\mathbf{P}_{\frac{k}{T}}$ to an ϵ -ball around π_1 in the space of probability distributions with respect to the total variation norm. We desire a stronger notion of stability in this paper to match the description of the quantum adiabatic theorem mentioned in [5]. We want to select T large enough so that starting at π_0 , the distribution will evolve under consecutive applications of $\mathbf{P}_{\frac{k}{T}}$ within an ϵ -corridor of $\pi_{\frac{k}{T}}$ for $1 \leq k \leq T$. This leads us to the following definition.

Definition 4. For $\epsilon > 0$ the stable adiabatic time of a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between the irreducible and aperiodic \mathbf{P}_0 and \mathbf{P}_1 , written as $t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon)$, is defined as follows:

$$t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = \inf\{T \in \mathbb{N} : \|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} < \epsilon \text{ for } 1 \leq k \leq T\}. \quad (5)$$

The main goal of this paper is finding a bound for the stable adiabatic time with respect to the maximum mixing time over all the transition probability matrices. For $\epsilon > 0$ we let

$$t_{mix}(\epsilon) = \sup_{s \in [0,1]} \{t_{mix}(\mathbf{P}_s, \epsilon)\}$$

and we seek our bound in terms of this $t_{mix}(\epsilon)$.

We divide our result into two main theorems to highlight the nature of this bound. Our first theorem gives us insight into the nature of the stable adiabatic time. Its proof is given in Section 4.

Theorem 2. *Given a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between the irreducible and aperiodic \mathbf{P}_0 and \mathbf{P}_1 and given $\delta \in (0, 1]$, for any $\epsilon > 0$,*

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon$$

for

$$T \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon\delta},$$

and $\delta \leq k/T \leq 1$.

We now state our main result. The proof is given in Section 3.

Theorem 3. *Given a time-inhomogeneous, discrete-time Markov chain governed by an adiabatic evolution between two time-homogeneous, discrete-time, n -state, irreducible and aperiodic Markov chains with probability transition matrices \mathbf{P}_0 and \mathbf{P}_1 , for any $\epsilon > 0$,*

$$t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = O\left(\frac{t_{mix}^4(\epsilon/2)}{\epsilon^3}\right). \quad (6)$$

The remaining sections of the paper are organized as follows: in Section 2 we state the necessary tools for the proof of our main theorem, in Section 3 we prove our main theorem, and Section 4 is dedicated to proofs.

2. PRELIMINARIES

We begin this section with a result on the stability of time-homogeneous Markov chains. We will find a lower bound for the mixing time of a time-homogeneous, discrete-time, irreducible and aperiodic Markov chain governed by the probability transition matrix \mathbf{P} in terms of the inverse of the smallest nonzero singular value of $\mathbb{I} - \mathbf{P}$. There are similar results in [11], where the lower bound for the mixing time of a time-homogeneous, discrete-time, irreducible, aperiodic and reversible Markov chain governed by the probability transition matrix \mathbf{P} is found in terms of the inverse smallest nonzero eigenvalue of $\mathbb{I} - \mathbf{P}$, or rather in terms of the relaxation time for \mathbf{P} . One should note that our work is not limited to reversible Markov chains. Our work applies to a much larger class of Markov chains. The proof of the following Proposition is in Section 4.

Proposition 2. *For a time-homogeneous, discrete-time, n -state, irreducible and aperiodic Markov chain, if we are given $\epsilon > 0$, then if σ is the smallest nonzero singular value of $\mathbb{I} - \mathbf{P}$,*

$$\frac{1 - 2\sqrt{n}\epsilon}{\sigma} \leq t_{mix}(\mathbf{P}, \epsilon). \quad (7)$$

This Proposition will be vital for proving Theorem 3 and it gives us some intuition about the mixing time. There have been many results bounding the relaxation time for reversible Markov chains on weighted graphs for example conductance bounds and weighted path upper bounds. In both [1] and [4] the authors introduce the necessary spectral structure to find these bounds. They also define a Dirichlet form to help derive the well-known Rayleigh Theorem and the Perron-Frobenius Theorem, which also describe bounds on the relaxation time. Our work, however, does not employ these techniques directly.

We now find a bound of $\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV}$ in terms of $\|\pi_0 - \pi_{\frac{k}{T}}\|_{TV}$. We devote the following Proposition to this endeavor and its proof is in Section 4.

Proposition 3. *For $1 \leq k \leq T$*

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \|\pi_{\frac{k}{T}} - \pi_0\|_{TV} + \frac{(k+1)^2}{2T}. \quad (8)$$

Now we can use the continuity of π_s at $s = 0$ to find an appropriate bound for $\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV}$ for $0 \leq k/T \leq \delta$. We devote the following Proposition to the discovery of how π_s is continuous at $s = 0$. The spectral structure of \mathbf{P}_0 is crucial to this development. The proof is in Section 4.

Proposition 4. *π_s is continuous with respect to the total variation norm at $s = 0$. In particular, for $\epsilon > 0$ if we let σ be the smallest nonzero singular value of $\mathbb{I} - \mathbf{P}_0$, then if*

$$\delta = \frac{\epsilon\sigma}{2n^{3/2}} \quad (9)$$

we have for all $s \leq \delta$, $\|\pi_s - \pi_0\|_{TV} \leq \epsilon$.

Now we can use Proposition 2 along with the fact that $t_{mix}(\mathbf{P}_0, \epsilon) \leq t_{mix}(\epsilon)$ to derive the following Corollary to Proposition 4.

Corollary 1. *π_s is continuous with respect to the total variation norm at $s = 0$. In particular, for $0 < \epsilon < 1/\sqrt{n}$ if*

$$\delta = \frac{\epsilon(1 - \sqrt{n}\epsilon)}{4n^{3/2}t_{mix}(\epsilon/2)} \quad (10)$$

we have for all $s \leq \delta$, $\|\pi_s - \pi_0\|_{TV} \leq \epsilon/2$.

We now have all the necessary tools to find a bound for the stable adiabatic time. We find our result and conclude our paper in the following section.

3. A BOUND FOR THE STABLE ADIABATIC TIME

We devote this section to finding a bound for the stable adiabatic time entirely in terms of the largest mixing time. We state our main result in the following theorem.

Theorem 4. *Given a time-inhomogeneous, discrete-time Markov chain governed by adiabatic evolution between the irreducible and aperiodic \mathbf{P}_0 and \mathbf{P}_1 , for any $\epsilon > 0$,*

$$t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) = O\left(\frac{t_{mix}^4(\epsilon/2)}{\epsilon^3}\right). \quad (11)$$

PROOF.

We first provide a sketch of the proof followed by the technical details. Our proof is based on the results of Theorem 2, Proposition 3, and Corollary 1. Specifically, we divide our proof into two cases. In the first case, we will show how to select T and δ in order to satisfy the two conditions in Theorem 2, namely:

$$T \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon\delta},$$

and

$$\delta \leq k/T \leq 1.$$

Therefore, by Theorem 2, we have

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon.$$

However, the selected T is not yet $t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon)$ since this only holds for k such that

$$\delta \leq k/T \leq 1.$$

In the second case, we will use the results of Proposition 3 and Corollary 1 to show that for the same selected T and δ ,

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon,$$

even in the case when

$$k/T \leq \delta < 1.$$

Therefore, we conclude that the selected T is a sufficient condition for $t_{sad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon)$.

We now proceed with the details of the proof, starting with the first case. Let $\epsilon > 0$. For this fixed ϵ , we choose T be an integer such that

$$\begin{aligned} T &\geq \frac{4t_{mix}^4(\epsilon/2)}{\epsilon^3} + \frac{4t_{mix}^2(\epsilon/2)}{\epsilon^2} + \frac{1}{\epsilon} \\ &= \left(\frac{2t_{mix}^2(\epsilon/2)}{\epsilon\sqrt{\epsilon}} + \frac{1}{\sqrt{\epsilon}}\right)^2. \end{aligned}$$

This implies

$$\sqrt{T} \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon\sqrt{\epsilon}} + \frac{1}{\sqrt{\epsilon}}.$$

Multiplying either side by $\sqrt{\epsilon}$ and subtracting 1 from either side, we obtain

$$\sqrt{\epsilon}\sqrt{T} - 1 \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon}.$$

Notice that $\sqrt{\epsilon}\sqrt{T} - 1 > 0$ because

$$\frac{2t_{mix}^2(\epsilon/2)}{\epsilon} > 0.$$

Dividing either side of the above inequality by $\sqrt{\epsilon}\sqrt{T} - 1$, we obtain

$$1 \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon(\sqrt{\epsilon}\sqrt{T} - 1)}.$$

Multiplying either side by T , we obtain

$$T \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon(\sqrt{\frac{\epsilon}{T}} - \frac{1}{T})}.$$

Now, let

$$\delta = \sqrt{\frac{\epsilon}{T}} - \frac{1}{T},$$

then clearly

$$T \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon\delta}.$$

Next, let k be an integer such that

$$\delta = \sqrt{\frac{\epsilon}{T}} - \frac{1}{T} \leq \frac{k}{T} \leq 1.$$

Then by Theorem 2, we conclude that

$$\|\pi_{\mathbf{0}}\mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon.$$

Now in the second (complementary) case, i.e., when $k/T \leq \delta < 1$, we will show that for the same selected $\delta = \sqrt{\frac{\epsilon}{T}} - \frac{1}{T}$, and T , it is still true that:

$$\|\pi_{\mathbf{0}}\mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon,$$

Let k be an integer such that

$$0 \leq \frac{k}{T} \leq \sqrt{\frac{\epsilon}{T}} - \frac{1}{T} = \delta.$$

Then,

$$\frac{k+1}{T} \leq \sqrt{\frac{\epsilon}{T}}.$$

Using Proposition 3, we have

$$\begin{aligned} \|\pi_{\mathbf{0}} \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} &\leq \|\pi_{\frac{k}{T}} - \pi_{\mathbf{0}}\|_{TV} + \frac{(k+1)^2}{2T} \\ &= \|\pi_{\frac{k}{T}} - \pi_{\mathbf{0}}\|_{TV} + \frac{T}{2} \left(\frac{k+1}{T}\right)^2 \\ &\leq \|\pi_{\frac{k}{T}} - \pi_{\mathbf{0}}\|_{TV} + \frac{T}{2} \left(\sqrt{\frac{\epsilon}{T}}\right)^2 \\ &= \|\pi_{\frac{k}{T}} - \pi_{\mathbf{0}}\|_{TV} + \frac{\epsilon}{2}. \end{aligned}$$

Next, from Corollary 1, as long as $\epsilon < 1/\sqrt{n}$ and $\sqrt{\frac{\epsilon}{T}} - \frac{1}{T} \leq \frac{\epsilon(1-\sqrt{n\epsilon})}{4n^{3/2}t_{mix}(\epsilon/2)}$, we have

$$\|\pi_{\mathbf{0}} \mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon$$

for

$$0 \leq \frac{k}{T} \leq \sqrt{\frac{\epsilon}{T}} - \frac{1}{T}.$$

It should be clear that as $\epsilon \rightarrow 0$,

$$\sqrt{\frac{\epsilon}{T}} - \frac{1}{T} \leq \frac{\epsilon(1-\sqrt{n\epsilon})}{4n^{3/2}t_{mix}(\epsilon/2)}$$

when

$$T \geq \frac{4t_{mix}^4(\epsilon/2)}{\epsilon^3} + \frac{4t_{mix}^2(\epsilon/2)}{\epsilon^2} + \frac{1}{\epsilon}.$$

This tells us that as $\epsilon \rightarrow 0$,

$$t_{sad}(\mathbf{P}_{\mathbf{0}}, \mathbf{P}_{\mathbf{1}}, \epsilon) \leq \frac{4t_{mix}^4(\epsilon/2)}{\epsilon^3} + \frac{4t_{mix}^2(\epsilon/2)}{\epsilon^2} + \frac{1}{\epsilon}.$$

We conclude that

$$t_{sad}(\mathbf{P}_{\mathbf{0}}, \mathbf{P}_{\mathbf{1}}, \epsilon) = O\left(\frac{t_{mix}^4(\epsilon/2)}{\epsilon^3}\right).$$

We see that this result somewhat reaffirms what has been shown in the Quantum Adiabatic Theorem in [2], but a main difference is that the inverse spectral gap bound for the quantum system is replaced with a mixing time bound in our result. Our result also has an extra multiple of $1/\epsilon$. Notice that the inverse spectral gap was a natural choice for the Quantum Adiabatic Theorem due to the Hamiltonian matrix being self-adjoint. For general, not necessarily reversible, Markov Chains, the Adiabatic Theorem is expressed using mixing times.

4. PROOFS

4.1. Proof of Proposition 1

Recall the proof of Theorem 1 in [9]. We notice that

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) \leq K t_{mix}(\mathbf{P}_1, \epsilon/2)$$

where

$$1 + \left(\frac{\left(1 + \frac{1}{K-1}\right)^{K-1}}{e} \right)^{t_{mix}(\mathbf{P}_1, \epsilon/2)} \leq \epsilon/2.$$

After performing some basic algebra and taking the natural logarithm of either side of the equation, we see that

$$\begin{aligned} \ln(1 - \epsilon/2) &\leq t_{mix}(\mathbf{P}_1, \epsilon/2) \left(\ln \left(\left(1 + \frac{1}{K-1}\right)^{K-1} \right) - 1 \right) \\ &= t_{mix}(\mathbf{P}_1, \epsilon/2) \left((K-1) \ln \left(1 + \frac{1}{K-1}\right) - 1 \right) \\ &= t_{mix}(\mathbf{P}_1, \epsilon/2) \left((K-1) \left(\sum_{j=1}^{\infty} (-1)^{j+1} \frac{1}{j(K-1)^j} \right) - 1 \right) \\ &= t_{mix}(\mathbf{P}_1, \epsilon/2) \left(\sum_{j=2}^{\infty} (-1)^{j+1} \frac{1}{j(K-1)^{j-1}} \right) \\ &= t_{mix}(\mathbf{P}_1, \epsilon/2) \left(\sum_{j=1}^{\infty} (-1)^{j+1} \frac{1}{j(K-1)^j} \binom{-j}{j+1} \right). \end{aligned}$$

It is clear now that if we select K large enough so that

$$\begin{aligned} \ln(1 - \epsilon/2) &\leq -t_{mix}(\mathbf{P}_1, \epsilon/2) \left(\sum_{j=1}^{\infty} (-1)^{j+1} \frac{1}{j(K-1)^j} \right) \\ &= -t_{mix}(\mathbf{P}_1, \epsilon/2) \ln \left(1 + \frac{1}{K-1}\right) \end{aligned}$$

then K will be large enough to satisfy the previous inequality.

Exponentiating either side of the equation and performing the basic algebra

required to solve for K we see that

$$\begin{aligned}
K &\geq 1 + \left(e^{\left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)} - 1 \right)^{-1} \\
&= 1 + \left(\sum_{j=0}^{\infty} \left(\frac{1}{j!} \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)^j \right) - 1 \right)^{-1} \\
&= 1 + \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)^{j-1} \right)^{-1} \\
&= 1 + \frac{t_{mix}(\mathbf{P}_1, \epsilon/2)}{-\ln(1-\epsilon/2)} \left(\sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)^{j-1} \right)^{-1}.
\end{aligned}$$

Notice that the infinite sum that we have is the sum of positive terms and the first term in the sum is 1. This tells us that

$$1 \leq \sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)^{j-1}$$

therefore

$$1 \geq \left(\sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{-\ln(1-\epsilon/2)}{t_{mix}(\mathbf{P}_1, \epsilon/2)} \right)^{j-1} \right)^{-1}.$$

This tells us that if we select K such that

$$K \geq 1 + \frac{t_{mix}(\mathbf{P}_1, \epsilon/2)}{-\ln(1-\epsilon/2)}$$

then the above inequality will be satisfied.

Finally we can expand $\ln(1-\epsilon/2)$ to find that

$$K \geq 1 + \frac{2t_{mix}(\mathbf{P}_1, \epsilon/2)}{\epsilon} \left(\sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{\epsilon}{2} \right)^{j-1} \right)^{-1}.$$

Again the infinite sum is the sum of positive terms, and the first term in the sum is 1. This tells us that

$$1 \geq \left(\sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{\epsilon}{2} \right)^{j-1} \right)^{-1}.$$

We conclude that if we select K such that

$$K \geq \frac{2t_{mix}(\mathbf{P}_1, \epsilon/2)}{\epsilon}$$

then

$$1 + \left(\frac{\left(1 + \frac{1}{K-1}\right)^{K-1}}{e} \right)^{t_{mix}(\mathbf{P}_1, \epsilon/2)} \leq \epsilon/2.$$

Therefore, we see that

$$t_{ad}(\mathbf{P}_0, \mathbf{P}_1, \epsilon) \leq \frac{2t_{mix}^2(\mathbf{P}_1, \epsilon/2)}{\epsilon}.$$

4.2. Proof of Theorem 2

To develop the tools for this theorem, we consider the following treatment of our probability transition matrices. If we are given $s \in (0, 1]$, then we see that

$$\mathbf{P}_t = \left(1 - \frac{t}{s}\right) \mathbf{P}_0 + \frac{t}{s} \mathbf{P}_s$$

for all $t \in [0, s]$.

Defining $\mathbf{P}_t^{(s)} = \mathbf{P}_{st}$, we see that

$$\mathbf{P}_t^{(s)} = (1 - t) \mathbf{P}_0^{(s)} + t \mathbf{P}_1^{(s)}$$

for all $t \in [0, 1]$. We also define $\pi_t^{(s)} = \pi_{st}$.

We see that $\{\mathbf{P}_t^{(s)}\}_{t \in [0, 1]}$ is a class of probability transition matrices where $\mathbf{P}_0 = \mathbf{P}_0^{(s)}$ and $\mathbf{P}_s = \mathbf{P}_1^{(s)}$.

Since the time-homogeneous Markov chains determined by \mathbf{P}_0 and \mathbf{P}_s are irreducible and aperiodic, we can consider a time-inhomogeneous, discrete-time Markov chain governed by adiabatic evolution between these two time-homogeneous Markov chains. We can apply Theorem 1 to show that

$$t_{ad}(\mathbf{P}_0^{(s)}, \mathbf{P}_1^{(s)}, \epsilon) = O\left(\frac{t_{mix}^2(\mathbf{P}_1^{(s)}, \epsilon/2)}{\epsilon}\right).$$

Now let $\epsilon > 0$ and $\delta \in (0, 1]$.

For $s \in [\delta, 1]$ we have that $T^* = t_{ad}(\mathbf{P}_0^{(s)}, \mathbf{P}_1^{(s)}, \epsilon)$ is the adiabatic time between $\mathbf{P}_0^{(s)}$ and $\mathbf{P}_1^{(s)}$.

This tells us that

$$\max_{\nu} \|\nu \mathbf{P}_{\frac{1}{T^*}}^{(s)} \mathbf{P}_{\frac{2}{T^*}}^{(s)} \cdots \mathbf{P}_1^{(s)} - \pi_1^{(s)}\|_{TV} \leq \epsilon.$$

Because $\pi_0^{(s)}$ is a specific distribution, we have that

$$\begin{aligned}
\epsilon &\geq \|\pi_0^{(s)} \mathbf{P}_{\frac{1}{T^*}}^{(s)} \mathbf{P}_{\frac{2}{T^*}}^{(s)} \cdots \mathbf{P}_1^{(s)} - \pi_1^{(s)}\|_{TV} \\
&= \|\pi_0^{(s)} \mathbf{P}_{\frac{(1/s)}{(T^*/s)}}^{(s)} \mathbf{P}_{\frac{(2/s)}{(T^*/s)}}^{(s)} \cdots \mathbf{P}_{\frac{(T^*/s)}{(T^*/s)}}^{(s)} - \pi_1^{(s)}\|_{TV} \\
&= \|\pi_0 \mathbf{P}_{\frac{1}{(T^*/s)}} \mathbf{P}_{\frac{2}{(T^*/s)}} \cdots \mathbf{P}_{\frac{s(T^*/s)}{(T^*/s)}} - \pi_{\frac{s(T^*/s)}}\|_{TV}.
\end{aligned}$$

Clearly if $T = t_{ad}(\mathbf{P}_0^{(s)}, \mathbf{P}_1^{(s)}, \epsilon)/s$, then

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \mathbf{P}_{\frac{2}{T}} \cdots \mathbf{P}_{\frac{sT}{T}} - \pi_{\frac{sT}}\|_{TV} \leq \epsilon.$$

We showed in Proposition 1 that for $\epsilon > 0$

$$t_{ad}(\mathbf{P}_0^{(s)}, \mathbf{P}_1^{(s)}, \epsilon) \leq \frac{2t_{mix}^2(\mathbf{P}_s, \epsilon/2)}{\epsilon}.$$

It follows that for $\epsilon > 0$

$$t_{ad}(\mathbf{P}_0^{(s)}, \mathbf{P}_1^{(s)}, \epsilon) \leq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon}.$$

For $\epsilon > 0$ if we let T any integer such that

$$T \geq \frac{2t_{mix}^2(\epsilon/2)}{\epsilon\delta}$$

we have

$$\|\pi_0 \mathbf{P}_{\frac{1}{T}} \mathbf{P}_{\frac{2}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} \leq \epsilon$$

for all $\delta \leq k/T \leq 1$.

4.3. Proof of Proposition 2

We know that an irreducible, aperiodic time-homogeneous Markov chain governed by a probability transition matrix \mathbf{P} has a unique stationary distribution, making the nullity of $(\mathbb{I} - \mathbf{P})$ equal to one when $\lambda = 1$. This would necessarily imply that the rank of $(\mathbb{I} - \mathbf{P})$ is $n - 1$.

Let $\sigma_1 \geq \cdots \geq \sigma_{n-1} = \sigma$ be the positive singular values of $(\mathbb{I} - \mathbf{P})$ with respect to the Euclidean inner product, which we will denote $\|\cdot\|_2$ throughout this paper. This implies that there exists an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ such that $\mathbf{v}_j(\mathbb{I} - \mathbf{P})(\mathbb{I} - \mathbf{P})^T = \sigma_j^2 \mathbf{v}_j$ for $1 \leq j \leq n - 1$ and $\mathbf{v}_n(\mathbb{I} - \mathbf{P})(\mathbb{I} - \mathbf{P})^T = \mathbf{0}$. Clearly $\mathbf{v}_n = \pi/\|\pi\|_2$.

For $t \in \mathbb{N}$ define $\mathbf{M}_{t-1} = \mathbb{I} + \mathbf{P} + \mathbf{P}^2 + \cdots + \mathbf{P}^{t-1}$.

Also define π to be the stationary distribution of \mathbf{P} .

Notice that $\mathbb{I} - \mathbf{P}^t = (\mathbb{I} - \mathbf{P})\mathbf{M}_{t-1}$.

For irreducible, aperiodic Markov chains we have that if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of \mathbf{P} such that $1 = \lambda_1 > |\lambda_2| \geq \dots \geq |\lambda_n|$, then

$$t, \frac{1 - \lambda_2^t}{1 - \lambda_2}, \dots, \frac{1 - \lambda_n^t}{1 - \lambda_n}$$

are the eigenvalues of \mathbf{M}_{t-1} . Notice that \mathbf{M}_{t-1} must be invertible because all eigenvalues are nonzero and also notice that t is the largest eigenvalue in modulus.

This implies that $\mathbb{I} - \mathbf{P} = (\mathbb{I} - \mathbf{P}^t)\mathbf{M}_{t-1}^{-1}$ and we see that

$$\sigma = \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P})\|_2 = \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t)\mathbf{M}_{t-1}^{-1}\|_2.$$

We see that if $\|\cdot\|_*$ is the standard matrix norm, then

$$\begin{aligned} \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t)\|_2 &= \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t)\mathbf{M}_{t-1}^{-1}\mathbf{M}_{t-1}\|_2 \\ &\leq \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t)\mathbf{M}_{t-1}^{-1}\|_2 \|\mathbf{M}_{t-1}\|_* \\ &\leq t \|\mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t)\mathbf{M}_{t-1}^{-1}\|_2 \\ &= t\sigma. \end{aligned}$$

If we let \mathbf{u} be a vector such that for $1 \leq i \leq n$, $\mathbf{u}(i) = 0$ whenever $\mathbf{v}_{n-1}(i) \geq 0$ and $\mathbf{u}(i) = -\mathbf{v}_{n-1}(i)$ whenever $\mathbf{v}_{n-1}(i) < 0$, then we have that $\nu_1 = \mathbf{u}/\|\mathbf{u}\|_1$ and $\nu_2 = (\mathbf{v}_{n-1} + \mathbf{u})/\|\mathbf{v}_{n-1} + \mathbf{u}\|_1$ are probability distributions and

$$\begin{aligned} \mathbf{v}_{n-1}(\mathbb{I} - \mathbf{P}^t) &= \mathbf{v}_{n-1} + (\|\mathbf{u}\|_1 - \|\mathbf{v}_{n-1} + \mathbf{u}\|_1)\pi \\ &\quad - (\|\mathbf{v}_{n-1} + \mathbf{u}\|_1(\nu_2 - \pi)\mathbf{P}^t - \|\mathbf{u}\|_1(\nu_1 - \pi)\mathbf{P}^t). \end{aligned}$$

For $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that \mathbf{x} and \mathbf{y} are probability measures, we see that

$$\frac{1}{2}\|\mathbf{x} - \mathbf{y}\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_{TV} \leq \frac{\sqrt{n}}{2}\|\mathbf{x} - \mathbf{y}\|_2.$$

Through the triangle inequality we see that if we select $t = t_{mix}(\mathbf{P}, \epsilon)$, then

$$\begin{aligned} \left\| (\|\mathbf{v}_{n-1} + \mathbf{u}\|_1(\nu_2 - \pi)\mathbf{P}^t - \|\mathbf{u}\|_1(\nu_1 - \pi)\mathbf{P}^t) \right\|_2 &\leq \|\mathbf{v}_{n-1} + \mathbf{u}\|_1 \cdot \|(\nu_2 - \pi)\mathbf{P}^t\|_2 \\ &\quad + \|\mathbf{u}\|_1 \cdot \|(\nu_1 - \pi)\mathbf{P}^t\|_2 \\ &\leq 2\|\mathbf{v}_{n-1} + \mathbf{u}\|_1 \cdot \|(\nu_2 - \pi)\mathbf{P}^t\|_{TV} \\ &\quad + 2\|\mathbf{u}\|_1 \cdot \|(\nu_1 - \pi)\mathbf{P}^t\|_{TV} \\ &\leq 2(\|\mathbf{v}_{n-1} + \mathbf{u}\|_1 + \|\mathbf{u}\|_1)\epsilon \\ &= 2\|\mathbf{v}_{n-1}\|_1\epsilon \\ &\leq 2\sqrt{n}\|\mathbf{v}_{n-1}\|_2\epsilon \\ &= 2\sqrt{n}\epsilon. \end{aligned}$$

Because $\mathbf{v}_{\mathbf{n}-1}$ and π are orthogonal, we see that

$$\begin{aligned} \|\mathbf{v}_{\mathbf{n}-1} + (\|\mathbf{u}\|_1 - \|\mathbf{v}_{\mathbf{n}-1} - \mathbf{u}\|_1)\pi\|_2 &= \sqrt{1 + (\|\mathbf{u}\|_1 - \|\mathbf{v}_{\mathbf{n}-1} - \mathbf{u}\|_1)^2 (\|\pi\|_2)^2} \\ &\geq 1. \end{aligned}$$

Now through the reverse triangle inequality, meaning that for vectors \mathbf{x} and \mathbf{y} , $\|\mathbf{x} - \mathbf{y}\|_2 \geq \|\mathbf{x}\|_2 - \|\mathbf{y}\|_2$, we see that if $t = t_{mix}(\mathbf{P}, \epsilon)$, then

$$\|\mathbf{v}_{\mathbf{n}-1}(\mathbb{I} - \mathbf{P}^t)\|_2 \geq 1 - 2\sqrt{n}\epsilon.$$

This now implies that

$$\frac{1 - 2\sqrt{n}\epsilon}{\sigma} \leq t_{mix}(\mathbf{P}, \epsilon).$$

4.4. Proof of Proposition 3

Because $\pi_0 \mathbf{P}_{\frac{j}{T}} = \pi_0 + \frac{j}{T} \pi_0 (\mathbf{P}_1 - \mathbf{P}_0)$ for $1 \leq j \leq k$ we notice that

$$\pi_0 \mathbf{P}_{\frac{j}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}} = \pi_0 \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}} + \frac{j}{T} \pi_0 (\mathbf{P}_1 - \mathbf{P}_0) \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}$$

for $1 \leq j \leq k-1$ and

$$\pi_0 \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}} = (\pi_0 - \pi_{\frac{k}{T}}) + \frac{k}{T} \pi_0 (\mathbf{P}_1 - \mathbf{P}_0).$$

Using the convention $\mathbf{P}_{j+1} \cdots \mathbf{P}_k = \mathbb{I}$ when $j \geq k$, we would see that

$$\pi_0 \mathbf{P}_{\frac{j}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}} = (\pi_0 - \pi_{\frac{k}{T}}) + \sum_{j=1}^k \frac{j}{T} \pi_0 (\mathbf{P}_1 - \mathbf{P}_0) \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}.$$

Taking the total variation norm to either side of the inequality, using the triangle inequality and pulling out constants, we see that

$$\begin{aligned} \|\pi_0 \mathbf{P}_{\frac{j}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} &= \|(\pi_0 - \pi_{\frac{k}{T}}) + \sum_{j=1}^k \frac{j}{T} \pi_0 (\mathbf{P}_1 - \mathbf{P}_0) \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}\|_{TV} \\ &\leq \|\pi_0 - \pi_{\frac{k}{T}}\|_{TV} + \sum_{j=1}^k \frac{j}{T} \|\pi_0 (\mathbf{P}_1 - \mathbf{P}_0) \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}\|_{TV}. \end{aligned}$$

Notice that for $1 \leq j \leq k-1$

$$\pi_0 (\mathbf{P}_1 - \mathbf{P}_0) \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} = \pi_0 \mathbf{P}_1 \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_0 \mathbf{P}_0 \mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}$$

is the difference between two probability distributions and

$$\pi_0 (\mathbf{P}_1 - \mathbf{P}_0) = \pi_0 \mathbf{P}_1 - \pi_0 \mathbf{P}_0$$

is also the difference between two probability distributions.

Because we are taking the total variation norm to the difference of two probability distributions we see that $\|\cdot\|_{TV} = \frac{1}{2}\|\cdot\|_1$ where $\|\cdot\|_1$ is the l_1 -norm.

We have that for probability distributions μ and ν , $\|\mu - \nu\|_{TV} = \frac{1}{2}\|\mu - \nu\|_1 \leq \frac{1}{2}(\|\mu\|_1 + \|\nu\|_1) \leq 1$.

This tells us that $\|\pi_0(\mathbf{P}_1 - \mathbf{P}_0)\mathbf{P}_{\frac{j+1}{T}} \cdots \mathbf{P}_{\frac{k}{T}}\|_{TV} \leq 1$ for $1 \leq j \leq k$.

We see then that

$$\begin{aligned} \|\pi_0\mathbf{P}_{\frac{1}{T}} \cdots \mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} &\leq \|\pi_{\frac{k}{T}} - \pi_0\|_{TV} + \sum_{j=1}^k \frac{j}{T} \\ &= \|\pi_{\frac{k}{T}} - \pi_0\|_{TV} + \frac{1}{T} \sum_{j=1}^k j \\ &= \|\pi_{\frac{k}{T}} - \pi_0\|_{TV} + \frac{k(k+1)}{2T} \\ &\leq \|\pi_{\frac{k}{T}} - \pi_0\|_{TV} + \frac{(k+1)^2}{2T}. \end{aligned}$$

4.5. Proof of Proposition 4

We begin with the creation of an orthonormal basis of eigenvectors associated with $(\mathbb{I} - \mathbf{P}_0)(\mathbb{I} - \mathbf{P}_0)^T$ by a singular value decomposition similar to the process we mentioned in Proposition 2.

Here we let $\sigma_1 \geq \cdots \geq \sigma_{n-1} = \sigma$ be the positive singular values of $(\mathbb{I} - \mathbf{P}_0)$ with respect to the Euclidean inner product. This implies that there exists an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ such that $\mathbf{v}_j(\mathbb{I} - \mathbf{P}_0)(\mathbb{I} - \mathbf{P}_0)^T = \sigma_j^2 \mathbf{v}_j$ for $1 \leq j \leq n-1$ and $\mathbf{v}_n(\mathbb{I} - \mathbf{P}_0)(\mathbb{I} - \mathbf{P}_0)^T = \mathbf{0}$.

Here $\mathbf{v}_n = \pi_0 / \|\pi_0\|_2$.

To show continuity at $s = 0$ let $\epsilon > 0$ and first notice that for any $s \in [0, 1]$, $(\pi_s - \pi_0)(\mathbb{I} - \mathbf{P}_0) = s\pi_s(\mathbf{P}_1 - \mathbf{P}_0)$.

Using the Euclidean norm, we see that if $\mathbf{P}_0 \neq \mathbf{P}_1$ and $s \neq 0$, then

$$\frac{\|(\pi_s - \pi_0)(\mathbb{I} - \mathbf{P}_0)\|_2}{\|\pi_s - \pi_0\|_2} = s \frac{\|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_2}{\|\pi_s - \pi_0\|_2}.$$

Throughout this proof we will use $\langle \cdot, \cdot \rangle$ as the Euclidean inner product.

For $1 \leq j \leq n$ let $c_j = \langle \pi_s - \pi_0, \mathbf{v}_j \rangle$. Then we see that $\pi_s - \pi_0 = \sum_{j=1}^n c_j \mathbf{v}_j$.

We have that

$$\begin{aligned}
\frac{\|(\pi_{\mathbf{s}} - \pi_{\mathbf{0}})(\mathbb{I} - \mathbf{P}_{\mathbf{0}})\|_2^2}{\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2^2} &= \frac{\langle (\pi_{\mathbf{s}} - \pi_{\mathbf{0}})(\mathbb{I} - \mathbf{P}_{\mathbf{0}}), (\pi_{\mathbf{s}} - \pi_{\mathbf{0}})(\mathbb{I} - \mathbf{P}_{\mathbf{0}}) \rangle}{\langle \pi_{\mathbf{s}} - \pi_{\mathbf{0}}, \pi_{\mathbf{s}} - \pi_{\mathbf{0}} \rangle} \\
&= \frac{\langle \pi_{\mathbf{s}} - \pi_{\mathbf{0}}, (\pi_{\mathbf{s}} - \pi_{\mathbf{0}})(\mathbb{I} - \mathbf{P}_{\mathbf{0}})(\mathbb{I} - \mathbf{P}_{\mathbf{0}})^T \rangle}{\langle \pi_{\mathbf{s}} - \pi_{\mathbf{0}}, \pi_{\mathbf{s}} - \pi_{\mathbf{0}} \rangle} \\
&= \frac{\langle \sum_{j=1}^n c_j \mathbf{v}_j, \sum_{j=1}^{n-1} \sigma_j^2 c_j \mathbf{v}_j \rangle}{\langle \sum_{j=1}^n c_j \mathbf{v}_j, \sum_{j=1}^n c_j \mathbf{v}_j \rangle} \\
&= \frac{\sum_{j=1}^{n-1} \sigma_j^2 c_j^2}{\sum_{j=1}^n c_j^2} \\
&\geq \sigma_{n-1}^2 \frac{\sum_{j=1}^{n-1} c_j^2}{\sum_{j=1}^n c_j^2} \\
&= \sigma_{n-1}^2 \left(1 - \frac{c_n^2}{\sum_{j=1}^n c_j^2} \right) \\
&= \sigma_{n-1}^2 \left(1 - \left(\frac{\langle \pi_{\mathbf{s}} - \pi_{\mathbf{0}}, \mathbf{v}_n \rangle}{\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2} \right)^2 \right).
\end{aligned}$$

If we let $\mathbf{w}(s) = (\pi_{\mathbf{s}} - \pi_{\mathbf{0}})/\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2$ then we see that

$$\sigma_{n-1}^2 \left(1 - (\langle \mathbf{w}(s), \mathbf{v}_n \rangle)^2 \right) \leq s^2 \frac{\|\pi_{\mathbf{s}}(\mathbf{P}_1 - \mathbf{P}_0)\|_2^2}{\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2^2}.$$

Because $\mathbf{w}(s)$ and \mathbf{v}_n are unit vectors, we can use the fact that

$$\|\mathbf{w}(s)\|_2^2 - 2 \langle \mathbf{w}(s), \mathbf{v}_n \rangle + \|\mathbf{v}_n\|_2^2 = \|\mathbf{w}(s) - \mathbf{v}_n\|_2^2$$

to show that

$$1 - \langle \mathbf{w}(s), \mathbf{v}_n \rangle = \frac{1}{2} \|\mathbf{w}(s) - \mathbf{v}_n\|_2^2$$

and we can use the fact that

$$\|\mathbf{w}(s)\|_2^2 + 2 \langle \mathbf{w}(s), \mathbf{v}_n \rangle + \|\mathbf{v}_n\|_2^2 = \|\mathbf{w}(s) + \mathbf{v}_n\|_2^2$$

to show that

$$1 + \langle \mathbf{w}(s), \mathbf{v}_n \rangle = \frac{1}{2} \|\mathbf{w}(s) + \mathbf{v}_n\|_2^2.$$

From this we see that $1 - (\langle \mathbf{w}(s), \mathbf{v}_n \rangle)^2 = \|\mathbf{w}(s) - \mathbf{v}_n\|_2^2 \cdot \|\mathbf{w}(s) + \mathbf{v}_n\|_2^2 / 4$. Plugging this into our previous equation, we can see that

$$\frac{\sigma_{n-1}^2}{4} \|\mathbf{w}(s) - \mathbf{v}_n\|_2^2 \cdot \|\mathbf{w}(s) + \mathbf{v}_n\|_2^2 \leq s^2 \frac{\|\pi_{\mathbf{s}}(\mathbf{P}_1 - \mathbf{P}_0)\|_2^2}{\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2^2}.$$

After performing some basic algebra we see that

$$\|\pi_{\mathbf{s}} - \pi_{\mathbf{0}}\|_2 \leq \frac{2s \|\pi_{\mathbf{s}}(\mathbf{P}_1 - \mathbf{P}_0)\|_2}{\sigma_{n-1} \|\mathbf{w}(s) - \mathbf{v}_n\|_2 \cdot \|\mathbf{w}(s) + \mathbf{v}_n\|_2}.$$

Notice that $\langle \mathbf{w}(s), \mathbf{1} \rangle / \sqrt{n} = 0$ and $\langle \mathbf{v}_n, \mathbf{1} \rangle / \sqrt{n} = 1 / (\sqrt{n} \|\pi_0\|_2)$ for all $s \in [0, 1]$. Because these are the scalar components of the projections of $\mathbf{w}(s)$ and \mathbf{v}_n onto $\mathbf{1}$ respectively, we see that the minimum possible value for $\|\mathbf{w}(s) - \mathbf{v}_n\|_2$ and $\|\mathbf{w}(s) + \mathbf{v}_n\|_2$ is at least $1 / (\sqrt{n} \|\pi_0\|_2)$.

We now have that

$$\begin{aligned} \|\pi_s - \pi_0\|_2 &\leq \frac{2sn \|\pi_0\|_2^2 \cdot \|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_2}{\sigma_{n-1}} \\ &\leq \frac{2sn \|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_2}{\sigma_{n-1}} \\ &= \frac{2sn \|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_2}{\sigma}. \end{aligned}$$

Again for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that \mathbf{x} and \mathbf{y} are probability measures, we see that

$$\frac{1}{2} \|\mathbf{x} - \mathbf{y}\|_2 \leq \|\mathbf{x} - \mathbf{y}\|_{TV} \leq \frac{\sqrt{n}}{2} \|\mathbf{x} - \mathbf{y}\|_2.$$

This will imply that

$$\|\pi_s - \pi_0\|_{TV} \leq \frac{2sn^{3/2} \|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_{TV}}{\sigma_{n-1}}.$$

Because $\pi_s(\mathbf{P}_1 - \mathbf{P}_0) = \pi_s \mathbf{P}_1 - \pi_s \mathbf{P}_0$ is the difference of two probability distributions, we see that $\|\cdot\|_{TV} = \frac{1}{2} \|\cdot\|_1$ where $\|\cdot\|_1$ is the l_1 -norm. This implies that

$$\|\pi_s(\mathbf{P}_1 - \mathbf{P}_0)\|_{TV} = \frac{1}{2} \|\pi_s \mathbf{P}_1 - \pi_s \mathbf{P}_0\|_1 \leq \frac{1}{2} (\|\pi_s \mathbf{P}_1\|_1 + \|\pi_s \mathbf{P}_0\|_1) \leq 1.$$

This shows that

$$\|\pi_s - \pi_0\|_{TV} \leq \frac{2sn^{3/2}}{\sigma}.$$

Clearly if $\epsilon > 0$, then

$$s \leq \delta = \frac{\epsilon \sigma}{2n^{3/2}}$$

implies $\|\pi_s - \pi_0\|_{TV} \leq \epsilon$.

This shows that π_s is continuous at $s = 0$.

References

- [1] D. Aldous and J.A. Fill, *Reversible Markov chains*, stat-www.berkeley.edu, 2002.
- [2] A. Ambainis and O. Regev, *An Elementary Proof of the Quantum Adiabatic Theorem*, arXiv:quant-ph/0411152v2

- [3] K. Bradford and Y. Kovchegov, *Adiabatic times for Markov chains and their applications*, Journal of Statistical Physics, Vol. **143**, 2011, pp. 955-969.
- [4] P. Brémaud, *Markov Chains: Gibbs fields, Monte Carlo Simulation, and Queues*, Springer Science+Business Media Inc. Texts in Applied Mathematics, Vol. **31**, 2010.
- [5] V.A. Fock, *Selected Works: Quantum Mechanics and Quantum Field Theory*, Chapman & Hall/CRC, 2004.
- [6] D.L. Isaacson and R.W. Madsen, *Markov Chains: Theory and Applications*, John Wiley, New York, 1976.
- [7] S. Karlin and H.M. Taylor, *A first course in Stochastic Processes*, Academic Press, 1975.
- [8] T. Kato, *On the Adiabatic Theorem of Quantum Mechanics*, Journal of the Physical Society of Japan, Vol. **5**, pp. 435 - 439, 1950.
- [9] Y. Kovchegov, *A note on adiabatic theorem for Markov chains*, Statistics and Probability Letters, Vol. **80**, 2010, pp. 186-190.
- [10] H. Krovi, M. Ozols and J. Roland, *Adiabatic condition and the quantum hitting time of Markov chains*, Physical Review A **82**, 022333, 2010.
- [11] D. Levin, Y. Peres and E. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.
- [12] S. Rajagopalan, D. Shah and J. Shin, *Network Adiabatic Theorem: An efficient Randomized Protocol for Contention Resolution*, ACM 978-1-60558-511-6/09/06, 2009.
- [13] S. Ross, *Simulation: Fourth Ed.*, Elsevier Academic Press, 2006.
- [14] L. Saloff-Coste and J. Zúñiga, *Covergence of some time inhomogeneous Markov chains via spectral techniques*, Stochastic Processes and their Applications, Vol. **117**, 2007, pp. 961 - 979.
- [15] L. Saloff-Coste and J. Zúñiga, *Merging for time-inhomogeneous finite Markov chains, Part I: singular values and stability*, Electronic Journal of Probability, Vol. **14**, 2009, pp. 1456-1494.
- [16] L. Saloff-Coste and J. Zúñiga, *Merging for time-inhomogeneous finite Markov chains, Part II: Nash and log-Sobolev inequalities*, Annals of Probability, Vol. **39** No. **3**, 2011, pp. 1161 - 1203.
- [17] L. Zacharias, T. Nguyen, Y. Kovchegov, K. Bradford, *An Adiabatic Approach to Analysis of Time-Inhomogeneous Markov chains: a Queueing Policy Application*, GLOBECOM, 2012.