ROCKY MOUNTAIN JOURNAL OF MATHEMATICS
Vol. , No. , YEAR
https://doi.org/rmj.YEAR..PAGE

STABLE ADIABATIC TIMES FOR A CONTINUOUS EVOLUTION OF MARKOV CHAINS

KYLE BRADFORD

ABSTRACT. This paper extends the concept of the adiabatic time to the stable adiabatic time. As $\varepsilon \searrow 0$ our paper shows that, for a continuous evolution of particular Markov chains, the stable adiabatic time is bounded above by a constant multiple of the square of the biggest mixing time over the evolution divided by ε. Furthermore, we showed that this bound is optimal.

1. INTRODUCTION

In the realm of Markov chains the mixing time serves as a key quantity for assessing the convergence of irreducible and aperiodic time-homogeneous Markov chains [1, 13, 15]. In this context, denote $|\cdot|_{TV}$ as the total variation norm and $|\cdot|_k$ as the $\ell^k(\mathbb{R}^n)$ norm.

Definition 1. *For $\varepsilon > 0$ 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:*

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

While time-homogeneous Markov chains have undergone extensive scrutiny, the stability of time-inhomogeneous Markov chains have less attention. Some literature mentions stability metrics for time-inhomogeneous Markov chains[16, 17, 18]. This paper discusses specific time-inhomogeneous Markov chains and a related metric of stability. These Markov chains take inspiration from the simulated annealing algorithm for finding the global extremum of a function. A greedy algorithm will search locally to make improved estimates of the extremum, but this algorithm may only find a local extremum. Algorithms that use metaheuristics allow for the immediate estimates to be worse in an effort to find the global extremum, but they typically take longer to stabilize to the solution. The simulated annealing algorithm uses a temperature parameter to transition between these two types of algorithms. Annealing starts with a high temperature and is decreased at each step following an annealing schedule. The transition that takes place during the allotted time period determines the global extremum.

This paper defines a continuous path function in a matrix space consisting of finite dimensional, irreducible and aperiodic probability transition matrices. A parameter $T \in \mathbb{N}$ will be defined to be sufficiently large comparable to the temperature parameter of an annealing schedule. The time-inhomogeneous Markov chain at time k has a probability transition matrix defined by evaluating the

1991 *Mathematics Subject Classification.*
Key words and phrases. time-inhomogeneous Markov chain, mixing time, stability, adiabatic time.

continuous path function at time k/T for $0 \le k \le T$ and by evaluating the function at 1 for $k > T$. The initial transition matrix corresponds to the annealing schedule at a high temperature, and the final transition matrix corresponds to the annealing schedule at the end of the time budget. The appropriate selection of T helps define the related metric of stability that is called the stable adiabatic time.

This article continues the effort in [3, 4, 11] to bound the stable adiabatic time of an evolving, time-inhomogeneous Markov chain by a function of the largest mixing time over the entire evolution. Specifically, this paper makes three important contributions: 1) finding an exact bound rather than an asymptotic bound, 2) finding a tighter, optimal bound of the stable adiabatic time and 3) expanding the types of evolutions to include all continuous transitions in the appropriate matrix space. Some of the strongest applications of the adiabatic time and the stable adiabatic time come from quantum physics and quantum computation, namely, the quantum adiabatic theorem from physics [7, 10] and quantum adiabatic computing [12]. There is a strong presence of adiabatic processes in optimization algorithms in queueing systems [6], network design [14] and network performance [19]. There is also an application to the stability of an Ising model with Glauber dynamics [3]. Many of these applications were discussed in detail in previous works. For example, the quantum adiabatic theorem was discussed in [2, 3, 4, 11] and the quantum computation applications were discussed in [4]. In [4] the time-inhomogeneous Markov chain was specifically governed by a convex-combination evolution of two irreducible, aperiodic probability transition matrices. In particular there were matrices $\mathbf{P}_0$ and $\mathbf{P}_1$ and $\mathbf{P}_t = (1-t)\mathbf{P}_0 + t\mathbf{P}_1$. Given a large integer T the probability transition matrix at time $k \le T$ for the time-inhomogeneous Markov chain was $\mathbf{P}_{\frac{k}{T}}$. Naturally, if stochastic matrices $\mathbf{P}_0$ and $\mathbf{P}_1$ are both irreducible and aperiodic, then $\mathbf{P_s}$ is both irreducible and aperiodic for $s \in [0,1]$. This allows for a definition of the mixing time for each $s \in [0,1]$. Taking the supremum of all of these mixing times is one of the ways that one can discuss stability for the time-inhomogeneous Markov chains with probability transition matrices $\mathbf{P}_{\frac{\mathbf{k}}{\mathbf{T}}}$. The following definition makes this formal.

Definition 2. *For $\varepsilon > 0$ the <u>largest mixing time</u> of a time-inhomogeneous, discrete-time Markov chain governed by <u>a convex-combination evolution</u> between the irreducible and aperiodic $\mathbf{P_0}$ and $\mathbf{P_1}$*

$$t_{mix}(\mathbf{P}_0, \mathbf{P}_1, \varepsilon) = \sup_{s \in [0,1]} \{t_{mix}(\mathbf{P}_s, \varepsilon)\}. \tag{2}$$

This paper has, thus far, mentioned the stable adiabatic time without stating the formal definition. Now there is enough background information to make this definition for convex-combination evolutions. This was the main object of study in [4] and will motivate the analogue that we will use in this paper.

Definition 3. *For $\varepsilon > 0$ the <u>stable adiabatic time</u> of a time-inhomogeneous, discrete-time Markov chain governed by <u>a convex-combination evolution</u> between the irreducible and aperiodic $\mathbf{P_0}$ and $\mathbf{P_1}$, which has unique stationary distribution $\pi_{\frac{\mathbf{k}}{\mathbf{T}}}$ for the probability transition matrix $\mathbf{P}_{\frac{\mathbf{k}}{\mathbf{T}}}$, is defined as :*

$$t_{sad}(\mathbf{P_0}, \mathbf{P_1}, \varepsilon) = \inf\{T \in \mathbb{N} : \|\pi_\mathbf{0}\mathbf{P}_{\frac{\mathbf{1}}{\mathbf{T}}} \cdots \mathbf{P}_{\frac{\mathbf{k}}{\mathbf{T}}} - \pi_{\frac{\mathbf{k}}{\mathbf{T}}}\|_{TV} < \varepsilon \text{ for } 1 \le k \le T\}. \tag{3}$$

The stable adiabatic time is another type of stability for these kinds of time-inhomogeneous Markov chains. It is natural to ask how the two previous definitions compare. This was discussed in [4] for these specific convex-combination evolutions. The following asymptotic result was discovered in [4] relating the stable adiabatic time and the largest mixing time.

Theorem 1. *Given a time-inhomogeneous, discrete-time Markov chain governed by a convex-combination evolution between the irreducible and aperiodic* $\mathbf{P_0}$ *and* $\mathbf{P_1}$*, for any* $\varepsilon > 0$,

$$t_{sad}(\mathbf{P_0}, \mathbf{P_1}, \varepsilon) = O\left(\frac{t_{mix}^4(\mathbf{P}_0, \mathbf{P}_1, \varepsilon/2)}{\varepsilon^3}\right) \text{ as } \varepsilon \searrow 0. \tag{4}$$

The main goal of this paper is to expand the types of evolutions that can take place. To elaborate, first let $\mathscr{M}_n([0,1])$ be the collection of all $n \times n$ matrices with entries in $[0,1]$. Define $\mathscr{P}_n = \{\mathbf{P} \in \mathscr{M}_n([0,1]) : \mathbf{P1} = \mathbf{1}\}$ where $\mathbf{1}$ is the n dimensional column vector with all entries 1 and define

$$\mathscr{P}_n^{ia} = \{\mathbf{P} \in \mathscr{P}_n : \mathscr{P} \text{ is irreducible and aperiodic}\}.$$

To describe continuity in this matrix space the standard matrix norm will be used. Specifically for a matrix $\mathbf{M}$ the matrix norm is defined as $\|\mathbf{M}\| = \max_\nu \|\nu M\|_1$ where the maximum is taken over all probability distributions ν. This paper considers continuous functions $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ with respect to the matrix norm to build the more general types of evolutions. One can now allow the time-inhomogeneous Markov chains to be governed by a continuous evolution defined through the function $\mathbf{P}$. Given a large integer T the probability transition matrix at time $k \le T$ for the time-inhomogeneous Markov chain was $\mathbf{P}\left(\frac{k}{T}\right)$. Because all probability transition matrices are in $\mathscr{P}_n^{ia}$ a mixing time exists for all $s \in [0,1]$. The supremum can be taken again to make a metric for stability for time-inhomogeneous Markov chains governed by these continuous evolutions. Note the difference between this definition and Definition 2.

Definition 4. *For* $\varepsilon > 0$ *the largest mixing time of a time-inhomogeneous, discrete-time Markov chain governed by a continuous evolution in* $\mathscr{P}_n^{ia}$*, written as* $t_{mix}(\mathbf{P}_{sup}, \varepsilon)$*, is defined as follows:*

$$t_{mix}(\mathbf{P}_{sup}, \varepsilon) = \sup_{s \in [0,1]} \{t_{mix}(\mathbf{P}(s), \varepsilon)\}. \tag{5}$$

Finally the version of the stable adiabatic time used in this paper can be introduced. The key difference in Definition 3 is the type of evolution. This version of the stable adiabatic time allows for a more general, continuous evolution in $\mathscr{P}_n^{ia}$.

Definition 5. *For* $\varepsilon > 0$ *the stable adiabatic time of a time-inhomogeneous, discrete-time Markov chain governed by a continuous evolution in* $\mathscr{P}_n^{ia}$*, written as* $t_{sad}(\mathbf{P}, \varepsilon)$*, is defined as follows:*

$$t_{sad}(\mathbf{P}, \varepsilon) = \inf\left\{T \in \mathbb{N} : \left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\cdots\mathbf{P}\left(\frac{k}{T}\right) - \pi\left(\frac{k}{T}\right)\right\|_{TV} < \varepsilon \right. \\ \left. \text{for } 1 \le k \le T\right\}. \tag{6}$$

With these definitions formally laid out it can be said that the purpose of this paper is to find a relationship between $t_{sad}(\mathbf{P}, \varepsilon)$ and $t_{mix}(\mathbf{P}_{sup}, \varepsilon)$ in an analogous way as Theorem 1. The machinery in this paper allows for an optimal result. The rest of the paper is organized as follows: Section 2 introduces the necessary background information to allow for a succinct proof of the main result. Section 3 gives the main result of the paper and gives a detailed proof of the main result. Section 4 gives a context of the importance of the result and additional proofs and argumentation is outlined in Section 5.

2. Supporting Material

When one makes a proof about continuous function spaces, a common technique involves using a dense subset known as the Lipschitz continuous function with finite Lipschitz constant. This section introduces two important propositions that aid the proof of the main result. For both the first proposition and the main result in Section 3 using Lipschitz continuous functions allows for a keen insight as to what commands these time-inhomogeneous Markov chains governed by a continuous evolution. In this matrix space the following definition of a Lipschitz continuous function is used.

Definition 6. *A function $\mathbf{P}^* : [0,1] \to \mathscr{M}_n([0,1])$ is Lipschitz if there exists a positive constant L, called the Lipschitz constant, so that*

$$\|\mathbf{P}^*(x) - \mathbf{P}^*(y)\| \leq L|x-y| \tag{7}$$

for $x,y \in [0,1]$.

The function $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ creates a function $\pi : [0,1] \to \mathbb{R}^n$. By definition $\mathbf{P}$ is continuous with respect to the matrix norm, so a natural question is whether π is a continuous function with respect to the total variation norm. The following proposition declares that it is. This in and of itself is not that surprising, but the nature of how it is continuous gives information that will be necessary to prove the main result. The following two propositions were introduced in [4].

Proposition 1. *Let $\sigma = \inf_{s\in[0,1]}\{\sigma(s)\}$ where $\sigma(s)$ is the smallest nonzero singular value of $\mathbb{I} - \mathbf{P}(s)$. If $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ is a continuous function with respect to the matrix norm, then $\pi : [0,1] \to \mathbb{R}^n$ is uniformly continuous with respect to the total variation norm. In particular, for $\varepsilon > 0$ there exists a positive constant L such that for $s \in [0,1]$ and*

$$\delta = \frac{\varepsilon\sigma}{3Ln^{3/2}}, \tag{8}$$

$t \in \{[0,1] : |t-s| \leq \delta\}$ implies that $\|\pi(t) - \pi(s)\|_{TV} \leq \varepsilon$.

Notice that in the above proposition the continuity depends on the smallest nonzero singular value of the function $\mathbf{P}$ throughout the entire evolution. This value σ has information relating to the largest mixing time of $\mathbf{P}$ throughout the entire evolution. The following proposition makes this point.

Proposition 2. *Let $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ be a continuous function with respect to the matrix norm. Let $\sigma = \inf_{s\in[0,1]}\{\sigma(s)\}$ where $\sigma(s)$ is the smallest nonzero singular value of $\mathbb{I} - \mathbf{P}(s)$. Given $\varepsilon > 0$,*

$$\frac{1 - 2\sqrt{n}\varepsilon}{\sigma} \leq t_{mix}(\mathbf{P}_{sup}, \varepsilon). \tag{9}$$

Instead of including a proof of Proposition 2 note that the proof follows rather directly from a similar argument in [4]. In this paper one can find a similar relationship between the smallest nonzero singular value of a matrix and its mixing time. Here the only thing to note is that the mixing time of time-homogeneous Markov chains associated with the smallest nonzero singular value is smaller than the supremum of all mixing times throughout the entire evolution. This provides all the necessary background to approach our main result. This result is now addressed in Section 3.

3. Main Result

The main result of this paper is given in the following theorem and proven in this section. It will provide the necessary analogue for the bound on the stable adiabatic time for time-inhomogeneous Markov chains governed by a continuous evolution by a function of the largest mixing time over the entire evolution. Note that this result differs from Theorem 1 by not being an asymptotic result and having a lower power of the largest mixing time bound the stable adiabatic time. After this theorem is proven, the impact of the result will be discussed in Section 4.

Theorem 2. *Given a time-inhomogeneous, discrete-time Markov chain governed by a continuous evolution in $\mathscr{P}_n^{ia}$, for $0 < \varepsilon < \frac{1}{2\sqrt{n}}$ and $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ a continuous function with respect to the matrix norm we have that*

$$t_{sad}(\mathbf{P}, \varepsilon) \leq \frac{3n^{3/2} L t_{mix}^2(\mathbf{P}_{sup}, \varepsilon)}{(1 - 2\sqrt{n}\varepsilon)\varepsilon} \tag{10}$$

Proof. Recall that the space of Lipschitz continuous functions from $[0,1]$ to $\mathscr{P}_n^{ia}$ with finite Lipschitz constant is dense in the space of continuous functions from $[0,1]$ to $\mathscr{P}_n^{ia}$. This implies that one can find a Lipschitz continuous function $\mathbf{P}^* : [0,1] \to \mathscr{P}_n^{ia}$ with Lipschitz constant L such that

$$\|\mathbf{P}(t) - \mathbf{P}^*(t)\| \leq \frac{\varepsilon}{4t_{mix}(\mathbf{P}_{sup}, \varepsilon/2)}$$

for all $t \in [0,1]$.

The goal of this proof is to select a value of T large enough so that

$$\left\| \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right) - \pi\left(\frac{k}{T}\right) \right\|_{TV} \leq \varepsilon$$

for $1 \leq k \leq T$.

Let

$$T = \frac{3n^{3/2} L t_{mix}^2(\mathbf{P}_{sup}, \varepsilon/2)}{(1 - 2\sqrt{n}\varepsilon)\varepsilon}.$$

At this point the proof is decomposed into two parts.

<u>Part 1. Assume that $k \geq t_{mix}(\mathbf{P}_{sup}, \varepsilon/2)$</u>

Let $N = k - t_{mix}(\mathbf{P}_{sup}, \varepsilon/2)$.

Observe that

$$\begin{aligned}
&\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right) \\
&\qquad = \nu_N\left(\mathbf{P}\left(\frac{k}{T}\right) + \left(\mathbf{P}\left(\frac{N+1}{T}\right) - \mathbf{P}\left(\frac{k}{T}\right)\right)\right)\mathbf{P}_{N+2}^{\circ} \\
&\qquad = \nu_N\mathbf{P}\left(\frac{k}{T}\right)\mathbf{P}_{N+2}^{\circ} + \nu_N\left(\mathbf{P}\left(\frac{N+1}{T}\right) - \mathbf{P}\left(\frac{k}{T}\right)\right)\mathbf{P}_{N+2}^{\circ}.
\end{aligned}$$

where $\nu_N = \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{N}{T}\right)$, $\mathbf{P}^{\circ}_{\ell} = \mathbf{P}\left(\frac{\ell}{T}\right)\cdots\mathbf{P}\left(\frac{k}{T}\right)$.
By continuing this process for $\mathbf{P}\left(\frac{i}{T}\right)$ for $i \geq N+2$, it can be shown that

$$\begin{aligned}
&\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)\\
&\quad = \nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{k-N}\\
&\quad + \sum_{\ell=0}^{k-N-2}\nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{\ell}\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)\mathbf{P}^{\circ}_{N+2+\ell}.
\end{aligned}$$

By the triangle inequality, it can be shown that

$$\begin{aligned}
&\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\
&\quad \leq \left\|\nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{k-N}-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\
&\quad + \sum_{\ell=0}^{k-N-2}\left\|\nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{\ell}\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)\mathbf{P}^{\circ}_{N+2+\ell}\right\|_{TV}.
\end{aligned}$$

Because $2\|\mu-\nu\|_{TV} = \|\mu-\nu\|_1$ whenever μ and ν are probability distributions,

$$\begin{aligned}
&\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\
&\quad \leq \left\|\nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{k-N}-\pi\left(\frac{k}{T}\right)\right\|_{TV} + \frac{1}{2}\sum_{\ell=0}^{k-N-2}\|\nu^{\circ}_{\ell}\mathbf{P}^{\circ}_{N+2+\ell}\|_1
\end{aligned}$$

where $\nu^{\circ}_{\ell} = \nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{\ell}\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)$.
Notice that for $0 \leq \ell \leq k-N-2$, $\mathbf{P}^{\circ}_{N+2+\ell}$ is a probability transition matrix. This will imply that

$$\begin{aligned}
\|\nu^{\circ}_{\ell}\mathbf{P}^{\circ}_{N+2+\ell}\|_1 &= \sum_{j=1}^{n}\left|\sum_{i=1}^{n}\nu^{\circ}_{\ell}(i)\,\mathbf{P}^{\circ}_{N+2+\ell}(i,j)\right|\\
&\leq \sum_{j=1}^{n}\sum_{i=1}^{n}\left|\nu^{\circ}_{\ell}(i)\right|\mathbf{P}^{\circ}_{N+2+\ell}(i,j)\\
&= \sum_{i=1}^{n}\left|\nu^{\circ}_{\ell}(i)\right|\sum_{j=1}^{n}\mathbf{P}^{\circ}_{N+2+\ell}(i,j)\\
&= \sum_{i=1}^{n}\left|\nu^{\circ}_{\ell}(i)\right|\\
&= \|\nu^{\circ}_{\ell}\|_1.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\left\| \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad \le \left\| \nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{k-N}-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad + \frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\| \nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{\ell}\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)\right\|_1 .
\end{aligned}$$

It is clear that $\nu_N\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{\ell}$ is a probability vector for $0\le\ell\le k-N-2$, so naturally

$$\begin{aligned}
&\left\| \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad \le \max_{\nu}\left\| \nu\left(\mathbf{P}\left(\frac{k}{T}\right)\right)^{k-N}-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad + \frac{1}{2}\sum_{\ell=0}^{k-N-2}\max_{\nu}\left\| \nu\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)\right\|_1
\end{aligned}$$

where the maximum is taken over all probability vectors ν.
Because $k-N=t_{mix}(\mathbf{P}_{sup},\varepsilon/2)\ge t_{mix}(\mathbf{P}\left(\frac{k}{T}\right),\varepsilon/2)$, it is easy to see that

$$\begin{aligned}
&\left\| \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad \le \frac{\varepsilon}{2}+\frac{1}{2}\sum_{\ell=0}^{k-N-2}\max_{\nu}\left\| \nu\left(\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right)\right\|_1 .
\end{aligned}$$

Observe that the terms in the sum of the right hand side of the inequality are now the matrix norms for the matrices $\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)$ for $0\le\ell\le k-N-2$. This would imply that

$$\begin{aligned}
&\left\| \pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\
&\quad \le \frac{\varepsilon}{2}+\frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\| \mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right\| .
\end{aligned}$$

By adding and subtracting the same value to the above inequality and then using the triangle inequality

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\ \leq \frac{\varepsilon}{2}+\frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\|\mathbf{P}^*\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}^*\left(\frac{k}{T}\right)\right\| \\ +\frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\|\mathbf{P}\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}^*\left(\frac{N+1+\ell}{T}\right)\right\| \\ +\frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\|\mathbf{P}^*\left(\frac{k}{T}\right)-\mathbf{P}\left(\frac{k}{T}\right)\right\|.\end{aligned}$$

Using the density of the Lipschitz continuous functions with finite Lipschitz constant in the continuous function space

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\ \leq \frac{\varepsilon}{2}+\frac{1}{2}\sum_{\ell=0}^{k-N-2}\left\|\mathbf{P}^*\left(\frac{N+1+\ell}{T}\right)-\mathbf{P}^*\left(\frac{k}{T}\right)\right\| \\ +\sum_{\ell=0}^{k-N-2}\frac{\varepsilon}{4t_{mix}\left(\mathbf{P}_{sup},\varepsilon/2\right)}.\end{aligned}$$

Because $\mathbf{P}^*:[0,1]\to\mathscr{P}_n^{ia}$ is a Lipschitz continuous function with Lipschitz constant L, it can be shown that

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\ \leq \frac{\varepsilon}{2}+\frac{L}{2}\sum_{\ell=0}^{k-N-2}\left|\frac{N+1+\ell}{T}-\frac{k}{T}\right| \\ +\sum_{\ell=0}^{k-N-2}\frac{\varepsilon}{4t_{mix}\left(\mathbf{P}_{sup},\varepsilon/2\right)}.\end{aligned}$$

After relabeling the sum

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV} \\ \leq \frac{\varepsilon}{2}+\frac{L}{4T}(k-N-1)(k-N) \\ +\frac{\varepsilon}{4t_{mix}\left(\mathbf{P}_{sup},\varepsilon/2\right)}(k-N-1).\end{aligned}$$

Because $k-N=t_{mix}(\mathbf{P}_{sup},\varepsilon/2)$

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\ \leq\frac{3\varepsilon}{4}+\frac{L}{4T}t_{mix}^{2}\left(\mathbf{P}_{sup},\varepsilon/2\right).\end{aligned}$$

T was selected to be large enough. In fact,

$$T=\frac{3n^{3/2}Lt_{mix}^{2}\left(\mathbf{P}_{sup},\varepsilon/2\right)}{(1-2\sqrt{n}\varepsilon)\varepsilon}\geq\frac{Lt_{mix}^{2}(\mathbf{P}_{sup},\varepsilon/2)}{\varepsilon}.$$

Finally it is shown that

$$\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\leq\varepsilon.$$

Part 2. Assume that $k<t_{mix}\left(\mathbf{P}_{sup},\varepsilon/2\right)$

First notice that

$$\begin{aligned}\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\\ =\left(\pi(0)-\pi\left(\frac{1}{T}\right)\right)\mathbf{P}_{1}^{\circ}+\pi\left(\frac{1}{T}\right)\mathbf{P}_{2}^{\circ}-\pi\left(\frac{k}{T}\right).\end{aligned}$$

Repeating this process, it can be shown that

$$\begin{aligned}\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\\ =\sum_{j=1}^{k}\left(\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right)\mathbf{P}_{j}^{\circ}.\end{aligned}$$

By the triangle inequality

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\ \leq\sum_{j=1}^{k}\left\|\left(\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right)\mathbf{P}_{j}^{\circ}\right\|_{TV}.\end{aligned}$$

Because $\mathbf{P}_{j}^{\circ}$ is a probability transition matrix

$$\left\|\left(\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right)\mathbf{P}_{j}^{\circ}\right\|_{TV}\leq\left\|\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right\|_{TV}.$$

This will imply that

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\ \leq \sum_{j=1}^{k}\left\|\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right\|_{TV}.\end{aligned}$$

Using Proposition 1 it is clear that as long as

$$T \geq \frac{3Ln^{3/2}t_{mix}(\mathbf{P}_{sup},\varepsilon/2)}{\varepsilon\sigma}$$

one has that

$$\left\|\pi\left(\frac{j-1}{T}\right)-\pi\left(\frac{j}{T}\right)\right\|_{TV} \leq \frac{\varepsilon}{t_{mix}(\mathbf{P}_{sup},\varepsilon/2)}.$$

This would imply that

$$\begin{aligned}\left\|\pi(0)\mathbf{P}\left(\frac{1}{T}\right)\mathbf{P}\left(\frac{2}{T}\right)\cdots\mathbf{P}\left(\frac{k-1}{T}\right)\mathbf{P}\left(\frac{k}{T}\right)-\pi\left(\frac{k}{T}\right)\right\|_{TV}\\ \leq \frac{k\varepsilon}{t_{mix}(\mathbf{P}_{sup},\varepsilon/2)}\\ \leq \varepsilon.\end{aligned}$$

Proposition 2 implies that

$$T = \frac{3Ln^{3/2}t_{mix}^{2}(\mathbf{P}_{sup},\varepsilon/2)}{(1-2\sqrt{n}\varepsilon)\varepsilon} \geq \frac{3Ln^{3/2}t_{mix}(\mathbf{P}_{sup},\varepsilon/2)}{\varepsilon\sigma}.$$

This completes our proof. □

4. Conclusion

Notice that an immediate consequence of Theorem 2 is that there is a tighter asymptotic bound when compared to the previous result in Theorem 1. Also convex-combination evolutions are a specific type of continuous evolution, so the class of evolutions is much broader. The following corollary sums up these two points.

Corollary 1. *Given a time-inhomogeneous, discrete-time Markov chain governed by a continuous evolution in $\mathscr{P}_n^{ia}$, for $\varepsilon > 0$ and $\mathbf{P} : [0,1] \to \mathscr{P}_n^{ia}$ a continuous function with respect to the matrix norm we have that*

$$t_{sad}(\mathbf{P},\varepsilon) = O\left(\frac{t_{mix}^{2}(\mathbf{P}_{sup},\varepsilon/2)}{\varepsilon}\right) \text{ as } \varepsilon \searrow 0. \tag{11}$$

To answer a final question, this bound is optimal. To show this, it suffices to find one specific function $\mathbf{P}$ such that the stable adiabatic time is exactly a constant multiplied my the square of the largest mixing time divided by ε. For this one can consider a convex-combination evolution. Here let

$$\mathbf{P}_0 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{P}_1 = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Notice that $\mathbf{P}_0$ and $\mathbf{P}_1$ do not create irreducible Markov chains; however, if we were to slightly alter the matrices, then we can preserve the structure of the convex combination, the approximate mixing times throughout the evolution and the stable adiabatic time. For example, define for a small value δ the values $a = \delta/(n-1)$ and $b = 1-\delta-a$. Letting $\mathbf{J}_n$ be the square matrix of all ones of dimension n, we can define two new probability transition matrices $\mathbf{P}_0^* = a\mathbf{J}_n + b\mathbf{P}_0$ and $\mathbf{P}_1^* = a\mathbf{J}_n + b\mathbf{P}_1$. The convex combination of $\mathbf{P}_0^*$ and $\mathbf{P}_1^*$ form an evolution in $\mathscr{P}_n^{ia}$.

As shown in [3] the adiabatic time for a convex-combination evolution from $\mathbf{P}_0$ to $\mathbf{P}_1$ is of the asymptotic order of the square of the largest mixing time divided by ε. The only inequality that must hold for the adiabatic time, rather than the stable adiabatic time, is for $\|\pi_\mathbf{0}\mathbf{P}_{\frac{1}{T}}\cdots\mathbf{P}_{\frac{T}{T}} - \pi_{\frac{T}{T}}\|_{TV} < \varepsilon$. Naturally, for all the other inequalities to hold $\|\pi_\mathbf{0}\mathbf{P}_{\frac{1}{T}}\cdots\mathbf{P}_{\frac{k}{T}} - \pi_{\frac{k}{T}}\|_{TV} < \varepsilon$ where $1 \le k < T$ one must select a value of T at least as large as a constant multiplied by the square of the largest mixing time divided by ε. Because the small alteration of the convex-combination evolution preserves all the stability of the original evolution, the result in this paper guarantees that this value of T must be of the same asymptotic order for the convex-combination evolution in $\mathscr{P}_n^{ia}$. This shows that the result from Corollary 1 is optimal.

5. PROOFS

5.1. *Proof of Proposition 1.* To begin, consider the creation of an orthonormal basis of eigenvectors associated with $(\mathbb{I}-\mathbf{P}(s))(\mathbb{I}-\mathbf{P}(s))^T$ with respect to $\|\cdot\|_2$ through a singular value decomposition of $(\mathbb{I}-\mathbf{P}(s))$, where $s \in [0,1]$.
Here let $\sigma_1(s) \ge \cdots \ge \sigma_{n-1}(s) = \sigma(s)$ be the positive singular values of $(\mathbb{I}-\mathbf{P}(s))$ with respect to the Euclidean inner product. This implies that there exists an orthonormal basis $\{\mathbf{v_1}(s), \cdots, \mathbf{v_n}(s)\}$ such that $\mathbf{v_j}(s)(\mathbb{I}-\mathbf{P}(s))(\mathbb{I}-\mathbf{P}(s))^T = \sigma_j^2(s)\mathbf{v_j}(s)$ for $1 \le j \le n-1$ and $\mathbf{v_n}(s)(\mathbb{I}-\mathbf{P}(s))(\mathbb{I}-\mathbf{P}(s))^T = \mathbf{0}$.
Here $\mathbf{v_n}(s) = \pi(s)/\|\pi(s)\|_2$.
To show continuity at s let $\varepsilon > 0$ and first notice that for any $t \in [0,1]$, $(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s)) = \pi(t)(\mathbf{P}(t)-\mathbf{P}(s))$.
Using the Euclidean norm, it can easily be seen that if $\mathbf{P}(t) \ne \mathbf{P}(s)$ and $t \ne s$, then

$$\frac{\|(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s))\|_2}{\|\pi(t)-\pi(s)\|_2} = \frac{\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2}{\|\pi(t)-\pi(s)\|_2}.$$

Throughout this proof $\langle\cdot,\cdot\rangle$ we denote the Euclidean inner product. For $1 \leq j \leq n$ let $c_j(s,t) = \langle \pi(t) - \pi(s), \mathbf{v_j}(s)\rangle$. Then $\pi(t) - \pi(s) = \sum_{j=1}^n c_j(s,t)\mathbf{v_j}(s)$. This will imply that

$$\begin{aligned}
&\frac{\|(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s))\|_2^2}{\|\pi(t)-\pi(s)\|_2^2} \\
&\qquad = \frac{\langle(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s)),(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s))\rangle}{\langle\pi(t)-\pi(s),\pi(t)-\pi(s)\rangle} \\
&\qquad = \frac{\langle\pi(t)-\pi(s),(\pi(t)-\pi(s))(\mathbb{I}-\mathbf{P}(s))(\mathbb{I}-\mathbf{P}(s))^T\rangle}{\langle\pi(t)-\pi(s),\pi(t)-\pi(s)\rangle} \\
&\qquad = \frac{\langle\sum_{j=1}^n c_j(s,t)\mathbf{v_j}(s),\sum_{j=1}^{n-1}\sigma_j^2(s)c_j(s,t)\mathbf{v_j}(s)\rangle}{\langle\sum_{j=1}^n c_j(s,t)\mathbf{v_j}(s),\sum_{j=1}^n c_j(s,t)\mathbf{v_j}(s)\rangle} \\
&\qquad = \frac{\sum_{j=1}^{n-1}\sigma_j^2(s)c_j^2(s,t)}{\sum_{j=1}^n c_j^2(s,t)} \\
&\qquad \geq \sigma_{n-1}^2(s)\frac{\sum_{j=1}^{n-1}c_j^2(s,t)}{\sum_{j=1}^n c_j^2(s,t)} \\
&\qquad = \sigma_{n-1}^2(s)\left(1-\frac{c_n^2(s,t)}{\sum_{j=1}^n c_j^2(s,t)}\right) \\
&\qquad = \sigma_{n-1}^2(s)\left(1-\left(\frac{<\pi(t)-\pi(s),\mathbf{v_n}(s)>}{\|\pi(t)-\pi(s)\|_2}\right)^2\right).
\end{aligned}$$

Letting $\mathbf{w}(s,t) = (\pi(t)-\pi(s))/\|\pi(t)-\pi(s)\|_2$, the above inequality can be written as

$$\sigma_{n-1}^2(s)\left(1-(\langle\mathbf{w}(s,t),\mathbf{v_n}(s)\rangle)^2\right) \leq \frac{\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2^2}{\|\pi(t)-\pi(s)\|_2^2}.$$

Because $\mathbf{w}(s,t)$ and $\mathbf{v_n}(s)$ are unit vectors, the fact that

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

can be used to show that

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

and the fact that

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

can be used to show that

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

From this it is clear that $1-(\langle \mathbf{w}(s,t), \mathbf{v_n}(s)\rangle)^2 = \|\mathbf{w}(s,t)-\mathbf{v_n}(s)\|_2^2 \cdot \|\mathbf{w}(s,t)+\mathbf{v_n}(s)\|_2^2/4$. Plugging this into the previous equation we obtain

$$\|\pi(t)-\pi(s)\|_2 \le \frac{2\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2}{\sigma_{n-1}(s)\|\mathbf{w}(s,t)-\mathbf{v_n}(s)\|_2 \cdot \|\mathbf{w}(s,t)+\mathbf{v_n}(s)\|_2}.$$

Notice that $\langle \mathbf{w}(s,t), \mathbf{1}\rangle/\sqrt{n} = 0$ and $\langle \mathbf{v_n}(s), \mathbf{1}\rangle/\sqrt{n} = 1/(\sqrt{n}\|\pi(s)\|_2)$ for all $t \in [0,1]$. Because these are the scalar components of the projections of $\mathbf{w}(s,t)$ and $\mathbf{v_n}(s)$ onto $\mathbf{1}$ respectively, it can be shown that the minimum possible value for $\|\mathbf{w}(s,t)-\mathbf{v_n}(s)\|_2$ and $\|\mathbf{w}(s,t)+\mathbf{v_n}(s)\|_2$ is at least $1/(\sqrt{n}\|\pi(s)\|_2)$. This shows that

$$\begin{aligned}\|\pi(t)-\pi(s)\|_2 &\le \frac{2n\|\pi(s)\|_2^2 \cdot \|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2}{\sigma_{n-1}(s)} \\ &\le \frac{2n\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2}{\sigma_{n-1}(s)} \\ &= \frac{2n\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_2}{\sigma(s)}.\end{aligned}$$

Let $\sigma = \min_{s\in[0,1]}\{\sigma(s)\}$.
Again for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{x}$ and $\mathbf{y}$ are probability measures, it is understood that

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

This will imply that

$$\begin{aligned}\|\pi(t)-\pi(s)\|_{TV} &\le \frac{n^{3/2}\|\pi(t)(\mathbf{P}(t)-\mathbf{P}(s))\|_1}{\sigma} \\ &\le \frac{n^{3/2}\max_\nu \|\nu(\mathbf{P}(t)-\mathbf{P}(s))\|_1}{\sigma}\end{aligned}$$

where the maximum is taken over all vectors ν such that $\|\nu\|_1 = 1$.
Using the matrix norm notation one can conclude that

$$\|\pi(t)-\pi(s)\|_{TV} \le \frac{n^{3/2}\|\mathbf{P}(t)-\mathbf{P}(s)\|}{\sigma}.$$

Notice that the space of Lipschitz continuous functions mapping $[0,1]$ to $\mathscr{P}_n^{ia}$ are dense in the space of continuous functions mapping $[0,1]$ to $\mathscr{P}_n^{ia}$. This implies that there exists a Lipschitz function $\mathbf{P}^*$ with Lipschitz constant L such that

$$\|\mathbf{P}(t)-\mathbf{P}^*(t)\| \le \frac{\sigma\varepsilon}{3n^{3/2}}$$

for all $t \in [0,1]$.
One can use the triangle inequality along with this density argument to conclude that

$$\begin{aligned}
\|\pi(t)-\pi(s)\|_{TV} &\leq \frac{n^{3/2}\|\mathbf{P}(t)-\mathbf{P}(s)\|}{\sigma} \\
&= \frac{n^{3/2}}{\sigma}\left(\|\mathbf{P}^*(t)-\mathbf{P}^*(s)+\mathbf{P}(t)-\mathbf{P}^*(t)+\mathbf{P}^*(s)-\mathbf{P}(s)\|\right) \\
&\leq \frac{n^{3/2}}{\sigma}\left(\|\mathbf{P}^*(t)-\mathbf{P}^*(s)\|+\|\mathbf{P}(t)-\mathbf{P}^*(t)\|+\|\mathbf{P}^*(s)-\mathbf{P}(s)\|\right) \\
&\leq \frac{n^{3/2}}{\sigma}\left(\|\mathbf{P}^*(t)-\mathbf{P}^*(s)\|+\frac{\sigma\varepsilon}{3n^{3/2}}+\frac{\sigma\varepsilon}{3n^{3/2}}\right) \\
&= \frac{n^{3/2}\|\mathbf{P}^*(t)-\mathbf{P}^*(s)\|}{\sigma}+\frac{2\varepsilon}{3}
\end{aligned}$$

Because $\mathbf{P}^*$ is Lipschitz continuous with Lipschitz constant L, one has that $\|\mathbf{P}^*(t)-\mathbf{P}^*(s)\| \leq L\left|t-s\right|$ for all $t,s \in [0,1]$.
This shows that

$$\|\pi(t)-\pi(s)\|_{TV} \leq \frac{Ln^{3/2}\left|t-s\right|}{\sigma}+\frac{2\varepsilon}{3}.$$

Clearly if $\varepsilon > 0$, then having

$$\left|t-s\right| \leq \delta = \frac{\varepsilon\sigma}{3Ln^{3/2}}$$

implies $\|\pi(t)-\pi(s)\|_{TV} \leq \varepsilon$.
This shows that π is continuous at $s \in [0,1]$. Because one can do this for any $s \in [0,1]$, it is seen that π is continuous with respect to the total variation norm on $[0,1]$. Because δ does not depend on the value of $s \in [0,1]$, it is shown that π is uniformly continuous.

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] K. Bradford, Y. Kovchegov and T. Nguyen, *Stable adiabatic times for Markov chains*, arXiv:1207.4733, 2016.

[5] P. Brémaud, *Markov Chains: Gibbs fields, Monte Carlo Simulation, and Queues*, Springer Science+Business Media Inc. Texts in Applied Mathematics, Vol. **31**, 2010.

[6] D. Nguyen-Huu, T. Duong and T. Nguyen, *Network Protocol Designs: Fast Queuing Policies via Convex Relaxation*, IEEE Transactions on Communications, Vol. **62**, 2014, pp. 182-193.

[7] V.A. Fock, *Selected Works: Quantum Mechanics and Quantum Field Theory*, Chapman & Hall/CRC, 2004.

[8] D.L. Isaacson and R.W. Madsen, *Markov Chains: Theory and Applications*, John Wiley, New York, 1976.

[9] S. Karlin and H.M. Taylor, *A first course in Stochastic Processes*, Academic Press, 1975.

[10] T. Kato, *On the Adiabatic Theorem of Quantum Mechanics*, Journal of t he Physical Society of Japan, Vol. **5**, pp. 435 - 439, 1950.

[11] Y. Kovchegov, *A note on adiabatic theorem for Markov chains*, Statistics and Probability Letters, Vol. **80**, 2010, pp. 186-190.

[12] H. Krovi, M. Ozols and J. Roland, *Adiabatic condition and the quantum hitting time of Markov chains*, Physical Review A **82**, 022333, 2010.

[13] D. Levin, Y. Peres and E. Wilmer, *Markov Chains and Mixing Times*, American Mathematical Society, 2009.

[14] 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.

[15] S. Ross, *Simulation: Fourth Ed.*, Elsevier Academic Press, 2006.

[16] 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.

[17] 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.

[18] 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.

[19] L. Zacharias, T. Nguyen, Y. Kovchegov, K. Bradford, *An Adiabatic Approach to Analysis of Time-Inhomogeneous Markov chains: a Queueing Policy Application*, GLOBECOM, 2012.

Department of Mathematics and Statistics, Georgia Southern University, Statesboro, GA 30458, USA

Email address: kbradford@georgiasouthern.edu