Hoeffding's inequality for general Markov chains with its applications to statistical learning

Jianqing Fan[*], Bai Jiang[*] and Qiang Sun[†]

Abstract

This paper establishes Hoeffding's lemma and inequality for bounded functions of general-state-space and not necessarily reversible Markov chains. The sharpness of these results is characterized by the optimality of the ratio between variance proxies in the Markov-dependent and independent settings. The boundedness of functions is shown necessary for such results to hold in general. To showcase the usefulness of the new results, we apply them for non-asymptotic analyses of MCMC estimation, respondent-driven sampling and high-dimensional covariance matrix estimation on time series data with a Markovian nature. In addition to statistical problems, we also apply them to study the time-discounted rewards in econometric models and the multi-armed bandit problem with Markovian rewards arising from the field of machine learning.

Keywords: Hoeffding's inequality, Markov chain, general state space, Markov chain Monte Carlo.

1 Introduction

Concentration inequalities bound the deviation of the sum of independent random variables from its expectation. They have found numerous applications in statistics, econometrics, machine learning and many other fields. One of the important and fundamental concentration inequalities was discovered by Hoeffding (1963). Hoeffding's lemma asserts that a bounded random variable $Z \in [a, b]$ is sub-Gaussian[1] with variance proxy $\frac{(b-a)^2}{4}$. It follows that the sum of n independent, bounded random variables $Z_i \in [a_i, b_i],\ i = 1, \ldots, n$ is sub-Gaussian with variance proxy $\sum_{i=1}^n \frac{(b_i-a_i)^2}{4}$. Specifically, for any $t \in \mathbb{R}$,

$$\mathbb{E}\left[e^{t\sum_{i=1}^n (Z_i - \mathbb{E}Z_i)}\right] \leq \exp\left(\sum_{i=1}^n \frac{(b_i - a_i)^2}{4} \cdot \frac{t^2}{2}\right). \tag{1}$$

[*]Department of Operations Research and Financial Engineering, Princeton University, 205 Sherred Hall, Princeton, NJ 08544; E-mail: jqfan@princeton.edu, baij@princeton.edu.

[†]Department of Statistical Sciences, University of Toronto, 100 St. George Street, Toronto, ON M5S 3G3, Canada; E-mail: qsun@utstat.toronto.edu.

[1]A random variable Z is sub-Gaussian with variance proxy σ^2 if Z has a finite mean $\mathbb{E}Z$ and $\mathbb{E}[e^{t(Z-\mathbb{E}Z)}] \leq \exp\left(\frac{\sigma^2 t^2}{2}\right)$ for any $t \in \mathbb{R}$.

From (1), the Cramér-Chernoff method (Boucheron et al., 2013) derives Hoeffding's inequality as follows. For any $\epsilon > 0$,

$$\mathbb{P}\left(\sum_{i=1}^{n} Z_i - \sum_{i=1}^{n} \mathbb{E}Z_i > \epsilon\right) \leq \exp\left(-\frac{\epsilon^2}{2\sum_{i=1}^{n}(b_i - a_i)^2/4}\right). \tag{2}$$

Similar bounds apply to the lower deviation $\sum_{i=1}^{n} Z_i - \sum_{i=1}^{n} \mathbb{E}Z_i < -\epsilon$ as well as the two-sided deviation $|\sum_{i=1}^{n} Z_i - \sum_{i=1}^{n} \mathbb{E}Z_i| > \epsilon$, with an additional factor of two in the latter case.

However, the independence assumption on random variables limits the applicability of Hoeffding's inequality and other concentration inequalities in many statistical, econometric and machine learning problems involving Markovian dependence. These problems include Markov chain Monte Carlo (MCMC), time series analyses, economic decision making and reinforcement learning. In this paper, we generalize the classical Hoeffding's lemma and inequality at their full strength for functions of general (general-state-space and not necessarily reversible) Markov chains. For a stationary Markov chain $\{X_i\}_{i \geq 1}$ and bounded functions $f_i : x \mapsto [a_i, b_i]$, we establish analogues of (1) and (2) for Markov-dependent random variables $Z_i = f_i(X_i)$. Our main result, formally stated as Theorem 1, asserts that $\sum_{i=1}^{n} f_i(X_i)$ is sub-Gaussian with variance proxy

$$\frac{1+\lambda}{1-\lambda} \cdot \sum_{i=1}^{n} \frac{(b_i - a_i)^2}{4},$$

where $\lambda \in [0, 1]$ is the norm of the Markov operator acting on the Hilbert space consisting of all squared-integrable and mean-zero functions with respect to the invariant distribution of the Markov chain. The quantity λ measures the temporal dependence of the Markov chain, as $1 - \lambda$ measures the speed of the Markov chain converging from non-stationarity towards stationarity (Rudolf, 2012).

This theorem simplifies to the classical results (1) and (2) in the independent setting, because independent variables can be viewed as functions of i.i.d. uniformly distributed variables on $[0, 1]$, which form a Markov chain with $\lambda = 0$. The corresponding Hoeffding-type inequality generalizes that in Miasojedow (2014) from the time-independent function case, in which $f_1 = f_2 f = \cdots = f_n = f$ are identical, to the time-dependent function case, in which f_i's can be different.

Interestingly, for the time-independent function case in which $f_i = f$, $a_i = a$ and $b_i = b$, we find that the ratio $\frac{1+\lambda}{1-\lambda}$ between variance proxies in the Markov-dependent and independent settings can be sharpened by replacing λ with a smaller quantity $\max\{\lambda_r, 0\}$, where $\lambda_r \in [-\lambda, +\lambda]$ is the maximum spectrum value of the additive reversiblization of the Markov operator (Fill, 1991).

This result, formally stated as Theorem 2, generalizes the Hoeffding-type inequality established by León and Perron (2004) for simple (finite-state-space and reversible) Markov chains to that for general Markov chains. For simple Markov chains considered by León and Perron (2004), λ_r is merely the second largest eigenvalue of the transition probability matrix. This sharper variance proxy ratio $\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$ is then shown to be optimal by appealing to the Central Limit Theorem of Markov chains (Geyer, 1992; Roberts and Rosenthal, 1997; Rosenthal, 2003).

It is also worth noting that the our theorems discover the counterpart of the classical Hoeffding's lemma for Markov chains, which has been missing in the previous literature. This counterpart immediately derives another interesting result that $[f(X_1), \ldots, f(X_n)]^{\mathrm{T}}$ is a sub-Gaussian random vector[2] with variance proxy $\frac{1+\lambda}{1-\lambda} \cdot \frac{(b-a)^2}{4}$.

The general and sharp inequalities presented by Theorems 1-2 enable non-asymptotic analyses of a few statistical, econometric and machine learning problems involving Markovian dependence. We showcase their utilities in five such problems, namely, MCMC estimation, respondent-driven sampling, high-dimensional covariance matrix estimation using Markov-dependent observations, time-discounted reward in econometric models, and the multi-armed bandit problem with Markovian rewards.

First, we derive a non-asymptotic error bound for MCMC estimation. MCMC estimates a complicated integral by averaging Markov chain samples after a burn-in period. This is arguably one of the central techniques of the Bayesian computational methods (Geyer, 1992; Gilks, 2005; Liu, 2008). A sharp concentration inequality is crucial to understand how long the Markov chain should run in order to control the estimation error, and how to allocate computational resources on the burn-in period and the sampling period. Our non-asymptotic analysis suggests that, roughly speaking, $\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}} \cdot n$ Markov chain samples are needed to achieve the same accuracy with n independent samples in the naive Monte Carlo method. We also find that a burn-in period of length $n_0 \gtrsim \log n$ suffices to get a tight concentration of MCMC estimate around the target integral, and thereafter justify the ad-hoc choice of $n_0 = 0.05n$ or $n_0 = 0.1n$ in practice. This insight is obtainable by neither the asymptotic theories of Markov chains nor non-asymptotic bounds on the mean squared error of MCMC (Rudolf, 2012; Łatuszyński et al., 2013).

Next, we conduct a non-asymptotic analysis for respondent-driven sampling (RDS) estimates (Goel and Salganik, 2009, 2010). RDS is a popular method to collect data from a hidden population in disease (e.g., HIV) infection studies (Heckathorn, 1997, 2002). RDS estimate is shown consistent for disease prevalence by the asymptotic theory (Law of Large Number) of Markov chains. Our non-asymptotic analysis reveals that a finite sample size of the logarithmic order of the total population suffices for the success of RDS. To the best of our knowledge, this insight has not been provided by the literature in RDS before.

Third, we consider the high-dimensional covariance matrix estimation problem (Fan et al., 2008; Bickel and Levina, 2008; Rothman et al., 2009; Lam and Fan, 2009; Cai et al., 2010; Cai and Liu, 2011; Fan et al., 2011, 2013) on time series data with a Markovian nature. High-dimensional covariance matrix estimation has found applications in economics,

[2] An n-dimensional random vector $\boldsymbol{Z}$ is sub-Gaussian with variance proxy σ^2 if $\boldsymbol{Z}$ has a finite mean $\mathbb{E}\boldsymbol{Z}$ and $\mathbb{E}[e^{\boldsymbol{t}'(\boldsymbol{Z}-\mathbb{E}\boldsymbol{Z})}] \leq \exp(\sigma^2\|\boldsymbol{t}\|_2^2/2)$ for any $\boldsymbol{t} \in \mathbb{R}^n$.

finance, biology, social networks, and health sciences (Fan et al., 2014, 2016). Although many real high-dimensional data, especially time series data like functional magnetic resonance imaging (fMRI) data and asset return data, consist of temporally dependent samples, high-dimensional covariance matrix estimation methods usually assume the sample independence for convenience of theoretical analyses. To fill in this gap, we suggest a Markov chain model for such temporally dependent data and provide a non-asymptotic analysis for the thresholding high-dimensional covariance matrix estimator in the Markov-dependent setting.

Fourth, we construct a tight non-asymptotic confidence interval for the time-discounted reward, which is commonly-seen in econometric models. In these models, a discount function describes the weights placed on rewards received at different time points and reflects the tendency of people to favor rewards received now over those received in future (Frederick et al., 2002). We consider two commonly-used discount functions, i.e., exponential and hyperbolic discount functions, in the vineyard model of Paulin (2015, Example 3.17). Note that the exponential discount function has also been widely used in reinforcement learning (Sutton and Barto, 2018).

Our last example is the multi-armed bandit problem with Markovian rewards. Researchers have recognized three fundamental forms of multi-armed bandit problems depending on the nature of the reward process: independently and identically distributed (i.i.d.), adversarial, and Markovian (Bubeck and Cesa-Bianchi, 2012). The Markovian form is much less studied than the other two (Tekin and Liu, 2010), primarily because bandit algorithms rely on the concentration of reward draws to identify the optimal arm, but unfortunately powerful concentration inequalities for Markov chains were in lack.

The rest of this paper proceeds as follows. We discuss related literature in the remaining of this section. Section 2 presents our main results in Theorem 1, for time-dependent functions, and Theorem 2, for a time-independent function. Section 3 discusses extensions to unbounded functions and inhomogeneous Markov chains. An impossibility result is given for unbounded functions. Section 4 outlines the proofs of Theorems 1-2. Section 5 applies our theorems to five problems in statistics, econometrics and machine learning. Section 6 concludes with a brief discussion. All technical proofs are collected in the appendix.

Related Literature

Several Hoeffding-type inequalities for Markov chains have been produced by spectral methods, e.g., Gillman (1993); Dinwoodie (1995); León and Perron (2004); Chung et al. (2012); Miasojedow (2014), to name a few. Most of these results except for Miasojedow (2014) hold for finite-state-space Markov chains only. This limitation hinders their applicabilities to many important problems, such as MCMC estimation (Geyer, 1992; Gilks, 2005; Liu, 2008) in which the state space is usually a multi-dimensional real space. In addition, all the aforementioned inequalities do not allow the function to vary in time. However, decision problems in economics and reinforcement learning usually involve temporal discounting and thus need to deal with various functions (Frederick et al., 2002; Sutton and Barto, 2018).

Spectral methods have also been used to derive Bernstein-type inequalities for Markov chains (Lezaud, 1998; Paulin, 2015). Their inequalities mainly work for finite-state-space Markov chains[3]. It is also worth noting that Hoeffding-type and Bernstein-type inequalities are different. The former requires the endpoints of value ranges of random variables only and obtains a subgaussian tail, in contrast the latter takes variances of random variables into account and obtains a mixture of subgaussian and subexponential tails.

Other exponential concentration inequalities exploit the minorization and drift conditions; see e.g. Glynn and Ormoneit (2002); Douc et al. (2011); Adamczak and Bednorz (2015). Other included Kontoyiannis et al. (2005, 2006). These inequalities do not assume non-zero spectral gaps, unlike those by spectral methods, but have less explicit constants or more complicated expressions.

There is another and less related research line for a function of n variables under Markov or other dependent structures. Marton (1996, 1998) pioneered the concentration of measure phenomenon for contracting Markov chains. Further progresses are made by Samson (2000); Chazottes et al. (2007); Redig and Chazottes (2009); Kontorovich and Ramanan (2008), among others.

2 Main Results

As a preparation for presenting theorems, we introduce notation as follows. Let $\mathcal{X}$ and π denote the state space and the invariant measure of a Markov chain $\{X_i\}_{i\geq 1}$. For function $h : \mathcal{X} \to \mathbb{R}$, write $\pi(h) := \int h(x)\pi(dx)$. Let $\mathcal{L}_2(\pi) = \{h : \pi(h^2) < \infty\}$ be the Hilbert space consisting of π-square-integrable functions, and $\mathcal{L}_2^0(\pi) = \{h \in \mathcal{L}_2(\pi) : \pi(h) = 0\}$ be its subspace of π-mean-zero functions. The transition probability kernel of the Markov chain, denoted by P, is viewed as an operator acting on $\mathcal{L}_2(\pi)$. Let $\lambda \in [0, 1]$ be the operator norm of P acting on $\mathcal{L}_2^0(\pi)$. We refer to $1 - \lambda$ as the *absolute spectral gap* of the Markov chain. Let P^* be the adjoint operator of P. We follow Fill (1991) and call $R = (P + P^*)/2$ the additive reversiblization[4] of P. It is known that the spectrum of R acting on $\mathcal{L}_2^0(\pi)$ is contained in the interval $[-\lambda, +\lambda]$ on the real line (Rudolf, 2012). Let $\lambda_r \in [-\lambda, +\lambda]$ denote the rightmost value of this spectrum. We refer to $1 - \lambda_r$ as the *right spectral gap* of the Markov chain.

In Theorems 1-2, we put the subscripts π in the probability and expectation notation $\mathbb{P}_\pi(\cdot)$ and $\mathbb{E}_\pi[\cdot]$ to emphasize that the Markov chain starts from stationarity. We will show

[3] Lezaud (1998) proved a Chernoff-type inequality for finite-state-space Markov chain. Paulin (2015) used Lezaud's method to derive Bernstein-type inequalities. But his original proof for general state spaces in the published version were incomplete. Inspired by an earlier version of our manuscript, Dr. Paulin corrected his proof and updated his paper on arXiv.

[4] It is called *real part* of P in the general operator theory (Conway, 2013).

later in the applications to MCMC and RDS how to extend results to non-stationary Markov chains.

Theorem 1. Let $\{X_i\}_{i\geq 1}$ be a Markov chain with invariant measure π and absolute spectral gap $1-\lambda > 0$. For any $t \in \mathbb{R}$, uniformly for all bounded functions $f_i : \mathcal{X} \to [a_i, b_i]$,

$$\mathbb{E}_\pi \left[e^{t\left(\sum_{i=1}^n f_i(X_i) - \sum_{i=1}^n \pi(f_i)\right)} \right] \leq \exp\left(\frac{1+\lambda}{1-\lambda} \cdot \sum_{i=1}^n \frac{(b_i - a_i)^2}{4} \cdot \frac{t^2}{2} \right). \tag{3}$$

It follows that for any $\epsilon > 0$

$$\mathbb{P}_\pi \left(\sum_{i=1}^n f_i(X_i) - \sum_{i=1}^n \pi(f_i) > \epsilon \right) \leq \exp\left(-\frac{1-\lambda}{1+\lambda} \cdot \frac{\epsilon^2}{2\sum_{i=1}^n (b_i - a_i)^2/4} \right). \tag{4}$$

The absolute spectral gap $1-\lambda$ quantifies the converging speed of the Markov chain towards its invariant distribution π (Rudolf, 2012). A smaller λ indicates less temporal dependence and a faster convergence speed. The ratio of the variance proxy of (3)-(4) compared to that of (1)-(2) is $\frac{1+\lambda}{1-\lambda}$, strictly increasing with λ. This is consistent with the intuition that a faster Markov chain $\{X_i\}_{i\geq 1}$ with less temporal dependence results in a smaller variance proxy and a tighter concentration of $\sum_{i=1}^n f_i(X_i)$.

Inequalities (3)-(4) given by Theorem 1 reduce to the classical Hoeffding's lemma (1) and inequality (2) in the independent setting. Indeed, independent random variables $Z_i \in [a_i, b_i]$ can be seen as transformations of i.i.d. random variables $U_i \sim \text{Uniform}[0,1]$ via the inverse cumulative distribution functions $F_{Z_i}^{-1} : [0,1] \to [a_i, b_i]$, i.e., $Z_i = F_{Z_i}^{-1}(U_i)$. The i.i.d. sequence $\{U_i\}_{i\geq 1}$ forms a stationary Markov chain on the state space $[0,1]$ with invariant measure $\pi(dy) = dy$, transition kernel $P(x, dy) = dy$ and $\lambda = 0$.

Inequality (4) generalizes that by Miasojedow (2014) from the time-independent function case, in which $f_1 = f_2 = \cdots = f_n = f$ are identical, to the time-dependent function case, in which f_i's may be different. For the time-independent function case, we find that it is possible to sharpen the variance proxy ratio in Theorem 1 by substituting λ with a smaller quantity $\max\{\lambda_r, 0\}$ (see Theorem 2).

We point out the relation of the absolute spectral gap of a Markov chain to the mixing time. The mixing time, denoted by $t_{\text{mix}}(\epsilon)$, is defined as the first occurrence time when the total variation distance between the distribution of the Markov chain and the stationary distribution π falls below a small threshold ϵ (Levin et al., 2009, Sections 3.5-3.6).

Definition 1 (Mixing Time)**.** Let $\{X_i\}_{i\geq 1}$ be a Markov chain on the state space $\mathcal{X}$ and with the invariant measure π, and the transition kernel P. Let $P^t(x, \cdot)$ be the t-step transition kernel, and let $\|\nu_1 - \nu_2\|_{\text{tv}}$ be the total variation distance between two probability measures ν_1, ν_2 on $\mathcal{X}$. Then the mixing time for some small $\epsilon > 0$ is defined as

$$t_{\text{mix}}(\epsilon) := \min\{t \geq 0 : \sup_{x \in \mathcal{X}} \|P^t(x, \cdot) - \pi\|_{\text{tv}} \leq \epsilon\}.$$

A Markov chain is uniformly ergodic if and only if $t_{\text{mix}}(\epsilon)$ is finite for some $\epsilon < 1/2$ (Roberts and Rosenthal, 2004, Proposition 7). If this Markov chain is also reversible, then it admits a positive absolute spectral gap (Rudolf, 2012, Figure 2), which has been made precise by Paulin (2015, Equations 1.3 and 3.5) as follows.

$$1 - \lambda \geq \frac{1}{1 + t_{\text{mix}}(\epsilon)/\log(1/2\epsilon)}, \quad \forall\, \epsilon \in [0, 1/2).$$

On the other hand, for an ergodic and reversible Markov chain on a finite state space, Paulin (2015, Equation 3.7) showed

$$1 - \lambda \leq \frac{2\log(1/2\epsilon) + \log(1/\pi_{\text{min}})}{2t_{\text{mix}}(\epsilon)}, \quad \forall\, \epsilon \in [0, 1/2),$$

where $\pi_{\text{min}} := \min_{x \in \mathcal{X}} \pi(x)$. To our best knowledge, there are few results on upper bound of similar form for reversible Markov chains on general state spaces.

Theorem 2. Let $\{X_i\}_{i\geq 1}$ be a Markov chain with invariant measure π and right spectral gap $1 - \lambda_r > 0$. For any $t \in \mathbb{R}$, uniformly for all bounded functions $f : \mathcal{X} \to [a, b]$,

$$\mathbb{E}_\pi \left[e^{t\left(\sum_{i=1}^n f(X_i) - n\pi(f)\right)} \right] \leq \exp\left(\frac{1 + \max\{\lambda_r, 0\}}{1 - \max\{\lambda_r, 0\}} \cdot \frac{n(b-a)^2}{4} \cdot \frac{t^2}{2} \right). \tag{5}$$

It follows that for any $\epsilon > 0$,

$$\mathbb{P}_\pi \left(\sum_{i=1}^n f(X_i) - n\pi(f) > \epsilon \right) \leq \exp\left(-\frac{1 - \max\{\lambda_r, 0\}}{1 + \max\{\lambda_r, 0\}} \cdot \frac{\epsilon^2}{2n(b-a)^2/4} \right). \tag{6}$$

Inequality (6) extends the result of León and Perron (2004) from finite-state-space and reversible Markov chains to general-state-space and not necessarily reversible Markov chains. Note that, for finite-state-space and reversible Markov chains, λ_r is merely the second largest eigenvalue of the transition probability matrix.

We show the sharpness of the variance proxy ratio in the regime of $\lambda_r \geq 0$ (which is the case for most Markov chains arising from applications), by resorting to the Central Limit Theorem for Markov chains. Consider a function $f : \mathcal{X} \to \{+1, -1\}$ such that $\pi(\{x : f(x) = +1\}) = \pi(\{x : f(x) = -1\}) = 1/2$ and a transition probability kernel $P(x, B) = \lambda_r \mathbb{I}(x \in B) + (1 - \lambda_r)\pi(B)$ for any state $x \in \mathcal{X}$ and any subset of the state space $B \subseteq \mathcal{X}$, where $\mathbb{I}(\cdot)$ is the indicator function. This transition probability kernel admits right spectral gap $1 - \lambda_r$. Suppose for the sake of contradiction that (5) holds with some strictly smaller ratio constant $C < \frac{1+\lambda_r}{1-\lambda_r}$. Then $nC \geq \text{Var}\left[\sum_{i=1}^n f(X_i)\right]$ for any $n \geq 1$, because the variance proxy of a sub-Gaussian random variable upper bounds its variance. However, Geyer (1992, Theorem 2.1), Roberts and Rosenthal (1997, Corollary 2.1) and Rosenthal (2003, Lemma 1) derive that

$$\lim_{n\to\infty} \frac{1}{n} \text{Var}\left(\sum_{i=1}^n f(X_i) \right) = \frac{1 + \lambda_r}{1 - \lambda_r},$$

implying that $C \geq \frac{1+\lambda_r}{1-\lambda_r}$. This contradicts to the initial assumption on constant C.

A simulation experiment is conducted to study the extent to which the replacement of λ with $\max\{\lambda_r, 0\}$ sharpens the variance proxy ratio. Of $1,000,000$ transition probability matrices of size 3×3, whose rows are independently drawn from the Dirichlet distribution with the shape parameters $(1/3, 1/3, 1/3)^{\mathrm{T}}$, the quantity $\frac{1+\lambda}{1-\lambda} / \frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$ is 7.99 on average.

3 Extensions to Unbounded Functions and Inhomogenous Markov Chains

It is known that the classical Hoeffding's inequality can be extended to some unbounded (sub-Gaussian) independent random variables. A natural question arises

Is it possible to relax the boundedness assumption on functions in Theorems 1-2?

Surprisingly, the answer is no. We provide such a counter-example as follows. Consider the case of a Markov chain $\{X_i\}_{i\geq 1}$ on the state space $\mathcal{X} = \mathbb{R}$ and the function $f : x \mapsto x$, in which $f(X_i) = X_i$ individually follows the standard Gaussian distribution, but no finite variance proxy exists for (3)-(6) to hold.

Theorem 3. Consider a stationary Markov chain $\{X_i\}_{i\geq 0}$ on the state space $\mathcal{X} = \mathbb{R}$ with invariant distribution π being standard Gaussian and transition kernel

$$P(x, B) = \lambda_r \mathbb{I}(x \in B) + (1 - \lambda_r)\pi(B), \quad \forall x \in \mathcal{X},\ \forall \text{ measurable } B \subseteq \mathcal{X},$$

for some $0 < \lambda_r < 1$. There exists no finite constant C such that

$$\mathbb{E}_\pi \left[e^{t \sum_{i=1}^n X_i} \right] \leq e^{nCt^2/2}$$

hold for any $t \in \mathbb{R}$ and any $n \geq 1$.

Theorem 1 implicitly assumes that the Markov chain is homogeneous, i.e., its transition probability kernel does not vary in time. The following result provides similar inequalities to (3)-(4) for inhomogenous Markov chains. Its proof is omitted as it is similar to that of Theorem 1.

Theorem 4. Consider an time-inhomogeneous, π-stationary Markov chain $\{X_i\}_{i\geq 1}$. If its transition probability kernel P_i admits absolute spectral gap $1 - \lambda_i$ for each $i \geq 1$, then, for any bounded functions $f_i : \mathcal{X} \mapsto [a_i, b_i]$, the sum $\sum_{i=1}^n f_i(X_i)$ is sub-Gaussian with variance proxy

$$\frac{(b_1 - a_1)^2}{8} + \sum_{i=1}^{n-1} \frac{1+\lambda_i}{1-\lambda_i} \cdot \frac{(b_i - a_i)^2 + (b_{i+1} - a_{i+1})^2}{8} + \frac{(b_n - a_n)^2}{8}$$
$$\leq \frac{1 + \max_{i=1}^n \lambda_i}{1 - \max_{i=1}^n \lambda_i} \cdot \sum_{i=1}^n \frac{(b_i - a_i)^2}{4}.$$

4 Proofs of Theorems 1-2

Theorems 1-2 are proven in the framework of operator theory in the Hilbert spaces $\mathcal{L}_2(\pi)$ and $\mathcal{L}_2^0(\pi)$. This section will outline their proofs first and highlight the technical novelty of our technique later.

Let us first define notation used in the proofs. Denote by E^f the multiplication operator of function $e^f : x \mapsto e^{f(x)}$, i.e., $E^f : h \in \mathcal{L}_2(\pi) \mapsto = e^f h$, and $\mathit{\Pi}$ the projection operator of π, i.e., $\mathit{\Pi} : h \in \mathcal{L}_2(\pi) \mapsto \pi(h)$. Let

$$\widehat{P}_\gamma = \gamma I + (1-\gamma)\mathit{\Pi}$$

be the convex combination of the identity operator I and the projection operator $\mathit{\Pi}$ with a coefficient $\gamma \in [0,1)$. Let $|||T|||_\pi$ be the norm of operator T acting on $\mathcal{L}_2(\pi)$.

The proof of Theorem 1 consists of four major steps, which are presented as Lemmas 1-4, respectively. Lemma 1 upper bounds the moment generating function of $\sum_{i=1}^n f_i(X_i)$ by the product of $|||E^{tf_i/2}\widehat{P}_\lambda E^{tf_i/2}|||_\pi$. Then the task reduces to bound the norm of operators of the form $E^{tf/2}\widehat{P}_\gamma E^{tf/2}$. Lemma 2 characterizes this operator norm as the limit of the $(1/n)$-scaled cumulant generating function (CGF) of $\sum_{i=1}^n f(X_i)$ for a Markov chain $\{\widehat{X}_i\}_{i=1}^n$ driven by $\widehat{P}_\gamma$. Lemma 3 constructs a two-state Markov chain $\{\widehat{Y}_i\}_{i=1}^n$, whose sum's CGF upper bounds that of $\{f(\widehat{X}_i)\}_{i=1}^n$. Lemma 4 investigates into the two-state chain system and directly compute the norm of the analogue of $E^{tf/2}\widehat{P}_\gamma E^{tf/2}$ in the two-state chain system.

Next, we refine Lemma 1 as Lemma 5 for the time-independent function case of $f_1 = \cdots = f_n = f$. Substituting Lemma 1 in the proof of Theorem 1 with its refinement Lemma 5 proves Theorem 2.

Lemma 1. Let $\{X_i\}_{i\geq 1}$ be a Markov chain with invariant measure π and absolute spectral gap $1-\lambda > 0$. Then for any bounded functions $f_i : \mathcal{X} \to [a_i, b_i]$ and any $t \in \mathbb{R}$,

$$\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f_i(X_i)}\right] \leq \prod_{i=1}^n |||E^{tf_i/2}\widehat{P}_\lambda E^{tf_i/2}|||_\pi .$$

Lemma 2. Let $\{\widehat{X}_i\}_{i\geq 1}$ be a Markov chain driven by Markov operator $\widehat{P}_\gamma$ with some $\gamma \in [0,1)$. Then for any bounded function $f : \mathcal{X} \to [a,b]$ and any $t \in \mathbb{R}$,

$$\log |||E^{tf/2}\widehat{P}_\gamma E^{tf/2}|||_\pi = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(\widehat{X}_i)}\right].$$

Definition 2 (Two-State Markov Chain)**.** Let $\{\widehat{Y}_i\}_{i\geq 1}$ be a stationary Markov chain on the state space $\mathcal{Y} = \{a, b\}$ with a transition probability matrix $\widehat{\boldsymbol{Q}}_\gamma$ determined by two parameters $\gamma \in [0,1)$ and $\mu \in (0,1)$ in the way

$$\widehat{\boldsymbol{Q}}_\gamma = \gamma \boldsymbol{I} + (1-\gamma)\begin{bmatrix} \boldsymbol{\mu}^{\mathrm{T}} \\ \boldsymbol{\mu}^{\mathrm{T}} \end{bmatrix},$$

where $\boldsymbol{\mu} = (1-\mu, \mu)^{\mathrm{T}}$ is the invariant distribution of the Markov chain.

Lemma 3 (León and Perron (2004),Theorem 2)**.** Let $\{\widehat{X}_i\}_{i\geq 1}$ be a Markov chain driven by the León-Perron operator $\widehat{P}_\gamma = \gamma I + (1-\gamma)\varPi$ with some $\gamma \in [0,1)$. For a bounded function $f : \mathcal{X} \to [a,b]$, let $\{\widehat{Y}_i\}_{i\geq 1}$ be the two-state Markov chain in Definition 2 with $\mu = \frac{\pi(f)-a}{b-a}$. Then for any convex function $\Psi : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}_\pi \left[\Psi \left(\sum_{i=1}^n f(\widehat{X}_i) \right) \right] \leq \mathbb{E}_{\boldsymbol{\mu}} \left[\Psi \left(\sum_{i=1}^n \widehat{Y}_i \right) \right].$$

In particular, for $\Psi : z \mapsto \exp(tz)$,

$$\mathbb{E}_\pi \left[e^{t\sum_{i=1}^n f(\widehat{X}_i)} \right] \leq \mathbb{E}_{\boldsymbol{\mu}} \left[e^{t\sum_{i=1}^n \widehat{Y}_i} \right].$$

Lemma 4. Let $\boldsymbol{E}^{ty/2}$ denote the diagonal matrix with elements $e^{ta/2}$, $e^{tb/2}$. Let $|||\boldsymbol{T}|||_{\boldsymbol{\mu}}$ denote the operator norm induced by the $\boldsymbol{\mu}$-weighted vector norm[5] for a 2×2 matrix $\boldsymbol{T}$. Recall notations $\widehat{\boldsymbol{Q}}_\gamma, \mu, \boldsymbol{\mu}$ from Definition 2. Write $\boldsymbol{\mu}(y) = (1-\mu)a + \mu b$. Then

$$|||\boldsymbol{E}^{ty/2} \widehat{\boldsymbol{Q}}_\gamma \boldsymbol{E}^{ty/2}|||_{\boldsymbol{\mu}} \leq \exp\left(\boldsymbol{\mu}(y)t + \frac{1+\gamma}{1-\gamma} \cdot \frac{(b-a)^2}{4} \cdot \frac{t^2}{2} \right)$$

Combing these four lemmas completes the proof of Theorem 1.

Proof of Theorem 1. With Lemma 1 in place, it suffices to show

$$\|E^{tf/2} \widehat{P}_\gamma E^{tf/2}\|_\pi \leq \exp\left(\pi(f) \cdot t + \frac{1+\gamma}{1-\gamma} \cdot \frac{(b-a)^2}{4} \cdot \frac{t^2}{2} \right).$$

Then taking $\gamma = \lambda$ and $f = f_i$ in the above display for $i = 1, \ldots, n$ and plugging them into Lemma 1 completes the proof of Theorem 1. To show the above display, we recall all notations in Lemmas 2-4 and write

$$\begin{aligned}
\log |||E^{tf/2} \widehat{P}_\gamma E^{tf/2}|||_\pi &= \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_\pi \left[e^{t\sum_{i=1}^n f(\widehat{X}_i)} \right] && \text{[Lemma 2]} \\
&\leq \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}_{\boldsymbol{\mu}} \left[e^{t\sum_{i=1}^n \widehat{Y}_i} \right] && \text{[Lemma 3]} \\
&= \log |||\boldsymbol{E}^{ty/2} \widehat{\boldsymbol{Q}}_\gamma \boldsymbol{E}^{ty/2}|||_{\boldsymbol{\mu}} && \text{[to be explained later]} \\
&\leq \boldsymbol{\mu}(y) \cdot t + \frac{1+\gamma}{1-\gamma} \cdot \frac{(b-a)^2}{4} \cdot \frac{t^2}{2} && \text{[Lemma 4]} \\
&= \pi(f) \cdot t + \frac{1+\gamma}{1-\gamma} \cdot \frac{(b-a)^2}{4} \cdot \frac{t^2}{2} && \text{[to be explained later]}
\end{aligned}$$

In the third line, we apply Lemma 2 to the two-state Markov chain $\{\widehat{Y}_i\}_{i\geq 1}$ and function $f : y \mapsto y$. In the fifth line, we use the fact that $\boldsymbol{\mu}(y) = (1-\mu)a + \mu b = \pi(f)$. □

Lemma 5. Let $\{X_i\}_{i\geq 1}$ be a Markov chain with invariant measure π and right spectral gap $1 - \lambda_r > 0$. Then for any bounded function $f : \mathcal{X} \to [a,b]$ and any $t \in \mathbb{R}$,

$$\mathbb{E}_\pi \left[e^{t\sum_{i=1}^n f(X_i)} \right] \leq |||E^{tf/2} \widehat{P}_{\max\{\lambda_r, 0\}} E^{tf/2}|||_\pi^n.$$

Proof of Theorem 2. Just substitute λ with $\max\{\lambda_r, 0\}$ and Lemma 1 with Lemma 5 in the proof of Theorem 1. □

[5] For a bivariate vector $y = (y_1, y_2)$, its $\boldsymbol{\mu}$-weighted vector norm is defined as $\|y\|_{\boldsymbol{\mu}} = \sqrt{\mu_1 y_1^2 + \mu_2 y_2^2}$.

Technical Novelty

In addition to several elements borrowed from León and Perron (2004); Miasojedow (2014), we exploit three techniques to improve upon their results.

Lemma 4 carries out a direct computation of the operator norm in the two-state chain system by a **convex analysis** argument. This is the key to establish the counterpart of the classical Hoeffding's lemma in the Markov-dependent setting:

$$\|E^{tf/2}\widehat{P}_\gamma E^{tf/2}\|_\pi \leq \exp\left(\pi(f)\cdot t + \frac{1+\gamma}{1-\gamma}\cdot\frac{(b-a)^2}{4}\cdot\frac{t^2}{2}\right).$$

Letting $\gamma = 0$ recovers Hoeffding's lemma for a single random variable. To the best of our knowledge, this extension of Hoeffding's lemma has never been discovered before.

Lemma 5 obtains a sharper bound in the time-independent case with constant $\max\{\lambda_r, 0\}$ in place of λ. This finally leads to the sharper Hoeffing-type inequality in Theorem 2. In the heart of this lemma lies the **power inequality on numerical range of operators** (Berger, 1965; Pearcy, 1966), which has never been used in the literature on Hoeffding-type and Bernstein-type inequalities for Markov chains.

Lemma 2 wraps up (Miasojedow, 2014, Lemma 3.9): the former considers a general bounded function and the latter considers a function taking finitely many values. For a function taking finitely many values, we get the same result with (Miasojedow, 2014, Lemma 3.9), but in a simpler way using **Weyl's theorem on the essential spectrum of self-adjust operators** (Weyl, 1909).

Lemma 1 refines (Miasojedow, 2014, Lemma 3.5) in order to cover the time-dependent function case. Lemma 3 is a restatement of León and Perron (2004, Theorem 2).

5 Applications

We showcase the utilities of our results by applying them to five problems in statistics, econometrics and machine learning. These applications are the Markov chain Monte Carlo estimation, Respondent-driven Sampling, high-dimensional covariance matrix estimation for Markov-dependent samples, time-discounted reward, and multi-arm bandit problems.

5.1 Markov Chain Monte Carlo

Suppose the task is to compute the integral $\pi(f) = \int f(x)\pi(x)dx$ of function f with respect to probability density function π, but π cannot be directly sampled. An MCMC algorithm generates a Markov chain $\{X_i\}_{i\geq 1}$, which converges to the invariant distribution π, and estimates $\pi(f)$ by averaging the realized values of function f on n Markov chain samples $\{X_i\}_{i=n_0+1}^{n_0+n}$ after a burn-in period of length n_0.

There are mainly two factors influencing the estimation error of MCMC methods. The first is how fast the Markov chain converges from the initial measure ν to the invariant measure π, which directly determines the length n_0 of the burn-in period. The second one is how the average of $f(X_i)$ fluctuates after the Markov chain reaches the stationarity. This has been investigated by Theorems 1 and 2.

For the first factor, we adopt the notation and the analysis of Rudolf (2012). Suppose the initial measure ν is absolutely continuous with respect to the invariant measure π. Denote by $\frac{d\nu}{d\pi}$ the Radon-Nikodym derivative of ν with respect to π. Let $\mathcal{L}_p(\pi)$ consist of all functions h such that $\pi(h^p) < \infty$. For any function $h \in \mathcal{L}_p(\pi)$, define

$$\|h\|_{\pi,p} := (\pi(h^p))^{1/p} = \left(\int h^p(x)\pi(x)dx\right)^{1/p}.$$

We have the following error bound for MCMC estimation.

Theorem 5. Let $\{X_i\}_{i\geq 1}$ be a Markov chain with invariant measure π, absolute spectral gap $1-\lambda$ and right spectral gap $1-\lambda_r$. Suppose the initial measure ν is absolutely continuous with respect to the invariant measure π and its derivative $\frac{d\nu}{d\pi}$ has a finite p-moment for some $p \in (1,\infty]$. Let $q = p/(p-1) \in [1,\infty)$ and

$$C = C(\nu, n_0, p) := \begin{cases} 1 + 2^{2/p}\lambda^{2n_0/q}\left\|\frac{d\nu}{d\pi} - 1\right\|_{\pi,p} & \text{if } p \in (1,2), \\ 1 + \lambda^{n_0}\left\|\frac{d\nu}{d\pi} - 1\right\|_{\pi,2} & \text{if } p = 2, \\ 1 + 2^{2/q}\lambda^{2n_0/p}\left\|\frac{d\nu}{d\pi} - 1\right\|_{\pi,p} & \text{if } p \in (2,\infty), \\ \|\frac{d\nu}{d\pi}\|_{\pi,\infty} = \operatorname{ess\,sup}\left|\frac{d\nu}{d\pi}\right| & \text{if } p = \infty. \end{cases}$$

Then, for any $t \in \mathbb{R}$, uniformly for all bounded function $f : \mathcal{X} \to [a,b]$,

$$\mathbb{E}_\nu\left[e^{t\sum_{i=n_0+1}^{n_0+n} f(X_i) - n\pi(f)}\right] \leq C\exp\left(q \cdot \frac{1+\max\{\lambda_r, 0\}}{1-\max\{\lambda_r, 0\}} \cdot \frac{n(b-a)^2}{4} \cdot \frac{t^2}{2}\right).$$

It follows that for any $\epsilon > 0$

$$\mathbb{P}_\nu\left(\frac{1}{n}\sum_{i=n_0+1}^{n_0+n} f(X_i) - \pi(f) > \epsilon\right) \leq C\exp\left(-\frac{1}{q} \cdot \frac{1-\max\{\lambda_r, 0\}}{1+\max\{\lambda_r, 0\}} \cdot \frac{n\epsilon^2}{2(b-a)^2/4}\right).$$

Note that the constant $C = C(\nu, n_0, p)$ depends on the initial measure ν and the length n_0 of the burn-in period. For finite $p \geq 1$, constant $C \to 1$ geometrically as $n_0 \to \infty$. Using the fact that $1 + x \leq e^x$, we observe that

$$\mathbb{P}_\nu\left(\frac{1}{n}\sum_{i=n_0+1}^{n_0+n} f(X_i) - \pi(f) > \epsilon\right) \leq e^{-O(n)+O(\kappa^{n_0})},$$

in which

$$1 > \kappa = \begin{cases} \lambda^{2/q} & \text{if } 1 < p \leq 2, \\ \lambda^{2/p} & \text{if } p > 2 \end{cases}.$$

In the exponent of the tail probability, the first term is linear in n and the second term is exponentially decaying in n_0. This observation suggests that $n_0 \gtrsim \log n$ suffices for the tail probability of MCMC estimates to exponentially decay in n, and thereafter justifies practitioners' ad-hoc choice of $n_0 = 0.05n$ or $n_0 = 0.1n$.

5.2 Respondent-driven Sampling

The case of $p = \infty$ in Theorem 5 applies to the respondent-driven sampling (RDS) problem. RDS methods, initially developed by Heckathorn (1997, 2002), have been widely used to collect data on disease prevalence and risk behaviors within some "hidden" populations. A population is "hidden" when the public acknowledgement of membership is unavailable, usually due to privacy concerns, so that traditional sampling methods do not work. For example, HIV infections concentrate on three hidden populations: men who have sex with men, injection drug users, and sex workers and their sexual partners (WHO and UNAIDS, 2009). RDS collects data through a chain-referral mechanism, in which current participants recruit their contacts to be new participants.

Goel and Salganik (2009, 2010) modeled this chain-referral mechanism as a random walk among the network of the hidden population under study, and proposed an MCMC-based estimator for the disease prevalence. Formally, let $\mathcal{X}$ denote the population under study, and let $\mathcal{E}$ denote the edge (contact) set among them. Let $|\mathcal{X}|$ and $|\mathcal{E}|$ denote their cardinalities. Assume that the network is connected and participants recruit their contacts uniformly at random. The chain-referral process is modeled as a random walk $\{X_i\}_{i\geq 1}$ on the graph $(\mathcal{X}, \mathcal{E})$ with uniform edge weights.

Let $f(x) = 1$ if node (member) $x \in \mathcal{X}$ is infected and 0 otherwise, let $d(x)$ be the degree of node x in the network. The prevalence of a disease

$$\bar{f} := \frac{\sum_{x\in\mathcal{X}} f(x)}{|\mathcal{X}|}$$

can be consistently estimated by the RDS estimate

$$\hat{f}_n := \sum_{i=1}^{n} \frac{f(X_i)}{d(X_i)} \Bigg/ \sum_{i=1}^{n} \frac{1}{d(X_i)}.$$

Indeed, the random walk has the invariant distribution $\pi(x) = d(x)/2|\mathcal{E}|$, and thus by the Law of Large Number for Markov chains

$$\frac{1}{n}\sum_{i=1}^{n} \frac{f(X_i)}{d(X_i)} \to \pi\left(\frac{f}{d}\right) = \frac{\sum_{x\in\mathcal{X}} f(x)}{2|\mathcal{E}|}$$
$$\frac{1}{n}\sum_{i=1}^{n} \frac{1}{d(X_i)} \to \pi\left(\frac{1}{d}\right) = \frac{|\mathcal{X}|}{2|\mathcal{E}|}$$

almost surely as $n \to \infty$. Combing the two asymptotic convergences obtains $\hat{f}_n \to \bar{f}$ almost surely as $n \to \infty$.

Theorem 5 with $p = \infty$ provides a non-asymptotical analysis for the RDS estimate. This analysis suggests that the minimum number of participants is $n = O(\log|\mathcal{X}|)$ for the success of RDS. To the best of our knowledge, this insight has not been observed in the previous literature. Specifically, if the chain-referral process starts from a specific node y, i.e., $\nu(y) = 1$ and $\nu(x) = 0$ for $x \neq y$, and all collected samples are used for estimation, i.e., $n_0 = 0$, then the constant in Theorem 5 is given by

$$C(\nu, 0, \infty) = \max_{x \in \mathcal{X}} \left| \frac{\nu(x)}{\pi(x)} \right| = \frac{2|\mathcal{E}|}{d(y)} = \frac{\bar{d}}{d(y)} |\mathcal{X}|,$$

where $\bar{d} = \sum_{x \in \mathcal{X}} d(x)/|\mathcal{X}| = 2|\mathcal{E}|/|\mathcal{X}|$ is the average degree of all nodes in the network. It follows that

$$\mathbb{P}_\nu \left(\left| \frac{1}{n} \sum_{i=1}^n \frac{f(X_i)}{d(X_i)} - \frac{\sum_{x \in \mathcal{X}} f(x)}{2|\mathcal{E}|} \right| > \epsilon \right) \leq \frac{2\bar{d}}{d(y)} |\mathcal{X}| \exp\left(-\frac{1 - \max\{\lambda_r, 0\}}{1 + \max\{\lambda_r, 0\}} \cdot 2n\epsilon^2 \right)$$

$$\mathbb{P}_\nu \left(\left| \frac{1}{n} \sum_{i=1}^n \frac{1}{d(X_i)} - \frac{|\mathcal{X}|}{2|\mathcal{E}|} \right| > \epsilon \right) \leq \frac{2\bar{d}}{d(y)} |\mathcal{X}| \exp\left(-\frac{1 - \max\{\lambda_r, 0\}}{1 + \max\{\lambda_r, 0\}} \cdot 2n\epsilon^2 \right).$$

We conclude that

$$\mathbb{P}_\nu \left(\hat{f}_n \in \left[\frac{\bar{f} - \epsilon\bar{d}}{1 + \epsilon\bar{d}}, \frac{\bar{f} + \epsilon\bar{d}}{1 - \epsilon\bar{d}} \right] \right) \geq 1 - \frac{4\bar{d}}{d(y)} |\mathcal{X}| \exp\left(-\frac{1 - \max\{\lambda_r, 0\}}{1 + \max\{\lambda_r, 0\}} \cdot 2n\epsilon^2 \right).$$

This bound of RDS estimates provides some guidance for the RDS practitioners. First, the degree $d(y)$ of the initial participant y needs to be comparable to the average degree $\bar{d}$ of all the population. Second, the number n of participants needed exceeds the logarithm of the total population.

5.3 High-dimensional Covariance Matrix Estimation

This subsection considers estimating high-dimensional covariance matrix with Markov-dependent samples by thresholding. In the setup of independent samples, this thresholding method has been intensively studied. See Bickel and Levina (2008); Lam and Fan (2009); Cai et al. (2010); Cai and Liu (2011); Fan et al. (2013), among others.

Suppose that $\{X_i\}_{i \geq 1}$ is a stationary Markov chain with invariant measure π and right spectral gap $1 - \lambda_r$, and $\boldsymbol{f} = (f_1, \ldots, f_p)^{\mathrm{T}}$ is a p-dimensional bounded feature mapping. Without loss of generality, we assume $\pi(f_j) = 0$ and $\sup_{x \in \mathcal{X}} |f_j(x)| \leq 1$ for all $j = 1, \ldots, p$. We are interested in estimating the covariance matrix

$$\boldsymbol{\Sigma} = \int \boldsymbol{f}(x) \boldsymbol{f}(x)^{\mathrm{T}} \pi(dx)$$

in the high-dimensional regime with $p \gg n$ but $\log p = o(n)$. This is potentially useful to model the fMRI data, in which $\{X_i\}_{i \geq 1}$ represents brain activities at various time points i, and $f_j(X_i)$ represents the BOLD (blood-oxygen-level dependent) signal from the region of interest j at time i, and $\boldsymbol{\Sigma}$ characterizes the functional connectivity between multiple regions.

We extend the classical analyses of the thresholding estimation method from the independent setting, in which mutual independence of $\{\boldsymbol{f}(X_i)\}_{i \geq 1}$ are assumed, to the Markov-dependent setting. Using Theorem 2 to deal with the Markov dependence, we obtain an analogous result compared to the classical result in the independent setting. The only difference is that the sample size should be increased by a factor $\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$ in the Markov-dependent setting, in order to achieve the same error rate as the independent setting.

Theorem 6. Consider the sparse covariance estimation problem in the Markov-dependent setting. Define the uniformity class of sparse covariance matrices as

$$\mathcal{M}(s,m)=\left\{\boldsymbol{M}\succeq 0:\ \boldsymbol{M}_{jj}\leq m,\sum_{k=1}^{p}\mathbb{I}(\boldsymbol{M}_{jk}\neq 0)\leq s,\forall j=1,\dots,p\right\}.$$

For a matrix $\boldsymbol{M}$, define the element-wise thresholding operator by

$$T_t(\boldsymbol{M})=\left[\boldsymbol{M}_{jk}\mathbb{I}(|\boldsymbol{M}_{jk}|>t)\right]_{1\leq j,k\leq p}.$$

Let $\widehat{\boldsymbol{\Sigma}}=\sum_{i=1}^{n}\boldsymbol{f}(X_i)\boldsymbol{f}(X_i)^{\mathrm{T}}/n$. If $\boldsymbol{\Sigma}\in\mathcal{M}(s,m)$ and the threshold t is chosen as

$$2\sqrt{2(2+\delta)\cdot\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}\cdot\frac{\log p}{n}}\leq t\leq C\sqrt{\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}\cdot\frac{\log p}{n}},$$

for a sufficiently large constant C, then with probability at least $1-2p^{-\delta}$

$$\begin{aligned}|||T_t(\widehat{\boldsymbol{\Sigma}})-\boldsymbol{\Sigma}|||_2\leq|||T_t(\widehat{\boldsymbol{\Sigma}})-\boldsymbol{\Sigma}|||_1&\leq s\left(2t+3\sqrt{2(2+\delta)\cdot\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}\cdot\frac{\log p}{n}}\right)\\&\leq\left(2C+3\sqrt{2(2+\delta)}\right)s\sqrt{\frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}\cdot\frac{\log p}{n}},\end{aligned}$$

where $|||\boldsymbol{S}|||_1$ and $|||\boldsymbol{S}|||_2$ denote the operator norms of matrix $\boldsymbol{S}$ induced by the ℓ_1-norm and ℓ_2-norm of vectors, respectively.

5.4 Time-discounted Reward

In this subsection, we construct confidence intervals for the time-discounted reward arising from economic models. In these models, discount functions are introduced to weight rewards received at different time points for modeling the tendency of favoring rewards received now over those received in future (Frederick et al., 2002).

Consider a vineyard model (Paulin, 2015, Example 3.17) as an example. The production of grapevine in a vineyard depends on the climate. For simplicity, we model the weather as a Markov chain $\{X_i\}_{i\geq 1}$ in the state space $\{0$ (bad weather), 1 (good weather)$\}$ and assume that the vineyard produces no grapevine in the bad weather and 1 dollar worth of grapevine in good weather. Suppose the Markov chain of the weather is stationary with invariant distribution $\pi=(\pi_0,\pi_1)^{\mathrm{T}}$ and admits absolute spectral gap $1-\lambda$.

With exponential discount of form ρ^i for some exponential discount coefficient $\rho\in(0,1]$, the present value of grapevine the vineyard produces during the next n years is

$$V_n=\sum_{i=1}^{n}\rho^i X_i,$$

It is not hard to see that $\mathbb{E}_\pi V_n = \frac{\rho(1-\rho^n)\pi_1}{1-\rho}$. Applying Theorem 1 yields

$$\mathbb{P}_\pi\left(|V_n - \mathbb{E}_\pi V_n| \geq \epsilon\right) \leq 2\exp\left(-\frac{1-\lambda}{1+\lambda}\cdot\frac{1-\rho^2}{\rho^2(1-\rho^{2n})}\cdot 2\epsilon^2\right).$$

That means, with probability at least $1-2\delta$,

$$\left|V_n - \frac{\rho(1-\rho^n)\pi_1}{1-\rho}\right| \leq \sqrt{\frac{1+\lambda}{1-\lambda}\cdot\frac{\rho^2(1-\rho^{2n})}{1-\rho^2}\cdot\frac{\log(1/\delta)}{2}}.$$

With hyperbolic discount of form $(1+\rho i)^{-1}$ for some degree of discounting $\rho \geq 0$, the similar $(1-2\delta)$-confidence interval of the present value is given by

$$\left|\sum_{i=1}^n \frac{X_i}{1+\rho i} - \sum_{i=1}^n \frac{\pi_1}{1+\rho i}\right| \leq \sqrt{\frac{1+\lambda}{1-\lambda}\cdot\sum_{i=1}^n\frac{1}{(1+\rho i)^2}\cdot\frac{\log(1/\delta)}{2}}.$$

5.5 Multi-armed Bandit with Markovian Rewards

The Multi-armed Bandit (MAB) problem has received much attention in decision theory, clinical trials and statistical machine learning. In this problem, there are a number, say K, of alternative arms, each with a stochastic reward with initially unknown expectation. The goal is to find the optimal strategy that maximizes the sum of rewards. Let $Z_j(t)$ be the reward from arm j played at round t. Let $j_\star(t) = \arg\max_{j=1}^K \mathbb{E}Z_j(t)$ be the index of the arm with highest expected reward at round t. Let $j(t) \in \{1,\dots,K\}$ be the arm that is chosen at round t. A large body of literature focus on minimization of the pseudo-regret on the first T rounds

$$\mathcal{R} = \mathbb{E}\left[\sum_{t=1}^T Z_{j_\star(t)}(t) - \sum_{t=1}^T Z_{j(t)}(t)\right].$$

Machine learners have recognized three fundamental formalizations of MAB problems depending on the nature of the reward process: i.i.d., adversarial, and Markovian (Bubeck and Cesa-Bianchi, 2012), but the last has been much less studied than the other two (Tekin and Liu, 2010). It is primarily because many bandit algorithms essentially rely on the concentration of average reward draws around its expectation to identify the optimal arm, but powerful concentration inequalities for Markov-dependent random variables were in lack. Let us showcase the utilities of our results by deriving bounds for the pseudo-regret of the celebrated Upper Confidence Bound (UCB) algorithm in the following Markovian MAB problem.

Suppose, in an MAB with K arms, each arm j has an underlying stationary Markov chain $\{X_{ji}\}_{i\geq 1}$ with invariant measure π_j and right spectral gap $1-\lambda_r$, and a reward function $f_j : x \mapsto [0,1]$. Whenever arm j is played, it returns a reward of the current state and let its Markov chain transition one step. Let $N_j(s) = \sum_{t=1}^s \mathbb{I}(j(t) = j)$ be the number of times arm j is played on the first s rounds, and $\Delta_j = \pi_j(f_j) - \pi_{j_\star}(f_{j_\star})$ be the gap between expected rewards of a suboptimal arm j and the optimal arm $j_\star$. Then the pseudo-regret in this set-up is given by

$$\mathcal{R} = \sum_{j=1}^{K} \Delta_j \mathbb{E} N_j(T).$$

Theorem 7 bounds the regret for the UCB algorithm for the MAB problem with Markovian rewards. If $\lambda_r = 0$ then this theorem recovers the classical regret bound for c-UCB algorithm.

Theorem 7 (c-UCB algorithm for Markovian MAB)**.** Consider the c-UCB algorithm with input parameter c for the above Markovian MAB. Let $\widehat{f}_{j,n} = \frac{1}{n}\sum_{i=1}^{n} f_j(X_{ji})$ be the sample mean of the first n rewards from arm j. At each round t, select

$$j(t) \in \arg\max_{j=1}^{K} \widehat{f}_{j,N_j(t-1)} + \sqrt{\frac{c \log t}{2N_j(t-1)}}.$$

If $c > 2 \cdot \frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$ then this c-UCB algorithm has pseudo-regret

$$\mathcal{R} \leq \sum_{j:\,\Delta_j>0} \left(\frac{2c}{\Delta_j} \log T + \frac{c\Delta_j}{c - 2 \cdot \frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}} \right).$$

6 Conclusion and Discussions

Markovian structure is frequently used in the literature to model dependence structure between samples. Yet the literature lacks powerful concentration inequalities for studying Markov chains. In this paper, we prove an optimal concentration inequality for possibly nonidentical functions of general Markov chains. It serves as the exact counterpart of the classical Hoeffding's inequality when samples are assumed to be independent. This new concentration inequality is optimal and involves the spectral gap as a coefficient in the exponent. It finds many applications in statistical learning problems when data are dependent, and we apply them to five different problems in statistics, economics and machine learning.

References

ADAMCZAK, R. and BEDNORZ, W. (2015). Exponential concentration inequalities for additive functionals of Markov chains. *ESAIM: Probability and Statistics* **19** 440–481.

BERGER, C. (1965). A strange dilation theorem. *Notices Amer. Math. Soc* **12** 590.

BICKEL, P. J. and LEVINA, E. (2008). Covariance regularization by thresholding. *The Annals of Statistics* **36** 2577–2604.

BOUCHERON, S., LUGOSI, G. and MASSART, P. (2013). *Concentration inequalities: a nonasymptotic theory of independence*. Oxford University Press.

BUBECK, S. and CESA-BIANCHI, N. (2012). Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning* **5** 1–122.

CAI, T. and LIU, W. (2011). Adaptive thresholding for sparse covariance matrix estimation. *Journal of the American Statistical Association* **106** 672–684.

CAI, T. T., ZHANG, C.-H. and ZHOU, H. H. (2010). Optimal rates of convergence for covariance matrix estimation. *The Annals of Statistics* **38** 2118–2144.

CHAZOTTES, J. R., COLLET, P., KÜLSKE, C. and REDIG, F. (2007). Concentration inequalities for random fields via coupling. *Probability Theory and Related Fields* **137** 201–225.

CHUNG, K.-M., LAM, H., LIU, Z. and MITZENMACHER, M. (2012). Chernoff-Hoeffding bounds for Markov chains: generalized and simplified. In *Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science*.

CONWAY, J. B. (2013). *A course in functional analysis*. Springer.

DINWOODIE, I. H. (1995). A probability inequality for the occupation measure of a reversible Markov chain. *The Annals of Applied Probability* **5** 37–43.

DOUC, R., MOULINES, E., OLSSON, J. and VAN HANDEL, R. (2011). Consistency of the maximum likelihood estimator for general hidden Markov models. *The Annals of Statistics* **39** 474–513.

FAN, J., FAN, Y. and LV, J. (2008). High dimensional covariance matrix estimation using a factor model. *Journal of Econometrics* **147** 186–197.

FAN, J., HAN, F. and LIU, H. (2014). Challenges of big data analysis. *National Science Review* **1** 293–314.

FAN, J., LIAO, Y. and LIU, H. (2016). An overview of the estimation of large covariance and precision matrices. *The Econometrics Journal* **19** C1–C32.

FAN, J., LIAO, Y. and MINCHEVA, M. (2011). High dimensional covariance matrix estimation in approximate factor models. *Annals of statistics* **39** 3320.

FAN, J., LIAO, Y. and MINCHEVA, M. (2013). Large covariance estimation by thresholding principal orthogonal complements. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75** 603–680.

FILL, J. A. (1991). Eigenvalue bounds on convergence to stationarity for non-reversible Markov chains, with an application to the exclusion process. *The Annals of Applied Probability* **1** 62–87.

FREDERICK, S., LOEWENSTEIN, G. and O'DONOGHUE, T. (2002). Time discounting and time preference: a critical review. *Journal of Economic Literature* **40** 351–401.

GEYER, C. J. (1992). Practical Markov chain Monte Carlo. *Statistical science* **7** 473–483.

GILKS, W. R. (2005). Markov chain Monte Carlo. *Encyclopedia of Biostatistics* **4**.

GILLMAN, D. W. (1993). *Hidden Markov chains: convergence rates and the complexity of inference*. Ph.D. thesis, Massachusetts Institute of Technology, Boston, US.

GLYNN, P. W. and ORMONEIT, D. (2002). Hoeffding's inequality for uniformly ergodic Markov chains. *Statistics & probability letters* **56** 143–146.

GOEL, S. and SALGANIK, M. J. (2009). Respondent-driven sampling as Markov chain Monte Carlo. *Statistics in Medicine* **28** 2202–2229.

GOEL, S. and SALGANIK, M. J. (2010). Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences* **107** 6743–6747.

HECKATHORN, D. D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* **44** 174–199.

HECKATHORN, D. D. (2002). Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems* **49** 11–34.

HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58** 13–30.

KONTOROVICH, A. and RAMANAN, K. (2008). Concentration inequalities for dependent random variables via the martingale method. *The Annals of Probability* **36** 2126–2158.

KONTOYIANNIS, I., LASTRAS-MONTANO, L. A. and MEYN, S. P. (2005). Relative entropy and exponential deviation bounds for general Markov chains. In *Proceedings of 2005 International Symposium on Information Theory*. IEEE.

KONTOYIANNIS, I., LASTRAS-MONTANO, L. A. and MEYN, S. P. (2006). Exponential bounds and stopping rules for MCMC and general Markov chains. In *Proceedings of the 1st International Conference on Performance Evaluation Methodologies and Tools*. ACM.

LAM, C. and FAN, J. (2009). Sparsistency and rates of convergence in large covariance matrix estimation. *The Annals of Statistics* **37** 4254.

ŁATUSZYŃSKI, K., MIASOJEDOW, B. and NIEMIRO, W. (2013). Nonasymptotic bounds on the estimation error of MCMC algorithms. *Bernoulli* **19** 2033–2066.

LEÓN, C. A. and PERRON, F. (2004). Optimal Hoeffding bounds for discrete reversible Markov chains. *The Annals of Applied Probability* **14** 958–970.

LEVIN, D. A., PERES, Y. and WILMER, E. (2009). *Markov chains and mixing times*. American Mathematical Society.

LEZAUD, P. (1998). Chernoff-type bound for finite Markov chains. *The Annals of Applied Probability* **8** 849–867.

LIU, J. S. (2008). *Monte Carlo strategies in scientific computing*. Springer Science & Business Media.

MARTON, K. (1996). Bounding $\bar{d}$-distance by informational divergence: a method to prove measure concentration. *The Annals of Probability* **24** 857–866.

MARTON, K. (1998). Measure concentration for a class of random processes. *Probability Theory and Related Fields* **110** 427–439.

MIASOJEDOW, B. (2014). Hoeffding's inequalities for geometrically ergodic Markov chains on general state space. *Statistics & Probability Letters* **87** 115–120.

PAULIN, D. (2015). Concentration inequalities for Markov chains by Marton couplings and spectral methods. *Electronic Journal of Probability* **20** 1–32.

PEARCY, C. (1966). An elementary proof of the power inequality for the numerical radius. *The Michigan Mathematical Journal* **13** 289–291.

REDIG, F. and CHAZOTTES, J. R. (2009). Concentration inequalities for Markov processes via coupling. *Electronic Journal of Probability* **14** 1162–1180.

ROBERTS, G. and ROSENTHAL, J. (1997). Geometric ergodicity and hybrid Markov chains. *Electronic Communications in Probability* **2** 13–25.

ROBERTS, G. O. and ROSENTHAL, J. S. (2004). General state space markov chains and mcmc algorithms. *Probability surveys* **1** 20–71.

ROSENTHAL, J. S. (2003). Asymptotic variance and convergence rates of nearly-periodic Markov chain Monte Carlo algorithms. *Journal of the American Statistical Association* **98** 169–177.

ROTHMAN, A. J., LEVINA, E. and ZHU, J. (2009). Generalized thresholding of large covariance matrices. *Journal of the American Statistical Association* **104** 177–186.

RUDOLF, D. (2012). *Explicit error bounds for Markov chain Monte Carlo*. Ph.D. thesis, Institute of Mathematics, University Jena, Jena, Germany.

SAMSON, P.-M. (2000). Concentration of measure inequalities for Markov chains and ϕ-mixing processes. *The Annals of Probability* **28** 416–461.

SUTTON, R. S. and BARTO, A. G. (2018). *Reinforcement learning: an introduction*. MIT press.

TEKIN, C. and LIU, M. (2010). Online algorithms for the multi-armed bandit problem with Markovian rewards. In *Proceedings of 48th Annual Allerton Conference on Communication, Control, and Computing*. IEEE.

WEYL, H. (1909). Über beschränkte quadratische formen, deren differenz vollstetig ist. *Rendiconti del Circolo Matematico di Palermo (1884-1940)* **27** 373–392.

WHO and UNAIDS (2009). *AIDS epidemic update: December 2009*. WHO Regional Office, Europe.

Supplementary Material to "Hoeffding's inequality for general Markov chains with its applications to statistical learning"

A Preliminaries

Throughout the paper, we assume that the state space $\mathcal{X}$ is Polish (separable, completely metrizable) and equipped with its sigma-algebra $\mathcal{B}$. Then $(\mathcal{X}, \mathcal{B})$ is a standard Borel space[1]. This is a common assumption to rigorously study Markov chains in the measure theory for probability, since several useful results such as the isomorphism to the real space, the existence of conditional probability kernels and generalizations of the classical Kolmogorov are true for standard Borel spaces (Mackey, 1957).

The distribution of a time-homogeneous Markov chain is uniquely determined by its initial measure ν and its transition kernel P. Let

$$\nu(B) = \mathbb{P}(X_1 \in B), \quad \forall\, B \in \mathcal{B};$$
$$P(X_i, B) = \mathbb{P}(X_{i+1} \in B | X_i), \quad \forall\, B \in \mathcal{B},\ \forall i \geq 0.$$

A transition kernel P is said invariant with a probability measure π on $(\mathcal{X}, \mathcal{B})$ if

$$\pi(B) = \int P(x, B)\pi(dx), \quad \forall\, B \in \mathcal{B}.$$

A Markov chain is said stationary if it starts from $\nu = \pi$.

Our analyses are conducted in the framework of operator theory on Hilbert spaces. The idea of using this framework originates from the fact that each transition kernel, if invariant with π, is viewed as a Markov operator on the Hilbert space $\mathcal{L}_2(\pi)$ consisting of all real-valued, $\mathcal{B}$-measurable, π-square-integrable functions on $\mathcal{X}$.

A.1 Hilbert Space $\mathcal{L}_2(\pi)$

Recall that we write $\pi(h) = \int h(x)\pi(dx)$ for any real-valued, $\mathcal{B}$-measurable function $h : \mathcal{X} \to \mathbb{R}$. Let $\mathcal{L}_p(\mathcal{X}, \mathcal{B}, \pi)$ be the set of real-valued, $\mathcal{B}$-measurable functions with finite p-moment, i.e.

$$\mathcal{L}_p(\mathcal{X}, \mathcal{B}, \pi) := \{h : \pi(|h|^p) < \infty\}.$$

[1]For formal definitions of Polish spaces and standard Borel spaces and more details, please refer to Preston (2008) and Orbanz (2015, Chapter 2.5).

Here $h_1, h_2 \in \mathcal{L}_p(\mathcal{X}, \mathcal{B}, \pi)$ are taken as identical if $h_1 = h_2$ π-almost everywhere (π-a.e.). For every $p \in [1, \infty]$, $\mathcal{L}_p(\mathcal{X}, \mathcal{B}, \pi)$ is a Banach space equipped with norm

$$\|h\|_{\pi,p} := \begin{cases} \pi(|h|^p)^{1/p} & \text{if } p < \infty, \\ \operatorname{ess\,sup} |h| & \text{if } p = \infty. \end{cases}$$

In particular, if $p = 2$ then

$$\mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi) := \left\{h : \pi(h^2) < \infty\right\}$$

is a Hilbert space[2] endowed with the following inner product

$$\langle h_1, h_2 \rangle_\pi = \int h_1(x) h_2(x) \pi(dx), \quad \forall\, h_1, h_2 \in \mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi),$$

since for every $h \in \mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi)$,

$$\|h\|_{\pi,2} = \sqrt{\langle h, h \rangle_\pi}.$$

By convention, the norm of a linear operator T on $\mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi)$ is defined as

$$|||T|||_{\pi,2} = \sup\{\|Th\|_{\pi,2} : \|h\|_{\pi,2} = 1\}.$$

Another important Hilbert space is the subspace of $\mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi)$ consisting of mean zero functions:

$$\mathcal{L}_2^0(\mathcal{X}, \mathcal{B}, \pi) := \{h \in \mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi) :\ \pi(h) = 0\}.$$

For simplicity of notation, we write $\|h\|_\pi$ and $|||T|||_\pi$ in place of $\|h\|_{\pi,2}$ and $|||T|||_{\pi,2}$, respectively. We also write $\mathcal{L}_2(\pi)$ and $\mathcal{L}_2^0(\pi)$ in place of $\mathcal{L}_2(\mathcal{X}, \mathcal{B}, \pi)$ and $\mathcal{L}_2^0(\mathcal{X}, \mathcal{B}, \pi)$ respectively, whenever the measurable space $(\mathcal{X}, \mathcal{B})$ is clear in the context.

A.2 Operators on $\mathcal{L}_2(\pi)$

Each transition kernel $P(x, B)$, if invariant with π, corresponds to a bounded linear operator on $\mathcal{L}_2(\pi)$. We call this operator a Markov operator and abuse P to denote it. That is,

$$Ph(x) = \int h(y) P(x, dy), \quad \forall\, x \in \mathcal{X},\ \forall\, h \in \mathcal{L}_2(\pi).$$

Next, we introduce five transition kernels and their associated Markov operators which appear frequently throughout the proof. We use the same notation for a transition kernel and its associated Markov operator.

[2] Here we consider this real Hilbert space instead of the complex Hilbert space, as the former is adequate for our proofs.

A.2.1 Identity operator I

The identity kernel given by

$$I(x, B) = \mathbb{I}(x \in B), \quad \forall\, x \in \mathcal{X},\ \forall\, B \in \mathcal{B},$$

generates a Markov chain, which never moves from its initial state. The identity kernel corresponds to the identity operator on $\mathcal{L}_2(\pi)$

$$I : h \in \mathcal{L}_2(\pi) \mapsto h.$$

A.2.2 Projection operator $\mathit{\Pi}$

The transition kernel given by

$$\mathit{\Pi}(x, B) = \pi(B), \quad \forall\, x \in \mathcal{X},\ \forall\, B \in \mathcal{B},$$

generates a Markov chain which consists of i.i.d. draws from the invariant measure π. Denote by the italicized symbol $\mathit{1}$ the constant function $x \in \mathcal{X} \mapsto 1$. The transition kernel $\mathit{\Pi}(x, B)$ corresponds to the following Markov operator

$$\mathit{\Pi} : h \in \mathcal{L}_2(\pi) \mapsto \pi(h)\mathit{1},$$

which is a projection operator of rank one since $\pi(h) = \langle h, \mathit{1} \rangle_\pi$. It is not hard to see that $P\mathit{\Pi} = \mathit{\Pi}P = \mathit{\Pi}$ if the transition kernel $P(x, B)$ is invariant with π.

A.2.3 Adjoint operator P^*

Definition A.1 (Time-reversal kernel)**.** A transition kernel P^* is said to be the time-reversal of a transition kernel P if

$$\int_{B_1} \pi(dx) P(x, B_2) = \int_{B_2} \pi(dx) P^*(x, B_1), \quad \forall\, B_1, B_2 \in \mathcal{B}.$$

Definition A.2 (Adjoint operator)**.** A linear operator T^* on a real Hilbert space $\mathcal{H}$ endowed with inner product $\langle \cdot, \cdot \rangle$ is said to be the adjoint of a linear operator T if

$$\langle Th_1, h_2 \rangle = \langle h_1, T^* h_2 \rangle, \quad \forall\, h_1, h_2 \in \mathcal{H}.$$

Definition A.3 (Self-adjoint operator)**.** A linear operator S on a real Hilbert space $\mathcal{H}$ endowed with inner product $\langle \cdot, \cdot \rangle$ is said to be self-adjoint if

$$\langle Sh_1, h_2 \rangle = \langle h_1, Sh_2 \rangle, \quad \forall\, h_1, h_2 \in \mathcal{H}.$$

The existence of the time-reversal kernel (Definition A.1) in a standard Borel space $(\mathcal{X}, \mathcal{B})$ is guarantteed by Breiman (1992, Theorem 4.34), Durrett (2010, Theorems 2.1.22 and 5.1.9) or Orbanz (2015, Theorem 3.6(1)). This time-reversal kernel is also unique up to differences on sets of probability zero (Orbanz, 2015, Theorem 3.6(2)). In most practical examples, probability measures $\pi(\cdot)$ and $\{P(x,\cdot) : x \in \mathcal{X}\}$ share a common reference measure. In these examples, let $\pi(x)$ and $P(x,y)$ denote their densities then the time-reversal transition kernel has a density of a simple closed form

$$P^*(x,y) = \frac{\pi(y)P(y,x)}{\pi(x)}.$$

The time-reversal kernel corresponds to the adjoint operator (Definition A.2) of P on $\mathcal{L}_2(\pi)$. If the transition kernel $P(x,B)$ is invariant with π, so is $P^*(x,B)$. And, $P^*\mathit{\Pi} = \mathit{\Pi}P^* = \mathit{\Pi}$. A Markov chain is said reversible if $P^* = P$. This condition is called the detailed balance condition when viewing P and P^* as transition kernels, or the self-adjointness (Definition A.3) when viewing P and P^* as Markov operators.

A.2.4 Additively-reversiblized operator R

Fill (1991) defined the additive reversiblization[3] of a Markov operator P as

$$R = \frac{P+P^*}{2},$$

which is self-adjoint and thus relates to a reversible Markov transition kernel $R(x,B)$. If the transition kernel $P(x,B)$ is invariant with π, so is $R(x,B)$. And, $R\mathit{\Pi} = \mathit{\Pi}R = \mathit{\Pi}$.

A.2.5 León-Perron operator $\widehat{P}$

Every convex combination of Markov transition kernels (operators) produces a Markov transition kernel (operator). We say a Markov operator is *León-Perron* if it is a convex combination of the identity operator I and the projection operator $\mathit{\Pi}$.

Definition A.4 (León-Perron operator)**.** A Markov operator $\widehat{P}_\gamma$ on $\mathcal{L}_2(\pi)$ is said León-Perron if it is a convex combination of operators I and $\mathit{\Pi}$ with some coefficient $\gamma \in [0,1]$, that is

$$\widehat{P}_\gamma = \gamma I + (1-\gamma)\mathit{\Pi}.$$

The associated transition kernel

$$\begin{aligned}\widehat{P}_\gamma(x,B) &= \gamma I(x,B) + (1-\gamma)\mathit{\Pi}(x,B) \\ &= \gamma \mathbb{I}(x \in B) + (1-\gamma)\pi(B), \quad \forall x \in \mathcal{X},\ \forall B \in \mathcal{B},\end{aligned}$$

characterizes a random-scan mechanism: at each step, the Markov chain either stays at the current state (with probability γ) or jumps to a new state drawn from π (with probability $1-\gamma$).

The key to prove the Hoeffding-type inequalities for Markov chains in (León and Perron, 2004) is the observation that a Markov chain driven by $\widehat{P}_\gamma$ is the extremal case of all Markov chains with $|||P - \mathit{\Pi}|||_\pi \le \gamma$. That is why we call this type of operators León-Perron.

[3] Operator $R = (P+P^*)/2$ is called the real part of P in the general operator theory (Conway, 2013). Here we follow Fill (1991) as we discuss this operator in the context of Markov chains.

A.3 Spectral Gaps

Definition A.5 (absolute spectral gap)**.** A Markov operator P admits an absolute spectral gap $1-\lambda$ if

$$\lambda = \lambda(P) := \|\!|P - \mathit{\Pi}|\!\|_\pi < 1.$$

It is elementary that $\lambda(I) = 1$, $\lambda(\mathit{\Pi}) = 0$, $\lambda(P) = \lambda(P^*) \geq \lambda(R)$, and $\lambda(\widehat{P}_\gamma) = \gamma$. Let $\mathcal{S}(T|\mathcal{H})$ denote the spectrum of an operator T acting on Hilbert space $\mathcal{H}$. It is known that the spectrum of self-adjoint Markov operator $R = (P + P^*)/2$ on $\mathcal{L}_2^0(\pi)$ is contained in $[-\lambda(R), +\lambda(R)]$ on the real line. Let

$$\begin{aligned} \lambda_r(R) &:= \sup\{s : s \in \mathcal{S}(R|\mathcal{L}_2^0(\pi))\}, \\ \lambda_l(R) &:= \inf\{s : s \in \mathcal{S}(R|\mathcal{L}_2^0(\pi))\}. \end{aligned} \tag{S.1}$$

Definition A.6 (Right spectral gap)**.** A Markov operator P admits a right spectral gap $1-\lambda_r$ if

$$\lambda_r = \lambda_r(R) < 1, \quad \text{where } R = (P + P^*)/2.$$

Since $R = (P + P^*)/2$ is self-adjoint,

$$\begin{aligned} \lambda_r(R) &:= \sup\{\langle Rh, h\rangle : \ \|h\|_\pi = 1,\ \pi(h) = 0\}, \\ \lambda_l(R) &:= \inf\{\langle Rh, h\rangle : \ \|h\|_\pi = 1,\ \pi(h) = 0\}, \\ \lambda(R) &= \sup\{|s| : s \in \mathcal{S}(R|\mathcal{L}_2^0(\pi))\} = \max\{\lambda_r(R), |\lambda_l(R)|\} \end{aligned}$$

It follows that

$$|\lambda_r(R)| \leq \lambda(R) \leq \frac{\lambda(P) + \lambda(P^*)}{2} = \lambda(P).$$

B Proof of Lemmas 1-5

Lemmas 1-2 invoke two auxiliary Lemmas B.1-B.2, respectively.

Lemma B.1. (i) For any León-Perron operator $\widehat{P}_\gamma = \gamma I + (1-\gamma)\mathit{\Pi}$ with $\gamma \in [0, 1)$ and any bounded function g,

$$\|g\|_\pi \leq \|\!|G\widehat{P}_\gamma G|\!\|_\pi^{1/2},$$

where G is the multiplication operator of the bounded function g, i.e.,

$$Gh(x) = g(x)h(x), \quad \forall x,\ \forall h \in \mathcal{L}_2(\pi).$$

(ii) Let P be a Markov operator with absolute spectral gap $1-\gamma$, and let $\widehat{P}_\gamma = \lambda I + (1-\gamma)\mathit{\Pi}$. For any $h_1, h_2 \in \mathcal{L}_2(\pi)$,

$$|\langle Ph_1, h_2\rangle_\pi| \leq \langle \widehat{P}_\gamma h_1, h_1\rangle_\pi^{1/2} \langle \widehat{P}_\gamma h_2, h_2\rangle_\pi^{1/2}.$$

(iii) Let P be a Markov operator with absolute spectral gap $1-\gamma$, and let $\widehat{P}_\gamma = \gamma I + (1-\gamma)\mathit{\Pi}$. For any self-adjoint operators S_1, S_2,

$$|\!|\!| S_1 P S_2 |\!|\!|_\pi \le |\!|\!| S_1 \widehat{P}_\gamma S_1 |\!|\!|_\pi^{1/2} |\!|\!| S_2 \widehat{P}_\gamma S_2 |\!|\!|_\pi^{1/2}.$$

Proof of Lemma B.1. (i) The case of $g \equiv 0$ trivially holds. For any non-zero g,

$$\begin{aligned}\|g\|_\pi^2 |\!|\!| G\widehat{P}_\gamma G |\!|\!|_\pi &\ge \langle g, G\widehat{P}_\gamma G g\rangle_\pi \ge \langle g^2, \widehat{P}_\gamma g^2\rangle_\pi \\ &= \pi(g^2)^2 + \gamma \|(I-\mathit{\Pi})g^2\|_\pi^2 \ge \pi(g^2)^2 = \|g\|_\pi^4.\end{aligned}$$

Dividing both sides by $\|g\|_\pi^2$ completes the proof.

(ii) Using the fact that $P\mathit{\Pi} = \mathit{\Pi} P = \mathit{\Pi}$ and the self-adjointness of $I - \mathit{\Pi}$,

$$\begin{aligned}|\langle Ph_1, h_2\rangle_\pi| &= |\langle (I-\mathit{\Pi})(P-\mathit{\Pi})(I-\mathit{\Pi})h_1, h_2\rangle_\pi + \langle \mathit{\Pi} h_1, h_2\rangle_\pi| \\ &= |\langle (P-\mathit{\Pi})(I-\mathit{\Pi})h_1, (I-\mathit{\Pi})h_2\rangle_\pi + \langle \mathit{\Pi} h_1, h_2\rangle_\pi| \\ &\le |\langle (P-\mathit{\Pi})(I-\mathit{\Pi})h_1, (I-\mathit{\Pi})h_2\rangle_\pi| + |\langle \mathit{\Pi} h_1, h_2\rangle_\pi| \\ &\le \gamma \|(I-\mathit{\Pi})h_1\|_\pi \|(I-\mathit{\Pi})h_2\|_\pi + |\pi(h_1)\pi(h_2)| \\ &\le \sqrt{\gamma\|(I-\mathit{\Pi})h_1\|_\pi^2 + \pi(h_1)^2} \cdot \sqrt{\gamma\|(I-\mathit{\Pi})h_2\|_\pi^2 + \pi(h_2)^2} \\ &= \langle \widehat{P}_\gamma h_1, h_1\rangle_\pi^{1/2} \cdot \langle \widehat{P}_\gamma h_2, h_2\rangle_\pi^{1/2}.\end{aligned}$$

(iii) Using part (ii) and the self-adjointness of S_1, S_2, $S_1\widehat{P}_\gamma S_1$ and $S_2\widehat{P}_\gamma S_2$,

$$\begin{aligned}|\!|\!| S_1 P S_2 |\!|\!|_\pi &= \sup_{h_1,h_2:\|h_1\|_\pi=\|h_2\|_\pi=1} |\langle S_1PS_2h_2, h_1\rangle_\pi| \\ &= \sup_{h_1,h_2:\|h_1\|_\pi=\|h_2\|_\pi=1} |\langle PS_2h_2, S_1h_1\rangle_\pi| \\ &\le \sup_{h_1,h_2:\|h_1\|_\pi=\|h_2\|_\pi=1} \langle \widehat{P}_\gamma S_1h_1, S_1h_1\rangle_\pi^{1/2} \cdot \langle \widehat{P}_\gamma S_2h_2, S_2h_2\rangle_\pi^{1/2} \\ &= \sup_{h_1,h_2:\|h_1\|_\pi=\|h_2\|_\pi=1} \langle S_1\widehat{P}_\gamma S_1h_1, h_1\rangle_\pi^{1/2} \cdot \langle S_2\widehat{P}_\gamma S_2h_2, h_2\rangle_\pi^{1/2} \\ &= \sup_{h_1:\|h_1\|_\pi=1} \langle S_1\widehat{P}_\gamma S_1h_1, h_1\rangle_\pi^{1/2} \cdot \sup_{h_2:\|h_2\|_\pi=1} \langle S_2\widehat{P}_\gamma S_2h_2, h_2\rangle_\pi^{1/2} \\ &= |\!|\!| S_1 \widehat{P}_\gamma S_1 |\!|\!|_\pi^{1/2} |\!|\!| S_2 \widehat{P}_\gamma S_2 |\!|\!|_\pi^{1/2}.\end{aligned}$$

□

Proof of Lemma 1. Note that $E^{tf} = (E^{tf/2})^2$, that $E^{tf/2}$ is self-adjoint, and that $E^{tf}\mathit{1} = e^{tf}$. By an elementary calculus,

$$\begin{aligned}\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f_i(X_i)}\right] &= \left\langle \mathit{1}, E^{tf_1}\left(\prod_{i=2}^n PE^{tf_i}\right)\mathit{1}\right\rangle_\pi \\ &= \left\langle \mathit{1}, E^{tf_1/2}\left(\prod_{i=1}^{n-1} E^{tf_i/2}PE^{tf_{i+1}/2}\right)E^{tf_n/2}\mathit{1}\right\rangle_\pi \\ &= \left\langle e^{tf_1/2}, \left(\prod_{i=1}^{n-1} E^{tf_i/2}PE^{tf_{i+1}/2}\right)e^{tf_n/2}\right\rangle_\pi.\end{aligned}$$

It follows that

$$\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f_i(X_i)}\right] \le \|e^{tf_1/2}\|_\pi \prod_{i=1}^{n-1} |||E^{tf_i/2}PE^{tf_{i+1}/2}|||_\pi \|e^{tf_n/2}\|_\pi.$$

Taking $g = e^{tf_1/2}$ or $e^{tf_n/2}$, $S_1 = E^{tf_i/2}$ and $S_2 = E^{tf_{i+1}/2}$ in Lemma B.1 yields

$$\begin{aligned}
\|e^{tf_1/2}\|_\pi &\le |||E^{tf_1/2}\widehat{P}_\gamma E^{tf_1/2}|||_\pi^{1/2},\\
\|e^{tf_n/2}\|_\pi &\le |||E^{tf_n/2}\widehat{P}_\gamma E^{tf_n/2}|||_\pi^{1/2},\\
|||E^{tf_i/2}PE^{tf_{i+1}/2}|||_\pi &\le |||E^{tf_i/2}\widehat{P}_\gamma E^{tf_i/2}|||_\pi^{1/2}|||E^{tf_{i+1}/2}\widehat{P}_\gamma E^{tf_{i+1}/2}|||_\pi^{1/2}.
\end{aligned}$$

Putting them together completes the proof. □

Lemma B.2. Let $\widehat{P}_\gamma = \gamma I + (1-\gamma)\mathit{\Pi}$ be a León-Perron operator. Let $f : \mathcal{X} \to \mathbb{R}$ be a simple function taking finitely many values. That is, there exists a finite set $\{\beta_1, \dots, \beta_k\}$ with $\beta_1 > \cdots > \beta_k$ such that $f^{-1}(\beta_j) \coloneqq \{x \in \mathcal{X} : f(x) = \beta_j\}$ satisfies

$$\pi(f^{-1}(\beta_j)) > 0,\ \forall 1 \le j \le k; \qquad \sum_{j=1}^k \pi(f^{-1}(\beta_j)) = 1.$$

Let

$$F(r) = \pi\left(\frac{(1-\gamma)e^f}{r-\gamma e^f}\right) = \sum_{j=1}^k \frac{(1-\gamma)e^{\beta_j}}{r-\gamma e^{\beta_j}}\pi(f^{-1}(\beta_j)).$$

The following statements hold.

(i) Each solution $r_\star$ to $F(r_\star) = 1$ is an eigenvalue of $E^{f/2}\widehat{P}_\gamma E^{f/2}$ associated with eigenfunction

$$h_\star = \frac{(1-\gamma)e^{f/2}}{r_\star - \gamma e^f}.$$

There are k such solutions $r_j \in (\gamma e^{\beta_j}, \gamma e^{\beta_{j-1}})$ for $j = 1, \dots, k$ (letting $\beta_0 = \infty$).

(ii) Regarding $\mathcal{S}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))$, the spectrum of operator $E^{f/2}\widehat{P}_\gamma E^{f/2}$ acting on $\mathcal{L}_2(\pi)$, we have

$$\begin{aligned}
\{r_j : 1 \le j \le k\} &\subseteq \mathcal{S}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))\\
&\subseteq \{\gamma e^{\beta_j} : 1 \le j \le k\} \cup \{r_j : 1 \le j \le k\}.
\end{aligned}$$

(iii) Regarding r_1, the largest eigenvalue of operator $E^{f/2}\widehat{P}_\gamma E^{f/2}$, we have

$$|||E^{f/2}\widehat{P}_\gamma E^{f/2}|||_\pi = r_1.$$

(iv) Let $\{\widehat{X}_i\}_{i\ge1}$ be a Markov chain driven by $\widehat{P}_\gamma$ then

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{E}_\pi\left[e^{\sum_{i=1}^n f(\widehat{X}_i)}\right] \ge \log r_1 = \log |||E^{f/2}\widehat{P}_\gamma E^{f/2}|||_\pi.$$

Proof of Lemma B.2. (i) Note that $\pi(e^{f/2}h_\star) = F(r_\star) = 1$. We have

$$\begin{aligned}
&E^{f/2}\widehat{P}_\gamma E^{f/2}h_\star(x) - r_\star h_\star(x) \\
&= \gamma e^{f(x)}h_\star(x) + (1-\gamma)e^{f(x)}\pi(e^{f/2}h_\star) - r_\star h_\star(x) \\
&= \gamma e^{f(x)}h_\star(x) + (1-\gamma)e^{f(x)}F(r_\star) - r_\star h_\star(x).
\end{aligned}$$

Plugging $F(r_\star) = 1$ and $h_\star = (1-\gamma)e^{f/2}/(r_\star - \gamma e^f)$ into the last line yields 0. This verifies that $r_\star$ is an eigenvalue of $E^{f/2}\widehat{P}_\gamma E^{f/2}$ with eigenfunction $h_\star$.

On each interval $(\gamma e^{\beta_j}, \gamma e^{\beta_{j-1}})$, function $F(r)$ continuously decreases to $-\infty$ (or 0 if $j = 1$) as $r \uparrow \gamma e^{\beta_{j-1}}$, and increases to $+\infty$ as $r \downarrow \gamma e^{\beta_j}$. Thus there exists $r_j \in (\gamma e^{\beta_j}, \gamma e^{\beta_{j-1}})$ such that $F(r_j) = 1$.

(ii) The operator $E^{f/2}\widehat{P}_\gamma E^{f/2}$ is self-adjoint, thus its spectrum consists of the discrete spectrum $\mathcal{S}_{\text{d}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))$, which consists of isolated eigenvalues of finite multiplicity, and the essential spectrum $\mathcal{S}_{\text{ess}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))$.

First, we show that $\mathcal{S}_{\text{d}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi)) \subseteq \{\gamma e^{\beta_j} : 1 \le j \le k\} \cup \{r_j : 1 \le j \le k\}$ by showing that any eigenvalue $r_\star$ belongs to either $\{\gamma e^{\beta_j} : 1 \le j \le k\}$ or $\{r_j : 1 \le j \le k\}$. Consider any pair of eigenvalue $r_\star$ and non-zero eigenfunction $h_\star$ such that

$$\begin{aligned}
r_\star h_\star(x) &= E^{f/2}\widehat{P}_\gamma E^{f/2}h_\star(x) \\
&= \gamma e^{f(x)}h_\star(x) + (1-\gamma)e^{f(x)/2}\pi(e^{f/2}h_\star), \quad \pi\text{-a.e. } x.
\end{aligned}$$

If $\pi(e^{f/2}h_\star) = 0$ then the above display implies that $(r_\star - \gamma e^{f(x)})h_\star(x) = 0$ for π-a.e. x. There exists at least one index j such that $h_\star(x)$ is not identically zero on the set $f^{-1}(\beta_j)$, rendering $r_\star = \gamma e^{\beta_j}$.

If $\pi(e^{f/2}h_\star) \neq 0$ and $r_\star \notin \{\gamma e^{\beta_j} : 1 \le j \le k\}$ then the last display is solved by the eigenfunction

$$h_\star(x) = \frac{(1-\gamma)e^{f(x)/2}}{r_\star - \gamma e^{f(x)}}\pi(e^{f/2}h_\star), \quad \pi\text{-a.e. } x.$$

Multiplying both sides with $e^{f(x)/2}$, taking expectation of both sides with respect to π, and dividing both sides by $\pi(e^{f/2}h_\star) \neq 0$ yields $F(r_\star) = 1$, that means $r_\star \in \{r_j : 1 \le j \le k\}$.

Next, we show that $\mathcal{S}_{\text{ess}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi)) \subseteq \{\gamma e^{\beta_j} : 1 \le j \le k\}$ by Weyl's theorem on essential spectrum (Weyl, 1909). Write

$$E^{f/2}\widehat{P}_\gamma E^{f/2} = \gamma E^f + (1-\gamma)E^{f/2}\mathit{\Pi} E^{f/2}$$

in the form of a self-adjoint operator cE^f perturbed by another self-adjoint operator $(1-\gamma)E^{f/2}\mathit{\Pi} E^{f/2}$. The perturbation $(1-\gamma)E^{f/2}\mathit{\Pi} E^{f/2}$ is of finite rank and thus compact. Weyl's theorem says that the essential spectrum of a self-adjoint operator is invariant to the perturbation of a self-adjoint, compact operator. Hence $\widehat{P}_\gamma$ shares the same essential spectrum with γE^f. Note that γE^f is the multiplication operator of function γe^f. Its spectrum is the essential range of γe^f, which is simply $\{\gamma e^{\beta_j} : 1 \le j \le k\}$. Thus

$$\mathcal{S}_{\text{ess}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi)) = \mathcal{S}_{\text{ess}}(\gamma E^f|\mathcal{L}_2(\pi)) \subseteq \{\gamma e^{\beta_j} : 1 \le j \le k\}.$$

Combining (i) and results for $\mathcal{S}_{\text{d}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))$ and $\mathcal{S}_{\text{ess}}(E^{f/2}\widehat{P}_\gamma E^{f/2}|\mathcal{L}_2(\pi))$ completes the proof.

(iii) By (i) and (ii), r_1 is the spectral radius of $E^{f/2}\widehat{P}_\gamma E^{f/2}$. Recall that $E^{f/2}\widehat{P}_\gamma E^{f/2}$ is self-adjoint. Thus $r_1 = |||E^{f/2}\widehat{P}_\gamma E^{f/2}|||_\pi$.

(iv) By (i), eigenvalue r_1 associates with eigenfunction

$$h_1 = \frac{(1-\gamma)e^{f/2}}{r_1 - \gamma e^f}$$

and $\langle h_1, e^{f/2}\rangle_\pi = F(r_1) = 1$. Let $\tilde{h}_1$ be the projection of $e^{f/2}$ onto h_1. It is elementary that

$$\tilde{h}_1 := \left\langle \frac{h_1}{\|h_1\|_\pi}, e^{f/2} \right\rangle_\pi \frac{h_1}{\|h_1\|_\pi} = \frac{h_1}{\|h_1\|_\pi^2},$$

and

$$\langle e^{f/2} - \tilde{h}_1, \tilde{h}_1\rangle_\pi = 0.$$

h_1 is the eigenfunction of the self-adjoint operator $E^{f/2}\widehat{P}_\gamma E^{f/2}$, thus

$$\begin{aligned} 0 &= \langle e^{f/2} - \tilde{h}_1, (E^{f/2}\widehat{P}_\gamma E^{f/2})^{n-1}\tilde{h}_1\rangle_\pi \\ &= \langle (E^{f/2}\widehat{P}_\gamma E^{f/2})^{n-1}(e^{f/2} - \tilde{h}_1), \tilde{h}_1\rangle_\pi. \end{aligned}$$

The self-adjoint operator $E^{f/2}\widehat{P}_\gamma E^{f/2}$ is positive semi-definite, so

$$\left\langle e^{f/2} - \tilde{h}_1, (E^{f/2}\widehat{P}_\gamma E^{f/2})^{n-1}(e^{f/2} - \tilde{h}_1)\right\rangle_\pi \ge 0.$$

Thus

$$\begin{aligned} \mathbb{E}_\pi\left[e^{\sum_{i=1}^n f(\widehat{X}_i)}\right] &= \left\langle e^{f/2}, (E^{f/2}\widehat{P}_\gamma E^{f/2})^{n-1} e^{f/2}\right\rangle_\pi \\ &\ge \left\langle \tilde{h}_1, (E^{f/2}\widehat{P}_\gamma E^{f/2})^{n-1}\tilde{h}_1\right\rangle_\pi \\ &= r_1^{n-1}/\|h_1\|_\pi^2, \end{aligned}$$

implying the desired result.

□

Proof of Lemma 2. Applying Lemma 1 yields

$$\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(\widehat{X}_i)}\right] \le |||E^{tf/2}\widehat{P}_\gamma E^{tf/2}|||_\pi^n.$$

It is left to show

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(\widehat{X}_i)}\right] \ge \log|||E^{tf/2}\widehat{P}_\gamma E^{tf/2}|||_\pi.$$

It is trivial for $t=0$. If $t\neq 0$, for $\epsilon>0$, define

$$f_\epsilon = \begin{cases} a+\epsilon\lceil (f-a)/\epsilon\rceil & \text{if } t>0,\\ a+\epsilon\lfloor (f-a)/\epsilon\rfloor & \text{if } t<0,\end{cases}$$

where $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ are the floor and ceiling function, respectively. Hence

$$tf_\epsilon \ge tf \ge tf_\epsilon - t\epsilon. \tag{S.2}$$

The self-adjoint operator $\widehat{P}_\gamma$ preserves the non-negativity of h (i.e. $\widehat{P}_\gamma h\ge 0$ if $h\ge 0$), so does self-adjoint operators $E^{tf/2}\widehat{P}_\gamma E^{tf/2}$ and $E^{tf_\epsilon/2}\widehat{P}_\gamma E^{tf_\epsilon/2}$. Thus

$$\begin{aligned}
|||E^{tf/2}\widehat{P}_\gamma E^{tf/2}|||_\pi &= \sup_{h:\ \|h\|_\pi=1} |\langle E^{tf/2}\widehat{P}_\gamma E^{tf/2}h,h\rangle_\pi| \\
&= \sup_{h\ge 0:\ \|h\|_\pi=1} \langle E^{tf/2}\widehat{P}_\gamma E^{tf/2}h,h\rangle_\pi \\
&\le \sup_{h\ge 0:\ \|h\|_\pi=1} \langle E^{tf_\epsilon/2}\widehat{P}_\gamma E^{tf_\epsilon/2}h,h\rangle_\pi & \text{[Using (S.2)]}\\
&= \sup_{h:\ \|h\|_\pi=1} |\langle E^{tf_\epsilon/2}\widehat{P}_\gamma E^{tf_\epsilon/2}h,h\rangle_\pi| \\
&= |||E^{tf_\epsilon/2}\widehat{P}_\gamma E^{tf_\epsilon/2}|||_\pi. & \text{(S.3)}
\end{aligned}$$

Note that f_ϵ takes finitely many values and thus fulfill the condition in Lemma B.2. Write

$$\begin{aligned}
\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(\widehat{X}_i)}\right] &\ge \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}_\pi\left[e^{\sum_{i=1}^n tf_\epsilon(\widehat{X}_i)-n\epsilon t}\right] & \text{[Using (S.2)]}\\
&= \liminf_{n\to\infty}\frac{1}{n}\log\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f_\epsilon(\widehat{X}_i)}\right]-\epsilon t \\
&\ge \log|||E^{tf_\epsilon/2}\widehat{P}_\gamma E^{tf_\epsilon/2}|||_\pi - \epsilon t & \text{[Lemma B.2(iv)]}\\
&\ge \log|||E^{tf/2}\widehat{P}_\gamma E^{tf/2}|||_\pi - \epsilon t. & \text{[Using (S.3)]}
\end{aligned}$$

Letting ϵ tend to 0 completes the proof. □

Proof of Lemma 4. Let $\theta(t)$ be the largest eigenvalue of matrix $\boldsymbol{E}^{ty/2}\widehat{\boldsymbol{Q}}_\gamma\boldsymbol{E}^{ty/2}$, then by Frobenius-Perron theorem,

$$|||\boldsymbol{E}^{ty/2}\widehat{\boldsymbol{Q}}_\gamma\boldsymbol{E}^{ty/2}|||_{\boldsymbol{\mu}} = \theta(t).$$

The largest eigenvalue $\theta(t)$ is the right solution to the following quadratic equation

$$\begin{aligned}0 &= \det(\theta \boldsymbol{I} - \boldsymbol{E}^{ty/2}\widehat{\boldsymbol{Q}}_\gamma \boldsymbol{E}^{ty/2})\\ &= \theta^2 - \left[(\gamma + (1-\gamma)(1-\mu))e^{ta} + (\gamma + (1-\gamma)\mu)e^{tb}\right]\theta + \gamma e^{ta+tb}\\ &= \theta^2 - (1+\gamma)\left[(1-p)e^{ta} + pe^{tb}\right]\theta + \gamma e^{ta+tb},\end{aligned}$$

where

$$p = \frac{\gamma + (1-\gamma)\mu}{1+\gamma}, \quad 1-p = \frac{\gamma + (1-\gamma)(1-\mu)}{1+\gamma}.$$

It suffices to show

$$\tilde{\theta}(t) = \exp\left(\boldsymbol{\mu}(y)\cdot t + \frac{1+\gamma}{1-\gamma}\cdot\frac{(b-a)^2}{4}\cdot\frac{t^2}{2}\right)$$

satisfies

$$\tilde{\theta}(t)^2 - (1+\gamma)\left[(1-p)e^{ta} + pe^{tb}\right]\tilde{\theta}(t) + \gamma e^{ta+tb} \geq 0, \text{ and} \tag{S.4}$$

$$\tilde{\theta}(t)^2 \geq \gamma e^{ta+tb}. \tag{S.5}$$

(S.4) is equivalent to

$$\frac{\tilde{\theta}(t) + \gamma e^{ta+tb}\tilde{\theta}(t)^{-1}}{1+\gamma} \geq (1-p)e^{ta} + pe^{tb}. \tag{S.6}$$

Let $\omega = (1+\gamma)/(1-\gamma)$. Using convexity of function $z \mapsto e^z$, the left-hand side of (S.6) is lower bounded as

$$\begin{aligned}&\frac{\tilde{\theta}(t) + \gamma e^{ta+tb}\tilde{\theta}(t)^{-1}}{1+\gamma}\\ =&\frac{\exp(t\boldsymbol{\mu}(y) + \omega(b-a)^2t^2/8) + \gamma\exp(at + bt - t\boldsymbol{\mu}(y) - \omega(b-a)^2t^2/8)}{1+\gamma}\\ \geq& \exp\left(\frac{t\boldsymbol{\mu}(y) + \omega(b-a)^2t^2/8 + \gamma at + \gamma bt - \gamma t\boldsymbol{\mu}(y) - \gamma\omega(b-a)^2t^2/8}{1+\gamma}\right)\\ =& \exp\left(t\cdot\frac{(1-\gamma)\boldsymbol{\mu}(y) + \gamma a + \gamma b}{1+\gamma} + \frac{(b-a)^2t^2}{8}\cdot\frac{(1-\gamma)\omega}{1+\gamma}\right)\\ =& \exp\left(t\cdot[(1-p)a + pb] + \frac{(b-a)^2t^2}{8}\right).\end{aligned}$$

The right hand side of (S.6) is the mgf of a Bernoulli random variable Z with $\mathbb{P}(Z = a) = 1-p$ and $\mathbb{P}(Z = b) = p$. By the classical Hoeffding's lemma, $\exp(t\cdot\mathbb{E}Z + (b-a)^2t^2/8) \geq \mathbb{E}e^{tZ}$. That is,

$$\exp\left(t\cdot[(1-p)a + pb] + \frac{(b-a)^2t^2}{8}\right) \geq (1-p)e^{ta} + pe^{tb}.$$

Concatenating the last two displays yields (S.6) (and thus (S.4)). On the other hand, (S.5) holds as

$$\begin{aligned}\log\left(\tilde{\theta}(t)^2e^{-ta-tb}\right) &= \frac{\omega(b-a)^2t^2}{4} + (2\mu-1)(b-a)t\\ &\geq -\frac{(2\mu-1)^2}{\omega} \geq -\frac{1}{\omega} = -\frac{1-\gamma}{1+\gamma} \geq \log\gamma.\end{aligned}$$

□

Proof of Lemma 5. By an elementary calculus,

$$\begin{aligned}\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(X_i)}\right] &= \langle e^{tf/2}, (E^{tf/2}PE^{tf/2})^{n-1}e^{tf/2}\rangle_\pi \\ &\leq \sup_{\|h\|_\pi=1} |\langle (E^{tf/2}PE^{tf/2})^{n-1}h, h\rangle| \|e^{tf/2}\|_2^2\end{aligned}$$

By the power inequality for the numerical radius of operators (Berger, 1965; Pearcy, 1966),

$$\sup_{\|h\|_\pi=1} |\langle (E^{tf/2}PE^{tf/2})^{n-1}h, h\rangle_\pi| \leq \left(\sup_{\|h\|_\pi=1} |\langle E^{tf/2}PE^{tf/2}h, h\rangle_\pi|\right)^{n-1}.$$

Note that $E^{tf/2}PE^{tf/2}$ preserves non-negativity (i.e., $E^{tf/2}PE^{tf/2}h \geq 0$ for any non-negative function h),

$$\sup_{\|h\|_\pi=1} |\langle E^{tf/2}PE^{tf/2}h, h\rangle_\pi| = \sup_{\|h\|_\pi=1} \langle E^{tf/2}PE^{tf/2}h, h\rangle_\pi.$$

Let $R = (P + P^*)/2$ be the additive reversiblization of P then

$$\langle E^{tf/2}PE^{tf/2}h, h\rangle_\pi = \langle E^{tf/2}RE^{tf/2}h, h\rangle_\pi.$$

From $R = (I - \varPi)R(I - \varPi) + \varPi$ (since $R\varPi = \varPi R = \varPi$) and self-adjointness of $E^{tf/2}, I - \varPi$, it follows that

$$\begin{aligned}\langle E^{tf/2}RE^{tf/2}h, h\rangle_\pi &= \langle E^{tf/2}(I - \varPi)h, R(I - \varPi)E^{tf/2}h\rangle_\pi + \pi(e^{tf/2}h)^2 \\ &\leq \lambda_r \|(I - \varPi)E^{tf/2}h\|_\pi^2 + \pi(E^{tf/2}h)^2 \\ &\leq \max\{\lambda_r, 0\}\|(I - \varPi)E^{tf/2}h\|_\pi^2 + \pi(E^{tf/2}h)^2 \\ &= \langle E^{tf/2}\widehat{P}_{\max\{\lambda_r,0\}}E^{tf/2}h, h\rangle_\pi \\ &\leq \left|\!\left|\!\left| E^{tf/2}\widehat{P}_{\max\{\lambda_r,0\}}E^{tf/2} \right|\!\right|\!\right|_\pi \|h\|_\pi^2.\end{aligned}$$

Putting all the above displays together yields

$$\mathbb{E}_\pi\left[e^{t\sum_{i=1}^n f(X_i)}\right] \leq \left|\!\left|\!\left| E^{tf/2}\widehat{P}_{\max\{\lambda_r,0\}}E^{tf/2} \right|\!\right|\!\right|_\pi^{n-1} \|e^{tf/2}\|_\pi^2.$$

We conclude the proof by applying Lemma B.1(i) with $g = e^{tf/2}$ to bound

$$\|e^{tf/2}\|_\pi^2 \leq \left|\!\left|\!\left| E^{tf/2}\widehat{P}_{\max\{\lambda_r,0\}}E^{tf/2} \right|\!\right|\!\right|_\pi.$$

□

C Proofs of Other Theorems

Proof of Theorem 3. Let $B_1 = 1$ and $\{B_i\}_{i\geq 2}$ be i.i.d. Bernoulli$(1 - \lambda)$ random variables, and let $\{W_i\}_{i\geq 1}$ be i.i.d. $\mathcal{N}(0, 1)$ random variables. We construct the Markov chain $\{X_i\}_{i\geq 0}$ in the following way.

$$X_i = (1 - B_i)X_{i-1} + B_iW_i, \ \forall i \geq 1.$$

By induction,

$$X_i = \sum_{j=1}^{i} \left(\prod_{k=j+1}^{i} (1 - B_k) \right) B_j W_j.$$

Let

$$N_i = \sum_{j=i}^{n} \left(\prod_{k=j+1}^{i} (1 - B_k) \right) B_j,$$

then

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} N_i W_i, \quad N_i \geq 0, \quad \sum_{i=1}^{n} N_i = n.$$

Further,

$$\begin{aligned}
\mathbb{E}_\pi \left[e^{t \sum_{i=1}^n X_i} \right] &= \mathbb{E}_\pi \left[\mathbb{E}_\pi \left(e^{t \sum_{i=1}^n N_i W_i} \middle| N_1, \ldots, N_n \right) \right] \\
&= \mathbb{E}_\pi \left[e^{t^2 \sum_{i=1}^n N_i^2 / 2} \right] \\
&\geq \mathbb{P}_\pi \left(N_1 = n \right) \mathbb{E}_\pi \left[e^{t^2 \sum_{i=1}^n N_i^2 / 2} \middle| N_1 = n \right] \\
&= \lambda^n e^{t^2 n^2 / 2},
\end{aligned}$$

which could not be bounded by $e^{O(n) t^2 / 2}$ uniformly for $n \geq 1$ and $t \in \mathbb{R}$. □

Proof of Theorem 5. Let νP^{n_0} denote the n_0-step transition of ν. Write

$$\begin{aligned}
&\mathbb{E}_\nu \left[e^{t \sum_{i=n_0+1}^{n_0+n} f(X_i) - n\pi(f)} \right] \\
&= \mathbb{E}_{\nu P^{n_0}} \left[e^{t \sum_{i=1}^n f(X_i) - n\pi(f)} \right] && \text{[Markov property]} \\
&= \mathbb{E}_\pi \left[\frac{d(\nu P^{n_0})}{d\pi}(X_1) \cdot e^{t \sum_{i=1}^n f(X_i) - n\pi(f)} \right] && \text{[Change measure]} \\
&\leq \left\| \frac{d(\nu P^{n_0})}{d\pi} \right\|_{\pi,p} \times \left\{ \mathbb{E}_\pi \left[e^{qt \sum_{i=1}^n f(X_i) - n\pi(f)} \right] \right\}^{1/q} && \text{[Hölder's inequality]} \\
&\leq \left\| \frac{d(\nu P^{n_0})}{d\pi} \right\|_{\pi,p} \times \exp \left(q \cdot \frac{1 + \max\{\lambda_r, 0\}}{1 - \max\{\lambda_r, 0\}} \cdot \frac{n(b-a)^2}{4} \cdot \frac{t^2}{2} \right). && \text{[Theorem 2]}
\end{aligned}$$

It remains to show $\|d(\nu P^{n_0})/d\pi\|_{\pi,p} \leq C(\nu, n_0, p)$. For $p \in (1, \infty)$, from (Rudolf, 2012, Lemma 3.16 and Lemma 3.17) and the fact that $\lambda(P) = \lambda(P^*)$, it follows that

$$\begin{aligned}
\left\| \frac{d(\nu P^{n_0})}{d\pi} \right\|_{\pi,p} &\leq \left\| \frac{d(\nu P^{n_0} - \pi)}{d\pi} \right\|_{\pi,p} + 1 \\
&= \left\| [(P^*)^{n_0} - \Pi] \left(\frac{d\nu}{d\pi} \right) \right\|_{\pi,p} + 1 \leq C(\nu, n_0, p).
\end{aligned}$$

For $p = \infty$, we have

$$\left\| \frac{d(\nu P^{n_0})}{d\pi} \right\|_{\pi,\infty} = \left\| (P^*)^{n_0} \left(\frac{d\nu}{d\pi} \right) \right\|_{\pi,\infty} = \operatorname{ess\,sup} (P^*)^{n_0} \left| \frac{d\nu}{d\pi} \right| \leq \operatorname{ess\,sup} \left| \frac{d\nu}{d\pi} \right|.$$

□

Proof of Theorem 6. Let $\epsilon_n = \max_{j,k} |\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk}|$. For each $1 \le j \le p$ and each $1 \le k \le p$,

$$\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk} = \frac{1}{n}\sum_{i=1}^{n} f_j(X_i) f_k(X_i) - \pi(f_j f_k).$$

For simplicity of notation, let $\omega = \frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$. By Theorem 2,

$$\mathbb{P}_\pi\left(\left|\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk}\right| > \epsilon\right) \le 2\exp\left(-\frac{n\epsilon^2}{2\omega}\right).$$

A union bound yields

$$\mathbb{P}_\pi\left(\epsilon_n > \epsilon\right) \le 2p^2 \exp\left(-\frac{n\epsilon^2}{2\omega}\right).$$

It follows that, with probability at least $1 - 2d^{-\delta}$,

$$\epsilon_n \le \sqrt{2(2+\delta)\omega \log p/n}.$$

Next, write

$$\begin{aligned}
|||T_t(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}|||_1 &= \max_k \sum_j |T_t(\widehat{\mathbf{\Sigma}}_{jk}) - \mathbf{\Sigma}_{jk}| \\
&\le \max_k \sum_j |\mathbf{\Sigma}_{jk}| \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| \le t, |\mathbf{\Sigma}_{jk}| \le t\right) \\
&\quad + \max_k \sum_j |\mathbf{\Sigma}_{jk}| \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| \le t, |\mathbf{\Sigma}_{jk}| > t\right) \\
&\quad + \max_k \sum_j |\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk}| \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| > t, |\mathbf{\Sigma}_{jk}| > t\right) \\
&\quad + \max_k \sum_j |\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk}| \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| > t, |\mathbf{\Sigma}_{jk}| \le t\right).
\end{aligned}$$

The first term is bounded by $\max_k \sum_j t\mathbb{I}(\mathbf{\Sigma}_{jk} \ne 0) \le ts$; the second term is bounded by $\max_k \sum_j (t+\epsilon_n)\mathbb{I}(\mathbf{\Sigma}_{jk} \ne 0) \le (t+\epsilon_n)s$; the third term is bounded by $\max_k \sum_j \epsilon_n \mathbb{I}(\mathbf{\Sigma}_{jk} \ne 0) \le \epsilon_n s$; and the fourth term is bounded by

$$\begin{aligned}
\max_k \sum_j \epsilon_n \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| > t, |\mathbf{\Sigma}_{jk}| \le t\right) &\le \epsilon_n \max_k \sum_j \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| > t, t/2 < |\mathbf{\Sigma}_{jk}| \le t\right) \\
&\quad + \epsilon_n \max_k \sum_j \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk}| > t, |\mathbf{\Sigma}_{jk}| \le t/2\right) \\
&\le \epsilon_n s + \epsilon_n \max_k \sum_j \mathbb{I}\left(|\widehat{\mathbf{\Sigma}}_{jk} - \mathbf{\Sigma}_{jk}| > t/2\right) \\
&\le \epsilon_n s + \epsilon_n \max_k \sum_j \mathbb{I}\left(\epsilon_n > t/2\right).
\end{aligned}$$

Collecting these pieces together yields that if $\epsilon_n \le t/2$ then

$$|||T_t(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}|||_1 \le s(2t + 3\epsilon_n).$$

Putting it together with the scaling of ϵ_n yields that, if

$$t \geq 2\sqrt{2(2+\delta)\omega \log p/n},$$

then with probability at least $1 - 2d^{-\delta}$,

$$|||T_t(\widehat{\mathbf{\Sigma}}) - \mathbf{\Sigma}|||_1 \leq s\left(2t + 3\sqrt{2(2+\delta)\omega \log p/n}\right).$$

$\square$

Proof of Theorem 7. We first argue by contradiction that at least one of three following events given $j(t) = j$ must be true

$$\begin{aligned}
\mathcal{E}_0(t) &= \left\{ \widehat{f}_{j_\star, N_{j_\star}(t-1)} + \sqrt{\frac{c \log t}{2N_{j_\star}(t-1)}} \leq \pi_{j_\star}(f_{j_\star}) \right\}, \\
\mathcal{E}_1(t) &= \left\{ \widehat{f}_{j, N_j(t-1)} - \sqrt{\frac{c \log t}{2N_j(t-1)}} > \pi_j(f_j) \right\}, \\
\mathcal{E}_2(t) &= \left\{ N_j(t-1) < \frac{2c \log T}{\Delta_j^2} \right\}.
\end{aligned}$$

Suppose for the sake of contradiction that all the three events are false. Then

$$\begin{aligned}
\widehat{f}_{j_\star, N_{j_\star}(t-1)} + \sqrt{\frac{c \log t}{2N_{j_\star}(t-1)}} &> \pi_{j_\star}(f_{j_\star}) \\
&= \pi_j(f_j) + \Delta_j \\
&\geq \pi_j(f_j) + \sqrt{\frac{2c \log t}{N_j(t-1)}} \\
&\geq \widehat{f}_{j, N_j(t-1)} + \sqrt{\frac{c \log t}{2N_j(t-1)}},
\end{aligned}$$

implying j cannot be selected at round t, i.e. $j(t) \neq j$.

For any integer $u \geq 1$

$$\begin{aligned}
u &\geq \sum_{t=1}^{T} \mathbb{I}(j(t) = j, N_j(t-1) < u) \\
&\geq \sum_{t=1}^{u} \mathbb{I}(j(t) = j, N_j(t-1) < u) + \sum_{t=u+1}^{T} \mathbb{I}(j(t) = j, N_j(t-1) < u) \\
&= \sum_{t=1}^{u} \mathbb{I}(j(t) = j) + \sum_{t=u+1}^{T} \mathbb{I}(j(t) = j, N_j(t-1) < u),
\end{aligned}$$

which implies

$$
\begin{aligned}
N_j(T) &= \sum_{t=1}^{T} \mathbb{I}(j(t)=j) = \sum_{t=1}^{u} \mathbb{I}(j(t)=j) + \sum_{t=u+1}^{T} \mathbb{I}(j(t)=j) \\
&\leq u - \sum_{t=u+1}^{T} \mathbb{I}(j(t)=j, N_j(t-1)<u) + \sum_{t=u+1}^{T} \mathbb{I}(j(t)=j) \\
&= u + \sum_{t=u+1}^{T} \mathbb{I}(j(t)=j, N_j(t-1)\geq u).
\end{aligned}
$$

In particular, let $u = \lceil 2c\log T/\Delta_j^2 \rceil$ then

$$
\begin{aligned}
\mathbb{E}N_j(T) &\leq u + \sum_{t=u+1}^{T} \mathbb{P}(j(t)=j, N_j(t-1)\geq u) \\
&\leq u + \sum_{t=u+1}^{T} \mathbb{P}(j(t)=j, \mathcal{E}_2^c(t)) \\
&\leq u + \sum_{t=u+1}^{T} \mathbb{P}(j(t)=j, \mathcal{E}_0(t)\cup\mathcal{E}_1(t)) \\
&\leq u + \sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_0(t)) + \sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_1(t)).
\end{aligned}
$$

Proceed to bound $\sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_0(t))$ and $\sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_1(t))$. For simplicity of notation, let $\omega = \frac{1+\max\{\lambda_r,0\}}{1-\max\{\lambda_r,0\}}$. By Theorem 2,

$$
\begin{aligned}
\mathbb{P}(\mathcal{E}_0(t)) &\leq \sum_{s=1}^{t} \mathbb{P}\left(\widehat{f}_{j,s} - \sqrt{\frac{c\log t}{2s}} > \pi_j(f_j)\right) \\
&\leq \sum_{s=1}^{t} \exp\left(-\frac{2s}{\omega} \times \frac{c\log t}{2s}\right) \\
&= t^{-c/\omega+1},
\end{aligned}
$$

thus

$$
\sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_0(t)) \leq \sum_{t=2}^{\infty} t^{-c/\omega+1} \leq \frac{1}{c/\omega-2}.
$$

The same argument applies for $\sum_{t=u+1}^{T} \mathbb{P}(\mathcal{E}_1(t))$. It follows that

$$
\begin{aligned}
\mathbb{E}N_j(T) &\leq u + \frac{2}{c/\omega-2} \\
&\leq \frac{2c\log T}{\Delta_j^2} + 1 + \frac{2}{c/\omega-2} \\
&= \frac{2c\log T}{\Delta_j^2} + \frac{c/\omega}{c/\omega-2},
\end{aligned}
$$

which completes the proof. □

References

BERGER, C. (1965). A strange dilation theorem. *Notices Amer. Math. Soc* **12** 590.

BREIMAN, L. (1992). *Probability*. 1st ed. Society for Industrial and Applied Mathematics.

CONWAY, J. B. (2013). *A course in functional analysis*. Springer.

DURRETT, R. (2010). *Probability: theory and examples*. 5th ed.

FILL, J. A. (1991). Eigenvalue bounds on convergence to stationarity for non-reversible Markov chains, with an application to the exclusion process. *The Annals of Applied Probability* **1** 62–87.

LEÓN, C. A. and PERRON, F. (2004). Optimal Hoeffding bounds for discrete reversible Markov chains. *The Annals of Applied Probability* **14** 958–970.

MACKEY, G. W. (1957). Borel structure in groups and their duals. *Transactions of the American Mathematical Society* **85** 134–165.

ORBANZ, P. (2015). Lecture notes in probability theory II.

PEARCY, C. (1966). An elementary proof of the power inequality for the numerical radius. *The Michigan Mathematical Journal* **13** 289–291.

PRESTON, C. (2008). Some notes on standard Borel and related spaces. *arXiv preprint arXiv:0809.3066* .

RUDOLF, D. (2012). *Explicit error bounds for Markov chain Monte Carlo*. Ph.D. thesis, Institute of Mathematics, University Jena, Jena, Germany.

WEYL, H. (1909). Über beschränkte quadratische formen, deren differenz vollstetig ist. *Rendiconti del Circolo Matematico di Palermo (1884-1940)* **27** 373–392.