Concentration Phenomenon for Random Dynamical Systems: An Operator Theoretic Approach

Muhammad Abdullah Naeem MUHAMMAD.ABDULLAH.NAEEM@DUKE.EDU
Department of Electrical and Computer Engineering
Duke University
Durham, NC 27708, USA

Editors: N. Matni, M. Morari, G. J. Pappas

Abstract

Via operator theoretic methods, we formalize the concentration phenomenon for a given observable 'r' of a discrete time Markov chain with 'μ_π' as invariant ergodic measure, possibly having support on an unbounded state space. The main contribution of this paper is circumventing tedious probabilistic methods with a study of a composition of the Markov transition operator P followed by a multiplication operator defined by e^r. It turns out that even if the observable/ reward function is unbounded, but for some for some $q > 2$, $\|e^r\|_{q\to 2} \propto \exp\left(\mu_\pi(r) + \frac{2q}{q-2}\right)$ and P is hyperbounded with norm control $\|P\|_{2\to q} < e^{\frac{1}{2}[\frac{1}{2}-\frac{1}{q}]}$, sharp non-asymptotic concentration bounds follow. *Transport-entropy* inequality ensures the aforementioned upper bound on multiplication operator for all $q > 2$. The role of *reversibility* in concentration phenomenon is demystified. These results are particularly useful for the reinforcement learning and controls communities as they allow for concentration inequalities w.r.t standard unbounded obersvables/reward functions where exact knowledge of the system is not available, let alone the reversibility of stationary measure.

Keywords: Transportation inequalities, Harris recurrent Markov chain, uniform ergodicity, Spectral gaps, Hyper-boundedness/contractivity, Operator theory, Sample complexity.

1. Introduction

Motivation. Successful use of control theory in high stake applications stems from its rigorous theoretical foundations; properties of dynamical system under consideration can be extracted from the spectral analysis (eigenvalues/eigenvector) of an associated matrix. Last decade has seen a tremendous surge in research activity on non-asymptotic analysis/concentration phenomenon of system identification and reinforcement learning for random dynamical systems (e.g., Tu and Recht (2018); Hao et al. (2020); Oymak (2019); Fazel et al. (2018); Simchowitz et al. (2018); Sarkar et al. (2019); Zahavy et al. (2019); Ziemann and Tu (2022)); tedious probabilistic techniques are employed making analysis intractable. If not, assumptions like boundedness of obervables, discounting the cost or discretizing the state space are employed. To address these limitations, there is a need to bring concentration phenomenon for dynamical system on a same theoretical footing with controls.

Related Literature and Outline of Paper. Whether it is a question of learning value function corresponding to a stabilizing policy for a Markov decision process(MDP) (Tu and Recht (2018)) or system identification (e.g., Sattar and Oymak (2020); Foster et al. (2020); Simchowitz et al. (2018)), the existing analysis techniques heavily rely on mixing arguments to conclude 'time-distant temporally dependent samples behave as independent identically distributed (iid). However, the notion of time-distant samples being uncorrelated is equivalent to L^2 *spectral gap or Poincare inequality* (see chapter 7 of Rosenblatt (2012)). Secondly, probabilistic analysis are tedious, opaque and leading to weak results (strong results are often limited to unrealistic assumptions).

Recognizing these limitation, *we study non-asymptotic analysis of policy evaluation for average reward based control on continuous state space: where expected value of the reward w.r.t stationary distribution of a Markov Decision Process (MDP) is approximated by its' empirical averages, with high probability.* Our approach is inspired from a conjecture put forth by Simon and Høegh-Krohn (1972) regarding spectral gaps for hyperbounded markov semi-groups, which was recently solved by Miclo (2015) for reversible case and by Glück (2020) in full generality. A remarkable advantage of this approach is uplifting a nonlinear random dynamical system into infinite dimensional space where it behaves as a linear system. We then, as described in Section 2, employ direct sum decomposition of the underlying Banach space, as in Lin (1974), to prove uniform ergodic theorem for Markov transition operator. Consequently, we show how hyperboundedness leads to the desired spectral properties (given in Luecke (1977)) for uncorellation of Poincare Inequality).

To the best of authors knowledge, Kloeckner (2017),Kloeckner (2019) interpreted the spectral gap in Markov transition operator into non-asymptotic concentration bounds but his analysis only works on bounded observables. In Section 3, we introduce Feynman-Kac semi-group related to unbounded observable, which based on spectral properties of the multiplication operator allows us to derive a sufficient condition for sharp concentration: *transport-entropy inequality.*; we then verify our results on linear Gaussian systems. Wang and Wu (2020) offers an operator theoretic treatment for concentration but pertinent functional inequalities are valid only reversibility assumption. Finally, concluding remarks are made in Section 4 along with discussion on future problems for consideration.

Notations. Space of probability measures on a metric space $(\mathcal{X}, d)$ is denoted by $\mathcal{P}(\mathcal{X})$. . $\mathbb{E}$ is used to denote expectation. For a function r, we use $< r >_\mu$ to denote expectation of r w.r.t μ. Given $\mu, \nu \in \mathcal{P}(\mathcal{X})$, Wasserstein metric of order $p \in [1, \infty)$ is defined as:

$$\mathcal{W}_d^p(\nu, \mu) = \left(\inf_{(X,Y)\in\Gamma(\nu,\mu)} \mathbb{E}\, d^p(X, Y) \right)^{\frac{1}{p}}; \tag{1}$$

here, $\Gamma(\nu, \mu) \in \mathcal{P}(\mathcal{X}^2)$, and $(X, Y) \in \Gamma(\nu, \mu)$ denotes that random variables (X, Y) follow some probability distributions on $\mathcal{P}(\mathcal{X}^2)$ with marginals ν and μ. Another way of comparing two probability distributions on $\mathcal{X}$ is via relative entropy which is defined as:

$$\begin{aligned} Ent(v||u) &= \int \log\left(\frac{d\nu}{d\mu}\right) d\nu, \qquad \text{if } \nu << \mu, \\ &= +\infty, \qquad\qquad\qquad \text{otherwise.} \end{aligned} \tag{2}$$

1.1. Problem Statement

Under the action of some state-dependent control policy π, we consider a closed loop random dynamical system of the form:

$$x_{k+1} = F\big(x_k, \pi(x_k), \epsilon_k\big), \quad \epsilon_k \quad iid,^{1} \tag{3}$$

1. For the sake of brevity, from now on we will exclude reference to π in state update equations as state dependent policy imlplies there exists some function G such that $F\big(x_k, \pi(x_k), \epsilon_k\big) = G(x_k, \epsilon_k)$.

where $x_k \in \mathbb{R}^n$ for all $k \in \mathbb{N}$ and $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$. In probabilistic language, closed loop dynamical system is a *Harris Ergodic Markov chain*. Assume, that Markov chain converges to a unique ergodic invariant measure μ_π. A question that is of utmost importance in average reward reinforcement learning (Zahavy et al. (2019)), when exact dynamics are unknown: if we have access to empirical averages of some unbounded reward function $r(x)$, we explore the following questions:

- *Concentration from simulating a single trajectory*

  When, how and why can we provide something similar to following exponential concentrations ?

$$\mu^N \left[\left| \frac{1}{N} \sum_{i=1}^{N} r(x_i) - < r >_{\mu_\pi} \right| > \epsilon \right] \leq 2 \exp\left(- \frac{N\epsilon^2}{K_{sys}(r)} \right), \tag{4}$$

  where r can be some unbounded function, in control theoretic or (RL) framework $r(x) := \|x\|$. $K_{sys}(r)$ is a constant dependent on system properties and 'smoothness' of r (related to Lipschitz constant)

- What sort of *regularity assumptions* on Markov transition operator are required to ensure concentration?

1.2. Contribution and Main Results

- This paper's fundamental contribution is reducing concentration phenomenon for Harris Ergodic Markov chain to a problem of bounding operator $e^r P : L^2(\mu_\pi) \longrightarrow L^2(\mu_\pi)$, a composition of Markov operator P and a multiplication operator e^r.
- A detailed mathematical and pedagogical overview of *Poincare/ Spectral gap inequality* is provided: turns out to be equivalent to temporally dependent samples having a concentration similar to iid samples.
- Even though multiplication operator e^r, associated with standard reward function (observable) used in continuous control and RL problems is *unbounded*, but *transport-entropy inequality* ensures that for all $q > 2$, $e^r : L^q(\mu_\pi) \longrightarrow L^2(\mu_\pi)$ is bounded. Combined with *hyperboundedness* of Markov kernel, concentration phenomenon is achieved.

2. Spectral gaps, Ergodic theorems and Poincare inequality

In order to mathematically express the phenomenon of *time-distant samples, although temporally dependent, behave as iid* : we will have to go to fundamentals of ergodic theory. $\mathbb{D}$ is used to denote unit disc in complex plane and $\mathbb{T}$ represents unit circle in complex plane. Consider the Banach space, with complex field, $L^p(\mu_\pi)$ of $p \in [1, \infty]$ integrable functions i.e., all measurable functions f with $\int |f(x)|^p \mu_\pi(dx) < \infty$. Conjugate power of p, $c(p)$ is defined as $c(p) \geq 1$ that satifies $\frac{1}{c(p)} + \frac{1}{p} = 1$. Every $g \in L^{c(p)}$ corresponds to bounded linear functional on $f \in L^p$ and its action is captured by $\langle f, g \rangle_{\mu_\pi} := \int f(x)\overline{g(x)} d\mu_\pi(x)$. We consider linearity in first argument of the inner product and boundedness follows by Cauchy-schwarz: $|\langle f, g \rangle_{\mu_\pi}| \leq \|f\|_p \|g\|_{c(p)}$. Markov transition operator $P : L^p \to L^p$ as $\|Pf\|_{L^p} = \left(\int |Pf(x)|^p \mu_\pi(dx) \right)^{\frac{1}{p}} = \left(\int | \int f(y)p(x, dy)|^p \mu_\pi(dx) \right)^{\frac{1}{p}} \leq$

$\|f\|_{L^p}$, where last inequality follows from Jensen inequality. P acts as an identity on constant functions: $P1 = 1$ and it is positive: $Pf \geq 0$ if $f \geq 0$. Markov chain is *reversible or satisfies detailed balance condition* if Markov transition operator P when viewed as an operator on Hilbert space $L^2(\mu_\pi)$ is equal to its' adjoint P^*: $\langle Pf, g\rangle := \langle f, P^*g\rangle_{\mu_\pi} = \langle f, Pg\rangle$. A simple observation reveals that P, P^*, PP^* and P^*P are all Markov operators with μ_π as invariant measure. Consider a sequence $(A_n)_{n\in\mathbb{N}}$ of bounded operators on some Banach space, there are different topologies (uniform,strong and weak) to study its convergence (see e.g., chapter 6 of Reed et al. (1980) for exact definitions).

Spectral decomposition

Definition 1 *Let A be a bounded operator on a Banach space $\mathcal{X}$, we will often denote it by $A \in \mathcal{B}(\mathcal{X})$. $\lambda \in \mathbb{C}$ is said to be in resolvent of A and denoted by $\lambda \in RS(A)$, if $\lambda I - A : \mathcal{X} \to \mathcal{X}$ is bijective. Bounded inverse theorem implies for $\lambda \in RS(A)$, $R_\lambda(A) := (\lambda I - A)^{-1} \in \mathcal{B}(\mathcal{X})$. Spectrum of A, denoted by $\sigma(A) := \mathbb{C} \setminus RS(A)$. Spectral radius of A is denoted by $\rho(A) := \sup|\lambda| : \lambda \in \mathbb{C}$. .Consequently, if there exists a $v \in \mathcal{X}$ and $\lambda \in \mathbb{C}$ such that $Av = \lambda v$ i.e., λ is an eigenvalue with corresponding eigenfunction v then $\lambda I - A$ is not injective which implies $\lambda \in \sigma(A)$. However, spectrum of A is not limited to eigenvalues, for a detailed monograph of spectral theory see e.g., Reed et al. (1980).*

Definition 2 *Let A be a bounded operator on some Banach space $\mathcal{X}$, it is called* Fredholm*: if (a) $dim[N(A)] < \infty$, (b) $dim[\mathcal{X} \setminus Im(A)] < \infty$ and (c) $Im(A)$ is closed. Condition (c) is redundant: it follows from (b). Where $dim[N(A)]$ is used to denote dimension of null space of A and $Im(A)$ means range of A. $dim[\mathcal{X} \setminus Im(A)]$ is often read as dimension of* cokernel *of A.*

Definition 3 *Index of a fredholm operator A , is defined as: $ind(A) = dim[N(A)] - dim[X \setminus Im(A)]$ and for some small perturbations $\Delta(A)$ to the operator $ind(A + \Delta(A)) = ind(A)$.*

Definition 4 *$\lambda \in \sigma(A)$ is said to be in essential spectrum of A, ($\lambda \in \sigma_{ess}(A)$) if and only if $\lambda I - A$ is not a Fredholm operator. $\lambda \in \sigma(A) \setminus \sigma_{ess}(A)$ is said to be in discrete spectrum, ($\lambda \in \sigma_{disc}(A)$) if and only if (a) λ is an isolated point of $\sigma(A)$ and (b) $\{\psi \in \mathcal{X} : A\psi = \lambda\psi\}$ is finite dimensional.*

Definition 5 *Markov operator is called* Hyperbounded*, if for some $1 \leq p < q \leq \infty$, P is a bounded operator from $L^p(\mu_\pi)$ to $L^q(\mu_\pi)$ (we will often call it $L^p - L^q$ Hyperboundedness). More stringent requirement of* Hypercontractivity*, requires P to be hyperbounded with $\|P\|_{L^p\to L^q} \leq 1$.*

Remark 6 *It follows from Jensen's inequality for all $r \geq 1$, $\|P\|_{L^r(\mu_\pi)} \leq 1$ and Riesz-Thorin interpolation implies for all $1 < p < q$, P is $L^p - L^q$ Hyperbounded.*

2.1. Ergodic theorems and consequences

Definition 7 *In functional analytic framework, the pair (P, μ_π) is said to be ergodic if any $f \in L^\infty(\mu_\pi)$ satisfies $Pf = f$, then f is a constant. In probabilistic language ergodic Markov chain is the one with unique invariant stationary distribution.*

Definition 8 *The pair (P, μ_π) is called aperiodic if for all $\lambda \in \mathbb{T} \setminus \{1\}$, $dim[N(\lambda I - P)] = 0$.*

Birkhoff Pointwise Ergodic theorem: Let $(x_n)_{n\in\mathbb{N}}$, be the samples from the ergodic Markov chain.

$$\frac{1}{N}\sum_{n=0}^{N-1} f(x_n) \longrightarrow \mu_\pi(f), \text{ almost everywhere } x_0 = x \ w.r.t \ \mu_\pi, \text{ and } f \text{ integrable w.r.t } \mu_\pi. \quad (5)$$

This is reminiscent of strong law of large numbers for iid sampling from distribution μ_π. However, a very natural requirement for having a concentration similar to iid setting is sharp decay of correlation i.e., for some $C < \infty$ and $\eta \in (0,1)$

$$|Cov_{\mu_\pi}[f(x_n), f(x_{n+m})]| \le C\eta^m Var_{\mu_\pi}(f), \quad \forall f \in L^2(\mu_\pi). \quad (6)$$

A first step towards formalizing preceding phenomenon is a concept related to *uniform ergodicity*.

Theorem 9 *If the pair (P,μ_π) is ergodic, let U be the orthogonal projection on $\{f \in L^2(\mu_\pi) : Pf = f\}$, we have a stronger version of mean ergodic theorem 'L^2- uniform ergodicity ', precisely said:*

$$\|\frac{1}{N}\sum_{n=0}^{N-1} P^n - U\|_{L^2(\mu_\pi)} \longrightarrow 0, \text{ that is the operator, } \frac{1}{N}\sum_{n=0}^{N-1} P^n \quad (7)$$

converges to a bounded linear projection, U in uniform operator topology

Proof If the pair (P,μ_π) is ergodic then so is the pair (P^*,μ_π). $Im(I-P)$ is closed, Why? *Fredholm argument*: because dimension of *cokernel* is finite: $dim[L^2(\mu_\pi) \setminus Im(I-P)] = dim[N(I-P^*)] = dim[N(I-P)] = 1$. Now consider the following decomposition:

$$(I-P) : N(I-P) \bigoplus N(I-P)^\perp \to Im(I-P) \bigoplus Im(I-P)^\perp. \quad (8)$$

Therefore, $I-P : N(I-P)^\perp \to Im(I-P)$ is bijective and by inverse mapping theorem $(I-P)^{-1} \in \mathcal{B}(Im(I-P), N(I-P)^\perp)$ so there exists a $K < \infty$ such that $\|(I-P)^{-1}\|_{Im(I-P)\to N(I-P)^\perp} \le K$ implying that the pair (P,μ_π) is *uniformly ergodic* because: given any $f \in Im(I-P)$ there exists a unique $\hat{g} \in N(I-P)^\perp$ such that $f = (I-P)\hat{g}$ and $\|\frac{1}{N}\sum_{n=0}^{N-1} P^n f\| = \|\frac{1}{N}\sum_{n=0}^{N-1} P^n(I-P)\hat{g}\| = \frac{\|(I-P^N)\hat{g}\|}{N} = \frac{\|(I-P^N)(I-P)^{-1}f\|}{N} \le 2\frac{\|(I-P)^{-1}f\|}{N} \le \frac{2K}{N}\|f\|$. So $\frac{1}{N}\sum_{n=0}^{N-1} P^n$ converges to 0 at a uniform rate on $Im(I-P)$. Notice that, when $Im(I-P)$ is closed: $Im(I-P) = N(I-P)^\perp$, because regardless of $Im(I-P)$ being closed : $Im(I-P) \subset N(I-P^*)^\perp$ and $Im(I-P)^\perp = N(I-P^*)$. Therefore, $Im(I-P)^\perp$ corresponds to $N(I-P)$ which by ergodicity assumption only comprises of constant function and $\frac{1}{N}\sum_{n=0}^{N-1} P^n_{|N(I-P)} = I$. ■

Remark 10 *Few points to note before moving to next theorem:*

1. *Under ergodicity and reversibility assumption, $I-P$ is an* Fredholm operator *with index 0, in simple words it means 1 is an isolated eigenvalue of P i.e., for some $\epsilon > 0$, $\sigma(P) = [-1, 1-\epsilon] \cup \{1\}$ and dimension of space of eigenfunctions associated to 1, is 1 .*

2. *we define $S := P_{|Im(I-P)}$ (when $Im(I-P)$ is closed), and since $Im(I-P)$ is invariant under P, implying $S : Im(I-P) \to Im(I-P)$. Because $(I-P)^{-1} \in \mathcal{B}(Im(I-P))$, $(I-S)^{-1}$ exists as well so $1 \notin \sigma(S)$ and from previous bullet point as S will contain subset of spectrum of P, indeed $\sigma(S) \subset [-1, 1-\epsilon]$*

Theorem 11 *If the pair (P, μ_π) is reversible, ergodic and aperiodic, then :*

$$\|P^n - U\|_{L^2(\mu_\pi)} \longrightarrow 0, \text{ in uniform operator topology} \tag{9}$$

Proof Consider the operator $I + P$, as a map:

$$I + P : N(I+P) \bigoplus N(I+P)^\perp \to Im(I+P) \bigoplus Im(I+P)^\perp. \tag{10}$$

$N(I+P) = \{f \in L^2(\mu_\pi) : Pf = -f\}$ which corresponds to *periodic* behavior but by hypothesis $dim[N(I+P)] = 0$ and $dim[L^2(\mu_\pi) \setminus Im(I+P)] = dim[N(I+P)] = 0$ (thanks to reversibility) so $I + P$ is a Fredholm operator implying that $-1 \notin \sigma_{ess}(P)$ (see e.g., Wu (2004)). Moreover, $-1 \notin \sigma(P)$ either. Since the spectrum of a bounded operator is always closed, implying there exists $\delta > 0$ such that $\sigma(S) \subset [-1+\delta, 1-\epsilon]$. Consequently $\rho(S)$, spectral radius of S satisfies $\rho(S) \leq max(|-1+\delta|, |1-\epsilon|) < 1$ and $P^n = S^n \bigoplus I_{N(I-P)}$ but $\rho(S) < 1$ implies $S^n \to 0$, because by Gelfand's formula for any $\rho \in (\rho(S), 1)$ there exists $M(\rho) < \infty$ such that $\|S^n\| \leq M(\rho)\rho^n$. ■

Spectral analysis of non-reversible P is more involved as it can lies inside a unit disc and this is where *hyperboundedness* comes into play.

Theorem 12 *Hyperboundedness of $L^2 - L^p$, where $p = 2+\epsilon$ and $\epsilon > 0$ implies that (a) for all $\lambda \in \mathbb{T}$, $dim[N(\lambda I - P)_2] < \infty$ and (b) $dim[N(\lambda I - P^*)_2] < \infty$.$(b)$independently implies $Im(\lambda I - P)_2$ is closed. Consequently:*

1. *$\sigma(P) \cap \mathbb{T}$ may only comprise of finite number of distinct eigenvalues and each distinct eigenvalue has a finite dimensional eigenspan.*
2. *$\sigma_{ess}(P)$ is contained inside some closed disc of radius $\alpha < 1$ in complex plane.*
3. *$\sigma(P)$ is 'gapped' i.e., it comprises of two disjoint sets :$\sigma(P) = \{\sigma(P) \cap \mathbb{T}\} \cup \{\sigma(P) \cap \alpha\mathbb{D}\}$.*

Proof Our proof is based on a result of Glück (2020): infinite dimensional L^p spaces are not isomorphic for $p \neq q$. If $f \in N[\lambda I - P]$, for $\lambda \in \mathbb{T}$, it must satisfy $|Pf(x)| = |f(x)|$. P is $L^2 - L^{p:=2+\epsilon}$ hyperbounded, for some $\epsilon > 0$. Then duality implies that P^* is $L^{c(p)} - L^2$ hyperbounded, where $c(p)$ is the conjugate power of p. Now recall, $P \in \mathcal{B}(L^p)$ and $dim[L^p \setminus Im(\lambda I - P)_p] = dim[N(\overline{\lambda} I - P^*)_{c(p)}]$. If $dim[N(\overline{\lambda} I - P^*)_{c(p)}] = \infty$, strict inclusion implies there exists a $\hat{g} \in L^{c(p)} \setminus L^2$ such that $P^*\hat{g} = \overline{\lambda}\hat{g}$ and $|P^*\hat{g}| = |\hat{g}|$ *which contradicts hyperboundedness*. Therefore, $dim[L^p \setminus Im(\lambda I - P)_p]$ is finite which implies $Im(\lambda I - P)$ is closed in $L^p$and as $L^2 \subsetneq L^{c(p)}$ we have $dim[L^2 \setminus Im(\lambda I - P)_2] = dim[N(\overline{\lambda} I - P^*)_2] < dim[N(\overline{\lambda} I - P^*)_{c(p)}] < \infty$ and we have that $Im(\lambda I - P)$ is closed when viewed as a map in L^2. Argument for $dim[N(\lambda I - P)_2]$ being finite follows the same hyperboundedness argument. Consequences follow from the disjoint nature of essential and discrete spectrum (see Theorem 7.9-7.11 of Reed et al. (1980)). ■

Theorem 13 *If the pair (P, μ_π) is ergodic, $L^2 - L^p$ hyperbounded for $p = 3, 4$ with norm condition $\|P\|_{L^2 - L^p} < 2^{\frac{1}{2} - \frac{1}{p}}$, then :*

$$\|P^n - U\|_{L^2(\mu_\pi)} \longrightarrow 0, \text{ in uniform operator topology} \tag{11}$$

Proof The norm condition $\|P\|_{L^2\to L^4} < 2^{\frac{1}{2}-\frac{1}{4}}$ or $\|P\|_{L^2\to L^3} < 2^{\frac{1}{2}-\frac{1}{3}}$ ensures aperiodicity (see theorem 5.1 and 5.3 Cohen et al. (2022)): only $\{1\} \in \sigma(P) \cap \mathbb{T}$ and corresponds to eigenspan of constant functions. Following the argument of theorem 11: $P^n = S^n \bigoplus I_{N(I-P)}$, where $S = P_{|Im(I-P)}$ and $\sigma(S) \subset \alpha\mathbb{D}$ for some $\alpha \in (0,1)$. Consequently $\rho(S) \leq \alpha$ and for all $\rho \in (\alpha, 1)$ there exists an $M(\rho) < \infty$ such that $\|S^n\| \leq M(\rho)\rho^n$ and result follows. ■

Remark 14 *A trivial corollary of* $\|P^n - U\|_{L^2(\mu_\pi)} \longrightarrow 0,$ *in uniform operator topology is that for some* $\rho \in (0,1)$*,* $C < \infty$ *and for all* $n \in \mathbb{N}$*:*

$$\|P^n\big(f - \mu_\pi(f)\big)\|_{L^2} \leq C\rho^n \|f - \mu_\pi(f)\|_{L^2}, \tag{12}$$

which is equivalent to sharp decay of correlation in (6).

3. Hyperboundedness and Transport-Entropy inequality implies concentration

After ensuring uncorrelation for time-distant samples, it is safe to relate to iid setting and remind ourselves that *sharp concentrations heavily rely on the ability to take exponential moments of the observable w.r.t underlying measure*. An attempt to analyse the uncorrelation and exponential integrability via a single linear operator brings us to:

Feynman-Kac semigroup plays an integral role in the study of fluctuations of time additive quantities of diffusion process in continuous time, Touchette (2018). The reader is referred to Wang and Wu (2020), for a detailed exposition on this topic in discrete time setting. We identify with the observable/reward function r, a Feynman-Kac semigroup which is a composition of Markov transition operator followed by an exponentiated multiplication operator w.r.t observable: $e^r P : D_{(\cdot)} \subseteq L^2(\mu_\pi) \to L^2(\mu_\pi)$ such that for all $g \in D_{(\cdot)}$ (domain on which semigroup operators can be viewed as bounded operator):

$$(e^r P)^n g(x) := \mathbb{E}_x[g(x_N) e^{\sum_{i=0}^{n-1} r(x_i)}] \ , \forall n \in \mathbb{N}. \tag{13}$$

Assume that $x_0 \sim \beta$ and $\beta << \mu_\pi$, we have the following upper bound on the deviation of the observable:

$$\mathbb{P}_\beta\left(\frac{1}{N}\sum_{i=0}^{N-1} r(x_i) - \mu_\pi(r) \geq \epsilon\right) \leq \left\|\frac{d\beta}{d\mu_\pi}\right\|_2 \inf_{s>0} \|(e^{sr}P)^N\|_{L^2-L^2} e^{-sN(\mu_\pi(r)+\epsilon)}. \tag{14}$$

So the task at hand is bounding the operator $e^r P$; that too in a meaningful way to get a favorable concentration result. P in itself is a positive contraction but exponentiated multiplication operator defined by unbounded observable require a short detour into its' spectral analysis and transport-entropy inequality.

Multiplication operator defined by e^r The operator $\big(e^r, D(e^r)_p\big)$ is closed and densely defined, see e.g., Proposition 3.10 from Chapter 1 in Engel and Nagel (2006). Essential range of e^r with μ_π as an underlying measure is defined as $e^r_{ess}(\mu_\pi) := \Big\{\lambda \in [1,\infty) : \mu_\pi\big(|e^r - \lambda| < \epsilon\big) \neq 0, \ \forall\, \epsilon > 0\Big\}$ and essential norm of e^r, $\|e^r\|_\infty := \sup\{\lambda \in e^r_{ess}(\mu_\pi)\}$. Multiplication operator e^r is bounded in some and hence all L^p, iff $\|e^r\|_\infty < \infty$. Consequently, $D(e^r)(p) = L^p$ and $\|e^r\|_{L^p\to L^p} = \|e^r\|_\infty$.

If the stationary measure μ_π is not compactly supported and is absolutely continuous w.r.t Lebesgue measure then $\|e^r\|_{L^p\to L^p} = \infty$. However, not all is lost, as suggested in previous section that we need some notion of hyperboundedness for uncorrelation. So if for some $p > 2$, $P : L^2 \to L^p$, in order to ensure $e^r P \in \mathcal{B}(L^2)$ it is sufficient to prove $e^r : L^p \to L^2$. This is where *transport-entropy inequality* comes into play.

Definition 15 *Consider metric space $(\mathcal{X}, d)$ and reference probability measure $\mu \in \mathcal{P}(\mathcal{X})$. Then we say that μ satisfies $\mathcal{T}_1^d(C)$ or to be concise $\mu \in \mathcal{T}_1^d(C)$ for some $C > 0$ if for all $\nu \in \mathcal{P}(\mathcal{X})$ and $\nu << \mu$, it holds that*

$$\mathcal{W}_d(\mu, \nu) \le \sqrt{2CEnt(\nu||\mu)} \tag{15}$$

Lemma 16 (Bobkov and Götze (1999)) *μ satisfies $\mathcal{T}_1^d(C)$ if and only if for all Lipschitz function f with $< f >_\mu := \mathbb{E}_\mu f$, it holds that*

$$\int e^{\lambda(f - <f>_\mu)} d\mu \le \exp(\frac{\lambda^2}{2} C\|f\|^2_{L(d)}), \quad \textit{where} \quad \|f\|_{L(d)} := \sup_{x\neq y} \frac{|f(x) - f(y)|}{d(x, y)}. \tag{16}$$

Theorem 17 *If $\mu_\pi \in \mathcal{T}_1^d(C_\pi)$ for some $C_\pi > 0$ and r is Lipschitz w.r.t metric d then we have for all $p > 2$, $e^r : L^p \longrightarrow L^2$ is a bounded operator with norm:*

$$\|e^r\|_{L^p\to L^2} \le \exp\left(\mu_\pi(r) + \frac{2pC_\pi\|r\|^2_{L(d)}}{(p-2)2}\right). \tag{17}$$

Proof Proof is a simple application of Cauchy-Schwarz and exponential moment inequality for Lipschitz functions under distribution μ_π satisfying Transport-entropy inequality. ■

Theorem 18 *Without any assumption of reversibility, If the invariant measure $\mu_\pi \in \mathcal{T}_1^d(C_\pi)$ for some $C_\pi > 0$ and r is Lipschitz w.r.t metric d and for some $q > 2$, $P : L^2(\mu_\pi) \longrightarrow L^q(\mu_\pi)$ is hyperbounded with norm $\|P\|_{L^2\to L^q} < e^{\frac{1}{2}\left[\frac{1}{2}-\frac{1}{q}\right]}$. Given $n \in \mathbb{Z}^+$ and $\delta \in (0, 1)$ and an initial distribution $\beta << \mu_\pi$, if $N \ge \ln\left(\|\frac{d\beta}{\mu_\pi}\|_2 \frac{1}{1-\delta}\right)\left(\frac{4n^2q}{(q-2)-4n^2q\ln\|P\|_{L^2\to L^q}}\right)$, we have that*

$$\mathbb{P}_\beta\left(\frac{1}{N}\sum_{i=0}^{N-1} r(x_i) - \mu_\pi(r) \ge \frac{\sqrt{C_\pi}\|r\|_{L(d)}}{n}\right) < 1 - \delta. \tag{18}$$

Proof Since, assumptions made in Theorem ensure $e^r P \in \mathcal{B}(L^2(\mu_\pi))$, we have that $\|(e^r P)^N\| \le \|e^r P\|^N$. Let $\epsilon(n) := \frac{\sqrt{C_\pi}\|r\|_{L(d)}}{n}$, where $\sqrt{C_\pi}\|r\|_{L(d)}$ is proportional to the square root of variance of r under distribution μ_π. Notice that as opposed to standard (ϵ, δ) probability arguments, here we have scaled $\epsilon = \epsilon(n)$ with a control variable n that we can increase to make ϵ arbitrarily small. From preceding discussion,

$$\|(e^{sr}P)^N\|_{L^2-L^2} e^{-sN(\mu_\pi(r)+\frac{\sqrt{C_\pi}\|r\|_{L(d)}}{n})} \le \|P\|^N_{2-q} e^{N\left(\left(\frac{2q}{q-2}\right)\frac{s^2}{2}C_\pi\|r\|^2_{L(d)} - s\sqrt{C_\pi}\frac{\|r\|_{L(d)}}{n}\right)}, \tag{19}$$

and a trivial calculation reveals

$$\inf_{s>0} \|P\|_{2-q}^{N} e^{N\left(\left(\frac{2q}{q-2}\right)\frac{s^2}{2}C_\pi\|r\|_{L(d)}^2 - s\sqrt{C_\pi}\frac{\|r\|_{L(d)}}{n}\right)} = e^{N\left(ln\|P\|_{2-q} - \frac{1}{2n^2}\left[\frac{q-2}{2q}\right]\right)}, \tag{20}$$

which we should be able to decrease as N increase; for every $n \in \mathbb{Z}^+$, which happens to be the case if $\|P\|_{L^2\to L^q} < e^{\frac{1}{2}\left[\frac{1}{2}-\frac{1}{q}\right]}$ and all other results follow. ■

Remark 19 *Since our norm control $\|P\|_{L^2\to L^q} < e^{\frac{1}{2}\left[\frac{1}{2}-\frac{1}{q}\right]}$ is in harmony with upper bound for $q = 3, 4$ given by Cohen et al. (2022) which implies implies aperiodicity and consequently Poincare'/ L^2- Spectral gap for Markov transition operator. We suspect that for $q \in (2, 3)$ our upper bound might imply aperiodicity and hence L^2- Spectral gap.*

Assuming hypercontractivity, a stronger theoretical result similar to Wang and Wu (2020), can be deduced.

Corollary 20 *Without any reversibility assumption on the pair (P, μ_π), if the stationary distribution satisfies T-E inequality i.e., $\mu_\pi \in \mathcal{T}_1^d(C_\pi)$ and for some $p > 2$ associated transition kernel is hypercontractive i.e., $\|P\|_{2\to p} \le 1$ then for any initial distribution $\beta << \mu_\pi$ and $N \in \mathbb{N}$:*

$$\mathbb{P}_\beta\left(\frac{1}{N}\sum_{i=0}^{N-1} r(x_i) - \mu_\pi(r) \ge \epsilon\right) \le \left\|\frac{d\beta}{d\mu_\pi}\right\|_2 \exp\left(-\frac{N\epsilon^2(p-2)}{4C_\pi\|r\|_{L(d)}^2 p}\right). \tag{21}$$

Consequently, given $\epsilon > 0$ and $\delta \in (0, 1)$ we have that for all $N \ge \frac{\ln\left(\left\|\frac{d\beta}{\mu_\pi}\right\|_2 \frac{1}{1-\delta}\right) 4C_\pi\|r\|_{L(d)}^2 p}{\epsilon^2(p-2)}$ chain satisfies $\mathbb{P}_\beta\left(\frac{1}{N}\sum_{i=0}^{N-1} r(x_i) - \mu_\pi(r) \ge \epsilon\right) \le 1 - \delta$.

Example 1 *Let us verify these result on one dimensional linear Gaussian dynamical system with $|\alpha| < 1$:*

$$x_{n+1} = \alpha x_n + w_n. \ , w_n \sim \mathcal{N}(0, 1) \text{ and iid.} \tag{22}$$

An easy check reveals stationary distribution of (22) is $\gamma_{1,\alpha} := \mathcal{N}\left(0, \frac{1}{1-\alpha^2}\right)$

Theorem 21 *Given α such that $|\alpha| < 1$, there exitst $p := p(\alpha) > 2$ such that $\|P\|_{2\to p,\gamma_{1,\alpha}} \le 1$ (Hypercontractivity) and $\|\left(e^r P\right)^N\| \le e^{N\left(\gamma_{1,\alpha}(r) + \left(\frac{2p}{p-2}\right)\frac{\|r\|_{L(d)}^2}{2}\right)}$.*

Proof *A trivial application of change of variable and Stein's lemma reveals:*

$$\|P\|_{2\to p} \le \frac{1}{(1-\alpha^4)^{\frac{1}{4}}} \frac{1}{\left(1 - \frac{\alpha^2 p}{(1+\alpha^2)}\right)^{\frac{1}{2p}}} \frac{1}{e^{\frac{2}{p}\left(1 - \frac{\alpha^2 p}{(1+\alpha^2)}\right)}} \tag{23}$$

It is a common knowledge (see Corollary 7.2 of Gozlan and Léonard (2010)), $\gamma_{1,\alpha} \in \mathcal{T}_1^d(1)$ *i.e., satisfies T-E inequality with constant 1. It follows from previous equation (23) that for all* $g \in L^2(\gamma_{1,\alpha})$ *such that* $\|g\|_2 \leq 1$*, there exists* $p > 2$ *such that* $\|Pg\|_p \leq 1$ *by applying Cauchy-Schwarz to the powers* $\frac{p}{2}$ *and its' conjugate number* $\frac{p}{p-2}$ *we get:*

$$\|(e^r P)g\|_2 \leq \|Pg\|_p \left(\int e^{\frac{2p}{p-2} r(x)} d\gamma_{1,\alpha}(x) \right)^{\frac{p-2}{2p}} \leq \|P\|_{2\to p} \left(e^{\gamma_{1,\alpha}(r)} e^{\left(\frac{2p}{p-2}\right) \frac{\|r\|^2_{L(d)}}{2}} \right), \tag{24}$$

where (24) follows from hypercontractivity of transition operator and the fact that stationary distribution satisfies T-E inequality with $C = 1$ *and* d *is the euclidean metric. Conclusion follows from the trivial inequality* $\|(e^r P)^N\| \leq \|(e^r P)\|^N$ *and we have a shorp concentration as in Corollary 20.* ■

4. Conclusion and future work

1. In this work, we narrowed down the concentration phenomenon for Harris ergodic Markov chains to a study of composition of the Markov transition operator followed by a an exponentiated multiplication operator defined by observable under consideration.

2. Hyperboundedness and transport-entropy inequality suffices for concentration phenomen.

However, there are still unanswered questions that needs further exploration:

- *aperiodicity and hyperboundedness*: we believe that via continuity of *fredholm index* we can extend the norm control on $\|P\|_{L^2 - L^p}$ that implies *aperiodicity*. Currently, only result for $p = 3, 4$ is available.

- Although transport-entropy inequality can be verified via exponential type Lyapunov function introduced by authors in their previous work Naeem and Pajic (2022). Easily verifiable conditions for Hyperboundedness on continuous state space is still an open problem (besides the linear gaussian case).

References

Sergej G Bobkov and Friedrich Götze. Exponential integrability and transportation cost related to logarithmic sobolev inequalities. *Journal of Functional Analysis*, 163(1):1–28, 1999.

Guy Cohen et al. ℓ2-quasi-compact and hyperbounded markov operators. *arXiv preprint arXiv:2206.08003*, 2022.

Klaus-Jochen Engel and Rainer Nagel. *A short course on operator semigroups*. Springer Science & Business Media, 2006.

Maryam Fazel, Rong Ge, Sham Kakade, and Mehran Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. In *International Conference on Machine Learning*, pages 1467–1476, 2018.

Dylan Foster, Tuhin Sarkar, and Alexander Rakhlin. Learning nonlinear dynamical systems from a single trajectory. In *Learning for Dynamics and Control*, pages 851–861. PMLR, 2020.

Jochen Glück. Spectral gaps for hyperbounded operators. *Advances in Mathematics*, 362:106958, 2020.

Nathael Gozlan and Christian Léonard. Transport inequalities. a survey. *arXiv preprint arXiv:1003.3852*, 2010.

Botao Hao, Nevena Lazic, Yasin Abbasi-Yadkori, Pooria Joulani, and Csaba Szepesvari. Provably efficient adaptive approximate policy iteration. *arXiv preprint arXiv:2002.03069*, 2020.

Benoît Kloeckner. Effective limit theorems for markov chains with a spectral gap. *arXiv preprint arXiv:1703.09623*, 2017.

Benoît Kloeckner. Effective berry–esseen and concentration bounds for markov chains with a spectral gap. *The Annals of Applied Probability*, 29(3):1778–1807, 2019.

Michael Lin. On the uniform ergodic theorem. ii. *Proceedings of the American Mathematical Society*, 46(2):217–225, 1974.

Glenn R Luecke. Norm convergence of tn. *Canadian Journal of Mathematics*, 29(6):1340–1344, 1977.

Laurent Miclo. On hyperboundedness and spectrum of markov operators. *Inventiones mathematicae*, 200(1):311–343, 2015.

Muhammad Abdullah Naeem and Miroslav Pajic. Transportation-inequalities, lyapunov stability and sampling for dynamical systems on continuous state space. *arXiv preprint arXiv:2205.12448*, 2022.

Samet Oymak. Stochastic gradient descent learns state equations with nonlinear activations. In *Conference on Learning Theory*, pages 2551–2579, 2019.

Michael Reed, Barry Simon, et al. *I: Functional analysis*, volume 1. Gulf Professional Publishing, 1980.

Murray Rosenblatt. *Markov Processes, Structure and Asymptotic Behavior: Structure and Asymptotic Behavior*, volume 184. Springer Science & Business Media, 2012.

Tuhin Sarkar, Alexander Rakhlin, and Munther A Dahleh. Finite-time system identification for partially observed lti systems of unknown order. *arXiv preprint arXiv:1902.01848*, 2019.

Yahya Sattar and Samet Oymak. Non-asymptotic and accurate learning of nonlinear dynamical systems. 2020.

Max Simchowitz, Horia Mania, Stephen Tu, Michael I Jordan, and Benjamin Recht. Learning without mixing: Towards a sharp analysis of linear system identification. In *Conference On Learning Theory*, pages 439–473, 2018.

Barry Simon and Raphael Høegh-Krohn. Hypercontractive semigroups and two dimensional self-coupled bose fields. *Journal of Functional Analysis*, 9(2):121–180, 1972.

Hugo Touchette. Introduction to dynamical large deviations of markov processes. *Physica A: Statistical Mechanics and its Applications*, 504:5–19, 2018.

Stephen Tu and Benjamin Recht. Least-squares temporal difference learning for the linear quadratic regulator. In *International Conference on Machine Learning*, pages 5005–5014, 2018.

Neng-Yi Wang and Liming Wu. Transport-information inequalities for markov chains. *The Annals of Applied Probability*, 30(3):1276–1320, 2020.

Liming Wu. Essential spectral radius for markov semigroups (i): discrete time case. *Probability theory and related fields*, 128(2):255–321, 2004.

Tom Zahavy, Alon Cohen, Haim Kaplan, and Yishay Mansour. Average reward reinforcement learning with unknown mixing times. *arXiv preprint arXiv:1905.09704*, 2019.

Ingvar Ziemann and Stephen Tu. Learning with little mixing. *arXiv preprint arXiv:2206.08269*, 2022.