Proceedings of Machine Learning Research vol 211:1–12, 2023 5th Annual Conference on Learning for Dynamics and Control

Transportation-Inequalities, Lyapunov Stability and Sampling for Dynamical Systems on Continuous State Space

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

Miroslav Pajic MIROSLAV.PAJIC@DUKE.EDU
Department of Electrical and Computer Engineering
Duke University
Durham, NC 27708, USA

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

Abstract

We study the concentration phenomenon for discrete-time random dynamical systems with an unbounded state space. We develop a heuristic approach towards obtaining exponential concentration inequalities for dynamical systems using an entirely functional analytic framework. We also show that existence of exponential-type Lyapunov function, compared to the purely deterministic setting, not only implies stability but also exponential concentration inequalities for sampling from the stationary distribution, via *transport-entropy inequality* (T-E). These results have significant impact in *reinforcement learning* (RL) and *controls*, leading to exponential concentration inequalities even for unbounded observables (i.e., rewards), while neither assuming reversibility nor exact knowledge of the considered random dynamical system (assumptions at heart of concentration inequalities in statistical mechanics and Markov diffusion processes).

Keywords: Transportaion inequalities, Exponential Lyapunov function, Sample complexity and Nonlinear random dynamical systems

1. Introduction

Motivation. Last decade has seen tremendous advancements in non-asymptotic analysis of system identification and optimal control for linear time-invariant (LTI) 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)). Techniques and analysis developed in this paper have been initially motivated by non-asymptotic analysis of average reward-based optimal control of linear dynamical system (LDS) and switched linear dynamical system (SLDS) where expected value of the reward w.r.t stationary distribution of a Markov chain is approximated, with high probability, by its empirical averages. Although sample complexity for control/dynamical systems on continuous state spaces has been extremely popular recently, we still lack fundamental understanding of the factors leading to sharp concentration inequalities.

For example, every ergodic Markov chain is mixing (i.e., correlation decreases asymptotically), but as we will see in the following sections that does not necessarily imply exponential concentration (vaguely speaking: empirical averages concentrate sharply around expectation w.r.t to stationary

© 2023 M.A. Naeem & M. Pajic.

measure), especially from a single trajectory. To begin with, concentration with respect to (w.r.t) which class of functions? To expound upon these questions, in this work we leverage upon *concentration of measure phenomenon* for the process level law of a Markov chain via transport-entropy inequality (T-E) inequalities (for a good monograph on this topic see e.g., Marton et al. (2004); Djellout et al. (2004)). The T-E approach is compact, concise and clear; even leading to exponential concentration for unbounded reward function of the form $r(x) := \|x\|$ for dynamical systems.

Related work. Techniques used to prove concentration of a measure include martingale methods, exchangeable pairs (e.g., see Chatterjee (2007)) and functional inequalities (transportation-entropy, logarithmic Sobolev/hypercontractivity and Poincare) Ledoux (2001). However, exploring measure concentration for dependent random variables as in Markov chains on unbounded spaces limits the application of martingale methods and exchangeable pairs. The RL and system identification communities, pretty much exclusively, have used *independent block technique* Oymak (2019); Tu and Recht (2018) that seems to work when observables (i.e., rewards) are *bounded*. SLDS fall into the category of Harris ergodic Markov chain and following relevant work on them in Łatuszyński et al. (2013), by using martingale and independent block techniques, authors in Naeem and Pajic (2020) were able to bound the mean-squared error between the empirical average and the expected reward (unbounded) from stationary distribution of SLDS; however, the employed probabilistic methods are tedious, opaque and lead to weak results.

Stein methods (see e.g., Chatterjee and Dey (2010)), are well suited towards discrete setting in statistical physics but have limitations for the models under consideration. Only recently, Wang and Wu (2020) started a formal, functional analytic study for concentration inequalities in discrete-time setting using the transport-information (T-I) inequality. Verification of T-I is plausible either in the discrete state space setting, or when a Markov chain is reversible and posseses a spectral gap in space of square integrable functions w.r.t its stationary distribution (see Wang and Wu (2020) for more details). As we will see in concentration for nonlinear random dynamical systems, spectral gaps might only exist in Wasserstein sense. Reversibility assumption originates from study of Langevin-type stochastic differential equations used to model physical phenomenons in the nature; however, the scope of this paper is not limited to the reversibility assumption.

Note that Blower and Bolley (2005); Djellout et al. (2004) have studied T-E inequalities for the case of stable LDS. However, a general framework to provide exponential concentration inequalities for deviation of empirical averages of a Markov chain with respect to (w.r.t) some unbounded test function is missing. Concentration for nonlinear random dynamical systems is not well explored, partly because they are not necessarily contractive in trivial metric. However, leveraging on weighted transport-entropy inequalities introduced by Bolley and Villani (2005) and exponentially fast convergence of Harris ergodic Markov chains in Wasserstein metric (Hairer and Mattingly (2011)), in this work, we still manage to provide sharp concentration by introducing an exponential Lyapunov condition.

Notation. We use $I_n \in \mathbb{R}^{n\times n}$ to denote the n dimensional identity matrix. For random variables x and y, $Cov(x, y)$ denote the covariance. $\mathcal{B}_\alpha^n := \{x \in \mathbb{R}^n : \|x\| := \|x\|_2 \leq \alpha\}$ is the α-ball in $\mathbb{R}^n$. $\chi_{\{\}}()$ is the indicator function, whereas $\rho(A)$ and $\|A\|_2$ represent the spectral radius and the usual matrix 2-norm of $A \in \mathbb{R}^{n\times n}$, respectively. A sequence $\{a(N)\}_{N\in\mathbb{N}} \in \mathcal{O}(N)$, if it increases at most linearly in N (this is not limited to asymptotic results). Space of probability measure on $\mathcal{X}$(continuous space) is denoted by $\mathcal{P}(\mathcal{X})$ and space of its Borel subsets is represented by $\mathbb{B}\big(\mathcal{P}(\mathcal{X})\big)$.

For a function r and $\mu \in \mathcal{P}(\mathcal{X})$, we use $< r >_\mu$ to denote expectation of r w.r.t μ. Finally, for a set $\mathcal{K} \subseteq \{1, ..., M\}$, its complement is $\mathcal{K}^{\complement} := \{1, ..., M\} \setminus \mathcal{K}$.

On a metric space $(\mathcal{X}, d)$, for $\mu, \nu \in \mathcal{P}(\mathcal{X})$, we define Wasserstein metric of order $p \in [1, \infty)$ 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, $\Gamma(\nu, \mu) \in P(\mathcal{X}^2)$, and $(X, Y) \in \Gamma(\nu, \mu)$ implies that random variables (X, Y) follow some probability distributions on $P(\mathcal{X}^2)$ with marginals ν and μ. Another way of comparing two probability distributions on $\mathcal{X}$ is via relative entropy, which is defined as

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

1.1. Problem Statement

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

$$x_{k+1} = F\big(x_k, \pi(x_k), \epsilon_k\big), \quad \text{with } \epsilon_k \quad i.i.d,^{1} \tag{3}$$

where $x_k \in \mathbb{R}^n$ for all $k \in \mathbb{N}$ and $F : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$. We assume that the transition kernel converges to some stationary distribution μ_π under Wasserstein metric $\mathcal{W}_d$ equipped with some distance function d.

If we have access to empirical averages of some unbounded reward function $r(x)$, in this work, 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 a control-theoretic or RL framework (e.g., $r(x) := \|x\|$), and $K_{sys}(r)$ is a constant dependent on system properties and 'smoothness' of r (related to Lipschitz constant)?

- When explicit knowledge of the stationary distribution and system dynamics is not available, what are easy-to-verify sufficient conditions/functional inequalities to derive concentration from a single trajectory?

- In which ways stability of the considered dynamical system affects the concentration? and

- Is correlation between samples dependent on system stability?

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

1.2. Contribution and Main Results

Accordingly, the main contributions of this paper are as follows:

- This paper's fundamental contribution is a novel point of view towards getting exponential concentration inequalities for a random dynamical system, in an extremely tractable manner. To achieve this, we connect ideas and techniques *from particle methods, functional inequalities and convergence of Markov chains in Wasserstein distance*.
- *For contractive LDS, we show how to obtain exponential concentration of the form* (4) *from single trajectory.* The idea is that we can find a common metric on the space of probability measures, such that the transition kernels are contractive and uniformly satisfy transportation-entropy inequality.
- Leveraging upon *weighted T-E inequalities* developed by Bolley and Villani (2005), we introduce an *exponential-type Lyapunov condition for Markov chains*; if the Lyapunov condition holds, the stationary distribution satisfies the transport-entropy inequality. Consequently, empirical averages of the test function, evaluated on i.i.d samples from the stationary distribution, concentrate sharply around their mean.
- In case of non-linear random dynamical systems, such as Harris ergodic Markov chains (HEMCs), if exponential Lyapunov function exists, we show that one can simulate independent trajectories and, after a small burn-in period, average their rewards to obtain with high probability, a sharp estimate of the expected reward w.r.t the stationary distribution.

Outline of paper. In Section 2, we lay down a mathematical framework to obtain concentration for dependent random variables under the assumptions of uniform transport-entropy constants for a Markov transition kernel and Wasserstein contractivity. We conclude Section 2 with verification of the developed results on the problem of sample complexity in policy evaluation for average-reward based optimal control for LDS. Section 3 focuses on concentration for HEMCs that are not necessarily convergent in Wasserstein metric with the trivial euclidean distance. We show that if exponential-type Lyapunov function exists, empirical averages of test function evaluated on i.i.d samples from stationary distribution of Harris chain are sharply concentrated. Finally, this phenomenon is verified on an example of an SLDS.

2. Extending Concentration to Dependent Random Variables via Tensorization

2.1. Preliminaries

Before we introduce the mathematical framework to derive concentration for dependent random variables, we introduce the following results utilized later in this work.

Definition 1 *Consider metric space $(\mathcal{X}, d)$ and reference probability measure $\mu \in P(\mathcal{X})$. 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 $\nu \in P(\mathcal{X})$ it holds that*

$$\mathcal{W}_d(\mu, \nu) \leq \sqrt{2CEnt(\nu||\mu)}. \tag{5}$$

Lemma 2 (**Bobkov and Götze (1999)**) *μ satisfies $\mathcal{T}_1^d(C)$ if and only if for any Lipschitz function f with $< f >_\mu := \mathbb{E}_\mu f$, it holds that*

$$\int e^{\lambda(f-<f>_\mu)} d\mu \leq \exp(\frac{\lambda^2}{2} C \|f\|^2_{L(d)}), \quad \textit{where} \quad \|f\|_{L(d)} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x,y)}. \tag{6}$$

Remark 3 *(6) along with the Markov inequality implies that if we sample x from $\mu \in \mathcal{T}_1^d(C)$, then*

$$\mathbb{P}\left[\left| r(x) - < r >_\mu \right| > \epsilon\right] \leq 2 \exp\left(- \frac{\epsilon^2}{2C\|r\|^2_{L(d)}}\right). \tag{7}$$

Now, consider a Markov chain $x^N := (x_i)_{i=1}^N$ with distribution $\mu^N \in P\left(\mathcal{X}^N\right)$ and $P^m(x, \mathcal{B}) := \mathbb{P}(x_m \in \mathcal{B} | x_0 = x)$, for all Borel subsets $\mathcal{B}$ of $\mathcal{X}$. We can extend the metric d to $\mathcal{X}^N$ as

$$d_{(N)}(x^N, y^N) = \sum_{i=1}^N d(x_i, y_i). \tag{8}$$

If $\mu^N \in \mathcal{T}_1^{d_{(N)}}\big(\mathcal{O}(N)\big)$ and r is one Lipschitz, i.e., $\|r\|_{L(d)} \leq 1$, then $\Phi(x^N) := \frac{1}{N}\sum_{i=1}^N r(x_i)$ satisfies $\|\Phi\|_{L(d_{(N)})} \leq \frac{1}{N}$; plugging these results into (6), we obtain that

$$\mu^N\left[\left| \frac{1}{N}\sum_{i=1}^N r(x_i) - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N r(x_i)\right)\right| > \epsilon\right] \leq 2\exp\left(-\frac{N\epsilon^2}{2C}\right). \tag{9}$$

2.2. Contractivity and Uniform Transport Constants

As one would wonder from (8), when does the T-E for process level law of Markov chain, increases at worse linearly with dimension (in sample term)? Sufficient conditions (see e.g., Djellout et al. (2004); Bolley and Villani (2005)) are

$$(i) \quad P(x, \cdot) \in T_1^d(C), \qquad \text{for all } x \in \mathcal{X}, \text{and some } C > 0, \tag{10}$$

$$(ii) \quad \mathcal{W}_d(P(x,\cdot), P(y,\cdot)) \leq \hat{\lambda} d(x,y), \quad \text{for all } (x,y) \in \mathcal{X}^2 \text{and some } \hat{\lambda} \in [0,1). \tag{11}$$

Property (10) is often referred to as existence of a uniform transportation constant and (11) represents contractivity of the Markov Chain in the Wasserstein metric / *spectral gap in the Wasserstein sense*. Now, the following result holds.

Lemma 4 *If (10) and (11) hold, process level distribution of samples from a Markov chain $(x_1, \ldots, x_N)$, which we will denote as $Law(x_1, \ldots, x_N)$, denoted by μ^N satisfies $T_1^{d_{(N)}}\left(\frac{CN}{(1-\hat{\lambda})^2}\right)$, for all $N \in \mathbb{N}$.*

Proof See Theorem 2.5 of Djellout et al. (2004) for a detailed proof. ■

2.3. Sharp Deviation Inequalities for Average-Reward Based Optimal Control

In optimal control and average reward reinforcement learning for *continuing task*, it is often the case that under the action of some state dependent policy, the resulting closed-loop dynamical system under consideration, $(x_i)_{i=1}^N$, mixes to some stationary distribution $\mu_\pi \in P(\mathcal{X})$. When exact system parameters and state are unknown but time-averages of the reward function $r(x) := \|x\|$, although unbounded, are available, sharp concentration bounds of $\frac{1}{N}\sum_{i=1}^N r(x_i)$ around $< r >_{\mu_\pi}$ are of utmost importance to ensure policy update algorithm converges in finite time: as knowledge of $< r >_{\mu_\pi}$ is required in actor-critic algorithm for average reward case, see e.g.,Zhang and Ross (2021). Sufficient conditions for concentration of empirical averages of $r(x) := \|x\|$ or any Lipschitz function essentially boils down to showing that the process level law of the Markov chain $\mu^N \in \mathcal{T}_1^{d_{(N)}}\big(\mathcal{O}(N)\big)$.

Theorem 5 *Assume that under a metric d, a random dynamical system uniformly satisfies T-E inequality with constant $C > 0$ and is contractive in Wasserstein sense with constant $\gamma \in (0,1)$. If we start deterministically with $x_0 := x$, we have for any 1 Lipschitz function r, the following deviation inequality:*

$$\mu^N\left[\left|\frac{1}{N}\sum_{i=1}^N r(x_i) - < r >_{\mu_\pi}\right| > \frac{\mathcal{W}_d(P(x.\cdot),\mu_\pi)}{N(1-\gamma)} + \epsilon\right] \leq 2\exp\left(-\frac{N\epsilon^2(1-\gamma)^2}{2C}\right). \tag{12}$$

Proof Since the theorem assumptions satisfy the claim in (4), it holds that $\mu^N \in \mathcal{T}_1^{d_{(N)}}\big(\frac{CN}{(1-\gamma)^2}\big)$. Let $(y_i)_{i=1}^N$ be i.i.d samples from μ_π and assume that the Markov chain starts deterministicly with $x_1 = x$. Then, we have for all $\epsilon > 0$ it holds

$$\begin{aligned}
&\mu^N\left[\left|\frac{1}{N}\sum_{i=1}^N r(x_i) - < r >_{\mu_\pi}\right| > \frac{\mathcal{W}_d(P(x,\cdot),\mu_\pi)}{N(1-\gamma)} + \epsilon\right] \\
&\leq \mu^N\left[\left|\frac{1}{N}\sum_{i=1}^N r(x_i) - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N r(x_i)\right)\right| + \left|\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N r(y_i)\right) - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N r(x_i)\right)\right| > \frac{\mathcal{W}_d(P(x.\cdot),\mu_\pi)}{N(1-\gamma)} + \epsilon\right] \\
&\leq \mu^N\left[\left|\frac{1}{N}\sum_{i=1}^N r(x_i) - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N r(x_i)\right)\right| > \epsilon\right] \qquad (13) \\
&\leq 2\exp\left(-\frac{N\epsilon^2(1-\gamma)^2}{2C}\right), \qquad (14)
\end{aligned}$$

where (13) follows from contractive dynamics in Wasserstein distance. ■

Remark 6 *The structure of the aforementioned concentration inequality is inspired by the work in Malrieu (2001) on bounding deviations of empirical measure formed by interacting particle systems from their infinite particle limit (Mckean-Vlasov diffusion); see e.g., problem section in Villani (2003) for an involved discussion on this matter. However, there is a subtle difference in our approach, as we work with an ℓ^2 inspired metric on $\mathcal{X}^N$ as in Malrieu (2001), i.e.,*

$$d^2_{(N)}(x^N, y^N) := \sqrt{\sum\nolimits_{i=1}^N d^2(x_i, y_i)}. \tag{15}$$

Then, $\Phi(x^N) := \frac{1}{N}\sum_{i=1}^N r(x_i)$ *w.r.t* ℓ^2 *metric is only* $\frac{1}{\sqrt{N}}$ *Lipschitz — i.e.,* $\|\Phi\|_{L(d^2_{(N)})} \leq \frac{1}{\sqrt{N}}$ *and any hope for concentration would require* $\mu^N \in \mathcal{T}_1^{d^2_{(N)}}\big(\mathcal{O}(1)\big)$ *(i.e.,* dimension free concentration*), which is very difficult to check; this is feasible only in Markov diffusion processes with uniformly convex external potentials and symmetric interaction potentials, see e.g., Malrieu (2001).*

Decay of correlation. By combining conditions from (10) and (11), with Taylor's expansion for small λ (terms of order up to λ^2) appearing on both sides in Bobkov-Gotze dual form (6), for all $x \in \mathcal{X}$ it holds that

$$|Cov_{P_x}[f(x_n), f(x_{n+k})]| \leq \frac{\hat{\lambda}^k}{1-\hat{\lambda}^2} C\|f\|^2_{L(d)}. \tag{16}$$

Sharp Concentration via Single Trajectory of norm-Stable/ contractive Linear Dynamical Systems: Consider a discrete-time linear dynamical system (LDS) of the form

$$y_{t+1} = Ay_t + \epsilon_t, \quad \|A\|_2 = \hat{\lambda} < 1 \quad \text{and i.i.d } \epsilon_t \sim \mathcal{N}(0, \mathcal{I}_n). \tag{17}$$

With the trivial euclidean metric $d(x,y) := \|x-y\|$, the transition kernel from (17) satisfies $P(x,\cdot) \in T_1^d(1),\ \forall x \in \mathcal{X}$ Talagrand (1996) and $\mathcal{W}_d^2(P(x,\cdot), P(y,\cdot)) = \|Ax - Ay\| \leq \hat{\lambda} d(x,y)$ (see e.g., Givens et al. (1984)). Now, an application of Jensens' inequality reveals $\mathcal{W}_d(P(x,\cdot), P(y,\cdot)) \leq \mathcal{W}_d^2(P(x,\cdot), P(y,\cdot))$ and contractivity follows. As conditions (10) and (11) are satisfied, we can use coupling technique inspired by Marton et al. (2004) to prove that $\mathbb{P}^N$:=Law$(y_1, \ldots, y_N)$ of the LDS satisfies $T_1^{d_{(N)}}(\frac{N}{(1-\hat{\lambda})^2})$; and if ν_π is the invariant measure corresponding to (17), we have

$$\begin{aligned}
&\mathbb{P}^N\left[\left|\frac{1}{N}\sum_{i=1}^N \|y_i\| - <\|y\|>_{\nu_\pi}\right| > \frac{\mathcal{W}_d(P(x.\cdot), \nu_\pi)}{N(1-\hat{\lambda})} + \epsilon\right] \leq \\
&\leq \mathbb{P}^N\left[\left|\frac{1}{N}\sum_{i=1}^N \|y_i\| - \mathbb{E}\left(\frac{1}{N}\sum_{i=1}^N \|y_i\|\right)\right| > \epsilon\right] \leq 2\exp\left(-\frac{N\epsilon^2(1-\hat{\lambda})^2}{2}\right).
\end{aligned} \tag{18}$$

3. Concentration for Nonlinear Random Dynamical Systems: The Case of Harris Ergodic Markov Chains

The case of nonlinear random dynamical systems suffers from a lack of uniform transport-entropy constant related to a contractive metric. However, if we can show that the invariant measure satisfies transport-entropy inequality and Markov chain converges exponentially fast to its stationary distribution: it is plausible to run independent simulations of Markov chain and sample the averages after some burn-in period.

Since, the results developed in this paper are aimed at facilitating the RL and controls community. In the absence of exact dynamics, we develop easily verifiable/realistic conditions that ensures exponential integrability of invariant measure. This brings us to the weighted transportaion-inequalities introduced in Bolley and Villani (2005) and plays an integral role when studying concentration phenomenon for nonlinear random dynamical systems. Their work allows for adding different weigths to the underlying distance function, precisely said:

Lemma 7 *Let* ϕ *be a non-negative integrable function, such that* $\int e^{\phi(x)^2}\mu_\pi(dx) < \infty$*- we get an upper bound on weighted total variation distance*

$$\|\phi(\mu_\pi - \nu)\|_{TV} \leq \sqrt{2}\left(1 + \log \int e^{\phi(x)^2} \mu_\pi(dx)\right)^{\frac{1}{2}} \sqrt{Ent(\nu||\mu_\pi)}. \tag{19}$$

A remarkable advantage of this formulation is an Lyapunov condition for underlying Markov chain so as to ensure that at least its' stationary measure satisfies T-E inequality.

Exponential Lyapunov function. Inspired by the assumption made in (particular case 14 of Bolley and Villani (2005)), we proposed an exponential Lyapunov condition:

1. There exists $\hat{\alpha} > 0$, $\beta > 0$ and $C > 0$ such that $\beta < \hat{\alpha}$ and:

$$\int e^{\hat{\alpha}\|y\|^2} P(x, dy) \leq C e^{\beta\|x\|^2}, \quad \text{for all } \ x \in \mathcal{X}.$$

Theorem 8 *If the exponential Lyapunov condition is satisfied, define $W_{\hat{\alpha}}(x) := e^{\hat{\alpha}\|x\|^2}$, then $n-$ th step transition kernel $P^n(x, \cdot)$ satisfy the following transport entropy inequality:*

$$\mathcal{W}_d(P^n(x, \cdot), \nu) \leq \sqrt{2}\left(\frac{1 + \log P^n W_{\hat{\alpha}}(x)}{\hat{\alpha}}\right)^{\frac{1}{2}} \sqrt{Ent(\nu||P^n(x, \cdot))} \tag{20}$$

Moreover, if an ergodic invariant measure μ_π exists: then exists also a finite positive constant $L_{\hat{\alpha},\beta,C}$ such that:

$$\mathcal{W}_d(\mu_\pi, \nu) \leq \sqrt{2L_{\hat{\alpha},\beta,C} Ent(\nu||\mu_\pi)} \tag{21}$$

Proof We will only prove the result for the invariant measure as the result for $n-$ th step will follow the same argument. Since the condition in hypothesis can also be written as:

$$\int e^{\hat{\alpha}\|y\|^2} P(x, dy) \leq C e^{(\beta - \hat{\alpha})\|x\|^2} e^{\hat{\alpha}\|x\|^2} \tag{22}$$

As $(\beta - \hat{\alpha}) < 0$, we can find $\eta_{\hat{\alpha}} \in (0, 1)$ and $\hat{C}_{\hat{\alpha}} < \infty$ such that $W_{\hat{\alpha}}(x)$ satisfy:

$$PW_{\hat{\alpha}}(x) \leq \eta_{\hat{\alpha}} W_{\hat{\alpha}}(x) + \hat{C}_{\hat{\alpha}}, \text{ and consequently via recursion } \int e^{\hat{\alpha}\|x\|^2} \mu_\pi(dx) \leq \frac{\hat{C}_{\hat{\alpha}}}{1 - \eta_{\hat{\alpha}}}. \tag{23}$$

and by defining $\phi(x) = \sqrt{\hat{\alpha}}\|x\|$, upper bound on weighted total variation from Lemma 7 implies that :

$$\|\phi(\mu_\pi - \nu)\|_{TV} \leq \sqrt{2}\left(1 + \log\left[\frac{\hat{C}_{\hat{\alpha}}}{1 - \eta_{\hat{\alpha}}}\right]\right)^{\frac{1}{2}} \sqrt{Ent(\nu||\mu_\pi)}. \tag{24}$$

Since, Wasserstein distance is upper bounded by weighted total-variation with weight $\|x\|$, after scaling we conclude that $\mu_\pi \in \mathcal{T}_1^d\left(\frac{1+\log\left[\frac{\hat{C}_{\hat{\alpha}}}{1-\eta_{\hat{\alpha}}}\right]}{\hat{\alpha}}\right)$ and the result for n-th step transition kernel follows via same argument. ■

Harris chains. As the notion of running multiple independent trajectories is plausible when burn-in period is negligible : $n-$th step transition kernel of Markov chain converges exponentially fast to an invariant measure in Wasserstein metric, this brings us to Harris ergodic Markov chains that by definition satisfy following conditions:

Lyapunov condition with geometric drift: There exists a Lyapunov function $V : \mathcal{X} \to [0, +\infty)$, which satisfies:

$$PV(x) \leq \hat{\gamma} V(x) + K, \quad \text{for some } \hat{\gamma} \in (0,1)\,, K < \infty \ \textit{and} \tag{25}$$

minorization condition: A sufficiently large level set of V (ironically it is called 'small set'), satisfies the minorization condition: i.e., there exists a set $\mathcal{S} := \{x \in \mathcal{X} : V(x) \leq R\}$ for some $R > \frac{2K}{1-\hat{\gamma}}$, $\beta \in (0,1)$ and $\hat{\nu} \in \mathcal{P}(\mathcal{X})$ such that:

$$\mathcal{P}(x, \cdot) \geq \beta \chi_{\mathcal{S}}(x) \hat{\nu}(\cdot). \tag{26}$$

Under these conditions it was shown by Hairer and Mattingly (2011) that for some $\beta^* > 0$HEMC is contractive in Wasserstein metric $\mathcal{W}_d$ with distance function: $d(x,y) := (2+\beta^* V(x)+\beta^* V(y))\chi_{x \neq y}$, . A unique ergodic invariant measure μ_π exists and for some finite C and $\kappa \in (0,1)$

$$\mathcal{W}_d\big(P^n(x,\cdot), \mu_\pi\big) \leq C\kappa^n \mathcal{W}_d(P(x,\cdot), \mu_\pi). \tag{27}$$

3.1. Application to Concentration for SLDSs

Model specifications. We consider a discrete-time SLDS of the form

$$x_{t+1} = \sum\nolimits_{j=1}^{M} (A_j x_t + w_t^j) \chi_{\mathcal{M}_j}(x_t). \tag{28}$$

Here, $x_t \in \mathbb{R}^n$ denote the system's state and $A_j \in \mathbb{R}^{n \times n}$ for $j = 1, ..., M$ capture system dynamics in each of the M Borel measurable regions that decompose the state-space and are pairwise disjoint satisfying $\bigcup_{j=1}^{M} \mathcal{M}_j = \mathbb{R}^n$. In addition, for a fixed region j, noise vectors w_t^j are i.i.d, and satisfy $w_t^j \sim \mathcal{N}(0, I_n)$ and $Cov(w_t^j, w_s^k) = 0$, for all $t, s \geq 0$ and $j \neq k \in \{1, 2, ..., M\}$.

Lemma 9 *Assume that there exists $\varrho < \infty$ such that for all $l \in \mathcal{K}_{bdd} := \{k \mid (1 \leq k \leq M)$ such that $\mathcal{M}_k \subsetneq \mathcal{B}_\varrho^n\}$,it holds that $\|A_l\|_2 \leq L < \infty$ and $\forall j \in (\mathcal{K}_{bdd})^{\complement}$, $\|A_j\|_2 \leq \gamma < 1$. Then, the system (28) mixes geometrically to a unique ergodic invariant distribution μ_π.*

Proof Consider function $V(x) = \|x\|_2$. From (28), we have $P(x, \mathcal{A}) = \sum_{j=1}^{M} P_j(x, \mathcal{A}) \chi_{\mathcal{M}_j}(x)$, where $P_j(x, \cdot) \sim \mathcal{N}(A_j x, I_n)$. Assuming the initial state $x_0 := x \in \mathcal{M}_k$ for some $k \in \mathcal{K}_{bdd}$, then:

$$PV^2(x) = \mathbb{E}_{y \sim \mathcal{N}(A_k x, I_n)} \|y\|_2^2 = \mathbb{E}_{z \sim \mathcal{N}(0, I_n)} \|z\|_2^2 + \|A_k x\|_2^2 \leq (n + L^2 \varrho^2). \tag{29}$$

However, if the initial state is $x_0 := x \in \mathcal{M}_j$ such that $j \in (\mathcal{K}_{bdd})^{\complement}$, then:

$$PV^2(x) = \mathbb{E}_{y \sim \mathcal{N}(A_j x, I_n)} \|y\|_2^2 = \mathbb{E}_{z \sim \mathcal{N}(0, I_n)} \|z\|_2^2 + \|A_j x\|_2^2 \leq n + \gamma^2 \|x\|_2^2 = (n + \gamma^2 V^2(x)). \tag{30}$$

Therefore, starting from any initial condition in $\mathbb{R}^n$, from (29) and (30) it holds that $PV^2(x) \leq \gamma^2 V^2(x) + (n + L^2 \varrho^2)$ and a trivial application of Jensen inequality reveals

$$PV(x) \leq \gamma V(x) + \underbrace{\sqrt{n + L^2 \varrho^2}}_{K}. \tag{31}$$

Minorization condition can be verified from Naeem and Pajic (2020) and the result follows. ■

Theorem 10 *For any $\hat{\alpha} \in (0, \frac{1-\gamma^2}{2})$, we have $\int e^{\hat{\alpha}\|x\|^2}\mu_\pi(dx) < \infty$ and consequently there exists a finite positive constant L_γ such that the invariant measure of SLDS, $\mu_\pi \in \mathcal{T}_1^d(L_\gamma)$.*

Proof An application of Stein's lemma on transition kernel of (28), reveal:

$$\int e^{\alpha\|y\|^2}P(x,dy) = \frac{1}{(1-2\alpha)^{\frac{n}{2}}}e^{\|A_jx\|^2(\alpha+\frac{2\alpha^2}{(1-2\alpha)})} = \frac{1}{(1-2\alpha)^{\frac{n}{2}}}e^{\|A_jx\|^2\frac{\alpha}{(1-2\alpha)}}, \quad \alpha < \frac{1}{2},\, x \in \mathcal{M}_j.$$

A simple linear algebra exercise reveals existence of a $\beta < \hat{\alpha}$ and C when $\hat{\alpha} \in (0, \frac{1-\gamma^2}{2})$ as mentioned in Theorem 8 and conclusion follows. ■

3.2. Gaussian Tail Inequality for Stationary Distribution of SLDS/Harris Chain with Exponential Lyapunov Function and its Consequences for Sampling

Assume that we have access to sampling $(y_i)_{i=1}^N$ i.i.d from μ_π. Since, $r(x) := \|x\|$ is $1-$ Lipschitz w.r.t $d(x,y) := \|x-y\|_2$ and $\mu_\pi \in \mathcal{T}_1^d\left(\frac{1+\log\left[\frac{\hat{C}_{\hat{\alpha}}}{1-\eta_{\hat{\alpha}}}\right]}{\hat{\alpha}}\right)$, we have that

$$\mathbb{P}\left[\left|\frac{1}{N}\sum_{i=1}^N r(y_i) - < r >_{\mu_\pi}\right| > \epsilon\right] \leq 2\exp\left(-\frac{N\epsilon^2\hat{\alpha}}{2+2\log\left[\frac{\hat{C}_{\hat{\alpha}}}{1-\eta_{\hat{\alpha}}}\right]}\right). \tag{32}$$

As any valid $\hat{\alpha}$ can be written down in the form of $\frac{1-\gamma^2}{2n}$ for $n > 1$, comparing with Linear Gaussian case (18) it is reassuring to see how deviations for SLDS and LDS have similar dependence in terms of the norm of stable system matrix.

Remark 11 *Although, we tried our best to show concentration for process level law of HEMCs under exponential-type Lyapunov condition, via weigthed T-E inequality but it got intractable due to non-uniform transport constants.*

4. Conclusion and discussion

We have provided with a general framework for getting concentration inequalities for random dynamaical systems, with respect to empirical averages of unbounded test functions; validated our analysis on the example of LDS and SLDS. Summarizing few key observations and open problems: *Exponential-type Lyapunov functions are sufficient for having exponential concentration inequalities* (although, we might have to run multiple independent trajectories). *Stability is necessary not sufficient:* As it should be evident from the example of HEMCs, under existence of exponential-type Lyapunov functions we can ensure concentration only by running multiple independent trajectories. Concentration via single trajectory requires some notion of regularity for transition kernels as well: one of them is uniform transportation constant and we intend on exploring more such regularity conditions in our future work. An interesting open problem, *can we relax uniform T-E condition and still get exponential inequalities from a single trajectory?* We suspect it might be the case when samples of Markov chain interact symmetrically; convincing examples include, Kac's interacting particles model exhibiting propagation of chaos and ubiquity of assumption on reversibility of Markov chain w.r.t its' stationary maeasure in discrete and continuous time Large Deviation Principle, see e.g, Wang and Wu (2020) and Guillin et al. (2009).

Acknowledgments

This work is sponsored in part by the AFOSR under award number FA9550-19-1-0169, and by the NSF under CNS-1652544 award as well as the National AI Institute for Edge Computing Leveraging Next Generation Wireless Networks, Grant CNS-2112562.

References

Gordon Blower and François Bolley. Concentration inequalities on product spaces with applications to markov processes. *arXiv preprint math/0505536*, 2005.

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.

François Bolley and Cédric Villani. Weighted csiszár-kullback-pinsker inequalities and applications to transportation inequalities. In *Annales de la Faculté des sciences de Toulouse: Mathématiques*, volume 14, pages 331–352, 2005.

Sourav Chatterjee. Stein's method for concentration inequalities. *Probability theory and related fields*, 138(1-2):305–321, 2007.

Sourav Chatterjee and Partha S Dey. Applications of stein's method for concentration inequalities. *The Annals of Probability*, 38(6):2443–2485, 2010.

Hacene Djellout, Arnaud Guillin, Liming Wu, et al. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Annals of Probability*, 32(3B): 2702–2732, 2004.

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.

Clark R Givens, Rae Michael Shortt, et al. A class of wasserstein metrics for probability distributions. *The Michigan Mathematical Journal*, 31(2):231–240, 1984.

Arnaud Guillin, Christian Léonard, Liming Wu, and Nian Yao. Transportation-information inequalities for markov processes. *Probability theory and related fields*, 144(3-4):669–695, 2009.

Martin Hairer and Jonathan C Mattingly. Yet another look at harris' ergodic theorem for markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, pages 109–117. Springer, 2011.

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

Krzysztof Łatuszyński, Błażej Miasojedow, Wojciech Niemiro, et al. Nonasymptotic bounds on the estimation error of mcmc algorithms. *Bernoulli*, 19(5A):2033–2066, 2013.

Michel Ledoux. *The concentration of measure phenomenon*. Number 89. American Mathematical Soc., 2001.

Florient Malrieu. Logarithmic sobolev inequalities for some nonlinear pde's. *Stochastic processes and their applications*, 95(1):109–132, 2001.

Katalin Marton et al. Measure concentration for euclidean distance in the case of dependent random variables. *Annals of probability*, 32(3B):2526–2544, 2004.

Muhammad Abdullah Naeem and Miroslav Pajic. Learning expected reward for switched linear control systems: A non-asymptotic view. *arXiv preprint arXiv:2006.08105*, 2020.

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

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.

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.

Michel Talagrand. Transportation cost for gaussian and other product measures. *Geometric & Functional Analysis GAFA*, 6(3):587–600, 1996.

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.

Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.

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

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

Yiming Zhang and Keith W Ross. On-policy deep reinforcement learning for the average-reward criterion. In *International Conference on Machine Learning*, pages 12535–12545. PMLR, 2021.