TIME-AVERAGED FUNCTIONAL CONVERGENCE FOR UNIFORMLY ERGODIC MARKOV CHAINS ON BOREL SPACES

Ungureanu Viorica Mariela, Assoc. Prof. PhD

"Constantin Brâncuși" University of Târgu Jiu, ROMANIA

ABSTRACT: This paper studies the long-term behavior of time-averaged functionals of Markov chains defined on standard Borel state spaces, focusing on chains that admit transition densities with respect to a σ-finite reference measure and satisfy uniform ergodicity. We consider sequences of measurable functions that are uniformly bounded across all states and time steps, and analyze the convergence of their time-averaged conditional expectations given past states. Using the uniform ergodicity property, we rigorously show that these time-averaged expectations converge to corresponding expressions computed with respect to the chain's invariant density. This provides a general framework for evaluating steady-state behavior of functionals in stochastic systems, including those arising in H_2 norm computations for systems with Markovian jumps. The use of standard Borel spaces as the state space for the Markov chain ensures well-defined measurability and integrability, supporting the applicability of these results to a wide class of Markov processes and stochastic control problems.

KEY WORDS: Uniform ergodicity, Markov chains, Time-averaged functionals, Standard Borel spaces, Invariant measure, Transition density, Stochastic control

1. INTRODUCTION

The analysis of the H_2 norm in Markov processes defined over general Borel state spaces plays a fundamental role in understanding and designing stochastic dynamical systems. The H_2 norm quantifies a system's performance in response to white-noise disturbances and serves as a key tool in optimal control and stability analysis. In systems with Markovian jumps - where transitions between states are governed by a Markov chain taking values in a general Borel space - a rigorous theoretical framework is essential for evaluating and optimizing system behavior.

Recent studies have emphasized the importance of the asymptotic behavior of the H_2 norm in such systems. For instance, Costa and Figueiredo [2] investigated quadratic control with partial information for discrete-time jump systems with the Markov chain in a general Borel space, while Hou and Zhang [3] studied $H_2/H\infty$ control design for detectable periodic Markov jump systems. Further contributions include full-information H_2 control for Borel-measurable Markov jump systems with multiplicative noises [4], H_2 control for Markov jump linear systems (MJ-LSs) with detector-based partial information [5], and extensions to infinite-dimensional MJLs [8, 9]. Recent works also address robustness and stability analysis for discrete-time MJLS on Borel spaces [10, 11].

Despite these advances, a comprehensive understanding of the limits and asymptotic properties of the H_2 norm for general Borel-space Markov processes remains incomplete. This paper aims to contribute to this area by establishing rigorous convergence results for time-averaged functionals of Markov chains with general Borel state spaces, leveraging uniform ergodicity conditions [6]. These results provide a solid foundation for the computation and optimization of the H_2 norm in complex stochastic systems under uncertainty, bridging the gap between theoretical development and practical control applications.

2. PRELIMINARIES

Let (X, d) be a metric space with distance d, and let this space be endowed with the topology $T(d)$ generated by open balls

$$B_d(x, \varepsilon) = \{y \in X \ : \ d(x, y) < \varepsilon\}.$$

We recall that the *Borel σ -algebra* $B(X)$ is the smallest σ -algebra containing all open sets: $B(X) = \sigma(T(d))$.

A *Polish space* is a complete and separable metric space (X, d) and its Borel σ -algebra is countably generated.

A *standard Borel space* (S, B) is a measurable space, measurably isomorphic to a Borel subset of a Polish space. That is, there is a Polish space (X, d), a Borel subset $X_0 \in B(X)$ and a bijection $\phi \ : \ S \to X_0$ such that both ϕ and ϕ^{-1} are measurable and $B = \phi^{-1}(B(X_0))$.

The following result is known (see [7] and the references therein).

Proposition 1. *Every standard Borel space (S, B) has a countably generated σ -algebra.*

Proof. We know that the associated Polish space (X, d) is separable, which implies that its topology has a countable base $\{U_n\}_{n \in N}$. Hence the Borel σ algebra on $X_0 \subseteq X$, $B(X_0)$ is countably generated (as a sub – σ -algebra of $B(X)$) and we can write

$$B(X_0) = \sigma(\{U_n \cap X_0 \ : \ n \in N\}).$$

Taking preimages under ϕ, we have

$$B = \sigma(\{\phi^{-1}(U_n \cap X_0): \ n \in N\}),$$

which shows that B is also countably generated.

Let (Ω, F, P) be a probability space, $(S, B(S))$ be a *standard Borel space* and μ be a σ -finite measure defined on S. Let $\{\eta_n\}_{n \in N}$ be a time-homogeneous Markov chain with the state space S. We recall that a minimal requirement for defining a Markov chain on a general measurable state space (S, B) is that B is countably generated (see [6]) and this condition is satisfied when $(S, B(S))$ is *standard Borel*

We also assume that the following hypothesis holds

H1: $\{\eta_n\}_{n \in N}$ is a time-homogeneous Markov chain such that:

1. its initial probability distribution has a density $f_0 > 0$ that is absolutely integrable with respect to μ and $P[\eta_0 \in B] = \int_B f_0(s)\mu(ds)$;
2. it has a transition kernel

$$p(l, B) = P[\eta_{n+1} \in B \mid \eta_n = l], \quad l \in S, \ B \in B(S).$$

which admits a nonnegative density $g(s \mid l)$ with respect to μ, which is measurable on $S \times S$, i.e.

$$p(l, B) = \int_B g(s \mid l)\mu(ds), B \in B(S), l \in S,$$

and $0 \leq g(s \mid l) < A$ for some constant $A > 0$.

We note that $\int_S g(s \mid l)\mu(ds) = 1$ and we recall that a transition kernel $p \ : \ S \times B(S) \to [0,1]$ has the following properties: $p(s, \cdot)$ is a probability measure on $(S, B(S))$ for every $s \in S$, and $p(\cdot, B)$ is $B(S)$ -measurable for every $B \in B(S)$.

The n -step transition kernel satisfies the Chapman--Kolmogorov equation

$$p^{n+m}(l, B) = \int_S p^n(l, ds)p^m(s, B), \quad n, m \in N.$$

Let $L_1(\mu)$ be the linear space of all real valued, μ measurable functions f defined on S that satisfy the condition $\int_S |f(s)|\mu(ds) < \infty$. We define the linear operator

$$(Tf)(l) := \int_S g(l \mid s)f(s)\mu(ds),$$

for any $f \in L_1(\mu)$. According to [1], the operator $T \ : \ L_1(\mu) \to L_1(\mu)$ is linear and bounded and its adjoint P is defined on $L_\infty(\mu)$ as

$$(Pf)(s) := \int_S g(l \mid s)f(l)\mu(dl).$$

If we denote $\pi^{k+1}(l) = E[g(l \mid \eta_k)]$ (see [1-2]), we get

$$\begin{aligned}\pi^{k+1}(l) &= E\big[E[g(l \mid \eta_k)]|_{\eta_{k-1}}\big] \\ &= E\left[\int_S g(l \mid i)g(i \mid \eta_{k-1})\mu(di)\right] \\ &= \int_S g(l \mid i)E[g(i \mid \eta_{k-1})]\mu(di) \\ &= T(E[g(i \mid \eta_{k-1})])(l) = T(\pi^k)(l).\end{aligned}$$

In the above computation, the expectation and the integral can be interchanged by Tonelli's theorem, since the functions involved are nonnegative and measurable. It follows that

$$\pi^k(l) = T(\pi^{k-1})(l), \text{ for all } k \geq 1$$

We also note that $\pi^1(l) = E[g(l \mid \eta_0)] = \int_S g(l \mid i)\nu(i)\mu(di) = T(\nu)(l)$ which implies that $\pi^k(l) = T^k(\nu)(l) \geq 0$.

It can be proved inductively (see [2]) that $\int_S [\pi^k(l)]\mu(dl) = \int_S T^k(\nu)(l)\mu(dl) = \int_S \nu(l)\mu(dl) = 1$, which implies that $\pi^k(l)$ is a density. Since $P[\eta_{n+1} \in B] = E[P[\eta_{n+1} \in B|_{\eta_n}]] = \int_B E[g(i \mid \eta_n)]\mu(di) = \int_B \pi^{n+1}(i)\mu(di)$ and it follows that $\pi^{n+1}(l)$ is the density of η_{n+1} with respect to μ.

3. UNIFORM ERGODICITY

In this section, we recall some basic concepts related to ergodicity and uniform ergodicity of Markov chains on general state spaces. These notions provide the theoretical framework for studying the long-term behaviour and convergence properties of the chain.

Definition 1. [6] *For two probability measures ν_1 and ν on a measurable space $(S, B(S))$, the total variation distance between them is defined as*

$$\|\nu_1 - \nu\|_{TV} = \sup_{A \in B(S)} |\nu_1(A) - \nu(A)|. \quad (1)$$

If ν_1 and ν admit densities g_1 and g with respect to a common reference measure λ, then two equivalent expressions of (1) *are given by*

$$\|\nu_1 - \nu\|_{TV} = \int_S |g_1(x) - g(x)|\lambda(dx) = \sup_{f,|f|\leq 1} \left| \int_S f(x)\big(g_1(x) - g(x)\big)\lambda(dx) \right|.$$

Definition 2. *A probability measure Π on $(S, B(S))$ is said to be invariant for the transition kernel p if*

$$\Pi(B) = \int_S p(s, B)\Pi(ds), \quad \forall B \in B(S).$$

Definition 3. *The chain $\{\eta_n\}_{n\in N}$ is said to be ergodic if there is an invariant probability measure Π for the kernel p such that $p^n(l,\cdot) \to \Pi(\cdot)$ as $n \to \infty$ in total variation for all $l \in S$.*

Definition 4. *[6] The chain $\{\eta_n\}_{n\in N}$ is said to be uniformly ergodic if there is an invariant measure Π for the kernel p and the constants $C < \infty$ and $\rho \in (0,1)$ such that*

$$\|p^n(s,\cdot) - \Pi(\cdot)\|_{TV} \leq C\rho^n, \forall s \in S,\ n \geq 0.$$

For some equivalent formulations of uniform ergodicity, we refer to Theorem 16.02 in [6].

In the rest of this paper, we will assume that **H1** and the following hypothesis hold:

H2: The Markov chain $\{\eta_n\}_{n\in N}$ is *uniformly ergodic.*

The following result is known (see Proposition 4.1 from [1] and Theorem 16.2.1 from [6])

Lemma 1. There is an invariant probability measure Π for the kernel p which has the following properties:

1) Π is absolutely continuous with respect to μ, that is, there is a nonnegative function $\pi \in L_1(\mu)$ such that

$$\Pi(ds) = \pi(s)\mu(ds),$$
$$\int_S \pi(s)\mu(ds) = 1;$$

2) the density π is a fixed point of the operator T, i.e.

$$\pi(l) = T(\pi)(l), l \in S$$

3) there are positive constants $C_1 < \infty$ and $\rho_1 \in (0,1)$ such that

$$\int_S |T^n(\nu)(s) - \pi(s)|\mu(ds) \leq C_1\rho_1^n.$$

4. CONVERGENCE OF TIME-AVERAGED FUNCTIONALS UNDER UNIFORM ERGODICITY CONDITIONS

Time-averaged functionals of a Markov chain provide a natural way to quantify long-term system behaviour and steady-state performance. For chains on standard Borel spaces, uniform ergodicity ensures convergence that is independent of the initial state. This section establishes that, under uniform ergodicity, the time-averaged expectations of bounded measurable functionals converge to the corresponding averages with respect to the chain's invariant measure, providing a rigorous foundation for performance analysis in stochastic systems.

Lemma 2. *Assume that hypotheses H1 and H2 are satisfied. Then for any integers $n > k \geq 0$*

and $i \in S, T^{n-k}(g(\cdot \mid l))(i)$ is the density of the $(n-k+1)$ -step transition kernel $p^{n-k+1}(l,\cdot)$ evaluated at i. Moreover,

$$E[g(i \mid \eta_n) \mid \eta_k = l] = T^{n-k}(g(\cdot \mid l))(i).$$

Proof. We proceed by induction on $m = n - k$.

Base step: $m = 1$. By Chapman--Kolmogorov equation, the 2-step kernel is

$$p^2(l,B) = \int_B p(l,ds_1)p(s_1,B) = \int_S g(s_1 \mid l)\mu(ds_1)\int_B g(i \mid s_1)\mu(di) = \int_B\int_S g(i \mid s_1)g(s_1 \mid l)\mu(ds_1)\mu(di)$$

By definition of T, the inner integral is exactly $T(g(\cdot \mid l))(i)$. Hence,

$$p^2(l,B) = \int_B T(g(\cdot \mid l))(i)\mu(di),$$

showing that $T(g(\cdot \mid l))(i)$ is the density of $p^2(l,\cdot)$.

Inductive step: Assume that for some $m \geq 1$,

$$p^{m+1}(l,B) = \int_B T^m(g(\cdot \mid l))(s_1)\mu(ds_1).$$

Then, the $(m+2)$ -step transition kernel satisfies

$$p^{m+2}(l,B) = \int_S p(s,B)p^{m+1}(l,ds)$$
$$= \int_S \left(\int_B T^m(g(\cdot \mid l))(s_1)\mu(ds_1)\right) \cdot g(i \mid s_1)\mu(di)$$
$$= \int_B (\int_S T^m(g(\cdot \mid l))(s_1) \cdot g(i \mid s_1)\mu(ds_1))\mu(di).$$

By the definition of T iterated, the inner integral is $T^{m+1}(g(\cdot \mid l))(i)$, so that

$$p^{m+2}(l,B) = \int_B T^{m+1}(g(\cdot \mid l))(i)\mu(di).$$

Conclusion: By induction, for all $m \geq 1$,

$$p^{m+1}(l,B) = \int_B T^m(g(\cdot \mid l))(i)\mu(di),$$

i.e., $T^m(g(\cdot \mid l))(i)$ is the density of $p^{m+1}(l,\cdot)$ evaluated at i.

Finally, using the definition of conditional expectation,

$$E[g(i \mid \eta_n) \mid \eta_k = l] = \int_S g(i \mid s)p^{n-k}(l,ds) = T^{n-k}(g(\cdot \mid l))(i).$$

Usually, when computing time-averaged integrals involved in H_2 norms, we need to show the convergence of sequences such as

$$\lim_{\tau\to\propto}\frac{1}{\tau}\sum_{n=k}^{\tau}\int_S \rho(n,l,i)\int_S g(i|s)\cdot\left(T^{n-k}\big(g(.\mid l)\big)\right)(s)\mu(ds)\mu(di)$$

Our goal is to express the above limits in terms of the invariant probability measure of the kernel p.

Theorem 1 *Let* $\{\eta_n\}_{n\geq 0}$ *be a uniformly ergodic Markov chain on a standard Borel space* $(S, B(S))$ *with transition density* $g(. \mid .)$ *with respect to a* σ *-finite measure* μ *, and let* π *be its invariant density. Suppose that* $\rho(n,l,j)$ *is a nonnegative real sequence which is* $S \times S$ *measurable and uniformly bounded: i.e. there is* $M > 0$ *such that* $|\rho(n,l,j)| \leq M$ *for all* n,l,j. *Then*

$$\lim_{\tau\to\propto}\frac{1}{\tau}\sum_{n=k}^{\tau}\int_S \rho(n,l,i)\int_S g(i|s)\cdot\left(T^{n-k}\big(g(.\mid l)\big)\right)(s)\mu(ds)\mu(di) = \lim_{\tau\to\propto}\frac{1}{\tau}\sum_{n=k}^{\tau}\int_S \rho(n,l,i)\int_S g(i|s)(\pi)(s)\mu(ds)\mu(di),$$

uniformly with respect to k *and* l.

Proof. Define the difference

$$\Delta_{n,k,l} = \left|\int_S \rho(n,l,i)\int_S g(i \mid s)[T^{n-k}\big(g(.\mid l)\big)(s) - \pi(s)]\mu(ds)\mu(di)\right|.$$

We recall that $0 \leq g(i \mid s) < A$, $T^{n-k}\big(g(.\mid l)\big) > 0$ and $|\rho(n,l,j)| \leq M$ for all n,l,j. Then

$$\Delta_{n,k,l} \leq \int_S |\rho(n,l,i)|\left|\int_S g(i \mid s)[T^{n-k}\big(g(.\mid l)\big)(s) - \pi(s)]\mu(ds)\right|\mu(di) \leq$$

$$M\int_S \left|\int_S g(i \mid s)\left(T^{n-k}(g(.\mid l))(s) - \pi(s)\right)\mu(ds)\right|\mu(di) =$$

$$M\int_S \left|T^{n-k+1}(g(.\mid l))(i) - T(\pi)(i)\right|\mu(di)$$

We conclude that

$$|\Delta_{n,k,l}| \le M\int_S |T^{n-k+1}(g(.\mid l))(i) - \pi(i)|\mu(di) \qquad \textbf{(2)}$$

On the other hand

$$\|p^n(l,\cdot) - \Pi(\cdot)\|_{TV} = \int_S |p^n(l,di) - \pi(i)|\mu(di) = \int_S |T^{n-1}(g(.\mid l))(i) - \pi(i)|\mu(di),$$

where the last equality follows from Lemma 2.

By the uniform ergodicity of the Markov chain, it follows that there is $C > 0$ and $\rho \in (0,1)$ such that for every l

$$\int_S |T^n(g(.\mid l))(s) - \pi(s)|\mu(ds) < C\rho^n.$$

We deduce from (2) that

$$|\Delta_{n,k,l}| \le MC\rho^{n-k+1}$$

for every $l \in S$. Finally, we see that the time-averaged sum satisfies the following inequality:

$$\left|\frac{1}{\tau}\sum_{n=k}^{\tau} \Delta_{n,k,l}\right| \le \frac{1}{\tau}\sum_{n=k}^{\tau} MC\rho^{n-k} \le \frac{MC}{\tau}\frac{1}{1-\rho} \to 0 \quad \text{as } \tau \to \infty.$$

Therefore, the time-averaged sum converges to zero as $\tau \to \infty$, which completes the proof.

Remark 1. When the state space S is countably infinite (for instance $S = \{1,2,..\}$), we take as reference measure the *counting measure*. In this case, the transition densities with respect to the reference measure reduce to the standard transition probabilities of the Markov chain:

$$T^{n-k}(g(.\mid l))(j) = P(\eta_{n+1} = j \mid \eta_k = l) = p^{n-k+1}(l,j),$$

and the invariant measure π is a discrete probability distribution on S, represented by a vector $\pi(i)$. Accordingly, the following integrals over $S \times S$ that appear in the continuous formulation

$$\int_S \int_S \rho(n,l,i)\, g(i \mid s) \cdot \left(T^{n-k}(g(.\mid l))\right)(s)\mu(ds)\mu(di)$$

are replaced by double sums over S:

$$\sum_{i\in S}\sum_{s\in S} \rho(n,l,i)p^1(s,i)p^{n-k+1}(l,s).$$

Consequently, the main convergence result can be expressed in the discrete form

$$\lim_{\tau\to\infty}\frac{1}{\tau}\sum_{n=k}^{\tau}\sum_{i\in S}\sum_{s\in S} \rho(n,l,i)p^1(s,i)p^{n-k+1}(l,s)$$

$$= \lim_{\tau\to\infty}\frac{1}{\tau}\sum_{n=k}^{\tau}\sum_{i\in S}\sum_{s\in S} \rho(n,l,i)p^1(s,i)\pi(s).$$

We note that the above identity was established in [8] for the case of countably infinite S, under the weaker hypothesis that the Markov chain $\{\eta_n\}_{n\in N}$ is irreducible, aperiodic, and positive recurrent, and possesses an ergodic Markov subchain. In discrete spaces, these conditions suffice because positive recurrence guarantees the existence of a unique invariant probability measure, while irreducibility and aperiodicity ensure convergence to this stationary distribution from any initial state. In contrast, for Markov processes defined on general (possibly uncountable) state spaces, a stronger assumption such as uniform ergodicity is required. This condition ensures the existence of a global minorisation (Doeblin-type)[6] property providing uniform control over convergence rates across all initial states. Without such a condition, convergence may hold only pointwise in total variation but fail to be uniform, preventing the interchange of limits and integrals in expressions involving transition densities.

5. CONCLUSIONS

In this paper, we analysed the convergence of time-averaged functionals for uniformly ergodic Markov chains defined on standard Borel spaces. By leveraging the properties of transition densities and uniform ergodicity, we demonstrated that the time-averaged expectations of bounded measurable functionals converge to the corresponding expectations under the invariant measure. This result provides a rigorous foundation for evaluating long-term performance in stochastic systems, including applications to H_2 norm computation in systems with Markovian jumps. The framework established here highlights the importance of uniform ergodicity for ensuring uniform convergence across all initial states, and it offers a basis for further research on more general classes of Markov processes and control systems.

6. Further Research

An interesting direction for future research is to investigate sufficient conditions under which uniform ergodicity of a Markov chain can be equivalently expressed in terms of convergence of transition densities, and to explore extensions of the current results to broader classes of functionals or to continuous-time Markov processes on general state spaces.

REFERENCES

[1] Costa, O. L. V, & Figueiredo, D. Z. *Filtering S-coupled algebraic Riccati equations for discrete-time Markov jump systems*. Automatica, *83*, 47-57, 2017.

[2] Costa, O. L. V., & Figueiredo, D. Z., *Quadratic control with partial information for discrete-time jump systems with the Markov chain in a general Borel space*, Automatica, 66, 73–84, 2016.

[3] Hou, T., & Zhang, W., $H_2/H\infty$ *Control Design of Detectable Periodic Markov Jump Systems*, 2015.

[4] Ma, H., Cui, Y., & Wang, Y., *Full Information* H_2 *Control of Borel-Measurable Markov Jump Systems with Multiplicative Noises*, Mathematics, 10(1), 37, 2021.

[5] Rodrigues, C. C. G., & Silva, J. A., *Fast Switching Detector-Based* H_2 *Control of Markov Jump Systems*, 2021.

[6] Meyn, S. P., & Tweedie, R. L., *Markov Chains and Stochastic Stability*, Springer Science & Business Media, 2012.

[7] Preston, Chris. *Some notes on standard Borel and related spaces.* arXiv preprint arXiv:0809.3066 (2008).

[8] Ungureanu, V. M., H_2*-optimal control for periodic, discrete-time Markov-jump systems with multiplicative noise in infinite dimensions*, IMA Journal of Mathematical Control and Information, 33(3), 813–830, 2016.

[9] Ungureanu, V. M., & Ungureanu, I. I., *Optimal control for infinite-dimensional linear systems with Markovian jumps in Borel spaces*, Surveys in Mathematics and its Applications, 20, 173–192, 2025.

[10] Xiao, C., et al., *Small Gain Theorem-Based Robustness Analysis of Discrete-Time MJLSs with the Markov Chain on a Borel Space and Its Application to NCSs*, arXiv preprint arXiv:2502.14188, 2025.

[11] Xiao, C., Hou, T., & Zhang, W., *Stability and bounded real lemmas of discrete-time MJLSs with the Markov chain on a Borel space*, Automatica, 169, 1827, 2024.