AN EXCURSION-THEORETIC APPROACH TO STABILITY OF DISCRETE-TIME STOCHASTIC HYBRID SYSTEMS

DEBASISH CHATTERJEE AND SOUMIK PAL

Abstract. We address stability of a class of Markovian discrete-time stochastic hybrid systems. This class of systems is characterized by the state-space of the system being partitioned into a safe or target set and its exterior, and the dynamics of the system being different in each domain. We give conditions for $\boldsymbol{L}_1$-boundedness of Lyapunov functions based on certain negative drift conditions outside the target set, together with some more minor assumptions. We then apply our results to a wide class of randomly switched systems (or iterated function systems), for which we give conditions for global asymptotic stability almost surely and in $\boldsymbol{L}_1$. The systems need not be time-homogeneous, and our results apply to certain systems for which functional-analytic or martingale-based estimates are difficult or impossible to get.

1. Introduction

Increasing complexity of engineering systems in the modern world has led to the hybrid system paradigm in systems and control theory [vS00, Lib03]. A hybrid system consists of a number of domains in the state-space and a dynamical law corresponding to each domain; thus, at any instant of time the dynamics of the system depends on the domain that its state is in. This facilitates introduction of modular analysis and synthesis tools in the sense that one restricts attention to behavior of the system in individual domains, which is typically a simpler problem, and is an invaluable aid in analysis or synthesis of complex systems featuring hundreds of states. However, understanding how the dynamics in the individual domains interact among each other is necessary in order to ensure smooth operation of the overall system. This article is a step towards understanding the behavior of Markovian stochastic hybrid systems which undergo excursions into different domains infinitely often. Here we consider the simplest and perhaps the most important hybrid system, consisting of a compact target or safe set and its exterior, with different dynamics inside and outside the safe set. Our objective is to introduce a new method of analysis of such systems when the states excurse outside the safe set infinitely often in course of its evolution. The analysis carried out here provides a basis for controller synthesis of systems with control inputs—it gives clear indications about the type of controllers to be designed in order to ensure certain natural and basic stability properties in closed loop.

Let us look at two interesting and practically important examples of hybrid systems with two domains—a compact safe set and its exterior, with different dynamics in each. The first concerns optimal control of a Markov process with state constraints. Markov control processes have been extensively studied; we refer the reader to the excellent monographs and surveys [BS78, Bor91, ABFG+93, HLL96, HLL99, Piu97] for further information, applications and references. For our purposes here, consider the canonical example of a linear controlled system perturbed

Date: January 13, 2009.
2000 *Mathematics Subject Classification.* Primary: 93E15; Secondary: 60J05.
Key words and phrases. stochastic system, stability, excursion theory, Markov process.

by additive Gaussian noise and having probabilistic constraints on the states. A hybrid structure of the controlled system naturally presents itself in the following fashion. Except in the most trivial of cases, constrained optimal control over an infinite horizon is impossible, and one resorts to a a rolling-horizon controller. (Rolling-horizon controllers are considerably popular, for basic definitions, comparisons and references see e.g., [Mac01] in the deterministic context, and [AS92, HLL90, HLL93] in the stochastic context). Computational overheads restrict the size of the window in the rolling-horizon controller, and determine the maximal (typically bounded) region—called the safe set—in which this controller can be active. No matter how good the resulting controller is, the additive nature of the Gaussian noise ensures that the states are subjected to excursions away from the safe set infinitely often almost surely. Once outside the safe set, the rolling-horizon controller is switched off and a recovery strategy is activated, whose task is to bring the states back to the safe set quickly and efficiently. This problem is of great practical interest and a subject of current research, see e.g., [ACCL08, CCK07] and the references in them for possible strategies inside the safe set, and [CCCL08] for one possible recovery strategy. Evidently, stability of this hybrid system depends largely on the recovery strategy, since as long as the states stay inside the safe set, they are bounded. However, traditional methods of stability analysis do not work well precisely because of the unlimited number of excursions. Theorem 2 of this article addresses this issue, and provides a method of ensuring strong boundedness and stability properties of the hybrid system. Intuitively it says that under the recovery strategy there exists a well-behaved supermartingale until the states hit the safe set, then the system state is bounded in expectation uniformly over time. A complete picture of stability and ergodic properties of a general controlled hybrid system is beyond the scope of the present article, and will be reported elsewhere. We refer the reader to [CFM05, Chapter 3] for earlier work pertaining to stability of a class of hybrid systems, and to [Kus71, MT93] for stability of general discrete-time Markov processes.

The second example is one that we shall pursue further in this article, namely, a class of discrete-time Markov processes called iterated function systems [BDEG88, LM94] (IFS). They are widely applied, for instance, in the construction of fractals [LM94], in studies on the process of generation of red blood corpuscles [LM02, LS04], in statistical physics [Kif86], and simulation of important stochastic processes [Wer05]. Of late they are being employed in key problems of physical chemistry and computational biology, namely, the behavior of the chemical master equation [Wil06, Chapter 6] (CME), which governs the continuous-time stochastic (Markovian) reaction-kinetics at very low concentrations (of the order of tens of molecules). Invariant distributions, certain finite-time properties, and robustness properties with respect to disturbances of the underlying Markov process are of interest in modeling and analysis of unicellular organisms. It is well-known that the CME is analytically intractable (see [JH07, ACK08] for special cases), but the invariant distribution of the Markov process can be recovered from simulation of the embedded Markov chain in a computationally efficient way [MA08]. This embedded chain is an IFS taking values in a nonnegative integer lattice. From a biological perspective, good health of a cell corresponds to the IFS evolving in a safe region on an average, despite moderate disturbances to the numbers of molecules involved in the key reactions. However, in most cases compact invariant sets do not exist. It is therefore of interest to find conditions under which, even though there are excursions of the states away from a safe set infinitely often, the IFS is stochastically bounded, or some strong stability properties hold. Theorem 2 of this article leads to results (in §3) which address this issue.

This article unfolds as follows. §2 contains our main results—Theorem 2 and 4, which provide conditions under which a Lyapunov function of the states is $\boldsymbol{L}_1$-bounded. We establish this $\boldsymbol{L}_1$-boundedness under a certain negative drift condition outside a compact set, and some more minor conditions. The negative drift alone is not enough, as pointed out in [PR99], where the authors establish variants of our results for scalar (not necessarily Markovian) processes having increments with bounded p-th moments for $p > 2$. For our results to hold, the underlying Markov process need not be time-homogeneous. Conditions for existence of invariant measures for Feller processes follow from our main results as immediate corollaries. §3 contains some applications of the results in §2 to stability and robustness of IFS. The classical weak stability questions concerning the existence and uniqueness of invariant measures of IFS, addressed in e.g., [DF99, JT01, Sza06], revolve around average contractivity hypotheses of the constituent maps and continuity of the probabilities. In §3.1 we look at stronger stability properties of the IFS, namely, global asymptotic stability almost surely and in expectation, for which we give sufficient conditions. There are no assumptions of global contractivity or memoryless choice of the maps at each iterate; we just require a condition resembling average contractivity in terms of Lyapunov functions with a suitable coupling condition with the Markovian transition probabilities. We mention that although some of the assumptions in [JT01] resemble ours, the conditions needed to establish existence of invariant measures in [JT01] are stronger than what we employ; see §3.1 for a detailed comparison. We also demonstrate in §3.2 that under mild assumptions, iterated function systems possess strong stability and robustness properties with respect to bounded disturbances. In this subsection the exogenous bounded disturbance is not modelled as a random process. In addition to the examples considered here, the results in §2 will be of interest in queueing theory, along the lines of the works [HLR96, BKR+01].

Notations. Let $\mathbb{N} := \{1, 2, \ldots\}$, $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$, and $\mathbb{R}_{\geqslant 0} := [0, \infty[$. We let $\|\cdot\|$ denote the standard Euclidean norm on $\mathbb{R}^d$. Let $\mathcal{K}$ denote the collection of strictly increasing continuous functions $\alpha : \mathbb{R}_{\geqslant 0} \longrightarrow \mathbb{R}_{\geqslant 0}$ such that $\alpha(0) = 0$; we say that a function α belongs to class-$\mathcal{K}_\infty$ if $\alpha \in \mathcal{K}$ and $\lim_{r\to\infty} \alpha(r) = \infty$. A function $\beta : \mathbb{R}_{\geqslant 0} \times \mathbb{N}_0 \longrightarrow \mathbb{R}_{\geqslant 0}$ belongs to class-$\mathcal{KL}$ if $\beta(\cdot, n) \in \mathcal{K}$ for a fixed $n \in \mathbb{N}_0$, and if $\beta(r, n) \to 0$ as $n \to \infty$ for fixed $r \in \mathbb{R}_{\geqslant 0}$. Recall that a function $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is locally Lipschitz continuous if for every $x_0 \in \mathbb{R}^d$ and open set O containing x_0, there exists a constant $L > 0$ such that $\|f(x) - f(x_0)\| \leqslant L \|x - x_0\|$ whenever $x \in O$. We let $\bar{B}_r$ denote the closed Euclidean ball around 0, i.e., $\bar{B}_r := \{y \in \mathbb{R}^d \,|\, \|y\| \leqslant r\}$. For a vector $v \in \mathbb{R}^d$ let v^{T} denote its transpose, and for a $d \times d$ matrix P let $\mathrm{trace}(P)$ denote the trace of P, and $\|v\|_P$ denote $\sqrt{v^{\mathrm{T}} P v}$. The maximum and minimum of two real numbers a and b is denoted by $a \vee b$ and $a \wedge b$, respectively.

2. Generalities

Before we get into hybrid systems, it will be simpler to follow the arguments if we start by considering a discrete-time Markov chain $(X_t)_{t\in\mathbb{N}_0}$ with transition kernel Q, taking values in the state space $\mathcal{S}$. Suppose that there exists a compact subset $K \subseteq \mathcal{S}$ and a nonnegative measurable real-valued function $V : \mathcal{S} \longrightarrow \mathbb{R}$ such that the following condition holds.

Assumption 1. There exists a nonnegative function $\varphi : \mathbb{N}_0 \times \mathbb{R}_{\geqslant 0} \longrightarrow \mathbb{R}$ such that for every fixed t, the function $\varphi(t, \xi)$ is increasing and convex in ξ, and that if we start the process from a point in the complement of K then, the derived process $(Y_t)_{t\in\mathbb{N}_0}$ defined by $Y_t = \varphi(t, V(X_t))$ is a supermartingale until the first time it hits

K. That is to say, if $X_0 = x_0 \in \mathcal{S} \smallsetminus K$, and if we define

$$\tau_K = \inf\Bigl\{t > 0 \Bigm| X_t \in K\Bigr\},$$

then the process $\bigl(Y_{t\wedge\tau_K}\bigr)_{t\in\mathbb{N}_0}$ is a supermartingale. ♢

Our objective is to prove under the above condition (and some more minor assumptions) that no matter what X_0 is, there exists a bound on $\mathsf{E}_{x_0}\bigl[V(X_t)\bigr]$ uniformly over time (the bound, of course, might depend on X_0).

Theorem 2. *Let $\theta(t) := 1/\varphi'(t, 0+)$ denote the right-hand derivative of the convex function $\varphi(t,\cdot)$ at zero. Assume that*

$$C := \sum_{t\in\mathbb{N}_0} \theta(t) < \infty. \tag{1}$$

In addition assume

(i) $\delta := \sup_{x\in K} V(x) < \infty$, and $\gamma := \sup_{t\in\mathbb{N}_0} \theta(t) < \infty$.
(ii) $\beta := \sup_{x_0\in K} \mathsf{E}\bigl[\varphi(0, V(X_1))\mathbf{1}_{\{X_1\in\mathcal{S}\smallsetminus K\}} \bigm| X_0 = x_0\bigr] < \infty$.

Then, under Assumption 1 and the conditions preceding, we have

$$\sup_{t\in\mathbb{N}_0} \mathsf{E}_{x_0}\bigl[V(X_t)\bigr] \leqslant C\beta + \delta + \gamma\varphi(0, V(x_0)).$$

In the rest of this section we prove the above theorem. Fix a time t, and define

$$g_t := \sup\Bigl\{s \leqslant t,\ s\in\mathbb{N}_0 \Bigm| X_s \in K\Bigr\} \quad\text{and}\quad h_t := \inf\Bigl\{s \geqslant t,\ s\in\mathbb{N}_0 \Bigm| X_s \in K\Bigr\}.$$

We follow the standard custom of defining supremum over empty sets to be $-\infty$, and the infimum over empty sets to be $+\infty$.

Note that g_t is not a stopping time with respect to the natural filtration generated by the process $(X_t)_{t\in\mathbb{N}_0}$, although h_t is. The random interval $[g_t, h_t]$ is a singleton if and only if $X_t \in K$. Otherwise, we say that X_t is within an excursion outside K. Naturally the random variable $\zeta := t - g_t$ is important, where ζ depends on t although we have suppressed it in the notation to keep it clean; ζ keeps track of what step in the excursion we currently are.

Now, we have the following decomposition

$$\mathsf{E}_{x_0}\bigl[V(X_t)\bigr] = \mathsf{E}_{x_0}\bigl[V(X_t)\mathbf{1}_{\{g_t=-\infty\}}\bigr] + \sum_{s=0}^{t} \mathsf{E}_{x_0}\bigl[V(X_t)\mathbf{1}_{\{g_t=s\}}\bigr]. \tag{2}$$

Our first objective is to bound each of the expectations $\mathsf{E}_{x_0}\bigl[V(X_t)\mathbf{1}_{\{g_t=s\}}\bigr]$.

Before we move on, let us first prove a little lemma which follows readily from Assumption 1.

Lemma 3. *Let $X_0 = x_0 \in \mathcal{S}\smallsetminus K$. Then*

$$\mathsf{E}_{x_0}\bigl[V(X_s)\mathbf{1}_{\{\tau_K>s\}}\bigr] \leqslant \varphi(0, V(X_0))\theta(s) \qquad \forall\, s\in\mathbb{N}_0, \tag{3}$$

where θ is defined in Theorem 2. Moreover, we have

$$\mathsf{P}_{x_0}\bigl(\tau_K > s\bigr) \leqslant \frac{\varphi(0, V(X_0))}{\varphi(s, 0)}. \tag{4}$$

Proof. This is a straightforward application of Optional Sampling Theorem (OST) for supermartingales [Kal02, Theorem 7.29]. Applying OST for the bounded stopping time $s\wedge\tau_K$ to the supermartingale $\bigl(\varphi(t, V(X_t))\bigr)_{t\in\mathbb{N}_0}$, we get

$$\begin{aligned}\varphi(0, V(X_0)) &\geqslant \mathsf{E}_{x_0}\bigl[\varphi(s\wedge\tau_K, V(X_{s\wedge\tau_K}))\bigr]\\ &\geqslant \mathsf{E}_{x_0}\bigl[\varphi(s, V(X_s))\mathbf{1}_{\{\tau_K>s\}}\bigr], \qquad \text{since } \varphi \text{ is nonnegative.}\end{aligned}$$

Now, since φ is convex in the second variable, we can write

$$\varphi(s,\xi) \geqslant \varphi(s,0) + \varphi'(s,0+)\xi.$$

Thus, substituting back, one has

$$\varphi(0,V(X_0)) \geqslant \varphi(s,0)\mathsf{P}_{x_0}\big(\tau_K > s\big) + \varphi'(s,0+)\mathsf{E}_{x_0}\big[V(X_s)\mathbf{1}_{\{\tau_K>s\}}\big].$$

Looking at each term separately on the right-hand side of the inequality above, and noting that both ϕ and ϕ' are positive, we arrive at (3) and (4). □

We are ready for the proof of Theorem 2.

Proof of Theorem 2. Let us consider three separate cases:

Case 1. $(-\infty < g_t < t)$. In this case g_t can take values $\{0,1,2,\ldots,t-1\}$. Now, if $s \in \{0,1,2,\ldots,t-1\}$, then

$$\begin{aligned}\mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{g_t=s\}}\big] &= \mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{X_s\in K\}}\mathbf{1}_{\{X_i\notin K,\ i=s+1,\ldots,t\}}\big]\\ &= \int_K Q^s(x_0,\mathrm{d}x)\int_{\mathcal{S}\smallsetminus K} Q(x,\mathrm{d}y)\ \mathsf{E}_y\big[V(X_{t-s-1})\mathbf{1}_{\{\tau_K>t-s-1\}}\big],\end{aligned}$$

and by (3) it follows that the right-hand side is at most

$$\int_K Q^s(x_0,\mathrm{d}x)\int_{\mathcal{S}\smallsetminus K} Q(x,\mathrm{d}y)\ \varphi(0,V(y))\theta(t-s-1).$$

Thus, one has

$$\begin{aligned}\mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{g_t=s\}}\big] &\leqslant \theta(t-s-1)\int_K Q^s(x_0,\mathrm{d}x)\int_{\mathcal{S}\smallsetminus K} Q(x,\mathrm{d}y)\ \varphi(0,V(y))\\ &\leqslant \theta(t-s-1)\sup_{x\in K}\mathsf{E}_x\big[\varphi(0,V(X_1))\mathbf{1}_{\{X_1\in\mathcal{S}\smallsetminus K\}}\big]\\ &= \theta(t-s-1)\beta.\end{aligned}$$

Case 2. $(g_t = t)$. This is easy, since $X_t \in K$ implies $V(X_t) \leqslant \delta$. Thus

$$\mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{g_t=t\}}\big] \leqslant \delta.$$

Case 3. $(g_t = -\infty)$. This is the case when the chain started from outside K and has not yet hit K, and therefore,

$$\mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{g_t=-\infty\}}\big] = \mathsf{E}_{x_0}\big[V(X_t)\mathbf{1}_{\{\tau_K>t\}}\big] \leqslant \varphi(0,V(x_0))\theta(t).$$

Combining all three cases above, we get the bound:

$$\mathsf{E}_{x_0}\big[V(X_t)\big] \leqslant \sum_{s=0}^{t-1}\theta(t-s-1)\beta + \delta + \varphi(0,V(x_0))\theta(t). \tag{5}$$

Maximizing the right-hand side of (5) over t, we arrive at

$$\sup_{t\in\mathbb{N}_0}\mathsf{E}_{x_0}\big[V(X_t)\big] \leqslant \beta\sum_{s=0}^{\infty}\theta(s) + \delta + \varphi(0,V(x_0))\sup_{t\in\mathbb{N}_0}\theta(t),$$

which is the bound stated in the theorem. □

The classical Foster-Lyapunov type supermartingales [MT93] are immediate examples of the previous theorem. In that case $\varphi(t,\xi) = \mathrm{e}^{\alpha t}\xi$, for some positive α. Thus $\varphi(t,\cdot)$ is a linear (hence, convex) function for each fixed t. The partial derivative of φ with respect to ξ at $\xi = 0$ is $1/\theta(t) = \mathrm{e}^{\alpha t}$, which shows that the sequence $(\theta(t))_{t\in\mathbb{N}_0}$ is summable. See also [FK04] and the references therein for more general Foster-Lyapunov type conditions.

2.1. **A Class of Hybrid Processes.** The preceding analysis can be extended for processes which switch their behavior depending on whether the current value is within K or not. They constitute a particularly useful class of controlled processes in which a controller attempts to drive the system into a *safe set* $K \subseteq \mathcal{S}$ whenever the system gets out of K due to its inherent randomness. Below we give a rigorous construction of such a process.

Consider a pair of Markov chains (Y, Z) where Y is a time-homogeneous chain, and Z is a (possibly) time-inhomogeneous chain. We construct a *hybrid discrete-time stochastic process* X by the following recipe. Firstly, let the state space for the process be $\mathcal{S}^{\mathbb{N}_0}$ along with the natural filtration

$$\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$$

generated by the coordinate maps. Secondly, we define the sequence of stopping times $\sigma_0 := \tau_0 := -\infty$ and $\tau_1 \leqslant \sigma_1 \leqslant \tau_2 \leqslant \sigma_2 \leqslant \dots$ by

$$\begin{aligned} \tau_i &:= \inf\bigl\{t > \sigma_{i-1} \bigm| X_t \in K\bigr\} \quad \text{and} \\ \sigma_i &:= \inf\bigl\{t > \tau_i \bigm| X_t \notin K\bigr\} \end{aligned}$$

for $i \in \mathbb{N}_0$. Finally, we define the process X by

$$\begin{aligned} &\mathsf{P}\bigl(X_{t+1} \in B \bigm| \mathcal{F}_t\bigr) = \mathsf{P}\bigl(Y_1 \in B \bigm| Y_0 = x\bigr) && \text{when } X_t = x, \text{ and} \\ & && \tau_i \leqslant t < \sigma_i, \text{ for some } i, \text{ and} \\ &\mathsf{P}\bigl(X_{t+1} \in B \bigm| \mathcal{F}_t\bigr) = \mathsf{P}\bigl(Z_{t+1-\sigma_i} \in B \bigm| Z_{t-\sigma_i} = x\bigr) && \text{when } X_t = x, \text{ and} \\ & && \sigma_i \leqslant t < \tau_{i+1}, \end{aligned}$$

for a measurable $B \subseteq \mathcal{S}$. Clearly, the process defined above behaves as the homogeneous chain Y whenever it is inside K. Once the process X exits the set K, a controller alters the behavior of the chain which, until it enters K again, behaves as a copy of the inhomogeneous chain Z starting from a point outside K. The process X is in general non-Markovian due to the possible time inhomogeneity. Nevertheless, it is a natural class of examples of switching systems whose Markovian behavior switches in different regions on the state space. We say that X is a (Y, Z)*-hybrid with respect to* K.

The following generalization of Theorem 2 can be proved along lines of the original proof. Note that the time inhomogeneity of the chain Z does not affect the arguments in the proof at any step.

Theorem 4. *Consider a stochastic process X which is a (Y, Z)-hybrid with respect to K for some homogeneous Markov chain Y and some possibly inhomogeneous Markov chain Z.*

Let $\varphi : \mathbb{N}_0 \times \mathbb{R}^+ \longrightarrow \mathbb{R}$ be a function which is increasing and convex in the second argument, such that the process $\bigl(\varphi(t, V(Z_t))\bigr)_{t \in \mathbb{N}_0}$ is a supermartingale. Suppose that:

(i) the finiteness condition in (1) holds, i.e.,

$$C := \sum_{t=0}^{\infty} \frac{1}{\varphi'(t, 0+)} < \infty. \tag{6}$$

(ii) $\delta := \sup_{x \in K} V(x) < \infty$, and $\gamma := \sup_{t \in \mathbb{N}_0} \theta(t) < \infty$;
(iii) $\beta := \sup_{y_0 \in K} \mathsf{E}\bigl[\varphi(0, V(Y_1)) \mathbf{1}_{\{Y_1 \in \mathcal{S} \smallsetminus K\}} \bigm| Y_0 = y_0\bigr] < \infty$.

Then, if the hybrid model X starts from $x_0 \in \mathcal{S} \smallsetminus K$, one has

$$\sup_{t \in \mathbb{N}_0} \mathsf{E}_{x_0}\bigl[V(X_t)\bigr] \leqslant C\beta + \delta + \gamma\varphi(0, V(x_0)).$$

It is interesting to note that the right side of above bound is a total of individual contributions by the control (for C), the choice of K (for δ), and the initial configuration (for $V(x_0)$). We would like to stress that the conclusion holds even when X is no longer a Markov chain due to the time inhomogeneity of Z. This is important, especially because operator-theoretic bounds like Foster-Lyapunov, or martingale-based bounds do not work in such a case.

2.2. **Examples.** Let us consider a class of examples for hybrid systems that can be analyzed by our results. Suppose that the state space for all the Markov chains is $\mathbb{R}^d$ and the safe set K is a compact set. It is clear that the only difficulty in applying Theorem 4 is to find a suitable function φ given the Markov chain Z and the function V. In applications, a natural choice for the function V is given by square of the Euclidean norm, i.e., $V(x) = \sum_{i=1}^d x_i^2$. For this choice of V, we describe below a natural class of examples of a Markov chains for which one can construct a φ that is nonnegative and increasing and convex in the second argument as required in the assumption of Theorem 4.

Consider a diffusion, with a possibly time-inhomogeneous drift function, given by the d-dimensional stochastic differential equation

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \mathrm{dW_t}, \tag{7}$$

where $W_t = (W_t(1), W_t(2), \ldots, W_t(d))$ is a vector of d independent Brownian motions. We shall construct a function φ such that φ satisfies all the required assumptions of Theorem 4, including the one that $\varphi(t, V(X_t))$ is a supermartingale outside a compact set K. If we now define $Z_i = X_{i \wedge \tau_K}$, i.e., the diffusion sampled at integer time points before hitting K, it is clear that Z is Markov chain such that $\varphi(i, V(Z_i))$ is a supermartingale. Thus, all that remains to check in order to apply Theorem 4 is the condition (6).

To construct such a φ, let us consider a well known family of one-dimensional diffusion, known as the squared Bessel processes. This family is indexed by a single nonnegative parameter $\delta \geqslant 0$ and is described as the unique strong solution of the SDE

$$dY_t = 2\sqrt{Y_t}d\beta_t + \delta dt, \qquad Y_0 = y_0 \geqslant 0, \tag{8}$$

where β is a one-dimensional standard Brownian motion. We have the following lemma.

Lemma 5. *Let $F : \mathbb{R} \to \mathbb{R}^+$ be a nonnegative, increasing, and convex function. Fix any terminal time τ. Then consider the function*

$$\varphi(t, y) = \mathrm{E}\left[F(Y_\tau)\,|Y_t = y\right], \quad 0 \leqslant t \leqslant \tau, \tag{9}$$

where Y is a solution of the SDE (8).

The function φ is obviously nonnegative, being the expectation of a nonnegative function. We claim the following is true.

(1) φ is increasing in y.
(2) φ is convex in y.
(3) And, φ satisfies the differential equation

$$\frac{\partial \varphi}{\partial t} + \delta\varphi' + 2y\varphi'' = 0, \quad y > 0,\ \tau > t > 0, \tag{10}$$

where φ' and φ'' refers to taking first and second derivatives with respect to the second argument y.

Proof. The proof proceeds by coupling. Let us first show that φ is increasing. Consider any two starting points $0 \leqslant x < y$. Construct on the same sample space two copies of BESQ processes $Y^{(1)}$ and $Y^{(2)}$ such that both of them satisfy (8) with respect to the same Brownian motion β but

$$Y_0^{(1)} = x, \quad \text{and} \quad Y_0^{(2)} = y.$$

This is possible to do since the SDE (8) admits a strong solution (see Proposition 2.13 in [KS08]).

Hence, by Proposition 2.18 in [KS08], it follows that $Y_t^{(1)} \leqslant Y_t^{(2)}$ for all $t \geqslant 0$. Since F is an increasing function, we get

$$\varphi(t,x) = \mathrm{E}\left(F(Y^{(1)}_{\tau-t})\right) \leqslant \mathrm{E}\left(F(Y^{(2)}_{\tau-t})\right) = \varphi(t,y).$$

This proves that φ is increasing in the second argument.

For convexity, we use a different coupling. We follow arguments very similar to the one used in the proof of Theorem 3.1 in [Hob98]. Consider three initial points $0 < z < y < x$. And let $\hat{X}, \hat{Y}, \hat{Z}$ be three independent BESQ processes that start from x, y, and z respectively. Define the stopping times

$$\tau_x = \inf\left\{u : \hat{Y}_u = \hat{X}_u\right\}, \quad \tau_z = \inf\left\{u : \hat{Y}_u = \hat{Z}_u\right\}.$$

Fix a time t and let $T = \tau - t$ where τ is the terminal time. Define

$$\sigma = \tau_x \wedge \tau_z \wedge T.$$

Now, on the event $\sigma = \tau_x$, it follows from symmetry that

$$\begin{aligned} &\mathrm{E}\left[\left(\hat{X}_T - \hat{Z}_T\right) F(\hat{Y}_T); \sigma = \tau_x\right] = \mathrm{E}\left[\left(\hat{Y}_T - \hat{Z}_T\right) F(\hat{X}_T); \sigma = \tau_x\right], \\ &\mathrm{E}\left[\left(\hat{X}_T - \hat{Y}_T\right) F(\hat{Z}_T); \sigma = \tau_x\right] = 0. \end{aligned} \tag{11}$$

Similarly, on the event $\sigma = \tau_z$, we have

$$\begin{aligned} &\mathrm{E}\left[\left(\hat{X}_T - \hat{Z}_T\right) F(\hat{Y}_T); \sigma = \tau_x\right] = \mathrm{E}\left[\left(\hat{X}_T - \hat{Y}_T\right) F(\hat{Z}_T); \sigma = \tau_x\right], \\ &\mathrm{E}\left[\left(\hat{Z}_T - \hat{Y}_T\right) F(\hat{X}_T); \sigma = \tau_x\right] = 0. \end{aligned} \tag{12}$$

And finally, when $\sigma = T$, we must have $\hat{Z}_T < \hat{Y}_T < \hat{X}_T$. We use the convexity property of F itself to get

$$\mathrm{E}\left[(\hat{X}_T - \hat{Z}_T)F(\hat{Y}_T); \sigma = T\right] \leqslant \mathrm{E}\left[(\hat{X}_T - \hat{Y}_T)F(\hat{Z}_T); \sigma = T\right] + \mathrm{E}\left[(\hat{Y}_T - \hat{Z}_T)F(\hat{X}_T); \sigma = T\right]. \tag{13}$$

Combining the three cases in (11), (12), and (13) we get

$$\mathrm{E}\left[(\hat{X}_T - \hat{Z}_T)F(\hat{Y}_T)\right] \leqslant \mathrm{E}\left[(\hat{X}_T - \hat{Y}_T)F(\hat{Z}_T)\right] + \mathrm{E}\left[(\hat{Y}_T - \hat{Z}_T)F(\hat{X}_T)\right]. \tag{14}$$

We now use the fact that $\hat{X}, \hat{Y}$, and $\hat{Z}$ are independent. Also, it is not difficult to see from the SDE (8) that $\mathrm{E}(\hat{X}_T) - x = \mathrm{E}(\hat{Y}_T) - y = \mathrm{E}(\hat{Z}_T) - z = \delta t$.

Thus from (14) we infer that

$$(x - z)\varphi(t,y) \leqslant (x-y)\varphi(t,z) + (y-z)\varphi(t,x), \quad \text{for all } 0 < z < y < x.$$

This proves convexity of φ in its second argument.

The equation (10) is simply the classical relation between the expectations of diffusions and the solutions of partial differential equations. □

Let us now return to the multidimensional diffusion given by (7). We now consider the process $\zeta_t = \varphi(t, \|X_t\|^2)$ where φ is the same function as in (9). We prove the following.

Theorem 6. *Assume that the drift function $b = (b_1, b_2, \dots, b_d)$ in the SDE (7) is such that there is a compact set $K \subseteq \mathbb{R}^d$ such that*

$$\sum_{i=1}^{d} x_i b_i(t,x) < 0, \quad \text{for all } x \notin K,\ t \geqslant 0.$$

Consider the process $\zeta_t = \varphi(t, \|X_t\|^2)$. Fix any terminal time $\tau > 0$.

With the set-up as above the stopped process $\varphi(\zeta_{t\wedge\tau_K\wedge\tau})$ is a (local) supermartingale. Here φ is the nonnegative, increasing, convex function defined in (9) with δ being equal to the dimension d. And τ_K is the hitting time of the set K.

Proof. Applying Itô's rule to ζ_t, we get

$$\mathrm{d}\zeta_t = dM_t + \left[\frac{\partial\varphi}{\partial t} + \mathcal{L}\varphi\right] dt, \tag{15}$$

where M is a local martingale which is a martingale under additional assumptions of boundedness on the first derivative of φ. Here $\mathcal{L}$ is the generator of X, and thus

$$\begin{aligned}\frac{\partial\varphi}{\partial t} + \mathcal{L}\varphi &= \frac{\partial\varphi}{\partial t} + \sum_{i=1}^{d} b_i \frac{\partial\varphi}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{d}\frac{\partial^2\varphi}{\partial x_i^2}\\ &= \frac{\partial\varphi}{\partial t} + 2\varphi' \sum_{i=1}^{d} b_i x_i + \frac{1}{2}\left[2d\varphi' + \varphi''\sum_{i=1}^{d} 4x_i^2\right]\\ &= \frac{\partial\varphi}{\partial t} + d\phi' + 2\left(\sum_i x_i^2\right)\varphi'' + 2\varphi'\sum_{i=1}^{d} b_i x_i = 2\varphi'\sum_i b_i x_i.\end{aligned}$$

The final equality holds since φ satisfies (10) at $y = \sum_i x_i^2$.

Now, we know that $\varphi' > 0$ since φ is increasing. And, by our assumption, $\sum_i x_i b_i < 0$ outside K. Thus,

$$\frac{\partial\varphi}{\partial t} + \mathcal{L}\varphi \leqslant 0, \quad \text{for all } x \notin K,\ 0 \leqslant t \leqslant \tau.$$

Hence from the semimartingale decomposition given in (15), the result follows. □

As another class of examples consider the linear system

$$X_{t+1} = AX_t + w_t, \qquad X_0 = x_0, \quad t \in \mathbb{N}_0, \tag{16}$$

where $x_0 \in \mathbb{R}^d$, A is an $d \times d$ matrix with real entries, $(w_t)_{t\in\mathbb{N}_0}$ is a sequence of independent and identically distributed d-dimensional random vectors with mean 0 and covariance matrix Σ. We assume that the matrix A is Schur stable, i.e., its eigenvalues are strictly less than 1 in magnitude. As such, the matrix A is nonsingular. Note that the transition kernel Q for $x \in \mathbb{R}^d$ and $B \subseteq \mathbb{R}^d$ a Borel subset is given by $Q(x, B) = \tilde{\mathsf{P}}(B - Ax)$, where $\tilde{\mathsf{P}}$ is the probability distribution of w_0. (Here $B - Ax$ is the set $\{y \in \mathbb{R}^d \mid y + Ax \in B\}$.)

Let P be a symmetric positive definite $d \times d$ matrix with real entries that solves the Lyapunov equation $A^{\mathrm{T}}PA - P \leqslant -I_d$, where I_d is the $d \times d$ identity matrix. It is a standard result that if A is Schur stable, such a matrix P exists. We define the function $V(x) := x^{\mathrm{T}}Px = \|x\|_P^2$, and calculate

$$\begin{aligned}QV(x) &= \mathsf{E}\big[X_1^{\mathrm{T}}PX_1 \big| X_0 = x\big] = \mathsf{E}\big[(Ax + w_0)^{\mathrm{T}}P(Ax + w_0)\big]\\ &= \|x\|^2_{A^{\mathrm{T}}PA} + \mathsf{E}\big[w_0^{\mathrm{T}}Pw_0\big] = \|x\|^2_{A^{\mathrm{T}}PA} + \operatorname{trace}(P\Sigma).\end{aligned}$$

If $\lambda_{\max}(P)$ denotes the maximal eigenvalue of P (which is positive), then we have $\|x\|^2_{A^\top PA} \leqslant \|x\|^2_P - \|x\|^2 \leqslant (1-\theta)\|x\|^2_P - (1-\theta\lambda_{\max}(P))\|x\|^2$ for $\theta \in]0, 1/\lambda_{\max}(P)[$. Clearly, for $\lambda_\circ := 1-\theta$,

$$\|x\| > \sqrt{\frac{\operatorname{trace}(P\Sigma)}{1-\theta\lambda_{\max}(P)}} \implies QV(x) \leqslant \lambda_\circ V(x).$$

We define the compact set $K := \left\{ y \in \mathbb{R}^d \,\middle|\, \|y\| \leqslant \sqrt{\operatorname{trace}(P\Sigma)/(1-\theta\lambda_{\max}(P))} \right\}$. Since $x \in \mathbb{R}^d \smallsetminus K$ implies that $QV(x) \leqslant \lambda_\circ V(x)$, we see immediately that the process $\left(\mathrm{e}^{\alpha(t\wedge\tau_K)} V\big(X_{t\wedge\tau_K}\big)\right)_{t\in\mathbb{N}_0}$ is a supermartingale, where $\alpha := -\ln\lambda_\circ > 0$. Since w_t is square-integrable, the quantities β and δ defined in Theorem 2 are finite; also, the function $\varphi(t,\xi) = \mathrm{e}^{\alpha t}\xi$, $\gamma = 1$, and $C = \sum_{t\in\mathbb{N}_0} \mathrm{e}^{-\alpha t} < \infty$. Therefore, we see that Theorem 2 applies, and that $\sup_{t\in\mathbb{N}_0} \mathsf{E}_{x_0}\big[V(X_t)\big] = \sup_{t\in\mathbb{N}_0} \mathsf{E}_{x_0}\big[\|X_t\|^2_P\big] < \infty$.

The assertion about the linear system (16) can also be established via classical Foster-Lyapunov conditions, as follows. Recall that a Markov process $(X_t)_{t\in\mathbb{N}_0}$ with transition kernel Q taking values in $\mathbb{R}^d$, is said to satisfy the *Foster-Lyapunov condition* [Mey08, Chapter 8] if

$$QV(x) \leqslant \lambda V(x) \quad \text{whenever} \quad \|x\| > a. \tag{17}$$

Given the condition (17), it is not difficult to directly derive a uniform upper bound on $\mathsf{E}_x\big[V(X_t)\big]$. Indeed, from (17) it follows that

$$QV(x) \leqslant \lambda V(x) + b\mathbf{1}_{\bar{B}_a}(x) \qquad \forall\, x \in \mathbb{R}^d.$$

Iterating the above inequality we get for $t \in \mathbb{N}$,

$$\begin{aligned} \mathsf{E}_x\big[V(X_t)\big] = Q^t V(x) &\leqslant \lambda^t V(x) + b\sum_{i=0}^{t-1} \lambda^i Q^i \mathbf{1}_{\bar{B}_a}(x) \\ &= \lambda^t V(x) + b\sum_{i=0}^{t-1} \lambda^i \mathsf{P}_x\big(X_i \in \bar{B}_a\big) \leqslant V(x) + \frac{b}{1-\lambda}. \end{aligned}$$

Given a distribution μ_0 of the initial condition X_0, if V is μ_0-integrable, it follows that $\sup_{t\in\mathbb{N}_0} \mathsf{E}\big[V(X_t)\big] \leqslant \int_{\mathbb{R}^d} \mu_0(\mathrm{d}x)V(x) + \frac{b}{1-\lambda} < \infty$.

2.3. **Invariant Measures.** The existence of an invariant measure of the Markov process $(X_t)_{t\in\mathbb{N}_0}$ follows from Theorem 2. We need the additional assumption that the transition kernel $(Q^t)_{t\in\mathbb{N}_0}$ is Feller, and invoke the classical Krylov-Bogoliubov theorem.

Recall that a Markov semigroup $(Q^t)_{t\in\mathbb{N}_0}$ on $\mathbb{R}^d$ is *Feller* if Q^t maps the bounded continuous functions into bounded continuous functions for each $t \in \mathbb{N}_0$; we say that $(Q^t)_{t\in\mathbb{N}_0}$ is a *Markov-Feller* semigroup. An invariant measure of the process $(X_t)_{t\in\mathbb{N}_0}$ or of the semigroup $(Q^t)_{t\in\mathbb{N}_0}$ is a A family (μ_i) of Borel probability measures on $\mathbb{R}^d$ is *tight* if for each $\varepsilon > 0$ there exists a compact set $K \subseteq \mathbb{R}^d$ such that $\inf_i \mu_i(K) \geqslant 1-\varepsilon$.

Theorem 7 (Krylov-Bogoliubov, [LB07, Theorem 8.1.19])**.** *For $x_0 \in \mathbb{R}^d$ and $T \in \mathbb{N}$ let μ_T be the measure $\mu_T(A) := \frac{1}{T}\sum_{t=0}^{T-1} Q^t(x_0, A)$, where A is a measurable subset of $\mathbb{R}^d$. If for some $x_0 \in \mathbb{R}^d$ the family $(\mu_T)_{T\in\mathbb{N}}$ is tight, then there exists an invariant measure of the semigroup $(Q^t)_{t\in\mathbb{N}_0}$.*

The following lemma is quite standard, but we provide a short proof for completeness.

Lemma 8. *Let* $V : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$ *be a measurable function such that the sublevel sets* $\{V \leqslant c\}$ *are compact for all* c *large enough, and* $\mathcal{M}$ *be a set of probability measures on* $\mathbb{R}^d$*. If* $\sup_{\mu\in\mathcal{M}} \int_{\mathbb{R}^d} \mu(\mathrm{d}x)\, V(x) < \infty$*, then* $\mathcal{M}$ *is tight.*

Proof. From Markov inequality, for $c > 0$ we have

$$\mu\big(\{V > c\}\big) \leqslant \frac{1}{c}\int_{\mathbb{R}^d} \mu(\mathrm{d}x)\, V(x) \leqslant \frac{1}{c}\sup_{\mu\in\mathcal{M}} \int_{\mathbb{R}^d} \mu(\mathrm{d}x)\, V(x),$$

and therefore, $\lim_{c\to\infty} \sup_{\mu\in\mathcal{M}} \mu\big(\mathbb{R}^d \smallsetminus \{V \leqslant c\}\big) = 0$. This shows that $\mathcal{M}$ is tight. □

Proposition 9. *Consider a Markov process* $(X_t)_{t\in\mathbb{N}_0}$ *with a Markov-Feller transition semigroup* $(Q^t)_{t\in\mathbb{N}_0}$*, and let* $V : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$ *be a measurable function such that the sublevel sets* $\{V \leqslant k\} \subseteq \mathbb{R}^d$ *are compact for all* k *large enough. Suppose that the hypotheses of Theorem 2 hold. Then there exists an invariant measure of the process* $(X_t)_{t\in\mathbb{N}_0}$*.*

Proof. Theorem 2 guarantees the existence of a $C > 0$ such that

$$\sup_{t\in\mathbb{N}_0} \int_{\mathbb{R}^d} Q^t(x_0, \mathrm{d}y)\, V(y) = \sup_{t\in\mathbb{N}_0} \mathsf{E}_{x_0}\big[V(X_t)\big] \leqslant C < \infty.$$

Therefore, the family $\big(Q^t(x_0,\cdot)\big)_{t\in\mathbb{N}_0}$ is tight by Lemma 8. We claim that tightness of the measures $\big(Q^t(x_0,\cdot)\big)_{t\in\mathbb{N}_0}$ implies that the set of measures $(\mu_T)_{T\in\mathbb{N}}$ defined in Theorem 7 is also tight. Indeed, for a given $\varepsilon > 0$ we choose a compact $K \subseteq \mathbb{R}^d$ such that $\sup_{t\in\mathbb{N}_0} Q^t\big(x_0, \mathbb{R}^d \smallsetminus K\big) < \varepsilon$; therefore $\mu_T\big(\mathbb{R}^d \smallsetminus K\big) = \frac{1}{T}\sum_{i=0}^{T-1} Q^t\big(x_0, \mathbb{R}^d \smallsetminus K\big) < \varepsilon$ for all $T \in \mathbb{N}$. The Krylov-Bogoliubov Theorem 7 now guarantees the existence of an invariant measure of the semigroup $(Q^t)_{t\in\mathbb{N}_0}$. □

3. Applications

In this section we look at several cases in which Theorem 2 of §2 applies and gives useful uniform $\boldsymbol{L}_1$ bounds of Lyapunov functions. In §3.1 we give sufficient conditions for global asymptotic stability almost surely and in $\boldsymbol{L}_1$ of discrete-time randomly switched systems. Assumptions of global contractivity in its standard form or memoryless choice of the maps at each iterate are absent; we just require a condition resembling average contractivity in terms of Lyapunov functions with a suitable coupling condition with the Markovian transition probabilities. In §3.2 we demonstrate that under mild hypotheses iterated function systems possess strong stability and robustness properties with respect to bounded disturbances that are not modelled as random processes.

3.1. Stability of Discrete-Time Randomly Switched Systems.

Consider the system

$$X_{t+1} = f_{\sigma_t}(X_t), \qquad X_0 = x_0, \quad t \in \mathbb{N}_0. \tag{18}$$

Here $\sigma : \mathbb{N}_0 \longrightarrow \mathcal{P} := \{1, \ldots, \mathrm{N}\}$ is a discrete-time random process, the map $f_i : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is continuous and locally Lipschitz, and there are points $x_i^\star \in \mathbb{R}^d$ such that $f_i(x_i^\star) = 0$ for each $i \in \mathcal{P}$. The initial condition of the system $x_0 \in \mathbb{R}^d$ is assumed to be known. Our objective is to study stability properties of this system by extracting certain nonnegative supermartingales.

The system (18) can be viewed as an iterated function system: $X_{t+1} = f_{\sigma_t} \circ \cdots \circ f_{\sigma_1} \circ f_{\sigma_0}(x_0)$. Varying the point x_0 but keeping the same maps leads to a family of Markov chains initialized from different initial conditions. The article [DF99] treats basic results on convergence and stationarity properties of such systems with the process $(\sigma_t)_{t\in\mathbb{N}_0}$ being a sequence of independent and identically distributed random variables taking values in $\mathcal{P}$, and each map f_i is a contraction. These results were generalized in [JT01] with the aid of Foster-Lyapunov arguments.

The analysis carried out in [JT01] requires a Polish state-space, and employs the following three principal assumptions: (a) the maps are non-separating on an average, i.e., the average separation of the Markov chains initialized at different points is nondecreasing over time; (b) there exists a set C such that the Markov chains started at different initial conditions contract after the set C is reached; and (c) there exists a measurable real-valued function $V \geqslant 1$, bounded on C, and satisfying a Foster-Lyapunov drift condition $QV(x) \leqslant \lambda V(x) + b\mathbf{1}_C(x)$ for some $\lambda \in]0,1[$ and $b < \infty$. Under these conditions the authors establish the existence and uniqueness of an invariant measure which is also globally attractive, and the convergence to this measure is exponential. In particular, this showed that the main results of [DF99], which are primarily related to existence and uniqueness of invariant probability measures, continue to hold if the contractivity hypotheses on the family $\{f_i\}_{i\in\mathcal{P}}$ are relaxed. In this subsection we look at stronger properties, namely, $\boldsymbol{L}_1$ boundedness and stability, and almost sure stability of the system (18) under Assumption 1. We note that no contractivity inside a compact set is needed to establish existence of an invariant measure under Assumption 1; the latter follows immediately from Proposition 9.

Assumption 10. The process $(\sigma_t)_{t\in\mathbb{N}_0}$ is an irreducible Markov chain with initial probability distribution π° and a transition matrix $P := [p_{ij}]_{\mathrm{N}\times\mathrm{N}}$. ♢

It is immediately clear that the discrete-time process $(\sigma_t, X_t)_{t\in\mathbb{N}_0}$, taking values in the Borel space $\mathcal{P}\times\mathbb{R}^d$, is Markovian under Assumption 10. The corresponding transition kernel is given by

$$Q\big((i,x), \mathcal{P}'\times B\big) = \textstyle\sum_{j\in\mathcal{P}'} p_{ij}\mathbf{1}_B\big(f_j(x)\big) \quad \text{for } \mathcal{P}'\subseteq\mathcal{P}, B \text{ a Borel subset of } \mathbb{R}^d, \text{ and } (i,x)\in\mathcal{P}\times\mathbb{R}^d.$$

Our basic analysis tool is a family of Lyapunov functions, one for each subsystem, and at different times we shall impose the following two distinct sets of hypotheses on them.

Assumption 11. There exist a family $\{V_i\}_{i\in\mathcal{P}}$ of nonnegative measurable functions on $\mathbb{R}^d$, functions $\alpha_1, \alpha_2 \in \mathcal{K}$, numbers $\lambda_\circ \in]0,1[$, $r > 0$ and $\mu > 1$, such that

(V1) $\alpha_1(\|x - x_i^\star\|) \leqslant V_i(x) \leqslant \alpha_2(\|x - x_i^\star\|)$ for all x and i,
(V2) $V_i(x) \leqslant \mu V_j(x)$ whenever $\|x\| > r$, for all i, j, and
(V3) $V_i(f_i(x)) \leqslant \lambda_\circ V_i(x)$ for all x and i. ♢

Assumption 12. There exist a family $\{V_i\}_{i\in\mathcal{P}}$ of nonnegative measurable functions on $\mathbb{R}^d$, functions $\alpha_1, \alpha_2 \in \mathcal{K}$, a matrix $[\lambda_{ij}]_{\mathrm{N}\times\mathrm{N}}$ with nonnegative entries, and numbers $r > 0$, $\mu > 1$, such that (V1)-(V2) of Assumption 11 hold, and

(V3′) $V_i(f_j(x)) \leqslant \lambda_{ij} V_i(x)$ for all x and i, j. ♢

Let us examine the above conditions in detail. It will be useful to recall here that the deterministic system $x_{t+1} = f_i(x_{t+1})$, $t \in \mathbb{N}_0$, with initial condition x_0 is said to be *globally asymptotically stable* (in the sense of Lyapunov) if (a) for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|x_0 - x_i^\star\| < \delta$ implies $\|x_t - x_i^\star\| < \varepsilon$ for all $t \in \mathbb{N}_0$, and (b) for every $r, \varepsilon' > 0$ there exists a $T > 0$ such that $\|x_0 - x_i^\star\| < r$ implies $\|x_t - x_i^\star\| < \varepsilon$ for all $t > T$. The condition (a) goes by the name of Lyapunov stability of the dynamical system (or of the corresponding equilibrium point $x_i^\star$), and (b) is the standard notion of global asymptotic convergence to $x_i^\star$.

The condition (V1) in Assumption 11 is standard in deterministic system theory literature; it ensures that each function V_i is positive definite; if α_1 is assumed to be of class $\mathcal{K}_\infty$ then V_i is also radially unbounded from $x_i^\star$. (V2) stipulates that outside

$\bar{B}_r$ the functions $\{V_i\}_{i\in\mathcal{P}}$ are linearly comparable to each other. This is possible if, for instance, $V_1(x) = \|x - x_1^\star\|_{P_1}^4$ and $V_2(x) = \|x - x_2^\star\|_{P_2}^2 + \|x - x_2^\star\|_{P_3}^4$ for some $x_1^\star, x_2^\star \in \mathbb{R}^d$ and positive definite matrices $P_i,\ i = 1, 2, 3$. The conditions (V1) and (V3) together imply that each subsystem is globally asymptotically stable, with sufficient stability margin—the smaller the number $\lambda_\circ$, the greater is the stability margin. In fact, standard converse Lyapunov theorems show that (V1) and (V3) are necessary and sufficient conditions for each subsystem to be globally asymptotically stable.

The only difference between Assumption 11 and Assumption 12 is in the conditions (V3) and (V3′). In the latter we keep track of how each Lyapunov function evolves along trajectories of every subsystem. Note that λ_{ii} is not required to be less than 1 for each i, which corresponds to allowing unstable subsystems.

We define $\hat{p}$ and $\tilde{p}$ to be the maximum diagonal and off-diagonal entry of the matrix P, respectively, i.e., $\hat{p} := \max_{i\in\mathrm{N}} p_{ii}$ and $\tilde{p} := \max_{i,j\in\mathcal{P}, i\neq j} p_{ij}$.

Proposition 13. *Consider the system (18), and suppose that either of the following two conditions holds:*

(S1) *Assumptions 10 and 11 hold, and* $\lambda_\circ(\hat{p} + \mu\tilde{p}) < 1$.

(S2) *Assumptions 10 and 12 hold, and* $\mu \cdot \Big(\max_{i\in\mathcal{P}} \sum_{j\in\mathcal{P}} p_{ij}\lambda_{ji}\Big) < 1$.

Let $\tau_r := \inf\{t \in \mathbb{N}_0 \,|\, \|X_t\| \leqslant r\}$ *and* $V_i'(x) := V_i(x)\mathbf{1}_{\mathbb{R}^d\smallsetminus\bar{B}_r}(x)$. *Suppose that* $\|x_0\| > r$. *Then there exists* $\alpha > 0$ *such that the process* $\big(\mathrm{e}^{\alpha(t\wedge\tau_r)} V'_{\sigma_{t\wedge\tau_r}}(X_{t\wedge\tau_r})\big)_{t\in\mathbb{N}_0}$ *is a nonnegative supermartingale.*

Note that the condition (S1) of Proposition 13 exploits far less of the Markovian structure of $(\sigma_t)_{t\in\mathbb{N}_0}$ than the condition (S2)—just two elements of the transition matrix $[p_{ij}]_{\mathrm{N}\times\mathrm{N}}$ are involved in (S1), whereas all the elements of this matrix are involved in (S2). On the one hand, the inequality in (S1) is clearly conservative, since we employ the uniform bounds of the transition probability matrix and stability margins of the subsystems. None of the subsystems is permitted to be unstable or expansive, for $\lambda_\circ$ is assumed to be strictly less than 1. On the other hand, the inequality in (S2) is a weighted sum, and unstable or expansive subsystems are permissible as long as the corresponding weights are small. Finally, it is necessary to keep track of how each Lyapunov function evolves along the trajectories of every other subsystem if unstable or expansive subsystems are allowed, which is why we need all the entries of the matrix $[\lambda_{ij}]_{\mathrm{N}\times\mathrm{N}}$ in (S2), unlike (S1) where we employ uniform stability bounds captured by $\lambda_\circ$.

Corollary 14. *Consider the system (18), and assume that the hypotheses of Proposition 13 hold. Then there exists a constant* $C > 0$ *such that* $\sup_{t\in\mathbb{N}_0} \mathsf{E}\big[\alpha_1(\|X_t\|)\big] < C$.

It is possible to derive simple conditions for stability of the system (18) from Proposition 13. To this end we briefly recall two standard stability concepts.

Definition 15. If $\ker(f_i - \mathrm{id}) = \{0\}$ for each $i \in \mathcal{P}$, the system (18) is said to be

- *globally asymptotically stable almost surely* if

  (AS1) $\mathsf{P}\Big(\forall\,\varepsilon > 0\ \exists\,\delta > 0 \text{ s.t. } \sup_{t\in\mathbb{N}_0} \|X_t\| < \varepsilon \text{ whenever } \|x_0\| < \delta\Big) = 1$,

  (AS2) $\mathsf{P}\Big(\forall\, r, \varepsilon' > 0\ \exists\, T > 0 \text{ s.t. } \sup_{\mathbb{N}_0\ni t>T} \|X_t\| < \varepsilon' \text{ whenever } \|x_0\| < r\Big) = 1$;

- *α-stable in $\boldsymbol{L}_1$* for some $\alpha \in \mathcal{K}$ if

  (SM1) $\forall\,\varepsilon > 0\ \exists\,\delta > 0$ s.t. $\sup_{t\in\mathbb{N}_0} \mathsf{E}\big[\alpha(\|X_t\|)\big] < \varepsilon$ whenever $\|x_0\| < \delta$,

  (SM2) $\forall\, r, \varepsilon' > 0\ \exists\, T > 0$ s.t. $\sup_{\mathbb{N}_0\ni t>T} \mathsf{E}\big[\alpha(\|X_t\|)\big] < \varepsilon'$ whenever $\|x_0\| < r$. ◇

Corollary 16. *Suppose that* $\ker(f_i - \mathrm{id}) = \{0\}$ *for each* $i \in \mathcal{P}$, *and that either of the hypotheses* (S1) *and* (S2) *of Proposition* 13 *holds with* $r = 0$. *Then*

- *there exists* $\alpha > 0$ *such that* $\lim_{t\to\infty} \mathsf{E}\big[\mathrm{e}^{\alpha t} V_{\sigma_t}(X_t)\big] = 0$, *and*
- *the system* (18) *is globally asymptotically stable almost surely and* α_1*-stable in* $\boldsymbol{L}_1$ *in the sense of Definition* 15.

The proofs of Proposition 13, Corollary 14 and Corollary 16 are given after the following simple lemma. For $t \in \mathbb{N}$ let the random variable N_t denote the number of times the state of the Markov chain changes on the period of length t starting from 0, i.e., $N_t := \sum_{i=1}^{t} \mathbf{1}_{\{\sigma_{i-1}\neq\sigma_i\}}$.

Lemma 17. *Under Assumption* 10 *we have for* $s < t$, $s, t \in \mathbb{N}_0$,

$$\mathsf{P}\big(N_t - N_s = k\big|\sigma_s\big) \leqslant \begin{cases} \left(\binom{t-s}{k}\hat{p}^{(t-s-k)}\tilde{p}^k\right) \wedge 1 & \text{if } k = 0, 1, \ldots, t-s,\\ 0 & \text{else.}\end{cases}$$

Proof. Fix $s < t$, $s, t \in \mathbb{N}_0$, and let $\eta_k(s,t) := \mathsf{P}\big(N_t - N_s = k\big|\sigma_s\big)$. Then by the Markov property, for $k = 0, 1, \ldots, t-s$,

$$\begin{aligned}\eta_k(s,t) &= \mathsf{P}\big(N_t - N_s = k\big|\sigma_s\big)\\ &= \mathsf{P}\big(N_t - N_s = k\big|N_{t-1} - N_s = k, \sigma_s\big)\mathsf{P}\big(N_{t-1} - N_s = k\big|\sigma_s\big)\\ &\quad + \mathsf{P}\big(N_t - N_s = k\big|N_{t-1} - N_s = k-1, \sigma_s\big)\mathsf{P}\big(N_{t-1} - N_s = k-1\big|\sigma_s\big)\\ &= \eta_k(s,t-1)\mathsf{P}\big(N_t - N_s = k\big|N_{t-1} - N_s = k, \sigma_s\big)\\ &\quad + \eta_{k-1}(s,t-1)\mathsf{P}\big(N_t - N_s = k\big|N_{t-1} - N_s = k-1, \sigma_s\big)\\ &\leqslant \hat{p}\eta_k(s,t-1) + \tilde{p}\eta_{k-1}(s,t-1).\end{aligned}$$

The set of initial conditions $\eta_i(s,t) = 0$ for all $i \geqslant t-s$, follow from the trivial observation that there cannot be more than $t-s$ changes of σ on a period of length $t-s$. This gives a well-defined set of recursive equations, and a standard induction argument shows that $\eta_k(s,t) \leqslant \binom{t-s}{k}\hat{p}^{(t-s-k)}\tilde{p}^k$. This proves the assertion. □

The estimate in Lemma 17 resembles the distribution of a Binomial random variable, except that $\hat{p} + \tilde{p} \geqslant 1$ in our case. The smaller the difference between the maximal off-diagonal and diagonal elements of the transition matrix $[p_{ij}]_{\mathrm{N}\times\mathrm{N}}$, the tighter is the bound in Lemma 17. The bound also becomes tighter as the difference $(t-s)$ increases.

Proof of Proposition 13. First we look at the assertion under the condition (S1). Fix $s < t$, $s, t \in \mathbb{N}_0$. Given $(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})$, from (V3) we get $V'_{\sigma_{s\wedge\tau_r}}\big(X_{(s+1)\wedge\tau_r}\big) \leqslant \lambda_\circ V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big)$, and if $\sigma_{s+1} \neq \sigma_s$, we employ (V2) to get $V'_{\sigma_{(s+1)\wedge\tau_r}}\big(X_{(s+1)\wedge\tau_r}\big) \leqslant \mu V'_{\sigma_{s\wedge\tau_r}}\big(X_{(s+1)\wedge\tau_r}\big)$. Therefore,

$$\begin{aligned} &V'_{\sigma_{(s+1)\wedge\tau_r}}\big(X_{(s+1)\wedge\tau_r}\big) \leqslant \mu\lambda_\circ V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big) \quad \text{if } \sigma_{(s+1)\wedge\tau_r} \neq \sigma_{s\wedge\tau_r}, \quad \text{and}\\ &V'_{\sigma_{(s+1)\wedge\tau_r}}\big(X_{(s+1)\wedge\tau_r}\big) \leqslant \lambda_\circ V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big) \quad \text{otherwise.}\end{aligned}$$

Iterating this procedure we arrive at the pathwise inequality

$$V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big) \leqslant \mu^{N_{t\wedge\tau_r} - N_{s\wedge\tau_r}}\lambda_\circ^{t\wedge\tau_r - s\wedge\tau_r} V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big). \tag{19}$$

Since $s\wedge\tau_r = t\wedge s\wedge\tau_r$, and $t\wedge\tau_r$ is measurable with respect to $\mathfrak{F}_{t\wedge s\wedge\tau_r}$, we invoke the Markov property of $(\sigma_t, X_t)_{t\in\mathbb{N}_0}$ and take conditional expectations given

$(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})$ to arrive at

$$\mathsf{E}\Big[V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\Big|(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})\Big] \leqslant V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big)\lambda_\circ^{t\wedge\tau_r - s\wedge\tau_r}\mathsf{E}\Big[\mu^{N_{t\wedge\tau_r}-N_{s\wedge\tau_r}}\Big|(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})\Big]. \tag{20}$$

Since Assumption 10 holds, we apply the estimate in Lemma 17 to get

$$\mathsf{E}\Big[\mu^{N_{t\wedge\tau_r}-N_{s\wedge\tau_r}}\Big|(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})\Big] \leqslant \sum_{k=0}^{t\wedge\tau_r - s\wedge\tau_r}\binom{t\wedge\tau_r - s\wedge\tau_r}{k}\hat{p}^{(t\wedge\tau_r - s\wedge\tau_r - k)}\tilde{p}^k\mu^k = \big(\hat{p}+\mu\tilde{p}\big)^{t\wedge\tau_r - s\wedge\tau_r},$$

and this substituted in (20) gives

$$\mathsf{E}\Big[V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\Big|(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})\Big] \leqslant V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big)\big(\lambda_\circ(\hat{p}+\mu\tilde{p})\big)^{t\wedge\tau_r - s\wedge\tau_r}.$$

Since $\lambda_\circ(\hat{p}+\mu\tilde{p}) < 1$ by assumption, there exists $\alpha > 0$ such that $\alpha' := \lambda_\circ(\hat{p}+\mu\tilde{p})\mathrm{e}^\alpha < 1$, and the above inequality leads to

$$\mathsf{E}\Big[\mathrm{e}^{\alpha(t\wedge\tau_r - s\wedge\tau_r)}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\Big|(\sigma_{s\wedge\tau_r}, X_{s\wedge\tau_r})\Big] \leqslant V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big)(\alpha')^{t\wedge\tau_r - s\wedge\tau_r} \leqslant V'_{\sigma_{s\wedge\tau_r}}\big(X_{s\wedge\tau_r}\big). \tag{21}$$

This shows that the process $\big(\mathrm{e}^{\alpha(t\wedge\tau_r)}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big)_{t\in\mathbb{N}_0}$ is a nonnegative supermartingale.

Let us now look at the assertion of the Proposition under the condition (S2). Fix $t\in\mathbb{N}_0$. Then from (V3′),

$$V'_j\big(f_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big) \leqslant \lambda_{j\sigma_{t\wedge\tau_r}}V'_j\big(X_{t\wedge\tau_r}\big) \quad \forall\, j\in\mathcal{P},$$

and in particular, in view of (V2),

$$\begin{aligned} V'_{\sigma_{(t+1)\wedge\tau_r}}\big(f_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big) &\leqslant \lambda_{\sigma_{(t+1)\wedge\tau_r}\sigma_{t\wedge\tau_r}}V'_{\sigma_{(t+1)\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\\ &\leqslant \mu\lambda_{\sigma_{(t+1)\wedge\tau_r}\sigma_{t\wedge\tau_r}}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big). \end{aligned}$$

This leads to

$$\begin{aligned} \mathsf{E}\Big[V'_{\sigma_{(t+1)\wedge\tau_r}}\big(X_{(t+1)\wedge\tau_r}\big)\Big|(\sigma_{t\wedge\tau_r}, X_{t\wedge\tau_r})\Big] &\leqslant \mu\left(\sum_{j\in\mathcal{P}}p_{\sigma_{t\wedge\tau_r}j}\lambda_{j\sigma_{t\wedge\tau_r}}\right)V'_{\sigma_{t\wedge\tau_r}}(X_{t\wedge\tau_r})\\ &\leqslant \mu\left(\max_{i\in\mathcal{P}}\sum_{j\in\mathcal{P}}p_{ij}\lambda_{ji}\right)V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big). \end{aligned}$$

Since by hypothesis there exists $\alpha > 0$ such that $\mu\left(\max_{i\in\mathcal{P}}\sum_{j\in\mathcal{P}}p_{ij}\lambda_{ji}\right)\mathrm{e}^\alpha < 1$, the last inequality shows immediately that $\big(\mathrm{e}^{\alpha(t\wedge\tau_r)}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big)_{t\in\mathbb{N}_0}$ is a supermartingale. This concludes the proof. □

Remark 18. It is possible to modify the above proof of Proposition 13 to directly verify that the process $\big(\mathrm{e}^{\alpha(t\wedge\tau_r)}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big)_{t\in\mathbb{N}_0}$ is a uniformly integrable supermartingale. Indeed, minor modifications in the proof are needed to demonstrate that the process $\Big(\big(\mathrm{e}^{\alpha(t\wedge\tau_r)}V'_{\sigma_{t\wedge\tau_r}}\big(X_{t\wedge\tau_r}\big)\big)^{1+\delta}\Big)_{t\in\mathbb{N}_0}$ is a supermartingale for some $\delta > 0$ small enough; the key observation is that the maps $]0,\infty[\ \ni y \longmapsto \lambda_\circ^{1+y}\big(\hat{p}+\tilde{p}\mu^{1+y}\big) \in\ ]0,\infty[$ and $]0,\infty[\ \ni y \longmapsto \mu^{1+y}\Big(\sum_{j\in\mathcal{P}}p_{ij}\lambda_{ji}^{1+y}\Big) \in\ ]0,\infty[$ are continuous. ◁

Proof of Corollary 14. First observe that since each map f_i is locally Lipschitz, the diameter of the set $D_i := \{f_i(x) | x \in \bar{B}_r\}$ is finite, and since $\mathcal{P}$ is finite, so is the diameter of $\bigcup_{i\in\mathcal{P}} D_i$. Therefore, if Q is the transition kernel of the Markov process $(\sigma_t, X_t)_{t\in\mathbb{N}_0}$, then employing (V1) and the fact that f_i is locally Lipschitz for each i, we arrive at

$$\begin{aligned}\mathsf{E}\Bigl[V_{\sigma_1}(X_1)\mathbf{1}_{\{X_1\in\mathbb{R}^d\smallsetminus\bar{B}_r\}}\Big|(\sigma_0, X_0) = (i, x_0)\Bigr] &= \int_{\mathcal{P}\times(\mathbb{R}^d\smallsetminus\bar{B}_r)} Q\bigl((i, x_0), j\otimes \mathrm{d}y\bigr)V_j(y)\\ &= \sum_{j\in\mathcal{P}} p_{ij}\mathbf{1}_{\mathbb{R}^d\smallsetminus\bar{B}_r}(f_j(x_0))V_j(f_j(x_0)) \leqslant \sum_{j\in\mathcal{P}} p_{ij}\mathbf{1}_{\mathbb{R}^d\smallsetminus\bar{B}_r}(f_j(x_0))\alpha_2(\|f_j(x_0)\|)\\ &\leqslant \sum_{j\in\mathcal{P}} p_{ij}L\|x_0\| < Lr < \infty\end{aligned}$$

for $\|x_0\| < r$, where L is such that $\sup_{j\in\mathcal{P},y\in\bar{B}_r}\|f_j(y)\| \leqslant L\|y\|$. This shows that condition (ii) of Theorem 2 holds under our hypotheses, and by Proposition 13 we know that there exists $\alpha > 0$ such that $\Bigl(\mathrm{e}^{\alpha(t\wedge\tau_r)}V_{\sigma_{t\wedge\tau_r}}\bigl(X_{t\wedge\tau_r}\bigr)\mathbf{1}_{\mathbb{R}^d\smallsetminus\bar{B}_r}\bigl(X_{t\wedge\tau_r}\bigr)\Bigr)_{t\in\mathbb{N}_0}$ is a supermartingale. Theorem 2 now guarantees the existence of a constant $C' > 0$ such that $\sup_{t\in\mathbb{N}_0}\mathsf{E}\bigl[V_{\sigma_t}(X_t)\mathbf{1}_{\mathbb{R}^d\smallsetminus\bar{B}_r}(X_t)\bigr] \leqslant C'$, and finally, from (V1) it follows that there exists a constant $C > 0$ such that $\sup_{t\in\mathbb{N}_0}\mathsf{E}\bigl[\alpha_1(\|X_t\|)\bigr] \leqslant C < \infty$, as asserted. □

Proof of Corollary 16. We prove almost sure global asymptotic stability and α_1 stability in $\boldsymbol{L}_1$ of (18) under the condition (S1) of Proposition 13; the proofs under (S2) are similar.

First observe that since $\ker(f_i - \mathrm{id}) = \{0\}$ for each $i \in \mathcal{P}$, i.e., 0 is the equilibrium point of each individual subsystem, $\mathsf{P}_{x_0}\bigl(\tau_{\{0\}} < \infty\bigr) = 0$ for $x_0 \neq 0$, where $\tau_{\{0\}}$ is the first time that the process $(X_t)_{t\in\mathbb{N}_0}$ hits $\{0\}$. Indeed, since $\ker(f_i - \mathrm{id}) = \{0\}$ for each $i \in \mathcal{P}$ and $x_0 \neq 0$ we have

$$Q\bigl((i, x_0), \mathcal{P}\times\{0\}\bigr) = \sum_{j\in\mathcal{P}} p_{ij}\mathbf{1}_{\{0\}}(f_j(x_0)) = 0,$$

which shows that $Q^n\bigl((i, x_0), \mathcal{P}\times\{0\}\bigr) = 0$ whenever $x_0 \neq 0$. The observation now follows from $\mathsf{P}_{x_0}\bigl(\tau_{\{0\}} < \infty\bigr) = \mathsf{P}_{x_0}\bigl(\bigcup_{n\in\mathbb{N}}\{\tau_{\{0\}} = n\}\bigr) \leqslant \sum_{n\in\mathbb{N}}\mathsf{P}_{x_0}\bigl(\tau_{\{0\}} = n\bigr)$. Therefore, with $\tau_{\{0\}} = \tau_r = \infty$, proceeding as in the proof of Proposition 13 above, one can show that $\bigl(\mathrm{e}^{\alpha t}V_{\sigma_t}\bigl(X_t\bigr)\bigr)_{t\in\mathbb{N}_0}$ is a supermartingale for some $\alpha > 0$. In particular, With $s = 0$ and $\tau_r = \infty$ in (21), we apply (V1) to arrive at

$$\lim_{t\to\infty}\mathsf{E}\bigl[\mathrm{e}^{\alpha t}V_{\sigma_t}\bigl(X_t\bigr)\bigr] = \lim_{t\to\infty}\mathsf{E}\Bigl[\mathsf{E}\bigl[\mathrm{e}^{\alpha t}V_{\sigma_t}\bigl(X_t\bigr)\big|(\sigma_0, x_0)\bigr]\Bigr] \leqslant \lim_{t\to\infty}\alpha_2(\|x_0\|)(\alpha')^t = 0.$$

The standard supermartingale convergence theorem and the definition of $\tau_{\{0\}}$ imply that

$$\mathsf{P}\Bigl(\lim_{t\to\infty}V_{\sigma_t}(X_t) = 0\Bigr) = 1.$$

With $s = 0$ and $\tau_r = \tau_{\{0\}} = \infty$, the pathwise inequality (19) in conjunction with (V1) give $V_{\sigma_t}(X_t) \leqslant \alpha_2(\|x_0\|)\mu^{N_t}\lambda_\circ^t$. The foregoing inequality implies that for almost every sample path $(\sigma_t, X'_t)_{t\in\mathbb{N}_0}$ corresponding to initial condition $X_0 = x'_0$ with $\|x'_0\| < \|x_0\|$, one has

$$\lim_{t\to\infty}V_{\sigma_t}(X'_t) \leqslant \lim_{t\to\infty}\alpha_2(\|x'_0\|)\mu^{N_t}\lambda_\circ^t \leqslant \lim_{t\to\infty}\alpha_2(\|x_0\|)\mu^{N_t}\lambda_\circ^t = 0,$$

which proves (AS2). Since the family $\{f_i\}_{i\in\mathcal{P}}$ is finite, and each f_i is locally Lipschitz, there exists $L > 0$ such that $\sup_{i\in\mathcal{P}}\|f_i(x)\| \leqslant L\|x\|$ whenever $\|x\| \leqslant 1$. Fix $\varepsilon > 0$. By (AS2) we know that for almost all sample paths there exists a constant $T > 0$ such that $\sup_{t\geqslant T}\|X_t\| < \varepsilon$ whenever $\|x_0\| < 1$. Then the choice of $\delta = \bigl(\varepsilon L^{-T}\bigr)\wedge 1$ immediately gives us the (AS1) property.

It remains to verify (SM1) and (SM2). Both the properties follow from (21) in the proof of Proposition 13, with $s = 0$ and $\tau_r = \tau_{\{0\}} = 0$. Indeed, with these values of s and τ_r, (21) becomes

$$\begin{aligned}\mathsf{E}\big[\mathrm{e}^{\alpha t}\alpha_1(\|X_t\|)\big|(\sigma_0, X_0)\big] &\leqslant \mathsf{E}\big[\mathrm{e}^{\alpha t}V_{\sigma_t}(X_t)\big|(\sigma_0, X_0)\big] \leqslant V_{\sigma_0}(X_0)(\alpha')^t\\ &\leqslant \alpha_2(\|x_0\|)(\alpha')^t\end{aligned}$$

in view of (V1), where $\alpha' = \lambda_\circ(\hat{p} + \mu\tilde{p})\mathrm{e}^{\alpha} < 1$. Therefore, given $\varepsilon > 0$, we simply choose $\delta < \alpha_2^{-1}(\varepsilon)$ to get (SM1). Given $r, \varepsilon' > 0$, we simply choose $T = 0 \vee \big(\ln(\alpha_2(r)/\varepsilon')/\ln(\alpha')\big)$ to get (SM2). This completes the proof. □

3.2. **Robust Stability of Discrete-Time Randomly Switched Systems.** Conditions for the existence of the supermartingale $\big(\mathrm{e}^{\alpha(t\wedge\tau_K)}V\big(X_{t\wedge\tau_K}\big)\big)_{t\in\mathbb{N}_0}$ in §2 can be easily expressed in terms of the transition kernel Q. However, if Q is not known exactly, which may happen if the model of the underlying system generating the Markov process $(X_t)_{t\in\mathbb{N}_0}$ is uncertain, one needs different methods. We look at one such instance below.

Consider the system

$$X_{t+1} = f_{\sigma_t}(X_t, w_t), \qquad X_0 = x_0, \quad t \in \mathbb{N}_0, \tag{22}$$

where we retain the definition σ from §3.1, $f_i : \mathbb{R}^d \times \mathbb{R}^m \longrightarrow \mathbb{R}^d$ is locally Lipschitz continuous in both arguments with $f_i(0, 0) = 0$ for each $i \in \mathcal{P}$, and $(w_t)_{t\in\mathbb{N}_0}$ is a bounded and measurable $\mathbb{R}^m$-valued disturbance sequence. We do not model $(w_t)_{t\in\mathbb{N}_0}$ as a random process. As such, the transition kernel of (22) is not unique.

Definition 19. The system (22) is said to be *input-to-state stable in* $\boldsymbol{L}_1$ if there exist functions $\chi, \chi' \in \mathcal{K}_\infty$ and $\psi \in \mathcal{KL}$ such that

$$\mathsf{E}_{x_0}\big[\chi(\|X_t\|)\big] \leqslant \psi(\|x_0\|, t) + \sup_{s\in\mathbb{N}_0}\chi'(\|w_s\|)$$

for all $t \in \mathbb{N}_0$. ◇

Intuitively, the system (22) is input-to-state (ISS) stable in $\boldsymbol{L}_1$ if the $\boldsymbol{L}_1$ norm of the state is asymptotically bounded by a function of the size of the noise, and is uniformly bounded in expectation. Our motivation for this definition comes from the concept of ISS in the deterministic context [JW01]. Consider the i-th subsystem of (22) $x_{t+1} = f_i(x_t, w_t)$ for $t \in \mathbb{N}_0$ with initial condition x_0; note that $(x_t)_{t\in\mathbb{N}_0}$ is a deterministic sequence. This nonlinear discrete-time system is said to be ISS if there exist functions $\psi \in \mathcal{KL}$ and $\chi \in \mathcal{K}_\infty$ such that $\|x_t\| \leqslant \psi(\|x_0\|, t) + \sup_{s\in\mathbb{N}_0}\chi(\|w_s\|)$ for $t \in \mathbb{N}_0$. A sufficient set of conditions (cf. [JW01, Lemma 3.5]) for ISS of this system is that there exist a continuous function $V : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$, $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$, $\rho \in \mathcal{K}$, and a constant $\lambda \in \,]0, 1[$, such that $\alpha_1(\|x\|) \leqslant V(x) \leqslant \alpha_2(\|x\|)$ for all $x \in \mathbb{R}^d$, and $V(f_i(x, w)) \leqslant \lambda V(x)$ whenever $\|x\| > \rho(\|w\|)$.

In this framework we have the following Proposition; observe that its hypotheses require that the subsystems of (22) are sufficiently nice, in the sense that under zero-disturbance they reduce to Assumption 12 with $r = 0$ (since $f_i(0, 0) = 0$).

Proposition 20. *Consider the system* (22), *and suppose that*

(i) *Assumption* 10 *holds,*
(ii) *there exist continuous functions* $V_i : \mathbb{R}^d \longrightarrow \mathbb{R}_{\geqslant 0}$ *for* $i \in \mathcal{P}$, $\alpha_1, \alpha_2, \rho \in \mathcal{K}_\infty$, *a constant* $\mu > 1$ *and a matrix* $[\lambda_{ij}]_{\mathrm{N}\times\mathrm{N}}$ *of nonnegative entries, such that*
  (a) $\alpha_1(\|x\|) \leqslant V_i(x) \leqslant \alpha_2(\|x\|)$ *for all* x *and* i,
  (b) $V_i(x) \leqslant \mu V_j(x)$ *for all* x *and* i, j, *and*
  (c) $V_i(f_j(x)) \leqslant \lambda_{ij}V_i(x)$ *whenever* $\|x\| > \rho(\|w\|)$ *and all* i, j,
(iii) $\mu\Big(\max_{i\in\mathcal{P}}\sum_{j\in\mathcal{P}} p_{ij}\lambda_{ji}\Big) < 1$.

Then (22) *is input-to-state stable in* $\boldsymbol{L}_1$ *in the sense of Definition* 19*.*

Proof. We define the compact set $K := \{(i, y) \in \mathcal{P} \times \mathbb{R}^d \big| \|y\| \leqslant \sup_{s\in\mathbb{N}_0} \rho(\|w_s\|)\}$, and let $\tau_K := \inf\{t \in \mathbb{N}_0 \big| X_t \in K\}$. In this setting we know from the preceding analysis that $\varphi(t, \xi) = \mathrm{e}^{\alpha t}\xi$, $\theta(t) = \mathrm{e}^{-\alpha t}$, and $C = 1/(1 - \mathrm{e}^{-\alpha})$. We see from the estimate (5) in the proof of Theorem 2 that

$$\begin{aligned}\mathsf{E}_{x_0}\big[V_{\sigma_t}(X_t)\big] &\leqslant \varphi(0, V_{\sigma_0}(x_0))\theta(t) + \frac{\beta}{1 - \mathrm{e}^{-\alpha}} + \delta \\ &\leqslant \alpha_2(\|x_0\|)\mathrm{e}^{-\alpha t} + \frac{\beta}{1 - \mathrm{e}^{-\alpha}} + \delta.\end{aligned}$$

Standard arguments show that there exists some $\chi'' \in \mathcal{K}_\infty$ such that β and δ are each dominated by $\chi''\big(\sup_{s\in\mathbb{N}_0}\|w_s\|\big)$, and therefore, there exists some $\chi' \in \mathcal{K}_\infty$ such that $\beta/(1 - \mathrm{e}^{-\alpha}) + \delta$ is dominated by $\chi'\big(\sup_{s\in\mathbb{N}_0}\|w_s\|\big)$. Applying (ii)(a) on the left-hand side of the last inequality, we conclude that (22) is input-to-state stable with $\chi = \alpha_1$ and $\psi(r, t) = \alpha_2(r)\mathrm{e}^{-\alpha t}$. □

Acknowledgments

The authors thank Daniel Liberzon and John Lygeros for helpful comments, and Andreas Milias-Argeitis for useful discussions related to the chemical master equation.

References

[ACCL08] M. Agarwal, E. Cinquemani, D. Chatterjee, and J. Lygeros, *On convexity of stochastic optimization problems with constraints*, Submitted, 2008.

[AS92] J. M. Alden and R. L. Smith, *Rolling horizon procedures in nonhomogeneous Markov decision processes*, Operations Research **40** (1992), no. suppl. 2, S183–S194.

[ACK08] D. F. Anderson, G. Craciun, and T. G. Kurtz, *Product-form stationary distributions for deficiency zero chemical reaction networks*, http://arxiv.org/abs/0803.3042, 2008.

[ABFG+93] A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh, and S. I. Marcus, *Discrete-time controlled Markov processes with average cost criterion: a survey*, SIAM Journal on Control and Optimization **31** (1993), no. 2, 282–344.

[BDEG88] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo, *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities*, Annales de l'Institut Henri Poincaré. Probabilités et Statistique **24** (1988), no. 3, 367–394, Erratum in ibid., **24** (1989), no. 4, 589–590.

[BS78] D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: the Discrete-Time Case*, Mathematics in Science and Engineering, vol. 139, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1978.

[Bor91] V. S. Borkar, *Topics in Controlled Markov Chains*, Pitman Research Notes in Mathematics Series, vol. 240, Longman Scientific & Technical, Harlow, 1991.

[BKR+01] A. Borodin, J. Kleinberg, P. Raghavan, M. Sudan, and D. P. Williamson, *Adversarial queuing theory*, Journal of the ACM **48** (2001), no. 1, 13–38.

[CCK07] M. Cannon, P. Couchman, and B. Kouvaritakis, *MPC for stochastic systems*, Assessment and future directions of nonlinear model predictive control, Lecture Notes in Control and Information Sciences, vol. 358, Springer, Berlin, 2007, pp. 255–268.

[CCCL08] D. Chatterjee, E. Cinquemani, G. Chaloulos, and J. Lygeros, *Stochastic control up to a hitting time: optimality and rolling-horizon implementation*, http://arxiv.org/abs/0806.3008, 2008.

[CFM05] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov Jump Linear Systems*, Probability and its Applications (New York), Springer-Verlag, London, 2005.

[DF99] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Review **41** (1999), no. 1, 45–76 (electronic).

[FK04] S. Foss and T. Konstantopoulos, *An overview of some stochastic stability methods*, Journal of Operations Research Society of Japan **47** (2004), no. 4, 275–303.

[HLR96] J. Håstad, T. Leighton, and B. Rogoff, *Analysis of backoff protocols for multiple access channels*, SIAM Journal on Computing **25** (1996), no. 4, 740–774.

[HLL90] O. Hernández-Lerma and J. B. Lasserre, *Error bounds for rolling horizon policies in discrete-time Markov control processes*, IEEE Transactions on Automatic Control **35** (1990), no. 10, 1118–1124.

[HLL93] ———, *Value iteration and rolling plans for Markov control processes with unbounded rewards*, Journal of Mathematical Analysis and Applications **177** (1993), no. 1, 38–55.

[HLL96] ———, *Discrete-Time Markov Control Processes: Basic Optimality Criteria*, Applications of Mathematics, vol. 30, Springer-Verlag, New York, 1996.

[HLL99] ———, *Further Topics on Discrete-Time Markov Control Processes*, Applications of Mathematics, vol. 42, Springer-Verlag, New York, 1999.

[Hob98] D. G. Hobson, *Volatility misspecification, option pricing and superreplication via coupling*, The Annals of Applied Probability **8** (1998) no. 1, 193–205.

[JH07] T. Jahnke and W. Huisinga, *Solving the chemical master equation for monomolecular reaction systems analytically*, Journal of Mathematical Biology **54** (2007), no. 1, 1–26.

[JT01] S. F. Jarner and R. L. Tweedie, *Locally contracting iterated functions and stability of Markov chains*, Journal of Applied Probability **38** (2001), no. 2, 494–507.

[JW01] Z-P. Jiang and Y. Wang, *Input-to-state stability for discrete-time nonlinear systems*, Automatica **37** (2001), no. 6, 857–869.

[KS08] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2 ed., Graduate Texts in mathematics, Springer, 2008.

[Kal02] O. Kallenberg, *Foundations of Modern Probability*, 2 ed., Probability and Its Applications, Springer Verlag, 2002.

[Kif86] Y. Kifer, *Ergodic Theory of Random Transformations*, Progress in Probability and Statistics, vol. 10, Birkhäuser Boston Inc., Boston, MA, 1986.

[Kus71] H. Kushner, *Introduction to Stochastic Control*, Holt, Rinehart and Winston, Inc., New York, 1971.

[Lib03] D. Liberzon, *Switching in Systems and Control*, Systems & Control: Foundations & Applications, Birkhäuser, Boston, 2003.

[LM94] A. Lasota and M. C. Mackey, *Chaos, Fractals, and Noise*, 2 ed., Applied Mathematical Sciences, vol. 97, Springer-Verlag, New York, 1994.

[LM02] A. Lasota and J. Myjak, *On a dimension of measures*, Polish Academy of Sciences. Bulletin. Mathematics **50** (2002), no. 2, 221–235.

[LS04] A. Lasota and T. Szarek, *Dimension of measures invariant with respect to the Ważewska partial differential equation*, Journal of Differential Equations **196** (2004), no. 2, 448–465.

[LB07] L. Lorenzi and M. Bertoldi, *Analytical Methods for Markov Semigroups*, Pure and Applied Mathematics, vol. 283, Chapman & Hall/CRC, Boca Raton, FL, 2007.

[Mac01] J. M. Maciejowski, *Predictive Control with Constraints*, Prentice Hall, 2001.

[Mey08] S. P. Meyn, *Control Techniques for Complex Networks*, Cambridge University Press, Cambridge, 2008.

[MT92] S. P. Meyn and R. L. Tweedie, *Stability of Markovian processes I—criteria for discrete-time chains*, Advances in Applied Probability **24** (1992), no. 3, 542–574.

[MT93] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer-Verlag, London, 1993.

[MA08] A. Milias-Argeitis, *Fast simulation of the chemical master equation with iterated function systems*, In preparation, 2008.

[PR99] R. Pemantle and J. S. Rosenthal, *Moment conditions for a sequence with negative drift to be uniformly bounded in* L^r, Stochastic Processes and their Applications **82** (1999), no. 1, 143–155.

[Piu97] A. B. Piunovskiy, *Optimal Control of Random Sequences in Problems with Constraints*, Mathematics and its Applications, vol. 410, Kluwer Academic Publishers, Dordrecht, 1997, With a preface by V. B. Kolmanovskii and A. N. Shiryaev.

[RY99] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3 ed., Grundlehren der Mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin, 1999.

[Sza06] T. Szarek, *Feller processes on nonlocally compact spaces*, The Annals of Probability **34** (2006), no. 5, 1849–1863.

[vS00] A. van der Schaft and H. Schumacher, *An Introduction to Hybrid Dynamical Systems*, Lecture Notes in Control and Information Sciences, vol. 251, Springer-Verlag London Ltd., London, 2000.

[Wer05] I. Werner, *Contractive Markov systems*, Journal of the London Mathematical Society. Second Series **71** (2005), no. 1, 236–258.

[Wil06] D. J. Wilkinson, *Stochastic Modelling for Systems Biology*, Chapman & Hall/CRC Mathematical and Computational Biology Series, Chapman & Hall/CRC, Boca Raton, FL, 2006.

ETL I19, Physikstrasse 3, ETH Zürich, 8092 Zürich, Switzerland, Ph. +41-44-632-2326, Fax: +41-44-632-1211

E-mail address: chatterjee@control.ee.ethz.ch

C-547 Padelford Hall, Department of Mathematics, University of Washington, Seattle, WA 98195, Ph. +1-206-543-7832

E-mail address: soumik@math.washington.edu