

Contraction Analysis of Discrete-Time Stochastic Systems

Yu Kawano , *Member, IEEE*, and Yohei Hosoe , *Member, IEEE*

Abstract—In this article, we develop a novel contraction framework for the stability analysis of discrete-time nonlinear systems with parameters following stochastic processes. For general stochastic processes, we first provide a sufficient condition for uniform incremental exponential stability (UIES) in the first moment with respect to a Riemannian metric. Then, focusing on the Euclidean distance, we present a necessary and sufficient condition for UIES in the second moment. By virtue of studying general stochastic processes, we can readily derive UIES conditions for special classes of processes, e.g., independent and identically distributed processes and Markov processes, which are demonstrated as selected applications of our results.

Index Terms—Contraction, discrete-time systems, incremental stability, nonlinear systems, stochastic systems.

I. INTRODUCTION

STARTING with a seminal paper [1], contraction theory draws attention from the systems and control community as a new differential geometric framework for the stability analysis of nonlinear systems. Differently from the standard Lyapunov analysis of an equilibrium point (e.g., [2] and [3]), incremental stability (i.e., stability of a pair of trajectories) [4] is studied by lifting the Lyapunov function to the tangent bundle [5]. Revisiting nonlinear control theory from this new angle has resulted in so-called differential approaches to, for instance, control design [6], [7], [8], [9], [10], observer design [11], [12], [13], dissipativity theory [14], [15], [16], [17], and balancing theory [18], [19]. Along with them, contraction (stability) analysis itself is in the middle of development in various problem settings; see, e.g., [20], [21], [22], and [23] for monotone systems, e.g., [24] and [25] for switched systems, e.g., [26] and [27] for systems under stochastic input noise, and, e.g., [28] for stochastic switched impulsive systems, a kind of Markov jump systems.

Manuscript received 24 August 2022; revised 7 May 2023; accepted 27 May 2023. Date of publication 7 June 2023; date of current version 30 January 2024. The work of Yu Kawano was supported in part by the JSPS KAKENHI under Grant JP21K14185 and Grant JP21H04875. The work of Yohei Hosoe was supported in part by the JST PRESTO under Grant JPMJPR2127 and in part by the JSPS KAKENHI under Grant JP20K04546 and Grant JP23H01433. Recommended by Associate Editor P. G. Mehta. (*Corresponding author: Yu Kawano.*)

Yu Kawano is with the Graduate School of Advanced Science and Engineering, Hiroshima University, Higashi-Hiroshima 739-0046, Japan (e-mail: ykawano@hiroshima-u.ac.jp).

Yohei Hosoe is with the Department of Electrical Engineering, Kyoto University, Kyoto 615-8510, Japan (e-mail: hosoe@kuee.kyoto-u.ac.jp).

Digital Object Identifier 10.1109/TAC.2023.3283678

In this article, we aim at newly developing contraction theory for discrete-time nonlinear systems with parameters following stochastic processes. None of the aforementioned papers deals with this class of systems; most of them focus on continuous-time deterministic systems. The aforementioned papers [26], [27], [28] studying stochastic systems are also for continuous-time systems. In the discrete-time case, the authors in [1], [29], [30], and [31] have studied deterministic systems and systems under stochastic input noise, respectively. Other than input noise, randomness is not incorporated in the contraction analysis of discrete-time systems. In other words, there is no contraction framework to analyze discrete-time systems with random parameters, such as the systems with white parameters [32] and Markov jump systems [33]. This is in contrast to a massive amount of researches on discrete-time Markov jump linear/nonlinear systems in the history, e.g., [33], [34], [35], [36] and recent rapid increase in the number of researches for machine learning to construct stochastic models from discrete-time empirical data, e.g., [37], [38], [39], [40]. When studying stochastic systems, typically, we specify the class of stochastic processes into, for instance, independent and identically distributed (i.i.d.) and Markovian, which can be viewed as ad hoc approaches because depending on processes, different stability conditions are obtained. For developing unified theory to deal with each process simultaneously, recently, the paper [41] gives second moment stability conditions for general stochastic processes in the discrete-time linear case, which contains the existing conditions for i.i.d. [42], [43] and Markovian [33], [44] as special cases.

Inspired by [41], in this article, we deal with general stochastic processes. To begin with contraction analysis of discrete-time nonlinear stochastic systems, we introduce a new stability notion, uniform incremental exponential stability (UIES), in the p th moment with respect to the Riemannian metric, which reduces to the standard p th moment stability [45], [46] when the distance is Euclidean, and a trajectory is fixed on an equilibrium point. As the first main result of this article, we provide a sufficient condition for UIES in the first moment. Then, as the second main contribution, focusing on the Euclidean distance, we present a necessary and sufficient condition for UIES in the second moment; second moment stability is stronger than first moment stability. By virtue of developing unified theory for general stochastic processes, we show that specifying processes readily yields UIES conditions for i.i.d. processes or Markov processes. Even UIES conditions in each specialized case are new contributions of this article on their own, due to lack of contraction theory for discrete-time stochastic systems.

The rest of this article is organized as follows. To explain main ideas of this article, we start with analysis of scalar systems in Section II. Then, we consider n -dimensional systems with general stochastic processes and present UIES conditions in Section III. A proposed condition is tailored to i.i.d. processes and Markov processes as special applications in Section IV. Finally, Section V concludes this article.

Notations: The sets of real numbers and integers are denoted by \mathbb{R} and \mathbb{Z} , respectively. Also, define $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Subsets of \mathbb{Z} are defined by $\mathbb{Z}_{k_0+} := \mathbb{Z} \cap [k_0, \infty)$ and $\mathbb{Z}_{k_0-} := \mathbb{Z} \cap (-\infty, k_0]$ for $k_0 \in \mathbb{Z}$. Also, $\mathbb{Z}_{[k_0, k]} := \mathbb{Z} \cap [k_0, k]$ for $k_0 \in \mathbb{Z}$ and $k \in \mathbb{Z}_{k_0+}$, where $\mathbb{Z}_{[k_0, k_0]} := \{k_0\}$. The $n \times n$ identity matrix is denoted by I_n . The set of $n \times n$ symmetric matrices is denoted by $\mathbb{S}^{n \times n}$, and that of symmetric and positive (respectively, semi) definite matrices is denoted by $\mathbb{S}_{>0}^{n \times n}$ (respectively, $\mathbb{S}_{\geq 0}^{n \times n}$). For $P, Q \in \mathbb{R}^{n \times n}$, $P \succ Q$ (respectively, $P \succeq Q$) means $P - Q \in \mathbb{S}_{>0}^{n \times n}$ (respectively, $P - Q \in \mathbb{S}_{\geq 0}^{n \times n}$). The Euclidean norm of a vector $x \in \mathbb{R}^n$ is denoted by $|x|$. A function $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be a distance if the following holds:

- 1) $d(x, x') \geq 0$ for all $x, x' \in \mathbb{R}^n$;
- 2) $d(x, x') = 0$ if and only if $x = x'$ for all $x, x' \in \mathbb{R}^n$;
- 3) $d(x, x'') \leq d(x, x') + d(x', x'')$ for all $x, x', x'' \in \mathbb{R}^n$.

An example of a distance is the Euclidean distance $d(x, x') = |x - x'|$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, where Ω , \mathcal{F} , and \mathbb{P} denote a sample space, a σ -algebra, and a probability measure, respectively. For the sake of notational simplicity, an \mathcal{X} -valued random variable $\xi_0: (\Omega, \mathcal{F}) \rightarrow (\mathcal{X}, \mathcal{B}(\mathcal{X}))$ is described by $\xi_0: \Omega \rightarrow \mathcal{X}$, where $\mathcal{B}(\mathcal{X})$ denotes the Borel σ -algebra on \mathcal{X} . An $\mathcal{X}^{\mathcal{T}}$ -valued stochastic process $\xi := (\xi_k)_{k \in \mathcal{T}}$ on $\mathcal{T} \subset \mathbb{Z}$ is defined as a mapping $\xi: \Omega \times \mathcal{F} \rightarrow \mathcal{X}^{\mathcal{T}} \times \mathcal{B}(\mathcal{X}^{\mathcal{T}})$. For some $k_0 \in \mathbb{Z}$, subsequences of a stochastic process $\xi: \Omega \rightarrow \mathcal{X}^{\mathbb{Z}}$ are denoted by $\xi^{k_0+} = (\xi_k)_{k \in \mathbb{Z}_{k_0+}}: \Omega \rightarrow \mathcal{X}^{\mathbb{Z}_{k_0+}}$ and $\xi^{k_0-} = (\xi_k)_{k \in \mathbb{Z}_{k_0-}}: \Omega \rightarrow \mathcal{X}^{\mathbb{Z}_{k_0-}}$. The support of $\xi^{(k_0-1)-}$ is denoted by $\hat{\Xi}^{(k_0-1)-}$. Let \mathcal{F}_k be the σ -algebra generated by a subsequence $\xi_{k_0}, \xi_{k_0+1}, \dots, \xi_k$ of a stochastic process ξ under the initial condition $\xi^{(k_0-1)-} = \hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$. Then, $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ for each $\hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$, namely, an increasing family $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ of sub- σ -algebras of \mathcal{F} . For a stochastic process ξ , the time-shift operator $S_k: \mathcal{X}^{\mathbb{Z}_{k+}} \rightarrow \mathcal{X}^{\mathbb{Z}_{0+}}$ such that $\zeta^{0+} = S_k \xi^{k+}$ is defined by $\zeta_0 = \xi_k$ and $\zeta_1 = \xi_{k+1}, \dots$, where ζ_0 is \mathcal{F}_k -measurable. Let $\xi^{(k_0-1)-} = \hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$. Then, the conditional expectation of a function of ξ^{k_0+} given $\xi^{(k_0-1)-}$ is denoted by $\mathbb{E}_0[(\cdot)] := \mathbb{E}[(\cdot) | \xi^{(k_0-1)-} = \hat{\xi}^{(k_0-1)-}]$, which is also called the first moment of a function. Accordingly, its p th moment is defined by $\mathbb{E}_0[(\cdot)^p]$. The conditional expectation of a function of ξ^{k_0+} given \mathcal{F}_k is denoted by $\mathbb{E}_0[(\cdot) | \mathcal{F}_k]$. Since $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ is a filtration, the conditional expectation satisfies $\mathbb{E}_0[\mathbb{E}_0[(\cdot) | \mathcal{F}_{k_2}] | \mathcal{F}_{k_1}] = \mathbb{E}_0[(\cdot) | \mathcal{F}_{k_1}]$ for each $k_1 \in \mathbb{Z}_{k_0+}$ and every $k_2 \in \mathbb{Z}_{k_1+}$.

II. SCALAR CASES FOR BRINGING MAIN IDEAS

A. System Description

In this section, we explain the main ideas of this article through analysis of scalar systems. Let $\xi = (\xi_k)_{k \in \mathbb{Z}}: \Omega \rightarrow \mathbb{R}^{\mathbb{Z}}$

be a stochastic process. Consider the following system:

$$x_{k+1} = f(x_k, \xi_k), \quad k \in \mathbb{Z}_{k_0+} \quad (1)$$

where $f(x, \eta)$ and $\partial f(x, \eta)/\partial x$ are continuous with respect to $x \in \mathbb{R}$ at each $\eta \in \mathbb{R}$, and $f(x, \xi_k)$ and $\partial f(x, \xi_k)/\partial x$ are \mathcal{F}_k -measurable at each $x \in \mathbb{R}$.

In this article, the initial time $k_0 \in \mathbb{Z}$ and the initial state $x_{k_0} \in \mathbb{R}$ are both deterministic. Note that $(k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}$ is not enough information to describe the behavior of the system when dealing with a general stochastic process ξ . The behavior of ξ itself depends on its initial condition $\xi^{(k_0-1)-} = \hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$, which is the observation of an infinite past sequence up to $k_0 - 1$ of ξ . In summary, the deterministic triplet $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R} \times \hat{\Xi}^{(k_0-1)-}$ is the initial condition of the system (1).

Various stochastic systems can be described in the form of (1). For instance, we can deal with a system under an i.i.d. noise by specifying ξ_k into i.i.d. Another example is a Markov jump system by specifying ξ_k in a Markovian switching signal. These are main subclasses of stochastic systems studied in systems and control. Stability theory for each class has been developed separately, mostly under the linearity assumption. In this article, we take a different approach for developing stochastic contraction theory. We first derive contraction conditions for a general stochastic system and then specify problem structures in order to obtain conditions tailored to each problem. Taking this approach, we can avoid developing contraction theory for each case separately.

B. Stability Notion

In this article, we study stability of a pair of trajectories. To make the considered property clearer, we use the auxiliary system consisting of copies of the system (1)

$$\begin{cases} x'_{k+1} = f(x'_k, \xi_k) \\ x''_{k+1} = f(x''_k, \xi_k) \end{cases} \quad k \in \mathbb{Z}_{k_0+} \quad (2)$$

where the initial condition $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R} \times \mathbb{R}) \times \hat{\Xi}^{(k_0-1)-}$ is again deterministic. For the auxiliary system, the initial states x'_{k_0} and x''_{k_0} are allowed to be different, but the others, including system vector field f and stochastic process ξ , are the same.

We are interested in the following stability property. Given $p \in \mathbb{Z}_{1+}$, there exist $a > 0$ and $\lambda \in (0, 1)$ such that

$$\mathbb{E}_0[d^p(x'_k, x''_k)] \leq a^p \lambda^{p(k-k_0)} d^p(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+} \quad (3)$$

for all $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R} \times \mathbb{R}) \times \hat{\Xi}^{(k_0-1)-}$, where d is a distance function mentioned in the notation part. Note that the initial condition (x'_{k_0}, x''_{k_0}) is deterministic, and thus, the right-hand side is deterministic.

We call (3) *UIES in the p th moment* (with respect to d). This is an integrated concept of 1) moment stability for stochastic systems and 2) UIES (with respect to d) for the contraction analysis of nonlinear deterministic systems. For stochastic systems, moment stability is defined with respect to the Euclidean distance by assuming that the origin is an equilibrium point. Specifying $d(x', x'') = |x' - x''|$ and $(x'_k, x''_k) = (x_k, 0)$, $k \in \mathbb{Z}_{k_0+}$, in (3) yields

$$\mathbb{E}_0[|x_k|^p] \leq a^p \lambda^{p(k-k_0)} |x_{k_0}|^p \quad \forall k \in \mathbb{Z}_{k_0+}. \quad (4)$$

The left-hand side is nothing but the p th moment of $|x_k|$, and thus, this property is called exponential stability (ES) in the p th moment. When $p = 2$, this property is also referred to as mean square (exponential) stability [45], [46]. A similar concept is mean stability [46], which studies $\mathbb{E}_0[x_k]$ (without taking the Euclidean norm) instead of the first moment $\mathbb{E}_0[|x_k|]$.

Next, we consider the deterministic case, i.e., f is independent of ξ_k . Then, $d^p(x'_k, x''_k)$ in the left-hand side becomes deterministic. Taking the p th roots of both sides yields

$$d(x'_k, x''_k) \leq a\lambda^{k-k_0} d(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+}. \quad (5)$$

This is a discrete-time version of UIES (with respect to d); see, e.g., [5, Def. 1] for the definition in the continuous-time case.

In contraction analysis, it is standard to employ the distance function d induced by a Riemannian metric (or more generally Finsler metric [5]). Let $\Gamma(x', x'')$ denote the collection of class C^1 paths $\gamma: [0, 1] \rightarrow \mathbb{R}$ such that $\gamma(0) = x'$ and $\gamma(1) = x''$. Then, the following function induced by a class C^1 Riemannian metric $r(x) > 0$, $x \in \mathbb{R}$, is a distance function:

$$d_r(x', x'') = \inf_{\gamma \in \Gamma(x', x'')} \int_0^1 \sqrt{r(\gamma(s)) \left(\frac{d\gamma(s)}{ds} \right)^2} ds. \quad (6)$$

An intuitive interpretation is that we consider a kind of the weighted Euclidean distance with a weight $r(x)$ depending on x . When r is constant, this becomes a weighted Euclidean distance, $d_r(x', x'') = \sqrt{r}|x' - x''|$. As found in the Appendix, the essence of the proof is the same for the general Riemannian metric and the Euclidean distance, and thus, we consider the distance (6).

C. Main Ideas for Stochastic Contraction Analysis

The main idea of contraction analysis is to study the time-evolutions of a path $\gamma(s)$ and its derivative $d\gamma(s)/ds$. This can be proceeded via analysis of the prolongation of the system (1)

$$\begin{cases} x_{k+1} = f(x_k, \xi_k) \\ \delta x_{k+1} = \frac{\partial f(x_k, \xi_k)}{\partial x} \delta x_k \end{cases} \quad k \in \mathbb{Z}_{k_0+} \quad (7)$$

where the second system is called a variational system whose state is δx_k . Again, the initial condition $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R} \times \mathbb{R}) \times \hat{\Xi}^{(k_0-1)-}$ is deterministic.

Let $(\phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}))_{k \in \mathbb{Z}_{k_0+}}$, or simply $(\phi_k(\xi^{(k-1)-}; x_{k_0}))_{k \in \mathbb{Z}_{k_0+}}$ denote a stochastic process $(x_k)_{k \in \mathbb{Z}_{k_0+}}$ generated by the system (1) under the initial condition $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R} \times \hat{\Xi}^{(k_0-1)-}$. Then, it follows that

$$\begin{aligned} \phi_{k_0}(\xi^{(k_0-1)-}; x_{k_0}) &= x_{k_0} \\ \phi_{k_0+1}(\xi^{k_0-}; x_{k_0}) &= f(x_{k_0}, \xi_{k_0}), \dots \end{aligned}$$

We consider the time-evolution of a path $\gamma \in \Gamma(x'_{k_0}, x''_{k_0})$, namely, $\phi_k(\xi^{(k-1)-}; \gamma(s))$. This is a class C^1 path connecting two stochastic processes, as confirmed by

$$x'_k = \phi_k(\xi^{(k-1)-}; x'_{k_0}) = \phi_k(\xi^{(k-1)-}; \gamma(0)) \quad (8a)$$

$$x''_k = \phi_k(\xi^{(k-1)-}; x''_{k_0}) = \phi_k(\xi^{(k-1)-}; \gamma(1)). \quad (8b)$$

To evaluate the distance (6), we also consider the partial derivative of $\phi_k(\xi^{(k-1)-}; \gamma(s))$ with respect to s , namely, $\partial \phi_k(\xi^{(k-1)-}; \gamma(s)) / \partial s$. This satisfies, from the chain rule

$$\begin{aligned} & \frac{\partial \phi_{k+1}(\xi^{k-}; \gamma(s))}{\partial s} \\ &= \frac{\partial f(\phi_k(\xi^{(k-1)-}; \gamma(s)), \xi_k)}{\partial s} \\ &= \frac{\partial f(\phi_k(\xi^{(k-1)-}; \gamma(s)), \xi_k)}{\partial \phi_k} \frac{\partial \phi_k(\xi^{(k-1)-}; \gamma(s))}{\partial s} \end{aligned} \quad (9)$$

where

$$\frac{\partial f(\phi_k(\xi^{(k-1)-}; \gamma(s)), \xi_k)}{\partial \phi_k} := \left. \frac{\partial f(x_k, \xi_k)}{\partial x} \right|_{x_k = \phi_k(\xi^{(k-1)-}; \gamma(s))}.$$

Throughout this article, we use this kind of notations for partial derivatives. Equation (9) implies that

$$(x_k, \delta x_k) = \left(\phi_k(\xi^{(k-1)-}; \gamma(s)), \frac{\partial \phi_k(\xi^{(k-1)-}; \gamma(s))}{\partial s} \right) \quad k \in \mathbb{Z}_{k_0+}, s \in [0, 1] \quad (10)$$

satisfies the dynamics of the prolonged system (7) for each $\gamma \in \Gamma(x'_{k_0}, x''_{k_0})$ and every initial condition $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R} \times \mathbb{R}) \times \hat{\Xi}^{(k_0-1)-}$.

Therefore, it is expected that if the prolonged system (7) satisfies $r(x_k)(\delta x_k)^2 \rightarrow 0$ as $k \rightarrow 0$ in some sense, then for any $\gamma \in \Gamma(x'_{k_0}, x''_{k_0})$ and $s \in [0, 1]$

$$r(\phi_k(\xi^{(k-1)-}; \gamma(s))) \left(\frac{\partial \phi_k(\xi^{(k-1)-}; \gamma(s))}{\partial s} \right)^2 \rightarrow 0. \quad (11)$$

Thus, $d_r(x'_{k_0}, x''_{k_0}) \rightarrow 0$ as $k \rightarrow 0$ is further expected from (6). This interpretation can be formalized as follows.

Proposition 2.1: A scalar system (1) is UIES in the first moment [i.e., (3) for $p = 1$] if there exist $\lambda \in (0, 1)$ and $r: \mathbb{R} \rightarrow \mathbb{R}$ of class C^1 such that $r(x_{k_0}) > 0$ and

$$\mathbb{E}_0 \left[r(f(x_{k_0}, \xi_{k_0})) \left(\frac{\partial f(x_{k_0}, \xi_{k_0})}{\partial x} \right)^2 \right] \leq \lambda^2 r(x_{k_0}) \quad (12)$$

for all $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R} \times \hat{\Xi}^{(k_0-1)-}$. \triangleleft

Sketch of the proof: The main idea of the proof is to study the time evolution of the left-hand side of (11) utilizing the prolonged system (7).

It follows from (7) and (12) that

$$\begin{aligned} & \mathbb{E}_0[r(x_{k_0+1})(\delta x_{k_0+1})^2] \\ &= \mathbb{E}_0 \left[r(f(x_{k_0}, \xi_{k_0})) \left(\frac{\partial f(x_{k_0}, \xi_{k_0})}{\partial x} \delta x_{k_0} \right)^2 \right] \\ &\leq \lambda^2 r(x_{k_0})(\delta x_{k_0})^2. \end{aligned} \quad (13)$$

Note that $(x_{k_0}, \delta x_{k_0})$ is deterministic. The time-shift $k_0 \mapsto k$ leads to

$$\mathbb{E}_0[r(x_{k+1})(\delta x_{k+1})^2 | \mathcal{F}_{k-1}] \leq \lambda^2 r(x_k)(\delta x_k)^2 \quad \text{a.s.}$$

where recall that $\mathbb{E}_0[\cdot | \mathcal{F}_{k-1}]$ is the conditional expectation given \mathcal{F}_{k-1} , and $\mathbb{E}_0[\cdot] = \mathbb{E}_0[\cdot | \mathcal{F}_{k_0-1}]$. Since $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ is a filtration,

it follows that $\mathbb{E}_0[\mathbb{E}_0[\cdot|\mathcal{F}_{k-1}]] = \mathbb{E}_0[\cdot]$ for each $k \in \mathbb{Z}_{k_0+}$. Therefore, taking the conditional expectation $\mathbb{E}_0[\cdot]$ of both sides yields $\mathbb{E}_0[r(x_{k+1})(\delta x_{k+1})^2] \leq \lambda^2 \mathbb{E}_0[r(x_k)(\delta x_k)^2]$.

Recursively utilizing this inequality from k to k_0 leads to

$$\mathbb{E}_0[r(x_k)(\delta x_k)^2] \leq \lambda^{2(k-k_0)} r(x_{k_0})(\delta x_{k_0})^2.$$

Again note that $(x_{k_0}, \delta x_{k_0})$ is deterministic. Taking the square roots of both sides and applying the Cauchy–Schwarz inequality [47, Corollary 3.1.12] yield

$$\begin{aligned} \mathbb{E}_0 \left[\sqrt{r(x_k)(\delta x_k)^2} \right] &\leq \sqrt{\mathbb{E}_0[r(x_k)(\delta x_k)^2]} \\ &\leq \lambda^{k-k_0} \sqrt{r(x_{k_0})(\delta x_{k_0})^2}. \end{aligned} \quad (14)$$

Now, we connect the analysis of the prolonged system (7) with (14). As shown in (10), $(x_{k+1}, \delta x_{k+1}) = (\phi_k(\xi^{(k-1)-}; \gamma(s)), \partial\phi_k(\xi^{(k-1)-}; \gamma(s))/\partial s)$ satisfies (7). Thus, substituting this into (14) and taking the integration with respect to s lead to

$$\begin{aligned} &\int_0^1 \mathbb{E}_0 \left[\sqrt{r(\phi_k(\xi^{(k-1)-}; \gamma(s))) \left(\frac{\partial\phi_k(\xi^{(k-1)-}; \gamma(s))}{\partial s} \right)^2} \right] ds \\ &\leq \lambda^{k-k_0} \int_0^1 \sqrt{r(\gamma(s)) \left(\frac{d\gamma(s)}{ds} \right)^2} ds. \end{aligned} \quad (15)$$

As shown in Appendix A, taking the conditional expectation and the integral with respect to s is commutative. Also, recall that $\phi_k(\xi^{(k-1)-}; \gamma(s))$ is a class C^1 path connecting between x'_k and x''_k . Therefore, it follows from (6) that

$$\mathbb{E}_0[d(x'_k, x''_k)] \leq \lambda^{k-k_0} \int_0^1 \sqrt{r(\gamma(s)) \left(\frac{d\gamma(s)}{ds} \right)^2} ds$$

for any $\gamma \in \Gamma(x'_k, x''_k)$. Finally, we have (3) by taking the infimum of γ as in (6). ■

Remark 2.2: In this article, we focus on UIES with respect to the deterministic pair of initial states (x'_{k_0}, x''_{k_0}) . For random (x'_{k_0}, x''_{k_0}) being independent of ξ , our results are still applicable because we guarantee (3) for all (x'_{k_0}, x''_{k_0}) and, thus, can take its conditional expectation with respect to (x'_{k_0}, x''_{k_0}) . For obtaining less conservative conditions, one may expect to take a conditional expectation of (12) with respect to x_{k_0} , but this does not directly yield that of (13) because $(x_k, \delta x_k)$ are not independent by (10). ◁

Example 2.3: We apply Proposition 2.1 to an affine system with respect to a random variable

$$x_{k+1} = g(x_k) + h(x_k)\xi_k, \quad k \in \mathbb{Z}_{k_0+}$$

where $g, h : \mathbb{R} \rightarrow \mathbb{R}$ are of class C^1 , and $\mathbb{E}_0[\xi_k] = 0$ for all $k \in \mathbb{Z}_{k_0+}$. We select a metric r as a constant. Namely, we consider the Euclidean distance. According to Proposition 2.1, this system is UIES in the first moment with respect to the Euclidean distance if there exists $\lambda \in (0, 1)$ such that

$$\begin{aligned} &\mathbb{E}_0 \left[\left(\frac{\partial g(x_{k_0})}{\partial x} + \frac{\partial h(x_{k_0})}{\partial x} \xi_{k_0} \right)^2 \right] \\ &= \left(\frac{\partial g(x_{k_0})}{\partial x} \right)^2 + \mathbb{E}_0[\xi_{k_0}^2] \left(\frac{\partial h(x_{k_0})}{\partial x} \right)^2 \leq \lambda^2 \end{aligned}$$

for all $(k_0, x_{k_0}, \xi^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where we used $\mathbb{E}_0[\xi_{k_0}] = 0$. If ξ_k is i.i.d. Gaussian noise with variance $\sigma^2 > 0$, then the condition becomes

$$\left(\frac{\partial g(x_{k_0})}{\partial x} \right)^2 + \sigma^2 \left(\frac{\partial h(x_{k_0})}{\partial x} \right)^2 \leq \lambda^2 \quad \forall x_{k_0} \in \mathbb{R}.$$

Therefore, by our results, one can study UIES in the first moment under i.i.d. Gaussian noise. In fact, utilizing a non-constant $r(x_{k_0})$ leads to a necessary and sufficient condition for UIES in the second moment; see Corollary 4.4 in the following. ◁

D. Metrics Depending on Stochastic Processes

In Proposition 2.1, we employ $r(x)$, which is independent of a stochastic process $(\xi_k)_{k \in \mathbb{Z}}$. In fact, using such an $r(x)$ restricts the class of applications. To see this, we start this section by reviewing stability analysis of a Markov jump linear system.

Example 2.4: Let ξ_k be stationary (its characteristic does not depend on $k \in \mathbb{Z}$) and takes a value only in $\mathcal{M} := \{1, \dots, M\}$. Let us denote the transition probability from mode $i \in \mathcal{M}$ to $j \in \mathcal{M}$ by $\pi_{j,i} := \mathbb{P}(\xi_{k+1} = j | \xi_k = i) \geq 0$ for each $k \in \mathbb{Z}$. According to [33, Eq. (3.15)], a scalar Markov jump linear system $x_{k+1} = a(i)x_k$ is ES in the second moment if and only if there exist $\lambda \in (0, 1)$ and $\hat{r}_{0,i} > 0, i \in \mathcal{M}$, such that

$$\sum_{j \in \mathcal{M}} \pi_{j,i} \hat{r}_{0,j} a^2(j) \leq \lambda^2 \hat{r}_{0,i} \quad \forall i \in \mathcal{M}. \quad (16)$$

Therefore, even for Markov jump linear systems, we need mode-dependent $\hat{r}_{0,i}$ to describe a necessary and sufficient condition. ◁

As observed previously, we can derive a less conservative stability condition by taking r as a function of a stochastic process $(\xi_k)_{k \in \mathbb{Z}}$ in addition to x_{k_0} . When dealing with such r , its domain needs to be considered carefully. For instance, let us consider $r_- : \xi^{k_0+} \mapsto \xi_{k_0}$ and $r_+ : \xi^{k_0+} \mapsto \xi_{k_0+1}$. Then, one may expect $r_-(\xi^{(k_0+1)+}) = r_+(\xi^{k_0+})$. However, the left-hand side is not well defined, since the domain of r_- is $\mathbb{R}^{\mathbb{Z}_{k_0+}}$, and $\xi^{(k_0+1)+}$ belongs to $\mathbb{R}^{\mathbb{Z}_{(k_0+1)+}}$. To resolve this issue, we utilize the time-shift operator $S_{k_0} : \xi^{k_0+} \mapsto \zeta^{0+}$ mentioned in the notation part. For mappings $\hat{r}_- : \zeta^{0+} \mapsto \zeta_0$ and $\hat{r}_+ : \zeta^{0+} \mapsto \zeta_1$, the compositions of mappings are $\hat{r}_- \circ S_{k_0} : \xi^{k_0+} \mapsto \zeta_0$ and $\hat{r}_+ \circ S_{k_0} : \xi^{k_0+} \mapsto \zeta_1$, respectively. Then, $\hat{r}_-(S_{k_0} \xi^{(k_0+1)+}) = \zeta_1$ is well defined and satisfies $\hat{r}_-(S_{k_0} \xi^{(k_0+1)+}) = \hat{r}_+(S_{k_0} \xi^{k_0+})$.

Now, we use $r(x_{k_0}, S_{k_0} \xi^{k_0+})$ for contraction analysis. Since $S_{k_0} \xi^{k_0+}$ is a subsequence of a stochastic process, we utilize its conditional expectation $\mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})]$, and its forward time-shift $\mathbb{E}_0[r(x_{(k_0+1)}, S_{(k_0+1)} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}]$ to describe conditions. In our analysis, to induce a distance function, we require that the conditional expectation $\mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})]$ is lower and upper bounded on a metric $\tilde{r}(x_{k_0}) > 0$ depending on only x_{k_0} . Note that even if $\mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})]$ is bounded, $r(x_{k_0}, S_{k_0} \xi^{k_0+})$ can be nonpositive or infinite at some $(x_{k_0}, S_{k_0} \xi^{k_0+})$ (with measure zero). In fact, in the converse analysis proceeded later, we only show the boundedness of $\mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})]$. Therefore, the range of r is taken as

$\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ for the sake of formality. Noting this, we generalize Proposition 2.1 as follows.

Proposition 2.5: A scalar system (1) is UIES in the first moment if there exist $c_1, c_2 > 0, \lambda \in (0, 1), \tilde{r} : \mathbb{R} \rightarrow \mathbb{R}$ of class C^1 , and $r : \mathbb{R} \times \mathbb{R}^{\mathbb{Z}_{0+}} \rightarrow \bar{\mathbb{R}}$ such that $\tilde{r}(x_{k_0}) > 0$ and

$$c_1 \tilde{r}(x_{k_0}) \leq \mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})] \leq c_2 \tilde{r}(x_{k_0}) \quad (17a)$$

$$\mathbb{E}_0 \left[\mathbb{E}_0[r(f(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] \left(\frac{\partial f(x_{k_0}, \xi_{k_0})}{\partial x} \right)^2 \right] \leq \lambda^2 \mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})] \quad (17b)$$

for all $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R} \times \hat{\Xi}^{(k_0-1)-}$. \triangleleft

Example 2.6: Before mentioning the sketch of the proof, we consider understanding the condition (17) by applying it to a Markov jump linear system $x_{k+1} = a(i)x_k$ in Example 2.4. Since ξ_{k_0} (respectively, ξ_{k_0+1}) only depends on ξ_{k_0-1} (respectively, ξ_{k_0}), it follows that

$$\begin{aligned} \mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})] &= \mathbb{E}[r(x_{k_0}, S_{k_0} \xi_{k_0}) | \xi_{k_0-1} = i] =: \hat{r}_{0,i} \\ \mathbb{E}_0[r(f(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] \\ &= \mathbb{E}[r(a(j)x_{k_0}, S_{k_0+1} \xi_{k_0+1}) | \xi_{k_0} = j] = \hat{r}_{0,j}. \end{aligned}$$

Then, (17) reduces to $\hat{r}_{0,i} > 0, i \in \mathcal{M}$, and

$$\mathbb{E}[\hat{r}_{0,j} a^2(j) | \xi_{k_0-1} = i] \leq \lambda^2 \hat{r}_{0,i} \quad \forall i \in \mathcal{M}.$$

Since the transition probability is given by $\pi_{j,i}$, the above-mentioned inequality is equivalent to (16). Therefore, using $r(x_{k_0}, S_{k_0} \xi^{k_0+})$, we can handle a Markov jump system. \triangleleft

Sketch of the proof of Proposition 2.5: The proof procedure is similar to that of Proposition 2.1. It follows from (7) and (17b) that

$$\begin{aligned} \mathbb{E}_0[\mathbb{E}_0[r(x_{k_0+1}, S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] (\delta x_{k_0+1})^2] \\ \leq \lambda^2 \mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})] (\delta x_{k_0})^2. \end{aligned}$$

The time-shift $k_0 \mapsto k$ yields

$$\begin{aligned} \mathbb{E}_0[\mathbb{E}_0[r(x_{k+1}, S_{k+1} \xi^{(k+1)+}) | \mathcal{F}_k] (\delta x_{k+1})^2 | \mathcal{F}_{k-1}] \\ \leq \lambda^2 \mathbb{E}_0[r(x_k, S_k \xi^{k+}) | \mathcal{F}_{k-1}] (\delta x_k)^2 \quad \text{a.s.} \end{aligned}$$

Thus, by taking the conditional expectation $\mathbb{E}_0[\cdot]$ of both sides, we have

$$\begin{aligned} \mathbb{E}_0[\mathbb{E}_0[r(x_{k+1}, S_{k+1} \xi^{(k+1)+}) | \mathcal{F}_k] (\delta x_{k+1})^2] \\ \leq \lambda^2 \mathbb{E}_0[\mathbb{E}_0[r(x_k, S_k \xi^{k+}) | \mathcal{F}_{k-1}] (\delta x_k)^2]. \end{aligned}$$

Recursively utilizing this inequality from k to k_0 leads to

$$\begin{aligned} \mathbb{E}_0[\mathbb{E}_0[r(x_k, S_k \xi^{k+}) | \mathcal{F}_{k-1}] (\delta x_k)^2] \\ \leq \lambda^{2(k-k_0)} \mathbb{E}_0[r(x_{k_0}, S_{k_0} \xi^{k_0+})] (\delta x_{k_0})^2. \end{aligned}$$

From (17a), we have

$$\mathbb{E}_0[\tilde{r}(x_k) (\delta x_k)^2] \leq \frac{c_2}{c_1} \lambda^{2(k-k_0)} \tilde{r}(x_{k_0}) (\delta x_{k_0})^2.$$

The rest is similar to the proof of Proposition 2.1. \blacksquare

Remark 2.7: In our analysis, the partial differentiability requirement of f with respect to x can slightly be weakened. Let $D \subset \mathbb{R}$ be a Lebesgue measurable set with the Lebesgue density 1. Also, let $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be both continuous and continuously differentiable with respect to x on $D \times \mathbb{R}$. Then, Propositions 2.1 and 2.5 can be generalized by replacing $(\partial f(x_{k_0}, \xi_{k_0}) / \partial x)^2$ with

$$\begin{aligned} \partial f_x^2(x_{k_0}, \xi_{k_0}) \\ := \limsup_{v \in \mathbb{R}, h \rightarrow 0^+} \left(\frac{f(x_{k_0} + vh, \xi_{k_0}) - f(x_{k_0}, \xi_{k_0})}{h} \right)^2. \end{aligned}$$

On $D \times \mathbb{R}$, we have $\partial f_x^2(x_{k_0}, \xi_{k_0}) = (\partial f(x_{k_0}, \xi_{k_0}) / \partial x)^2$. Therefore, the difference is analysis on $\mathbb{R} \setminus D$ with the Lebesgue density 0. \triangleleft

III. MAIN RESULTS

In this section, we develop stochastic contraction theory for n -dimensional stochastic systems, which can be viewed as generalizations of the scalar case. First, we mention the considered class of systems and define UIES in the p th moment. Then, we derive UIES conditions. After that, we revisit the linear stochastic case [41] and the nonlinear deterministic case [1], [29] to show that our results contain results for these cases.

A. Incremental Stability in Moments

Let $\xi := (\xi_k)_{k \in \mathbb{Z}} : \Omega \rightarrow (\mathbb{R}^m)^{\mathbb{Z}}$ be a general stochastic process. In the scalar case, f is assumed to be time invariant. As a further generalization, we consider the time-varying case. Namely, we consider a discrete-time nonlinear stochastic system, described by

$$x_{k+1} = f_k(x_k, \xi_k), \quad k \in \mathbb{Z}_{k_0+} \quad (18)$$

where $f_k(x, \eta)$ and $\partial f_k(x, \eta) / \partial x$ are continuous with respect to $x \in \mathbb{R}^n$ at each $(k, \eta) \in \mathbb{Z} \times \mathbb{R}^m$, and $f_k(x, \xi_k)$ and $\partial f_k(x, \xi_k) / \partial x$ are \mathcal{F}_k -measurable at each $(k, x) \in \mathbb{Z} \times \mathbb{R}^n$. As mentioned in Remark 2.7 for the scalar case, f_k and $\partial f_k / \partial x$ are allowed to be piecewise continuous, which will be explained in Remark 3.5 later. The initial condition $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$ is deterministic, where k_0, x_{k_0} , and $\hat{\xi}^{(k_0-1)-}$, respectively, denote the initial time, the initial state, and the initial condition of the stochastic process $\xi := (\xi_k)_{k \in \mathbb{Z}} : \Omega \rightarrow (\mathbb{R}^m)^{\mathbb{Z}}$, namely, $\hat{\xi}^{(k_0-1)-}$ is the observation of an infinite past sequence up to $k_0 - 1$ of ξ .

Recalling the distance function d in the notation part, we extend the concept of UIES (3) in the p th moment to the n -dimensional case. The difference from the scalar case is only the dimensions of the variables (x'_k, x''_k) and ξ_k .

Definition 3.1: The system (18) is said to be UIES in the p th moment (with respect to distance d) if there exist $a > 0$ and $\lambda \in (0, 1)$ such that

$$\mathbb{E}_0[d^p(x'_k, x''_k)] \leq a^p \lambda^{p(k-k_0)} d^p(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+} \quad (19)$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. \triangleleft

As mentioned in the scalar case, UIES in the p th moment is an integrated concept of 1) ES in the p th moment (4) for stochastic systems and 2) UIES (5) for contraction analysis of nonlinear deterministic systems. On the other hand, our results provide sufficient conditions even if (x'_{k_0}, x''_{k_0}) is random being independent of ξ because (19) is preserved by taking its conditional expectation with respect to (x'_{k_0}, x''_{k_0}) in such a case.

In the n -dimensional case, a Riemannian metric is a class C^1 matrix-valued function $\tilde{P} : \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$. Let $\Gamma(x', x'')$ denote the collection of class C^1 paths $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ such that $\gamma(0) = x'$ and $\gamma(1) = x''$. Then, the distance function induced by a Riemannian metric $\tilde{P}(x)$ is

$$d_{\tilde{P}}(x', x'') := \inf_{\gamma \in \Gamma(x', x'')} \int_0^1 \sqrt{\frac{d^\top \gamma(s)}{ds} \tilde{P}(\gamma(s)) \frac{d\gamma(s)}{ds}} ds. \quad (20)$$

We study UIES in the p th moment with respect to this distance. If \tilde{P} is a constant matrix (that is symmetric and positive definite), then this is a weighted Euclidean distance. Applying the Hopf–Rinow theorem [48], [49], it is mentioned in [5] that for any given $\varepsilon > 0$ and $(x', x'') \in \mathbb{R}^n \times \mathbb{R}^n$, there exists $\gamma^* \in \Gamma(x', x'')$ such that

$$\int_0^1 \sqrt{\frac{d^\top \gamma^*(s)}{ds} \tilde{P}(\gamma^*(s)) \frac{d\gamma^*(s)}{ds}} ds \leq (1 + \varepsilon) d_{\tilde{P}}(x', x''). \quad (21)$$

We use this path γ^* for our analysis.

B. Incremental Stability Conditions

In this section, we present UIES conditions for n -dimensional nonlinear stochastic systems. First, we state a condition with respect to the Riemannian metric, which is a natural generalization of Proposition 2.5 for the scalar system.

Theorem 3.2: A system (18) is UIES [with respect to $d_{\tilde{P}}$ defined by (20)] in the first moment if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, $\tilde{P} : \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ of class C^1 , and $P : \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{0+}} \rightarrow \bar{\mathbb{R}}^{n \times n}$ such that

$$c_1^2 \tilde{P}(x_{k_0}) \preceq \mathbb{E}_0[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \preceq c_2^2 \tilde{P}(x_{k_0}) \quad (22a)$$

$$\mathbb{E}_0 \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \mathbb{E}_0[P(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \right] \preceq \lambda^2 \mathbb{E}_0[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \quad (22b)$$

for all $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where S_k is the time-shift operator introduced in the notation part. \triangleleft

Proof: The proof is provided in Appendix A. \blacksquare

Next, we focus on the Euclidean distance, which corresponds to specifying \tilde{P} in Theorem 3.2 into the identity matrix. In this case, it is possible to obtain a UIES condition for second moment stability, stronger than first moment stability because we can avoid to apply the Cauchy–Schwarz inequality in contrast to the general Riemannian metric case; for more details, see the proof

in Appendix B. Moreover, we have the converse proof, stated in the following.

Theorem 3.3: A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $P : \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{0+}} \rightarrow \bar{\mathbb{R}}^{n \times n}$ such that

$$c_1^2 I_n \preceq \mathbb{E}_0[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \preceq c_2^2 I_n \quad (22c)$$

and (22b) holds for all $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$.

Proof: The proof is provided in Appendix B. \blacksquare

In the converse proof, we show that the conditional expectation $\mathbb{E}_0[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})]$ is lower and upper bounded as in (22c). This holds even if an element of P is infinite at some $(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})$ (with measure zero). Therefore, the range of P is taken as $\bar{\mathbb{R}}^{n \times n}$ in Theorems 3.2 and 3.3 for the sake of formality.

Main ideas of the proofs of Theorems 3.2 and 3.3 are similar to the scalar case. Namely, by using the inequality (22b), we study the stability of the prolongation of the system (18)

$$\begin{cases} x_{k+1} = f_k(x_k, \xi_k) \\ \delta x_{k+1} = \frac{\partial f_k(x_k, \xi_k)}{\partial x} \delta x_k \end{cases} \quad k \in \mathbb{Z}_{k_0+}. \quad (23)$$

Then, we utilize (10) to connect analysis of the prolongation with the time evolutions of a path and its partial derivative. Finally, we take the integration of the obtained inequality with respect to s as done in (15) to conclude UIES.

It is not always easy to verify the conditions in Theorems 3.2 and 3.3, since we need to deal with an infinite sequence of random variables. These theorems are provided as a stepping stone for developing stochastic contraction theory. Their usage is to derive UIES conditions tailored to subclasses of the system (18) by utilizing special structures of the subclasses. This is illustrated in Section IV.

Remark 3.4: From the proofs of Theorems 3.2 and 3.3 in Appendices A and B, one notices that a (nonuniform) incremental exponential stability (IES) condition in the first or second moment can readily be obtained by replacing c_1, c_2, λ , and \tilde{P} with those depending on k_0 . By IES in the p th moment at $k_0 \in \mathbb{Z}$, we mean that there exist $a(k_0) > 0$ and $\lambda(k_0) \in (0, 1)$ such that

$$\mathbb{E}_0[d^p(x'_k, x''_k)] \leq a^p(k_0) \lambda^{p(k-k_0)}(k_0) d^p(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+}$$

for each $((x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$.

On the other hand, Theorems 3.2 and 3.3 can be generalized to incremental stability analysis on a convex subset $D \subset \mathbb{R}^n$ when $f_k : D \times \mathbb{R}^m \rightarrow D$, $k \in \mathbb{Z}$, because D is a (robustly) positively invariant set for such f_k . \triangleleft

Remark 3.5: Remark 2.7 for the partial differentiability of f_k in the scalar case can be generalized to the n -dimensional case.

C. Linear Stochastic Cases

In this section, we show that our results contain results for linear stochastic systems [41]. In the linear case, the system (18) becomes

$$x_{k+1} = A(\xi_k)x_k, \quad k \in \mathbb{Z}_{k_0+} \quad (24)$$

under the deterministic initial condition $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where $A : \mathbb{R}^m \rightarrow \mathbb{R}^{n \times n}$. Note that the system (24) is general in the sense of that this covers the linear system $x_{k+1} = \tilde{A}_k x_k$ with random $(\tilde{A}_k)_{k \in \mathbb{Z}} : \Omega \rightarrow (\mathbb{R}^{n \times n})^{\mathbb{Z}}$ because another representation of \tilde{A}_k is

$$\tilde{A}_k = \begin{bmatrix} \xi_{k,1} & \cdots & \xi_{k,n} \\ \vdots & \ddots & \vdots \\ \xi_{k,n(n-1)} & \cdots & \xi_{k,n^2} \end{bmatrix} =: A(\xi_k).$$

A similar discussion holds in the nonlinear case. For instance, a polynomial system with random coefficients can be represented by (18). Since f_k is time-varying, we can handle more wider classes of systems.

In the linear case, $\partial f_k(x_k, \xi_k)/\partial x = A(\xi_k)$ is independent of x_k . Thus, P and \tilde{P} in Theorems 3.2 and 3.3 can be chosen to be independent of x_k . Then, the conditions of these two theorems become the same, and thus, we only apply Theorem 3.3. Next, for the linear system (24), UIES in the p th moment reduces to ES in the p th moment (4) as mentioned in the scalar case. Therefore, applying Theorem 3.3 to the linear system (24) recovers [41, Th. 3]. In other words, our results can be viewed as natural extensions of the result for linear stochastic systems to nonlinear stochastic systems.

Corollary 3.6: A linear stochastic system (24) is ES in the second moment (i.e., (4) for $p = 2$) if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $P : (\mathbb{R}^m)^{\mathbb{Z}_{0+}} \rightarrow \mathbb{R}^{n \times n}$ such that

$$c_1^2 I_n \preceq \mathbb{E}_0[P(S_{k_0} \xi^{k_0+})] \preceq c_2^2 I_n$$

$$\mathbb{E}_0 \left[A^\top(\xi_{k_0}) \mathbb{E}_0[P(S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] A(\xi_{k_0}) - \lambda^2 P(S_{k_0} \xi^{k_0+}) \right] \preceq 0$$

for all $(k_0, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \hat{\Xi}^{(k_0-1)-}$. \triangleleft

D. Nonlinear Deterministic Cases

In this section, we apply our results to nonlinear deterministic systems, described by

$$x_{k+1} = g_k(x_k), \quad k \in \mathbb{Z}_{k_0+} \quad (25)$$

under the initial condition $(k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}^n$, where $g_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is of class C^1 for each $k \in \mathbb{Z}$.

In the deterministic case, P in Theorems 3.2 and 3.3 can be chosen to be independent of ξ_k as in the following corollaries.

Corollary 3.7: A nonlinear deterministic system is UIES (5) [with respect to $d_{\tilde{P}}$ defined by (20)] if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, $\tilde{P} : \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ of class C^1 , and $P : \mathbb{Z} \times \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$c_1^2 \tilde{P}(x_{k_0}) \preceq P(k_0, x_{k_0}) \preceq c_2^2 \tilde{P}(x_{k_0})$$

$$\frac{\partial^\top g_{k_0}(x_{k_0})}{\partial x} P(k_0 + 1, g_{k_0}(x_{k_0})) \frac{\partial g_{k_0}(x_{k_0})}{\partial x} \preceq \lambda^2 P(k_0, x_{k_0}) \quad (26)$$

for all $(k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}^n$. \triangleleft

Corollary 3.8: A nonlinear deterministic system is UIES with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $P : \mathbb{Z} \times \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that $c_1^2 I_n \preceq P(k_0, x_{k_0}) \preceq c_2^2 I_n$ and (26) hold for all $(k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}^n$. \triangleleft

Corollary 3.8 is equivalent to [29, Th. 15]. Corollary 3.7 is a new result of this article, which can be understood as a discrete-time version of the UIES condition for continuous-time systems; see, e.g., [5, Th. 1]. Therefore, our results can be viewed as natural extensions of the results for nonlinear deterministic systems to stochastic systems.

IV. APPLICATIONS

In the previous section, we have presented UIES conditions for general stochastic systems (18). In this section, we illustrate the utility of the obtained conditions by tailoring them to specific classes of processes. In particular, we study cases where ξ follows temporally independent processes or Markov processes. In most of the literature of stochastic control, e.g., [35], [36], [42], [44], stability conditions have been separately developed for each special class of processes. By virtue of studying the general stochastic process ξ , conditions for each special case are provided simply by restricting the class of ξ as in [41] about linear systems. Due to the lack of contraction analysis for stochastic systems, the obtained conditions in each special case are new contribution of this article on their own. In this section, we only tailor Theorem 3.3, but similar results corresponding to Theorem 3.2 can readily be obtained.

A. Temporally Independent Processes

In this section, we consider ξ satisfying the following assumption. Such ξ is called a *temporally independent process*.

Assumption 4.1: For $\xi = (\xi_k)_{k \in \mathbb{Z}}$, the random vectors $\xi_k, k \in \mathbb{Z}$, are independently distributed. \triangleleft

Under Assumption 4.1, conditions (22b) and (22c) in Theorem 3.3 are independent of $\hat{\xi}^{(k_0-1)-}$ for each $k_0 \in \mathbb{Z}$. Hence, the conditional expectation can be replaced with the (standard) expectation. Then, we have the following corollary of Theorem 3.3 by defining:

$$\hat{P}(k_0, x_{k_0}) := \mathbb{E}[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})], \quad (k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}^n. \quad (27)$$

Corollary 4.2: Suppose that Assumption 4.1 holds. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $\hat{P} : \mathbb{Z} \times \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$c_1^2 I_n \preceq \hat{P}(k_0, x_{k_0}) \preceq c_2^2 I_n$$

$$\mathbb{E} \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0})) \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \right] \preceq \lambda^2 \hat{P}(k_0, x_{k_0})$$

for all $(k_0, x_{k_0}) \in \mathbb{Z} \times \mathbb{R}^n$. \triangleleft

Proof: Substituting (27) into (22c) yields the first inequalities. Next, substituting (27) into (22b) leads to

$$\mathbb{E}_0 \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \mathbb{E}_0[\hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0})) | \mathcal{F}_{k_0}] \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \right] \preceq \lambda^2 \hat{P}(k_0, x_{k_0}).$$

Since $\mathbb{E}_0[\cdot|\mathcal{F}_{k_0}]$ is the conditional expectation given \mathcal{F}_{k_0} , it follows that

$$\mathbb{E}_0[\hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}))|\mathcal{F}_{k_0}] = \hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0})).$$

Therefore, we obtain the last inequality. \blacksquare

We further consider a stationary case, i.e., ξ_k and f_k are independent of k .

Assumption 4.3: The stochastic process ξ is stationary (in the strict sense), i.e., none of the characteristics of ξ_k changes with time k . Moreover, none of f_k changes with time k , i.e., $f_k = \hat{f}_0$ for all $k \in \mathbb{Z}$. \triangleleft

Note that the stochastic process satisfying Assumptions 4.1 and 4.3 is an i.i.d. process. Under Assumptions 4.1 and 4.3, \hat{P} in (27) can be chosen as a k_0 -independent function. Namely, we have the following corollary without the proof.

Corollary 4.4: Suppose that Assumptions 4.1 and 4.3 hold. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0, \lambda \in (0, 1)$, and $\hat{P}_0 : \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$c_1^2 I_n \preceq \hat{P}_0(x_0) \preceq c_2^2 I_n$$

$$\mathbb{E} \left[\frac{\partial^\top \hat{f}_0(x_0, \xi_0)}{\partial x} \hat{P}_0(\hat{f}_0(x_0, \xi_0)) \frac{\partial \hat{f}_0(x_0, \xi_0)}{\partial x} \right] \preceq \lambda^2 \hat{P}_0(x_0)$$

for all $x_0 \in \mathbb{R}^n$. \triangleleft

Remark 4.5: In Corollary 4.4, we consider the stationary case. As a more general case, Corollary 4.2 can be specialized to the periodic case where there exists a positive integer N such that $f_{\kappa N+i} = f_{\kappa+i}, i = 0, 1, \dots, N-1, \kappa \in \mathbb{Z}$, and none of the characteristics of $\xi_{\kappa N+i}, i = 0, 1, \dots, N-1$ changes with $\kappa \in \mathbb{Z}$. The generalized condition is described by using periodic $\hat{P}_i, i = 0, 1, \dots, N-1$; for more details, see a similar discussion in the linear case [41, Corollary 3]. \triangleleft

Example 4.6: We apply Corollary 4.4 to a linear system (24) with i.i.d. ξ . It is ES in the second moment if and only if there exist $\lambda \in (0, 1)$ and $\hat{P}_0 \in \mathbb{S}_{>0}^{n \times n}$ such that $\mathbb{E}[A^\top(\xi_0)\hat{P}_0 A(\xi_0)] \preceq \lambda^2 \hat{P}_0$. This condition is found in [43, Th. 2]. \triangleleft

B. General Markov Processes

In this section, we consider the case where ξ is a general Markov process.

Assumption 4.7: For each $\Theta_j \subset \mathbb{R}^m$, every $j \in \mathbb{Z}_{(i+1)+}$ and $i \in \mathbb{Z}$, it follows that

$$\mathbb{P}(\xi_j \in \Theta_j | \xi_i, \xi_{i-1}, \dots) = \mathbb{P}(\xi_j \in \Theta_j | \xi_i)$$

where $\mathbb{P}(\cdot)$ denotes the conditional probability. \triangleleft

Assumption 4.7 implies that the conditional expectation \mathbb{E}_0 can be simplified as

$$\mathbb{E}_0[\cdot] = \mathbb{E}[\cdot | \xi_{k_0-1} = \hat{\xi}_{k_0-1}] \quad (28a)$$

$$\mathbb{E}_0[\cdot | \mathcal{F}_k] = \mathbb{E}[\cdot | \xi_k], k \in \mathbb{Z}_{k_0+} \quad (28b)$$

for each $(k_0, \hat{\xi}_{k_0-1}) \in \mathbb{Z} \times \Theta_{k_0-1}$, where note that Θ_{k_0-1} is the support of ξ_{k_0-1} . Then, for P in Theorem 3.3, there exists $\hat{P} : \mathbb{Z} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$\mathbb{E}_0[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})]$$

$$= \mathbb{E}[P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+}) | \xi_{k_0-1} = \hat{\xi}_{k_0-1}]$$

$$= \hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1}) \quad (29)$$

for each $(k_0, x_{k_0}, \hat{\xi}_{k_0-1}) \in \mathbb{Z} \times \mathbb{R}^n \times \Theta_{k_0-1}$. Now, we have the following corollary of Theorem 3.3 for Markov processes.

Corollary 4.8: Suppose that Assumption 4.7 holds. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0, \lambda \in (0, 1)$, and $\hat{P} : \mathbb{Z} \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$c_1^2 I_n \preceq \hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1}) \preceq c_2^2 I_n$$

$$\mathbb{E} \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), \xi_{k_0}) \right.$$

$$\left. \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \Big|_{\xi_{k_0-1} = \hat{\xi}_{k_0-1}} \right] \preceq \lambda^2 \hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1})$$

for all $(k_0, x_{k_0}, \hat{\xi}_{k_0-1}) \in \mathbb{Z} \times \mathbb{R}^n \times \Theta_{k_0-1}$. \triangleleft

Proof: Substituting (29) into (22c) yields the first inequalities. Next, substituting (29) into (22b) leads to

$$\mathbb{E}_0 \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \mathbb{E}_0[\hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), \xi_{k_0}) | \mathcal{F}_{k_0}] \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x} \right] \preceq \lambda^2 \hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1}). \quad (30)$$

Since $\mathbb{E}_0[\cdot | \mathcal{F}_{k_0}]$ is the conditional expectation given \mathcal{F}_{k_0} , it follows that

$$\begin{aligned} & \mathbb{E}_0[\hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), \xi_{k_0}) | \mathcal{F}_{k_0}] \\ &= \hat{P}(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), \xi_{k_0}). \end{aligned}$$

Substituting this into (30) and applying (28a) leads to the last inequality. \blacksquare

In the stationary case, again \hat{P} can be chosen as a k_0 -independent function. Namely, we have the following corollary without the proof.

Corollary 4.9: Suppose that Assumptions 4.3 and 4.7 hold. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0, \lambda \in (0, 1)$, and $\hat{P}_0 : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{S}_{>0}^{n \times n}$ such that

$$c_1^2 I_n \preceq \hat{P}_0(x_0, \hat{\xi}_{-1}) \preceq c_2^2 I_n$$

$$\mathbb{E} \left[\frac{\partial^\top \hat{f}_0(x_0, \xi_0)}{\partial x} \hat{P}_0(\hat{f}_0(x_0, \xi_0), \xi_0) \frac{\partial \hat{f}_0(x_0, \xi_0)}{\partial x} \Big|_{\xi_{-1} = \hat{\xi}_{-1}} \right]$$

$$\preceq \lambda^2 \hat{P}_0(x_0, \hat{\xi}_{-1})$$

for all $(x_0, \hat{\xi}_{-1}) \in \mathbb{R}^n \times \Theta_{-1}$. \triangleleft

Remark 4.10: Again Corollary 4.8 can be specialized to the periodic case by using periodic $\hat{P}_i, i = 0, 1, \dots, N-1$. \triangleleft

C. Finite-Mode Markov Chains

In this section, we further consider the case where ξ is a finite-mode Markov chain, which is nonstationary (i.e., nonhomogeneous) unless the transition probability is time invariant.

Assumption 4.11: The process ξ is given by a finite-mode Markov chain defined on the mode set $\mathcal{M} := \{1, \dots, M\}$, i.e., ξ_k can take a value only in \mathcal{M} at each $k \in \mathbb{Z}$. \triangleleft

The process ξ satisfying this assumption is a special case of the general Markov process in Assumption 4.7, and the corresponding system (18) can be seen as a stochastic switched nonlinear system with the switching signal ξ given by a finite-mode Markov chain. Such a system is nothing but a standard Markov jump nonlinear system, e.g., [35], [36]; also see, e.g., [34] for Markov jump linear systems. This exemplifies the generality of the system class dealt with in this article.

Let us denote the transition probability from mode i to j by

$$\pi_{j,i}^k := \mathbb{P}(\xi_{k+1} = j | \xi_k = i) \geq 0$$

for each $k \in \mathbb{Z}$. By the definition, it satisfies

$$\sum_{j \in \mathcal{M}} \pi_{j,i}^k = 1, \quad k \in \mathbb{Z}$$

for all $i \in \mathcal{M}$. Then, by using the mode dependent function \hat{P}_i , $i \in \mathcal{M}$, Corollaries 4.8 and 4.9 are further simplified as stated in the following, where the latter is about the stationary (i.e., homogeneous) Markov chain.

Corollary 4.12: Suppose that Assumption 4.11 holds. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $\hat{P}_i : \mathbb{Z} \times \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$, $i \in \mathcal{M}$, such that

$$\begin{aligned} c_1^2 I_n &\leq \hat{P}_i(k_0, x_{k_0}) \leq c_2^2 I_n \\ \sum_{j \in \mathcal{M}} \pi_{j,i}^{k_0} \frac{\partial^\top f_{k_0}(x_{k_0}, j)}{\partial x} \hat{P}_j(k_0 + 1, f_{k_0}(x_{k_0}, j)) \\ &\quad \frac{\partial f_{k_0}(x_{k_0}, j)}{\partial x} \leq \lambda^2 \hat{P}_i(k_0, x_{k_0}) \end{aligned}$$

for all $(k_0, x_{k_0}, i) \in \mathbb{Z} \times \mathbb{R}^n \times \mathcal{M}$. \triangleleft

Corollary 4.13: Suppose that Assumptions 4.3 and 4.11 hold. A system (18) is UIES in the second moment with respect to the Euclidean distance if and only if there exist $c_1, c_2 > 0$, $\lambda \in (0, 1)$, and $\hat{P}_{0,i} : \mathbb{R}^n \rightarrow \mathbb{S}_{>0}^{n \times n}$, $i \in \mathcal{M}$, such that

$$\begin{aligned} c_1^2 I_n &\leq \hat{P}_{0,i}(x_0) \leq c_2^2 I_n \\ \sum_{j \in \mathcal{M}} \pi_{j,i} \frac{\partial^\top \hat{f}_0(x_0, j)}{\partial x} \hat{P}_{0,j}(\hat{f}_0(x_0, j)) \frac{\partial \hat{f}_0(x_0, j)}{\partial x} \\ &\leq \lambda^2 \hat{P}_{0,i}(x_0) \end{aligned} \quad (31)$$

for all $(x_0, i) \in \mathbb{R}^n \times \mathcal{M}$. \triangleleft

Remark 4.14: Again Corollary 4.12 can be specialized to the periodic case by using periodic $\hat{P}_{k,i}$, $k = 0, 1, \dots, N-1$. \triangleleft

Example 4.15: As a generalization of Examples 2.4 and 2.6, we apply Corollary 4.13 to a linear system (24) when ξ is given by a finite-mode Markov chain. This system can be represented

by $x_{k+1} = A(j)x_k$, and thus, (31) becomes $\hat{P}_{0,i} \succ 0$ and

$$\sum_{j \in \mathcal{M}} \pi_{j,i} A^\top(j) \hat{P}_{0,j} A(j) \leq \lambda^2 \hat{P}_{0,i} \quad \forall i \in \mathcal{M}.$$

This is equivalent to [33, Eq. (3.15)]. By restricting classes of the processes and systems, we finally establish the connection between our results and the well-known condition for Markov jump linear systems. \triangleleft

D. Observer Design for Markov Jump Systems

In this section, we apply Corollary 4.13 to the observer design of Markov jump systems, which is further utilized for the state estimation of an epidemic model.

Consider the following Markov jump system:

$$\begin{cases} x_{k+1} = g(x_k, i), & i \in \mathcal{M} \\ y_k = Cx_k \end{cases} \quad (32)$$

$$\mathbb{P}(\xi_{k+1} = j | \xi_k = i) = \pi_{j,i} \quad (32)$$

where the initial state at initial time $k_0 = 0$ is $x_0 \in \mathbb{R}^n$. As observer dynamics, we consider the following system:

$$\hat{x}_{k+1} = g(\hat{x}_k, i) + H_i(C\hat{x}_k - y) \quad (33)$$

where the initial state at initial time $k_0 = 0$ is $\hat{x}_0 \in \mathbb{R}^n$. Note that if $\hat{x}_0 = x_0$, then $\hat{x}_k = x_k$ for all $k \in \mathbb{Z}_{0+}$. In other words, the set of solutions $\hat{x}_k = \hat{x}_k(\hat{x}_0, y, i)$ to (33) contains a solution $x_k = x_k(x_0, i)$ to (32) as a special case where $\hat{x}_0 = x_0$. Therefore, the system (33) becomes an observer of (32) if (33) is IES in the second moment with respect to the Euclidean distance uniformly in x , i.e., if there exist $a, \lambda > 0$ such that

$$\mathbb{E}[|\hat{x}_k - \hat{x}'_k|^2] \leq a^2 \lambda^{2k} |\hat{x}_0 - \hat{x}'_0|^2 \quad \forall k \in \mathbb{Z}_{0+}$$

for each $((\hat{x}_0, \hat{x}'_0), x_0) \in (\mathbb{R}^n \times \mathbb{R}^n) \times \mathbb{R}^n$. In fact, for $\hat{x}'_0 = x_0$, IES implies that $\mathbb{E}[|\hat{x}_k - x_k|^2] \rightarrow 0$ as $t \rightarrow \infty$. Note that, in this discussion, the system (32) is not assumed to be IES. IES uniformly in x can be verified by applying Corollary 4.13. In general, Corollary 4.13 consists of an infinite family of linear matrix inequalities (LMIs), which can be relaxed to a finite one if $\partial g(\hat{x}, i)/\partial x$, $i \in \mathcal{M}$, is bounded. If $\partial g(\hat{x}, i)/\partial x$ is bounded, there exist $A_i^{(\ell)} \in \mathbb{R}^{n \times n}$ and $\theta_i^{(\ell)}(\hat{x}) \geq 0$, $i \in \mathcal{M}$, $\ell = 1, \dots, L$, such that

$$\frac{\partial g(\hat{x}, i)}{\partial x} = \sum_{\ell=1}^L \theta_i^{(\ell)}(\hat{x}) A_i^{(\ell)} \quad \text{and} \quad \sum_{\ell=1}^L \theta_i^{(\ell)}(\hat{x}) = 1 \quad (34)$$

for each $\hat{x} \in \mathbb{R}^n$ and every $i \in \mathcal{M}$. Based on this, we have the following corollary without the proof.

Corollary 4.16: There exist H_i , $i \in \mathcal{M}$ such that a system (33) with bounded $\partial g_0(\hat{x}, i)/\partial x$, $i \in \mathcal{M}$, is IES in the second moment with respect to the Euclidean distance uniformly in x if there exist $\lambda \in (0, 1)$, $\hat{P}_{0,i} \in \mathbb{S}_{>0}^{n \times n}$, and \hat{H}_i , $i \in \mathcal{M}$, such that

$$\begin{bmatrix} \lambda^2 \hat{P}_{0,i} & * & * & * \\ \sqrt{\pi_{1,i}} (\hat{P}_{0,1} A_1^{(\ell_1)} + \hat{H}_1 C) & \hat{P}_{0,1} & * & * \\ \vdots & 0 & \ddots & * \\ \sqrt{\pi_{m,i}} (\hat{P}_{0,m} A_m^{(\ell_m)} + \hat{H}_m C) & 0 & 0 & \hat{P}_{0,m} \end{bmatrix} \succeq 0 \quad (35)$$

for all $i \in \mathcal{M}$ and $\ell_1, \dots, \ell_m \in \{1, \dots, L\}$. Moreover, such H_i , $i \in \mathcal{M}$, can be designed as $H_i = \hat{P}_{0,i}^{-1} \hat{H}_i$. \blacksquare

In Corollary 4.16, we specify $\hat{P}_{0,i}$, $i \in \mathcal{M}$, in constants in order to obtain an LMI condition. This approach can be applied to other stochastic processes and may be generalized to polynomial $\hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1})$ with respect to x_{k_0} and $\hat{\xi}_{k_0-1}$ based on the sum-of-squares. In the polynomial case, $\mathbb{E}_0[\hat{P}(k_0, x_{k_0}, \hat{\xi}_{k_0-1})]$ is not a bounded function of x_{k_0} , and thus we need to apply Theorem 3.2 instead of Theorem 3.3.

Example 4.17: We consider the observer design of a network susceptible-infected-susceptible (SIS) model with two populations. Let $\beta > 0$ and $\gamma > 0$ be the infection rate and recovery rate, respectively. Then, population l 's fraction of infected individuals $x_{l,k}$ is described by

$$\begin{cases} x_{1,k+1} = x_{1,k} - \gamma x_{1,k} + \beta \xi_k (1 - x_{1,k}) x_{2,k} \\ x_{2,k+1} = x_{2,k} - \gamma x_{2,k} + \beta \xi_k (1 - x_{2,k}) x_{1,k} \\ y_k = x_{1,k} \end{cases} \quad k \in \mathbb{Z}_{0+}$$

where $\xi_k = i \in \{0, 1\}$ represents an edge-Markovian dynamic graph [50], which represents the existence of the contact between the populations. Since $x_{l,k} = 0$ and $x_{l,k} = 1$, respectively, mean disease free and all infected, $D = [0, 1]^2$ is positively invariant in practice.

We consider a scenario where population 1 estimates the fraction of infected individuals in population 2, i.e., $x_{2,k}$. This is an observer design problem, and we can apply Corollary 4.16. On D , there exist $\theta_i^{(\ell)}(\hat{x}) \geq 0$, $i \in \{0, 1\}$, $\ell = 1, \dots, 4$, such that (34) holds for

$$\begin{aligned} A_1^{(\ell)} &= A := (1 - \gamma)I_2, \ell = 1, \dots, 4 \\ A_2^{(1)} &= A + \alpha\beta \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_2^{(2)} = A + \alpha\beta \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} \\ A_2^{(3)} &= A + \alpha\beta \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}, A_2^{(4)} = A + \alpha\beta \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}. \end{aligned}$$

Then, (35) becomes

$$\begin{bmatrix} \lambda^2 \hat{P}_{0,i} & * & * \\ \sqrt{\pi_{1,i}}(\hat{P}_{0,1}A + \hat{H}_1 C) & \hat{P}_{0,1} & * \\ \sqrt{\pi_{2,i}}(\hat{P}_{0,2}A_2^{(\ell)} + \hat{H}_2 C) & 0 & \hat{P}_{0,2} \end{bmatrix} \succeq 0 \quad i = 1, 2, \ell = 1, \dots, 4 \quad (36)$$

where $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Let $\alpha = 1$, $\beta = 0.01$, $\gamma = 0.001$, $\pi_{1,1} = 1/3$, $\pi_{2,1} = 2/3$, $\pi_{1,2} = 1/4$, $\pi_{2,2} = 3/4$, and $\lambda = \sqrt{0.9}$. Using a set of solutions to (36) with $\hat{P}_{0,i} \succ 0$, $i = 1, 2$, an observer gain at each mode can be designed as

$$H_1 = \hat{P}_{0,1}^{-1} \hat{H}_1 = \begin{bmatrix} -0.940 \\ 0 \end{bmatrix}, H_2 = \hat{P}_{0,2}^{-1} \hat{H}_2 = \begin{bmatrix} 2.02 \\ 1.29 \end{bmatrix}.$$

V. CONCLUSION

In this article, we have studied moment UIES for discrete-time nonlinear stochastic systems in the contraction framework. In particular, we have presented a sufficient condition for UIES in the first moment with respect to the Riemannian metric and

a necessary and sufficient condition for UIES in the second moment with respect to the Euclidean distance. Then, the second moment UIES condition has been tailored to i.i.d. processes and Markov processes as specialized applications. The proposed approach can also be applied to hidden Markov models. Tailoring the proposed UIES conditions to different subclasses of stochastic systems is included in future work.

As mentioned in this article, our results can be understood as a generalization of the results for linear stochastic systems. The results of the linear case have been applied to networked control systems with randomly time-varying communication delays [51]. Furthermore, their practical usefulness has already been validated by remote control experiments of actual vehicles [52]. Due to the linearity assumption, the vehicle velocities have been assumed to be constants. Applying the results of this article, we are interested in remote control experiments of actual vehicles including their velocity control.

From theoretical viewpoints, other possible research directions are to deal with random initial conditions and to study incremental versions of different stability properties, such as almost sure convergence, stochastic convergence, convergence in law, and recurrence; a Lyapunov framework for recurrence analysis is referred to as Foster–Lyapunov analysis [53], [54]. The proposed framework can be expected as a first step for enlarging applications of contraction theory to various problems involving randomness.

APPENDIX A

PROOF OF THEOREM 3.2

Before providing the proof, we proceed with auxiliary analysis for the prolonged system (23). To emphasize that $(x_k)_{k \in \mathbb{Z}_{k_0+}} : \Omega \rightarrow (\mathbb{R}^n)^{\mathbb{Z}_{k_0+}}$ is a stochastic process under the initial condition $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, this is also denoted by $(\phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}))_{k \in \mathbb{Z}_{k_0+}}$ or simply $(\phi_k(\xi^{(k-1)-}))_{k \in \mathbb{Z}_{k_0+}}$. Namely, it follows that

$$x_k = \phi_k(\xi^{(k-1)-}) \quad (37)$$

$$\phi_{k+1}(\xi^{(k)-}) = f_k(\phi_k(\xi^{(k-1)-}), \xi_k), k \in \mathbb{Z}_{k_0+} \quad (38)$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where $\phi_{k_0} = x_{k_0}$.

Then, the solution to the variational system can be described as

$$\delta x_k = \Phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \delta x_{k_0}, k \in \mathbb{Z}_{k_0+} \quad (39)$$

or simply $\delta x_k = \Phi_k(\xi^{(k-1)-}) \delta x_{k_0}$, $k \in \mathbb{Z}_{k_0+}$, for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$, where using the solution $\phi_k(\xi^{(k-1)-})$ to the system (18), Φ_k is defined and computed by

$$\Phi_{k_0} := \frac{\partial \phi_{k_0}}{\partial x_{k_0}} = I_n \quad (40)$$

$$\begin{aligned} &\Phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \\ &:= \frac{\partial f_{k-1}(\phi_{k-1}(\xi^{(k-2)-}), \xi_{k-1})}{\partial \phi_{k-1}} \dots \end{aligned}$$

$$\begin{aligned} & \frac{\partial f_{k_0+1}(\phi_{k_0+1}(\xi_{k_0}), \xi_{k_0+1})}{\partial \phi_{k_0+1}} \frac{\partial f_{k_0}(\phi_{k_0}, \xi_{k_0})}{\partial \phi_{k_0}} \frac{\partial \phi_{k_0}}{\partial x_{k_0}} \\ &= \frac{\partial \phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-})}{\partial x_{k_0}}, k \in \mathbb{Z}_{(k_0+1)+} \end{aligned} \quad (41)$$

where $\phi_{k_0} = x_{k_0}$, (38), and the chain rule is used.

Remark 1.1: Note that ϕ_k , $k \in \mathbb{Z}_{k_0+}$, is a composition function of f_i , $i = k_0, \dots, k-1$. From the assumption for f_k , both $\phi_k(\xi^{(k-1)-})$ and $\delta x_k = \Phi_k(\xi^{(k-1)-})\delta x_{k_0}$ are \mathcal{F}_{k-1} -measurable functions for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. \triangleleft

Remark 1.2: The notations $\phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-})$ and $\Phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-})$ (or $\phi_k(\xi^{(k-1)-})$ and $\Phi_k(\xi^{(k-1)-})$ in shorthand) are used to emphasize that they are stochastic processes. When they are considered as mappings $\phi_k : \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k-1]}} \rightarrow \mathbb{R}^n$ and $\Phi_k : \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k-1]}} \rightarrow \mathbb{R}^{n \times n}$, the notations $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ with η (or ϕ_k and Φ_k in shorthand) are used, respectively. Note that as the mappings ϕ_k and Φ_k satisfy the counterparts of (38), (39), and (41), i.e.,

$$\phi_{k+1}(x_{k_0}, \eta^{k-}; k_0) = f_k(\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0), \eta_k) \quad (42)$$

$$\delta x_k = \Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)\delta x_{k_0}, k \in \mathbb{Z}_{k_0+} \quad (43)$$

$$\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0) = \frac{\partial \phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)}{\partial x_{k_0}} \quad k \in \mathbb{Z}_{(k_0+1)+} \quad (44)$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \eta^{k-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k]}}$, respectively. \triangleleft

Now, we are ready to prove Theorem 3.2.

Proof: (Step 1) The quadratic forms of both sides in (22a) with respect to (the deterministic) $\delta x_{k_0} \in \mathbb{R}^n$ satisfy

$$\begin{aligned} c_1^2 \delta x_{k_0}^\top \tilde{P}(x_{k_0}) \delta x_{k_0} &\leq \delta x_{k_0}^\top \mathbb{E}_0 [P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \delta x_{k_0} \\ &\leq c_2^2 \delta x_{k_0}^\top \tilde{P}(x_{k_0}) \delta x_{k_0} \end{aligned} \quad (45)$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Since $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-})$ is arbitrary in (45), and both $\phi_k(\xi^{(k-1)-})$ and $\delta x_k = \Phi_k(\xi^{(k-1)-})\delta x_{k_0}$ are \mathcal{F}_{k-1} -measurable as mentioned in Remark 1.1, the time-shift $k_0 \mapsto k$, $k \in \mathbb{Z}_{k_0+}$, of the first inequality yields

$$\begin{aligned} & c_1^2 \delta x_k^\top \tilde{P}(\phi_k(\xi^{(k-1)-})) \delta x_k \\ & \leq \delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k \text{ a.s.} \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Taking the conditional expectations $\mathbb{E}_0[\cdot]$ of both sides leads to

$$\begin{aligned} & c_1^2 \mathbb{E}_0 [\delta x_k^\top \tilde{P}(\phi_k(\xi^{(k-1)-})) \delta x_k] \\ & \leq \mathbb{E}_0 [\delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k] \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned} \quad (46)$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$.

(Step 2) The quadratic forms of both sides in (22b) with respect to (the deterministic) $\delta x_{k_0} \in \mathbb{R}^n$ satisfy

$$\begin{aligned} & \mathbb{E}_0 \left[\delta x_{k_0}^\top \frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} \right. \\ & \quad \mathbb{E}_0 [P(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] \\ & \quad \left. \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} \delta x_{k_0} \right] \\ & \leq \lambda^2 \delta x_{k_0}^\top \mathbb{E}_0 [P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \delta x_{k_0} \end{aligned}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. The time-shift $k_0 \mapsto k$, $k \in \mathbb{Z}_{k_0+}$, yields

$$\begin{aligned} & \mathbb{E}_0 \left[\delta x_k^\top \frac{\partial^\top f_k(\phi_k(\xi^{(k-1)-}), \xi_k)}{\partial \phi_k} \right. \\ & \quad \mathbb{E}_0 [P(k + 1, f_k(\phi_k(\xi^{(k-1)-}), \xi_k), S_{k+1} \xi^{(k+1)+}) | \mathcal{F}_k] \\ & \quad \left. \frac{\partial f_k(\phi_k(\xi^{(k-1)-}), \xi_k)}{\partial \phi_k} \delta x_k \middle| \mathcal{F}_{k-1} \right] \\ & \leq \lambda^2 \delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k \text{ a.s.} \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

or equivalently, from (23) and (38),

$$\begin{aligned} & \mathbb{E}_0 [\delta x_{k+1}^\top \mathbb{E}_0 [P(k + 1, \phi_{k+1}(\xi^{k-}), S_{k+1} \xi^{(k+1)+}) | \mathcal{F}_k] \\ & \quad \delta x_{k+1} | \mathcal{F}_{k-1}] \\ & \leq \lambda^2 \delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k \text{ a.s.} \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Recall that $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ for each $\hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$. Then, taking the conditional expectations $\mathbb{E}_0[\cdot]$ of both sides of the abovementioned inequality yields

$$\begin{aligned} & \mathbb{E}_0 [\delta x_{k+1}^\top \mathbb{E}_0 [P(k + 1, \phi_{k+1}(\xi^{k-}), S_{k+1} \xi^{(k+1)+}) | \mathcal{F}_k] \delta x_{k+1}] \\ & \leq \lambda^2 \mathbb{E}_0 [\delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k] \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. A recursive use of this from k to k_0 leads to

$$\begin{aligned} & \mathbb{E}_0 [\delta x_k^\top \mathbb{E}_0 [P(k, \phi_k(\xi^{(k-1)-}), S_k \xi^{k+}) | \mathcal{F}_{k-1}] \delta x_k] \\ & \leq \lambda^{2(k-k_0)} \delta x_{k_0}^\top \mathbb{E}_0 [P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \delta x_{k_0} \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned} \quad (47)$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$.

In summary, the second inequality of (45)–(47) leads to

$$\mathbb{E}_0 [\delta x_k^\top \tilde{P}(\phi_k(\xi^{(k-1)-})) \delta x_k]$$

$$\leq \frac{c_2^2}{c_1^2} \lambda^{2(k-k_0)} \delta x_{k_0}^\top \tilde{P}(x_{k_0}) \delta x_{k_0} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (48)$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Taking the square roots of both sides and applying the Cauchy–Schwarz inequality [47, Corollary 3.1.12] (with $\mathbb{P}(\Omega) = 1$) to the left-hand side yield

$$\begin{aligned} & \mathbb{E}_0 \left[\sqrt{\delta x_k^\top \tilde{P}(\phi_k(\xi^{(k-1)-})) \delta x_k} \right] \\ & \leq \frac{c_2}{c_1} \lambda^{k-k_0} \sqrt{\delta x_{k_0}^\top \tilde{P}(x_{k_0}) \delta x_{k_0}} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (49) \end{aligned}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$.

(Step 3) Here, we consider ϕ_k and Φ_k as the mappings $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$; recall Remark 1.2. For each pair $(x'_{k_0}, x''_{k_0}) \in \mathbb{R}^n \times \mathbb{R}^n$, let $\gamma^* \in \Gamma(x'_{k_0}, x''_{k_0})$ be the geodesic with respect to \tilde{P} , i.e., a path satisfying (21); note that γ^* is independent of k_0 and $\eta^{(k-1)-}$. Let $(x_{k_0}, \delta x_{k_0}) = (\gamma^*(s), d\gamma^*(s)/ds)$, $s \in [0, 1]$, be the initial states of $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$. Then, it follows from (42) and the chain rule that

$$\begin{aligned} & \frac{\partial \phi_{k+1}(\gamma^*(s), \eta^{k-}; k_0)}{\partial s} \\ & = \frac{\partial f_k(\phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0), \eta_k)}{\partial s} \\ & = \frac{\partial f_k(\phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0), \eta_k)}{\partial \phi_k} \frac{\partial \phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0)}{\partial s} \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \quad (50) \end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \eta^{k-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k]}}$ and every $s \in [0, 1]$. This implies that $\partial \phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0) / \partial s$ satisfies (44) for $(x_{k_0}, \delta x_{k_0}) = (\gamma^*(s), d\gamma^*(s)/ds)$, $s \in [0, 1]$. In this case, (43) becomes

$$\begin{aligned} & \frac{\partial \phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0)}{\partial s} \\ & = \Phi_k(\gamma^*(s), \eta^{(k-1)-}; k_0) \frac{d\gamma^*(s)}{ds} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (51) \end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \eta^{(k-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k-1]}}$ and every $s \in [0, 1]$.

(Step 4) Now, we consider stochastic processes. The equality (50) implies that $\partial \phi_k(\xi^{(k-1)-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-}) / \partial s$, $k \in \mathbb{Z}_{k_0+}$, is a solution to the variational system (23) and satisfies the counterpart of (51), i.e.,

$$\begin{aligned} & \frac{\partial \phi_k(\xi^{(k-1)-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-})}{\partial s} \\ & = \Phi_k(\xi^{(k-1)-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-}) \frac{d\gamma^*(s)}{ds} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (52) \end{aligned}$$

under the initial state $(x_{k_0}, \delta x_{k_0}) = (\gamma^*(s), d\gamma^*(s)/ds)$ for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$ and every $s \in [0, 1]$.

Substituting $(x_{k_0}, \delta x_{k_0}) = (\gamma^*(s), d\gamma^*(s)/ds)$, $s \in [0, 1]$, into (49) and applying (52) lead to

$$\begin{aligned} & \mathbb{E}_0 \left[\left(\frac{d^\top \gamma^*(s)}{ds} \Phi_k^\top(\xi^{(k-1)-}) \right. \right. \\ & \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \frac{d\gamma^*(s)}{ds} \right)^{1/2} \right] \\ & \leq \frac{c_2}{c_1} \lambda^{k-k_0} \sqrt{\frac{d^\top \gamma^*(s)}{ds} \tilde{P}(\gamma^*(s)) \frac{d\gamma^*(s)}{ds}} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (53) \end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$ and every $s \in [0, 1]$, where in the left-hand side, the argument $(k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-})$ is dropped from $\phi_k(\xi^{(k-1)-})$ and $\Phi_k(\xi^{(k-1)-})$.

(Step 5) We consider integrating both sides of (53) with respect to s in $[0, 1]$. In (20) and (21), the Riemann integrals are used. For the sake of formality, they need to be replaced with the Lebesgue integrals. To this end, we introduce a measurable space corresponding to s . Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ be the measurable space, where μ is the Lebesgue measure. Note that both $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ and $(\Omega, \mathcal{F}, \mathbb{P})$ are complete and σ -finite. Then, the product measurable space naturally induced by the Cartesian product $\mathbb{R} \times \Omega$, denoted by $(\mathbb{R} \times \Omega, \mathcal{L}, \lambda)$, is complete and σ -finite [47, Th. 5.1.2 and Remark 5.1.2].

To take the Lebesgue integrals for (53), we introduce the following functions:

$$(\bar{\gamma}(s), \overline{\partial \gamma}(s)) := \begin{cases} (\gamma^*(s), d\gamma^*(s)/ds), & s \in [0, 1] \\ (0, 0), & s \in \mathbb{R} \setminus [0, 1]. \end{cases} \quad (54)$$

Since γ^* is of class C^1 on $[0, 1]$, $(\bar{\gamma}, \overline{\partial \gamma})$ is piecewise continuous on \mathbb{R} . In (53), $(\gamma^*(s), d\gamma^*(s)/ds)$ can be replaced with $(\bar{\gamma}(s), \overline{\partial \gamma}(s))$, and the corresponding inequality holds for all $s \in \mathbb{R}$ instead of $s \in [0, 1]$. That is, we have

$$\begin{aligned} & \mathbb{E}_0 \left[\left(\overline{\partial \gamma}^\top(s) \Phi_k^\top(\xi^{(k-1)-}) \right. \right. \\ & \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \overline{\partial \gamma}(s) \right)^{1/2} \right] \\ & \leq \frac{c_2}{c_1} \lambda^{k-k_0} \sqrt{\overline{\partial \gamma}^\top(s) \tilde{P}(\bar{\gamma}(s)) \overline{\partial \gamma}(s)} \quad \forall k \in \mathbb{Z}_{k_0+} \quad (55) \end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$ and every $s \in \mathbb{R}$, where in the left-hand side, the argument $(k_0, \bar{\gamma}(s), \hat{\xi}^{(k_0-1)-})$ is dropped from $\phi_k(\xi^{(k-1)-})$ and $\Phi_k(\xi^{(k-1)-})$.

Now, we consider the left-hand side of (55). A piecewise continuous function is measurable, and the composition of measurable functions is again measurable [47, Prop. 2.1.1]. Then, according to Remark 1.1 and the continuity of $f_k(x, \eta)$ and $\partial f_k(x, \eta) / \partial x$ with respect to $x \in \mathbb{R}^n$ at each $(k, \eta) \in \mathbb{Z} \times \mathbb{R}^m$, the function $(\overline{\partial \gamma}^\top(s) \Phi_k^\top(\xi^{(k-1)-}) \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \overline{\partial \gamma}(s))^{1/2}$, $k \in \mathbb{Z}_{k_0+}$, is \mathcal{L} -measurable at each $(k_0, (x_{k_0}, x'_{k_0})) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n)$.

Taking the μ -integrations for both sides of (55) yield

$$\begin{aligned}
& \int_{\mathbb{R}} \mathbb{E}_0 \left[\left(\overline{\partial\gamma}^\top(s) \Phi_k^\top(\xi^{(k-1)-}) \right. \right. \\
& \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \overline{\partial\gamma}(s) \right)^{1/2} \right] d\mu \\
& \leq \frac{c_2}{c_1} \lambda^{k-k_0} \int_{\mathbb{R}} \sqrt{\overline{\partial\gamma}^\top(s) \tilde{P}(\overline{\gamma}(s)) \overline{\partial\gamma}(s)} d\mu \\
& = \frac{c_2}{c_1} \lambda^{k-k_0} \int_0^1 \sqrt{\frac{d^\top \gamma^*(s)}{ds} \tilde{P}(\gamma^*(s)) \frac{d\gamma^*(s)}{ds}} ds \\
& = \frac{c_2(1+\varepsilon)}{c_1} \lambda^{k-k_0} d_{\tilde{P}}(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+} \quad (56)
\end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$, where the first equality follows from (54) and the fact that the Lebesgue and Riemann integrals coincide with each other when $\sqrt{\frac{d^\top \gamma^*(s)}{ds} \tilde{P}(\gamma^*(s)) \frac{d\gamma^*(s)}{ds}}$ is bounded and Riemann integrable with respect to s on $[0, 1]$ at each $(x'_{k_0}, x''_{k_0}) \in \mathbb{R}^n \times \mathbb{R}^n$ (see, e.g., [47, Th. 2.4.1]); the last equality follows from (21).

In (56), the most right-hand side is bounded for each $(k_0, (x'_{k_0}, x''_{k_0})) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n)$. This implies that the most left-hand side is μ -integrable at each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Therefore, from the Fubini-Tonelli theorem [47, Sec. 5.2], the order of the integrals in the most left-hand side is commutative; recall that $(\mathbb{R} \times \Omega, \mathcal{L}, \lambda)$ is complete and σ -finite. Namely, it follows that

$$\begin{aligned}
& \int_{\mathbb{R}} \mathbb{E}_0 \left[\left(\overline{\partial\gamma}^\top(s) \Phi_k^\top(\xi^{(k-1)-}) \right. \right. \\
& \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \overline{\partial\gamma}(s) \right)^{1/2} \right] d\mu \\
& = \mathbb{E}_0 \left[\int_{\mathbb{R}} \left(\overline{\partial\gamma}^\top(s) \Phi_k^\top(\xi^{(k-1)-}) \right. \right. \\
& \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \Phi_k(\xi^{(k-1)-}) \overline{\partial\gamma}(s) \right)^{1/2} d\mu \right] \\
& = \mathbb{E}_0 \left[\int_0^1 \left(\frac{\partial^\top \phi_k(\xi^{(k-1)-})}{\partial s} \right. \right. \\
& \quad \left. \left. \tilde{P}(\phi_k(\xi^{(k-1)-})) \frac{\partial^\top \phi_k(\xi^{(k-1)-})}{\partial s} \right)^{1/2} ds \right] \\
& \geq \mathbb{E}_0 [d_{\tilde{P}}(\phi_k(\xi^{(k-1)-}; k_0, x'_{k_0}, \hat{\xi}^{(k_0-1)-}), \\
& \quad \phi_k(\xi^{(k-1)-}; k_0, x''_{k_0}, \hat{\xi}^{(k_0-1)-}))] \quad \forall k \in \mathbb{Z}_{k_0+} \quad (57)
\end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$, where the first equality follows from the Fubini-Tonelli theorem; the second one follows from (52), (54), and the fact that the Lebesgue and Riemann integrals coincide with each other by a similar reasoning as mentioned for (56); the last inequality follows from (20) and the fact that $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-})$

is a path connecting $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(0), \hat{\xi}^{(k_0-1)-})$, $\gamma^*(0) = x'_{k_0}$, to $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(1), \hat{\xi}^{(k_0-1)-})$, $\gamma^*(1) = x''_{k_0}$.

From (56) and (57), we obtain

$$\begin{aligned}
& \mathbb{E}_0 [d_{\tilde{P}}(\phi_k(\xi^{(k-1)-}; k_0, x'_{k_0}, \hat{\xi}^{(k_0-1)-}), \\
& \quad \phi_k(\xi^{(k-1)-}; k_0, x''_{k_0}, \hat{\xi}^{(k_0-1)-}))] \\
& \leq \frac{c_2(1+\varepsilon)}{c_1} \lambda^{k-k_0} d_{\tilde{P}}(x'_{k_0}, x''_{k_0}) \quad \forall k \in \mathbb{Z}_{k_0+}
\end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0})) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n)$. This implies that the system is UIES in the first moment. \blacksquare

APPENDIX B PROOF OF THEOREM 3.3

When $\tilde{P} = I$, solving the corresponding Euler-Lagrange equation [48, Eq. (5.3.2)] gives

$$|x' - x''|^2 = \inf_{\gamma \in \Gamma(x', x'')} \int_0^1 \left| \frac{d\gamma(s)}{ds} \right|^2 ds \quad (58)$$

and the geodesic is the line segment $\gamma^*(s) = (1-s)x'_{k_0} + sx''_{k_0}$. That is, when $\tilde{P} = I$, we can directly use (48) for the sufficiency proof, but this is not true for general \tilde{P} . Utilizing (58), we prove Theorem 3.3 in the following.

Proof: (Sufficiency) If \tilde{P} is identity, (48) reduces to

$$\mathbb{E}_0 [|\delta x_k|^2] \leq \frac{c_2^2}{c_1^2} \lambda^{2(k-k_0)} |\delta x_{k_0}|^2 \quad \forall k \in \mathbb{Z}_{k_0+}$$

for each $(k_0, (x_{k_0}, \delta x_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Substituting $(x_{k_0}, \delta x_{k_0}) = (\gamma^*(s), d\gamma^*(s)/ds)$ with $\gamma^*(s) = (1-s)x'_{k_0} + sx''_{k_0}$ (and consequently $d\gamma^*(s)/ds = x''_{k_0} - x'_{k_0}$) into this and taking the μ -integration as in the proof of Theorem 3.2 yield

$$\begin{aligned}
& \mathbb{E}_0 \left[\int_0^1 \left| \frac{\partial \phi_k(\xi^{k-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-})}{\partial s} \right|^2 ds \right] \\
& = \mathbb{E}_0 \left[\int_{\mathbb{R}} \left| \frac{\partial \phi_k(\xi^{k-}; k_0, \overline{\gamma}(s), \hat{\xi}^{(k_0-1)-})}{\partial s} \right|^2 d\mu \right] \\
& = \int_{\mathbb{R}} \mathbb{E}_0 \left[\left| \frac{\partial \phi_k(\xi^{k-}; k_0, \overline{\gamma}(s), \hat{\xi}^{(k_0-1)-})}{\partial s} \right|^2 \right] d\mu \\
& \leq \frac{c_2^2}{c_1^2} \lambda^{2(k-k_0)} |x'_{k_0} - x''_{k_0}|^2 \quad \forall k \in \mathbb{Z}_{k_0+}
\end{aligned}$$

for each $(k_0, (x'_{k_0}, x''_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$, where $\overline{\gamma}$ is defined in (54). From (58) and the fact that $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(s), \hat{\xi}^{(k_0-1)-})$ is a path connecting $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(0), \hat{\xi}^{(k_0-1)-})$, $\gamma^*(0) = x'_{k_0}$ to $\phi_k(\xi^{(k-1)-}; k_0, \gamma^*(1), \hat{\xi}^{(k_0-1)-})$, $\gamma^*(1) = x''_{k_0}$, the system is UIES in the second moment with respect to the Euclidean distance.

(Necessity) (Step 1) In fact, (52) holds for an arbitrary path $\gamma \in \Gamma(x_{k_0}, x'_{k_0})$. As for the sufficiency proof,

we choose $\gamma(s) = (1-s)x_{k_0} + sx'_{k_0}$ (and, thus, $d\gamma(s)/ds = x'_{k_0} - x_{k_0}$). Then, it follows from fundamental theorem of calculus, (37), and (52) that

$$\begin{aligned} & x'_k - x_k \\ &= \phi_k(\xi^{(k-1)-}; k_0, x'_{k_0}, \hat{\xi}^{(k_0-1)-}) \\ &\quad - \phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \\ &= \int_0^1 \frac{\partial \phi_k(\xi^{(k-1)-}; k_0, (1-s)x_{k_0} + sx'_{k_0}, \hat{\xi}^{(k_0-1)-})}{\partial s} ds \\ &= \int_0^1 \Phi_k(\xi^{(k-1)-}; k_0, (1-s)x_{k_0} + sx'_{k_0}, \hat{\xi}^{(k_0-1)-}) \\ &\quad (x'_{k_0} - x_{k_0}) ds \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, x'_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Substituting this into the definition (19) of the UIES in the second moment with respect to the Euclidean distance yields

$$\begin{aligned} & \mathbb{E}_0 \left[\left| \int_0^1 \Phi_k(\xi^{(k-1)-}; k_0, (1-s)x_{k_0} + sx'_{k_0}, \hat{\xi}^{(k_0-1)-}) \right. \right. \\ & \quad \left. \left. (x'_{k_0} - x_{k_0}) ds \right|^2 \right] \leq a^2 \lambda^{2(k-k_0)} |x'_{k_0} - x_{k_0}|^2 \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, x'_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Since $x'_{k_0} \in \mathbb{R}^n$ is arbitrary, we choose $x'_{k_0} = x_{k_0} + hv$ with $h \in \mathbb{R}$ and $v \in \mathbb{R}^n$. Substituting this yields

$$\begin{aligned} & \mathbb{E}_0 \left[\left| \int_0^1 \Phi_k(\xi^{(k-1)-}; k_0, x_{k_0} + shv, \hat{\xi}^{(k_0-1)-}) v ds \right|^2 \right] \\ & \leq a^2 \lambda^{2(k-k_0)} |v|^2 \quad \forall k \in \mathbb{Z}_{k_0+}. \end{aligned}$$

The change of the variables $\bar{s} = sh$ leads to

$$\begin{aligned} & \mathbb{E}_0 \left[\left| \frac{1}{h} \int_0^h \Phi_k(\xi^{(k-1)-}; k_0, x_{k_0} + \bar{s}v, \hat{\xi}^{(k_0-1)-}) v d\bar{s} \right|^2 \right] \\ & \leq a^2 \lambda^{2(k-k_0)} |v|^2 \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, (x_{k_0}, x'_{k_0}), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Note that this holds for an arbitrary $h \in \mathbb{R}$, which implies

$$\begin{aligned} & \liminf_{h \rightarrow 0} \mathbb{E}_0 \left[\left| \frac{1}{h} \int_0^h \Phi_k(\xi^{(k-1)-}; k_0, x_{k_0} + \bar{s}v, \hat{\xi}^{(k_0-1)-}) v d\bar{s} \right|^2 \right] \\ & \leq a^2 \lambda^{2(k-k_0)} |v|^2 \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned} \quad (59)$$

for each $(k_0, (x_{k_0}, v), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$. Applying Fatou's lemma [47, Th. 2.3.7] to the left-hand side yields

$$\mathbb{E}_0 \left[\liminf_{h \rightarrow 0} \left| \frac{1}{h} \int_0^h \Phi_k(\xi^{(k-1)-}) v d\bar{s} \right|^2 \right]$$

$$\leq \liminf_{h \rightarrow 0} \mathbb{E}_0 \left[\left| \frac{1}{h} \int_0^h \Phi_k(\xi^{(k-1)-}) v d\bar{s} \right|^2 \right] \quad \forall k \in \mathbb{Z}_{k_0+} \quad (60)$$

for each $(k_0, (x_{k_0}, v), \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times \hat{\Xi}^{(k_0-1)-}$.

We consider the left-hand side of (60). Here, we take ϕ_k and Φ_k as the mappings $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$; recall Remark 1.2. Applying product and sum rules of the limit and fundamental theorem of calculus in order lead to

$$\begin{aligned} & \lim_{h \rightarrow 0} \left| \frac{1}{h} \int_0^h \Phi_k(x_{k_0} + \bar{s}v, \eta^{(k-1)-}; k_0) v d\bar{s} \right|^2 \\ &= \left| \lim_{h \rightarrow 0} \frac{1}{h} \int_0^h \Phi_k(x_{k_0} + \bar{s}v, \eta^{(k-1)-}; k_0) v d\bar{s} \right|^2 \\ &= \left| \Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0) v \right|^2 \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned} \quad (61)$$

for each $(k_0, (x_{k_0}, v), \eta^{(k-1)-}) \in \mathbb{Z} \times (\mathbb{R}^n \times \mathbb{R}^n) \times (\mathbb{R}^m)^{\mathbb{Z}_{[k_0, k-1]}}$. These equalities imply the existence of the limit in the most left-hand side. Therefore, the limit inferior of the left-hand side of (60) is equivalent to the limit. Combining (59)–(61) leads to, for stochastic processes,

$$\begin{aligned} & \mathbb{E}_0 \left[\left| \Phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) v \right|^2 \right] \leq a^2 \lambda^{2(k-k_0)} |v|^2 \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$ and $v \in \mathbb{R}^n$. Since $v \in \mathbb{R}^n$ is arbitrary, it holds that

$$\begin{aligned} & \mathbb{E}_0 \left[\sigma^2 \left(\Phi_k(\xi^{(k-1)-}; k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \right) \right] \leq a^2 \lambda^{2(k-k_0)} \\ & \quad \forall k \in \mathbb{Z}_{k_0+} \end{aligned} \quad (62)$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where $\sigma(\cdot)$ denotes the largest singular value (more precisely, valued function).

(Step 2) Here, we again consider ϕ_k and Φ_k as the mappings $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$; recall Remark 1.2. Let us take λ_1 such that $\lambda < \lambda_1 < 1$, and define the K -dependent matrix-valued mapping $P_K : \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{k_0+}} \rightarrow \mathbb{S}_{>0}^{n \times n}$, $K \in \mathbb{Z}_{k_0+}$, such that

$$\begin{aligned} & P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+}) \\ &:= \frac{1}{\lambda_1^2} \sum_{k=k_0}^K \frac{1}{\lambda_1^{2(k-k_0)}} \Phi_k^\top(x_{k_0}, \eta^{(k-1)-}; k_0) \\ & \quad \Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0), \quad K \in \mathbb{Z}_{k_0+} \end{aligned} \quad (63)$$

for each $(k_0, x_{k_0}, \eta^{k_0+}) \in \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{k_0+}}$. Note that $(P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+}))_{K \in \mathbb{Z}_{k_0+}}$ is an increasing sequence with respect to the relation \preceq for each $(k_0, x_{k_0}, \eta^{k_0+}) \in \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{k_0+}}$, and thus, $P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+})$ has the (pointwise convergence) limit

$$P(k_0, x_{k_0}, S_{k_0} \eta^{k_0+})$$

$$:= \lim_{K \rightarrow \infty} P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+}) \in \bar{\mathbb{R}}^{n \times n} \quad (64)$$

for each $(k_0, x_{k_0}, \eta^{k_0+}) \in \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{k_0+}}$.

Now, we consider stochastic processes. According to Remark 1.1, $\phi_k(\xi^{(k-1)-})$, $k \in \mathbb{Z}_{k_0+}$, and $\Phi_k(\xi^{(k-1)-})$, $k \in \mathbb{Z}_{k_0+}$, are both \mathcal{F}_{k-1} -measurable for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, and thus, $P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})$, $K \in \mathbb{Z}_{k_0+}$, is \mathcal{F} -measurable for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. Since the limit of a sequence of measurable functions is again measurable [47, Prop. 2.1.5], $P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})$ is \mathcal{F} -measurable for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. Therefore, the monotone convergence theorem [47, Th. 2.3.4] is applicable for $P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})$, namely

$$\begin{aligned} & \mathbb{E}_0 [P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \\ &= \lim_{K \rightarrow \infty} \mathbb{E}_0 [P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \end{aligned} \quad (65)$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. Note that these conditional expectations are not necessarily to be finite.

(Step 3) From (40) and (63), it follows that

$$\frac{1}{\lambda_1^2} I_n \preceq P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+}) \text{ a.s. } \forall K \in \mathbb{Z}_{k_0+}$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. This inequality is preserved under taking the conditional expectations $\mathbb{E}_0[\cdot]$ of both sides. From (62) and (63), the conditional expectation $\mathbb{E}_0[P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})]$ is also upper bounded. That is, it holds that

$$c_1^2 I_n \leq \mathbb{E}_0[P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \leq c_2^2 I_n \quad \forall K \in \mathbb{Z}_{k_0+} \quad (66)$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$, where $c_1 := 1/\lambda_1$ and

$$c_2 := \frac{a}{\lambda_1} \lim_{K \rightarrow \infty} \sum_{k=k_0}^K \left(\frac{\lambda}{\lambda_1} \right)^{k-k_0} > 0.$$

Note that c_2 is a k_0 -independent positive constant because of $\lambda < \lambda_1$. Since (66) holds for an arbitrary $K \in \mathbb{Z}_{k_0+}$, taking $K \rightarrow \infty$ and using (65) yield

$$c_1^2 I_n \leq \mathbb{E}_0 [P(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \leq c_2^2 I_n \quad \forall K \in \mathbb{Z}_{k_0+}$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. This is nothing but (22c).

(Step 4) Here, we again consider ϕ_k and Φ_k as the mappings $\phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$ and $\Phi_k(x_{k_0}, \eta^{(k-1)-}; k_0)$; recall Remark 1.2. From (63), $P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+})$ satisfies

$$\begin{aligned} & \lambda_1^2 P_K(k_0, x_{k_0}, S_{k_0} \eta^{k_0+}) \\ & - \frac{\partial^\top f_{k_0}(x_{k_0}, \eta_{k_0})}{\partial x_{k_0}} \\ & P_K(k_0 + 1, f_{k_0}(x_{k_0}, \eta_{k_0}), S_{k_0+1} \eta^{(k_0+1)+}; \eta_{k_0}) \\ & \frac{\partial f_{k_0}(x_{k_0}, \eta_{k_0})}{\partial x_{k_0}} = I_n \quad \forall K \in \mathbb{Z}_{k_0+} \end{aligned} \quad (67)$$

for each $(k_0, x_{k_0}, \eta^{k_0+}) \in \mathbb{Z} \times \mathbb{R}^n \times (\mathbb{R}^m)^{\mathbb{Z}_{k_0+}}$, where the equality

$$\Phi_{k_0+1}(x_{k_0}, \eta_{k_0}; k_0) = \frac{\partial f_{k_0}(x_{k_0}, \eta_{k_0})}{\partial x_{k_0}}$$

following from (42), (44), and $\phi_{k_0} = x_{k_0}$ is used. Note that P_K in the second term of the left-hand side of (67) depends on η_{k_0} .

From (67), the corresponding stochastic processes satisfy

$$\begin{aligned} & \frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} P_K(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) \\ & \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} \preceq \lambda_1^2 P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+}) \text{ a.s. } \forall K \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. Taking the conditional expectations of $\mathbb{E}_0[\cdot | \mathcal{F}_{k_0}]$ and $\mathbb{E}_0[\cdot]$ in order for both sides and using the fact that $(\mathcal{F}_k)_{k \in \mathbb{Z}_{k_0+}}$ is a filtration on $(\Omega, \mathcal{F}, \mathbb{P})$ for each $\hat{\xi}^{(k_0-1)-} \in \hat{\Xi}^{(k_0-1)-}$ for the right-hand side, it follows that

$$\begin{aligned} & E_0 \left[\frac{\partial^\top f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} \right. \\ & \left. \mathbb{E}_0 [P_K(k_0 + 1, f_{k_0}(x_{k_0}, \xi_{k_0}), S_{k_0+1} \xi^{(k_0+1)+}) | \mathcal{F}_{k_0}] \right. \\ & \left. \frac{\partial f_{k_0}(x_{k_0}, \xi_{k_0})}{\partial x_{k_0}} \right] \preceq \lambda_1^2 \mathbb{E}_0 [P_K(k_0, x_{k_0}, S_{k_0} \xi^{k_0+})] \\ & \quad \forall K \in \mathbb{Z}_{k_0+} \end{aligned}$$

for each $(k_0, x_{k_0}, \hat{\xi}^{(k_0-1)-}) \in \mathbb{Z} \times \mathbb{R}^n \times \hat{\Xi}^{(k_0-1)-}$. Finally, taking $K \rightarrow \infty$ with (65) yields (22b). ■

REFERENCES

- [1] W. Lohmiller and J.-J. E. Slotine, "On contraction analysis for non-linear systems," *Automatica*, vol. 34, no. 6, pp. 683–696, 1998.
- [2] H. K. Khalil, *Nonlinear Systems*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.
- [3] W. M. Haddad and V. Chellaboina, *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton, NJ, USA: Princeton Univ. Press, 2008.
- [4] D. Angeli, "A Lyapunov approach to incremental stability properties," *IEEE Trans. Autom. Control*, vol. 47, no. 3, pp. 410–421, Mar. 2002.
- [5] F. Forni and R. Sepulchre, "A differential Lyapunov framework for contraction analysis," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 614–628, Mar. 2014.
- [6] R. Reyes-Báez, A. J. van der Schaft, and B. Jayawardhana, "Tracking control of fully-actuated port-hamiltonian mechanical systems via sliding manifolds and contraction analysis," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8256–8261, 2017.
- [7] R. Reyes-Báez, A. van der Schaft, B. Jayawardhana, and L. Pan, "A family of virtual contraction based controllers for tracking of flexible-joints port-Hamiltonian robots: Theory and experiments," *Int. J. Robust Nonlinear Control*, vol. 30, no. 8, pp. 3269–3295, 2020.
- [8] I. R. Manchester and J.-J. E. Slotine, "Control contraction metrics: Convex and intrinsic criteria for nonlinear feedback design," *IEEE Trans. Autom. Control*, vol. 62, no. 6, pp. 3046–3053, Jun. 2017.
- [9] Y. Kawano and T. Ohtsuka, "Nonlinear eigenvalue approach to differential Riccati equations for contraction analysis," *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6497–6504, Dec. 2017.
- [10] Y. Kawano and K. Kashima, "An LMI framework for contraction-based nonlinear control design by derivatives of Gaussian process regression," *Automatica*, vol. 151, 2023, Art. no. 110928.
- [11] S. Bonnabel and J.-J. E. Slotine, "A contraction theory-based analysis of the stability of the deterministic extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 60, no. 2, pp. 565–569, Feb. 2015.

- [12] V. Andrieu, B. Jayawardhana, and L. Praly, "Transverse exponential stability and applications," *IEEE Trans. Autom. Control*, vol. 61, no. 11, pp. 3396–3411, Nov. 2016.
- [13] A. P. Dani, S.-J. Chung, and S. Hutchinson, "Observer design for stochastic nonlinear systems via contraction-based incremental stability," *IEEE Trans. Autom. Control*, vol. 60, no. 3, pp. 700–714, Mar. 2014.
- [14] A. J. van der Schaft, "On differential passivity," *IFAC-PapersOnLine*, vol. 46, no. 23, pp. 21–25, 2013.
- [15] F. Forni and R. Sepulchre, "On differentially dissipative dynamical systems," *IFAC-PapersOnLine*, vol. 46, no. 23, pp. 15–20, 2013.
- [16] Y. Kawano, C. K. Kosaraju, and J. M. A. Scherpen, "Krasovskii and shifted passivity based control," *IEEE Trans. Autom. Control*, vol. 66, no. 10, pp. 4926–4932, Oct. 2021.
- [17] Y. Kawano, M. Cucuzzella, S. Feng, and J. M. A. Scherpen, "Krasovskii and shifted passivity based output consensus," 2022, *arXiv:2207.01430*.
- [18] Y. Kawano and J. M. A. Scherpen, "Model reduction by differential balancing based on nonlinear Hankel operators," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3293–3308, Jul. 2017.
- [19] Y. Kawano, "Controller reduction for nonlinear systems by generalized differential balancing," *IEEE Trans. Autom. Control*, vol. 67, no. 11, pp. 5856–5871, Nov. 2022.
- [20] F. Forni and R. Sepulchre, "Differentially positive systems," *IEEE Trans. Autom. Control*, vol. 61, no. 2, pp. 346–359, Feb. 2016.
- [21] Y. Kawano, B. Besselink, and M. Cao, "Contraction analysis of monotone systems via separable functions," *IEEE Trans. Autom. Control*, vol. 65, no. 8, pp. 3486–3501, Aug. 2020.
- [22] Y. Kawano and F. Forni, " \mathcal{K} -monotonicity and feedback synthesis for incrementally stable networks," 2022, *arXiv:2207.09828*.
- [23] Y. Kawano and M. Cao, "Contraction analysis of virtually positive systems," *Syst. Control Lett.*, vol. 168, 2022, Art. no. 105358.
- [24] M. Di Bernardo, D. Liuzza, and G. Russo, "Contraction analysis for a class of nondifferentiable systems with applications to stability and network synchronization," *SIAM J. Control Optim.*, vol. 52, no. 5, pp. 3203–3227, 2014.
- [25] D. Fiore, S. J. Hogan, and M. Di Bernardo, "Contraction analysis of switched systems via regularization," *Automatica*, vol. 73, pp. 279–288, 2016.
- [26] Q.-C. Pham, N. Tabareau, and J.-J. Slotine, "A contraction theory approach to stochastic incremental stability," *IEEE Trans. Autom. Control*, vol. 54, no. 4, pp. 816–820, Apr. 2009.
- [27] H. Tsukamoto, S.-J. Chung, and J.-J. E. Slotine, "Neural stochastic contraction metrics for learning-based control and estimation," *IEEE Contr. Syst. Lett.*, vol. 5, no. 5, pp. 1825–1830, Nov. 2021.
- [28] B. Liu, B. Xu, and T. Liu, "Almost sure contraction for stochastic switched impulsive systems," *IEEE Trans. Autom. Control*, vol. 66, no. 11, pp. 5393–5400, Nov. 2021.
- [29] D. N. Tran, B. S. Rüffer, and C. M. Kellett, "Convergence properties for discrete-time nonlinear systems," *IEEE Trans. Autom. Control*, vol. 63, no. 8, pp. 3415–3422, Aug. 2018.
- [30] Q.-C. Pham and J.-J. Slotine, "Stochastic contraction in Riemannian metrics," 2013, *arXiv:1304.0340*.
- [31] H. Tsukamoto and S.-J. Chung, "Convex optimization-based controller design for stochastic nonlinear systems using contraction analysis," in *Proc. IEEE 58th Conf. Decis. Control*, 2019, pp. 8196–8203.
- [32] W. L. De Koning, "Compensability and optimal compensation of systems with white parameters," *IEEE Trans. Autom. Control*, vol. 37, no. 5, pp. 579–588, May 1992.
- [33] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-Time Markov Jump Linear Systems*. London, U.K.: Springer-Verlag, 2005.
- [34] J.-W. Lee and G. E. Dullerud, "Uniform stabilization of discrete-time switched and Markovian jump linear systems," *Automatica*, vol. 42, no. 2, pp. 205–218, 2006.
- [35] C. A. C. Gonzaga and O. L. V. Costa, "Stochastic stability for discrete-time Markov jump Lur'e systems," in *Proc. IEEE 52nd Conf. Decis. Control*, 2013, pp. 5993–5998.
- [36] X. Zhong, H. He, H. Zhang, and Z. Wang, "Optimal control for unknown discrete-time nonlinear Markov jump systems using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 12, pp. 2141–2155, Dec. 2014.
- [37] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 2951–2959.
- [38] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
- [39] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. De Freitas, "Taking the human out of the loop: A review of Bayesian optimization," *Proc. IEEE Proc. IRE*, vol. 104, no. 1, pp. 148–175, Jan. 2016.
- [40] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [41] Y. Hosoe and T. Hagiwara, "On second-moment stability of discrete-time linear systems with general stochastic dynamics," *IEEE Trans. Autom. Control*, vol. 67, no. 2, pp. 795–809, Feb. 2022.
- [42] M. Ogura and C. Martin, "Generalized joint spectral radius and stability of switching systems," *Linear Algebra Appl.*, vol. 439, no. 8, pp. 2222–2239, 2013.
- [43] Y. Hosoe and T. Hagiwara, "Equivalent stability notions, Lyapunov inequality, and its application in discrete-time linear systems with stochastic dynamics determined by an i.i.d. process," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4764–4771, Nov. 2019.
- [44] O. L. V. Costa and D. Z. Figueiredo, "Stochastic stability of jump discrete-time linear systems with Markov chain in a general Borel space," *IEEE Trans. Autom. Control*, vol. 59, no. 1, pp. 223–227, Jan. 2014.
- [45] F. Kozin, "A survey of stability of stochastic systems," *Automatica*, vol. 5, no. 1, pp. 95–112, 1969.
- [46] Y. Fang, K. A. Loparo, and X. Feng, "Almost sure and δ moment stability of jump linear systems," *Int. J. Control*, vol. 59, no. 5, pp. 1281–1307, 1994.
- [47] K. B. Athreya and S. N. Lahiri, *Measure Theory and Probability Theory*. New York, NY, USA: Springer, 2006.
- [48] D. Bao, S. S. Chern, and Z. Shen, *An Introduction to Riemann-Finsler Geometry*. New York, NY, USA: Springer-Verlag, 2012.
- [49] D. Burago, Y. Burago, and S. Ivanov, *A Course in Metric Geometry*. Providence, Rhode Island, USA: Amer. Math. Soc., 2001.
- [50] L. Zino and M. Cao, "Analysis, prediction, and control of epidemics: A survey from scalar to dynamic network models," *IEEE Circuits Syst. Mag.*, vol. 21, no. 4, pp. 4–23, Oct.–Dec. 2021.
- [51] Y. Hosoe, "Stochastic aperiodic control of networked systems with i.i.d. time-varying communication delays," in *Proc. IEEE 61st Conf. Decis. Control*, 2022, pp. 3562–3567.
- [52] S. Kameoka and Y. Hosoe, "Remote control of vehicles in a random communication delay environment and experimental results," in *Proc. 10th IFAC Symp. Robust Control Des.*, 2022, pp. 214–217.
- [53] A. R. Teel, A. Subbaraman, and A. Sferlazza, "Stability analysis for stochastic hybrid systems: A survey," *Automatica*, vol. 50, no. 10, pp. 2435–2456, 2014.
- [54] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*. Berlin, Germany: Springer, 2012.



Yu Kawano (Member, IEEE) received the M.S. and Ph.D. degrees in engineering from Osaka University, Suita, Japan, in 2011 and 2013, respectively.

Since 2019, he has been an Associate Professor with the Graduate School of Advanced Science and Engineering, Hiroshima University, Higashi-Hiroshima, Japan. As a Postdoctoral Researcher, he then joined Kyoto University, Kyoto, Japan, and moved to the University of Groningen, Groningen, The Netherlands in

2016. He has held Visiting Research Positions with the Tallinn University of Technology, Tallinn, Estonia, University of Groningen, and University of Pavia, Pavia, Italy. His research interests include nonlinear systems, complex networks, model reduction, and privacy of control systems.

Dr. Kawano is an Associate Editor for *Systems and Control Letters* and a Member of the EUCA Conference Editorial Board.



Yohei Hosoe (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in electrical engineering from Kyoto University, Kyoto, Japan, in 2009, 2011, and 2013, respectively.

Since 2013, he has been with the Department of Electrical Engineering, Kyoto University, where, since 2020, he has been a Junior Associate Professor. From 2018 to 2019, he also held a Visiting Scalar Position with LAAS CNRS, Toulouse, France. His research focuses on theory and applications of stochastic control and

robust control.